An Efficient Deep Learning Approach To Detect Multiple Neurodegenerative Diseases Using Image Data

Shadman Rahman Sameen
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
shadman.rahman.sameen@g.bracu.ac.bd

Shahabuddin Ahmed Khan
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
shahabuddin.ahmed.khan@g.bracu.ac.bd

Mir Md. Muktasif Sakib
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
mir.muktasif.sakib@g.bracu.ac.bd

Tushar Chowdhury
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
tushar.chowdhury@g.bracu.ac.bd

Md. Ashraful Alam, PhD
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
ashraful.alam@bracu.ac.bd

***Abstract*—In today's time early and accurate diagnosis of the neurodegenerative diseases (such as Alzheimer Disease (AD), Parkinson Disease (PD), Frontotemporal Dementia (FTD)) is essential to timely intervention and is quite laborious since the Magnetic Resonance Imaging (MRI) scans of all of them carry same set of characteristics. The deep learning pipeline proposed in this study will utilize 2D sagittal sectioning of 3D MRI columns and will be focusing on such crucial regions as the hippocampus, substantia nigra, and fortal lobe. These pictures were downsized to the identical standard of 224x224x3 and heavily augmented to promote generalization. We benchmarked the transfer learning classification performance of a number of pre-learning CNN models. Among the standard models, MobileNetV2 exhibited superior performance on the test set (94.36%), as opposed to EfficientNetB0 (85.47%), ResNet50 (68.29%), DenseNet121 (62.65%), and VGG19 (33.33%). However, MobileNetV2 exhibited a little overfitting. To address the challenges of accuracy problem of multiclass disease detection and overfitting issue we proposed SadNetV1 that was constructed by integrating MobileNetV2, Squeeze-and-Excitation (SE) blocks and Spatial Attention mechanism to facilitate effective dimensionality reduction in capturing channel and spatial dependence. The proposed SadNetV1 model showed improved performance of 96.15% test, 96.84% train, and 97.11% validation accuracy and can offer generalization in complex MRI images and could differentiate between AD and PD, and FTD. These results give credence to the potential of attention-augmented lightweight networks in effective categorization of neurodegenerative diseases in clinics.**

Keywords— Deep Learning, Computer Vision,Neurodegenerative disease, Alzheimer, Parkinson, Fronto Temporal Dementia, MRI, Attention Mechanism, MobileNetV2, SadNetV1, Medical Imaging.

I. Introduction

A. Background Information

The problem of neurodegenerative disease conditions, including AD, PD, and FTD is one of the most complicated and incapacitating disorders in aging persons. Such diseases can be defined as progressive neural impairments of functionality and structure, which causes cognitive, motor and behavioral disorders. It is really important to make early and accurate diagnosis, however, similarity in clinical manifestations of the investigated diseases makes their differential diagnosis rather challenging and not all these pathologies can be distinguished within their initial stages of the disease.

The increase in the number of people affected by these diseases further brings a spotlight on the importance of new diagnostic tools. Recent discoveries in medical imaging and deep learning have proved positive in answering the question. Nevertheless, the diagnostic procedures that are known today, e.g., Magnetic Resonance Imaging (MRI) may be restricted by the lack of availability in low-resource settings, as well as, the costly nature of specialized tests. To that end, this study is proposing a new deep learning model which, using 3D MRIs and later converted into 2D MRIs, can be capable of simultaneously identifying AD, PD, and FTD. Incorporating MobileNetV2, Squeeze-and-Excitation (SE) blocks, and Spatial Attention, our model SadNetV1 represents an efficient network that allows enhancing the prediction accuracy, allowing its clinical use.

The primary objectives of this study are:

- To enable the creation of a lightweight, and efficient model, which would be able to detect AD, PD, and FTD based on converted 2D MRI slices.
- To improve preprocessing methods of MRI slices in order to guarantee improved generalization to the varying disease modes.
- To compare the given model with conventional CNNs and test its effectiveness by standard measures of accuracy, precision, and recall.
- To investigate Explainable AI (XAI) methods, in specific Grad-CAM, to understand model's prediction mechanism and increase clinical trust.

The proposed study will implement an efficient and accurate deep learning model within real-time neurodegenerative disease diagnosis speculating on interpretability of the model and practical application in clinical practice.

II. Related Work

Diseases like Alzheimer disease, Parkinson disease, and Frontotemporal Dementia, are degenerative diseases that result in neuronal structure and functional decline. The most prevalent basis of dementia is AD, which mainly targets memory and cognition [1]. PD is a disorder of the movement manifested by motor symptoms (e.g., tremor and rigidity) sometimes also featuring non-motor symptoms (e.g. cognitive decline) [2]. Instead, FTD impairs language and behavioral abilities because of the degeneration of the frontal and temporal lobes [3]. These conditions are clinically diverse yet often have overlapping symptoms, leaving diagnosis rather problematic.

It has been indicated in recent literature that deep learning effectively detects Alzheimer Disease (AD) based on MRI-data. Raees and Thomas analyzed (2021) [4] VGG-16 and AlexNet CNN-based models against conventional SVMs to classify AD, MCI, and NC by the ADNI dataset, and obtained higher accuracy. Murugan et al. (2021) [5] proposed DEMNET, with an accuracy rate of 95.23%, and focusing on the interpretability and class balancing, based on SMOTE. Ejaz Ul Haq et. al (2023) [6] were proposed with 92.3% accuracy and in real-time, a lightweight CNN LSTM fusion model. The three studies emphasize the opportunities of deep learning in early and correct AD diagnosis, at the same time, mentioning the necessity of data quality, computational performance, and prospect of clinical implications.

Recently, there was a discussion of the effectiveness of AI in detecting early signs of the Parkinson Disease. Wang et al. (2020) [7] obtained 96.68% accuracy in using premotor biomarkers and deep neural networks on PPMI data. Abdullah et al. (2023) [8] included genetic algorithms in ResNet50, VGG19, and InceptionV3 to process the data of handwriting, achieving the accuracy of 95.29%. Recently, Quan et al. (2021) [9] used BiLSTM to take into account dynamic speech features that detect PD in a more accurate way compared to using only natural speech. According to Ali et al. (2024) [10], with LOSO-CV, an L1-SVM and DNN hybrid model reached unparalleled 100 per cent accuracy and enhanced generalization on voice datasets. Collectively, the given pieces attest to the potential of DL to provide precise, non-invasive PD evaluation.

Hu et al. (2021) [11] presented a DL framework trained on raw 3D MRI to classify patients with AD and FTD as well as NC with an accuracy of 91.83%, the model was additionally interpretable. When using cross-sectional and longitudinal MRI, Pérez-Millan et al. (2023) [12] used PCA and SVM to increase diagnostic rates, particularly using longitudinal data. The study by Ma et al. (2020) [13] utilized multi-scale representation and augmentation driven by GAN that delivered an accuracy of 88.28% and noted the increased sensitivity of FTD. With these studies combined [11]–[13], the possible use of DL/ML in AD-FTD diagnostic applications is indicated, although a remark of the necessity of more abundant, heterogeneous datasets as well as multimodal fusion is made.

Overall, current deep learning solutions to Alzheimer Disease (AD), Parkinson Disease (PD), and Frontotemporal Dementia (FTD) have demonstrated valuable performance in terms of accuracy and interpretability. Nevertheless, other issues related to overlapping clinical symptoms, data imbalance, and poor generalizability across variating datasets are still present, which underlines the necessity to develop more integrated, scalable, and clinically proven models.

III. Dataset

A. Dataset Description

The data is composed of full-brain sagittal T2-weighted 3D MRI brain scans as retrieved from ADNI (Alzheimer) [14], PPMI (Parkinson) [15] and NIFTD (Frontotemporal Dementia) [16] through the (IDA) Image and Data Archive which is maintained at the LONI (Laboratory of Neuro Imaging) at the University of Southern California. A total of 255 distinct patients were chosen after quality filtering, with a balanced distribution of AD, PD and FTD. The number of successful slices obtained per class was 60-89 (30 disease related slices were contributed per subject) per image.
..

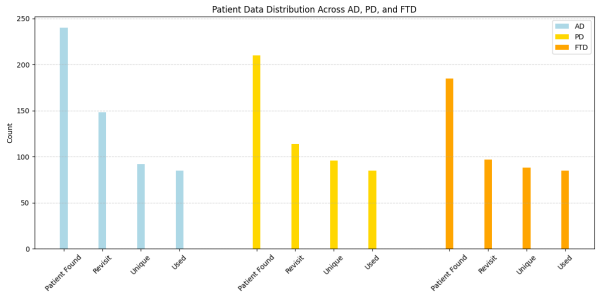

Fig. 1. Visual Representation of Distribution Per Class

The number of patients had several instances of revisit, including AD (240 total, 148 revisits, 85 used), PD (210 total, 114 revisits, 85 used) and FTD (185 total, 97 revisits, 85 used). Three patient per class was kept to check our model's credibility using model prediction and XAI. The full resolution scans (e.g., MPRAGE, MGRAPPA) were selected to scans on 3T Siemens (AD/FTD) and Philips (PD) scanners due to relative anatomical dimensional parallels.

B. Data Preprocessing

The preprocessing pipeline guarantees quality and consistency of data fed to train convolutional neural networks. Steps include:

- **Slice Extraction:** Per subject (30 sagittal slices/subject) were cut based on each one of their 3D MRI data focusing on the disease specific areas of concern; FTD: 60 to 69 (frontal lobe), AD 70 to 79 (hippocampus), and PD 80 to 89 (substantia nigra).
- **Quality Control of Images:** Corrupt slices, blank slices and low-resolution slices were discarded. The scans of Philips (PD) 3T and 3T Siemens (AD/FTD) with high-resolution maintenance only.
- **Pixel Intensity Normalization:** Because of scanner variation (AD: 2884, PD/FTD: 23072) the pixel values were normalized to [0,1] by means of the min-max normalization.
- **Gamma Correction:** Conditionally applied to equalize slice brightness between classes, PD (gamma=1.2), AD(gamma=1.0) and FTD(gamma=0.6).

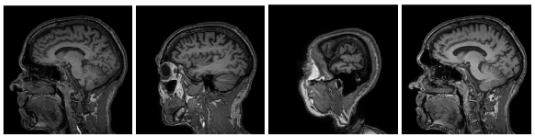

Fig. 2. Sample picture of extracted slices

- **Equalization of contrast:** Contrast was equalised as the PD slices were found to be of high contrast when compared to AD slices and FTD.
- **Resizing:** All photos were made 224x224 pixels.

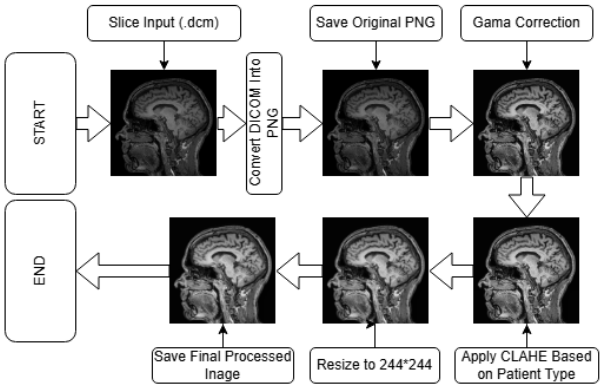

Fig. 3. Pre Processing Flow Chart

- **Data Augmentation (training set only):** Using the PyTorch, a pipeline of transformations was imposed. Here basic augmentation was used

Such an augmentation made it more robust to color, lighting and spatial shapes at the same time as maintaining clinical validity. The final data was split to 70% training, 15% validation and 15% test mentioned in fig. 4.

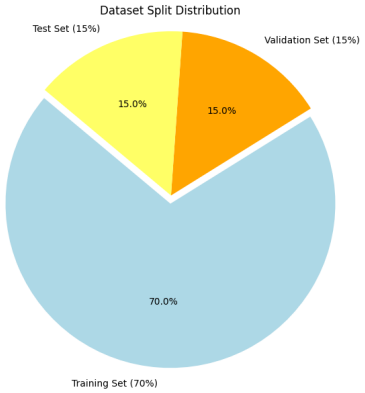

Fig. 4. Data Distribution Pie-Chart

IV. Methodology

The architecture of our proposed model will leverage on the power of MobileNetV2, Squeeze-and-Excitation (SE) blocks, and Spatial Attention mechanisms to ensure that our model can work on extracting spatial and temporal features, and be computationally efficient. The MobileNetV2 is used due to its small size and optimization properties; it utilises depthwise separable convolutions to cut the computational workload by significant amount.

The feature extraction is initiated by a Conv2D layer that has 32-filter of size 3x3, BatchNorm2D, and activation as ReLU6. This extracts low-level information in the input MRI scans. The network goes on with some more layers of convolution, one such layer consisting of 32 filters and bottom layer is 64 filters, to allow it to progressively extract more complex features. Squeeze-and-Excitation (SE) blocks are introduced after the feature extraction layers which recalibrate the channel-wise feature maps. These blocks do global average pooling and the model has two fully connected layers focusing on the most significant features to understand the difference between the diseases.

1) Global Average Pooling (GAP):

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{c,i,j} \tag{1}$$

2) Fully Connected Layers:

$$s = \sigma\left(W_2 \cdot \mathrm{ReLU}\left(W_1 \cdot z\right)\right) \tag{2}$$

3) Recalibration:

$$\hat{x}_c = x_c \cdot s_c \tag{3}$$

where $\hat{x}_c$ is the recalibrated feature map for each channel.

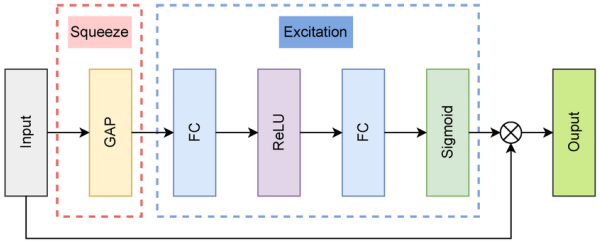

Fig. 5. Flow of SE Block

Moreover, the Spatial Attention mechanism is provided with the objective to assist the model in concentrating on the most important parts of the MRI scans. The Spatial Attention block learns attention maps with a 2D convolution with a kernel size of 7x7 and only one filter, and the model learns to focus on the significant parts of the brain, like the hippocampus in the case of AD or the substantia nigra in PD. This assists in the enhancement of classification accuracy as well as interpretation.

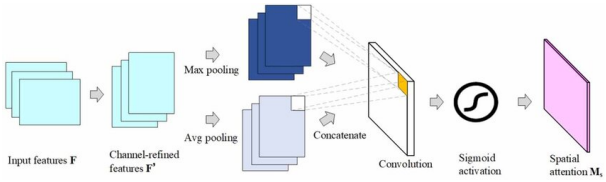

Fig. 6. Flow of Spatial Attention Block

The model then passes through the Dense layer of 1280 units following the feature extraction and attention mechanisms. The softmax layer of this fully-connected layer will be the decision-making layer where the features of individually convolutional and attention layers will be merged to generate the final classification. This is a Dense layer with the ReLU activation to generate non-linearity and make the model learn complicated relations among the extracted features. Lastly, there are the three-unit Softmax output layer to categorize the MRI scan into any of the three, i.e. AD, PD, or FTD.

The model employs Adam Optimizer and initial learning rate is fixed to 1×10^{-5} which is adjusted as per the training gradients. Learning rate is automatically optimized by ReduceLROnPlateau scheduler, that decreasing learning rate when the validation loss stagnates. In order to avoid overfitting, Early Stopping is implemented, in which the training is discontinued unless the validation loss is reduced after 10 epochs in a row. Since we are performing multi class classification, loss function being used is Cross-Entropy loss function since it gives the difference between predicted class labels and the actual ones.

$$L = -\sum_{i=1}^{C} y_i \log(p_i) \tag{4}$$

The model utilizes 70% training, 15% validation and rest 15% as a testing part of the data. Data augmentation is done through the use of rotation, flipping and translation to make the model more robust.

Layers	Output Shape	Parameters
Conv2d-1	(32, 112, 112)	864
InvertedResidual-153	(320, 7, 7)	0
Conv2d-154	(1280, 7, 7)	409,600
SEBlock-160	(1280, 1, 1)	0
SpatialAttention-163	(1280, 1, 1)	0
Linear-166	3	3,843
Total params:		2,433,974
Trainable params:		2,433,974
Non-trainable params:		0

TABLE I. Model Parameters and Output Shapes

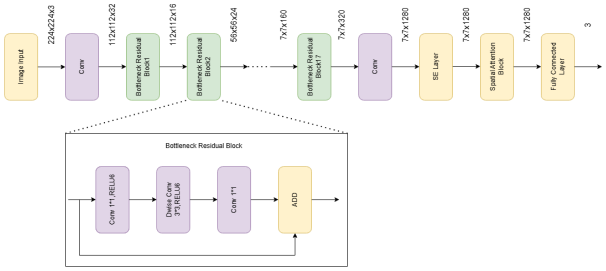

Fig. 7. SadNetv1 Architecture Diagram

To boost interpretability, Grad-CAM (Gradient-weighted Class Activation Mapping) technique is used to visualize the significant parts of the MRI images the model uses when making its choice. Grad-CAM estimates the gradient of the target class to the feature maps of the final convolutional layer and apply weight to the feature maps according to the gradients and this process will be followed by ReLU activation to emphasize on only the positive contribution. The heatmap obtained is superimposed on the initial MRI scan to indicate the areas that had the most significant impacts on the predictions made by the model.

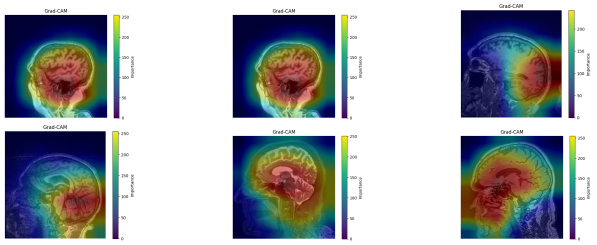

Fig. 8. Heatmaps From Grad Cam

V. Results

A. Summary of Model Evaluation

Six CNNs were trained to evaluate the capacity of deep learning models to recognize neurodegenerative diseases based on the 3D sagittal MRI slice annotated for three diseases, including AD, PD, and FTD. The data sample was characterized by training (70%), validation (15%), and testing (15%) divisions; all of the models were trained similarly.

B. Comparison of CNN Architectures Performance

These are the five pre training models: VGG19, ResNet50, DenseNet121, EfficientNetB0, MobileNetV2, and a designed model called SadNetV1, which combines MobileNetV2 with Squeeze-and-Excitation (SE) blocks and Spatial Attention.

As seen in Fig. 9, VGG19 was demonstrated to be performing poorly on all splits (train, validation, and test) which means that it is underfitting. ResNet50 and DenseNet121 showed better training accuracies, however, they have overfitting as their validation/test accuracies are much lower. The pretrained models with the higher scores are EfficientNetB0 and MobileNetV2, which had better balance (94.36 and 94.01 test accuracy, correspondingly).

The SadNetV1 model performed considerably higher than the rest with 96.15 percent test, 96.84 percent train and 97.11 percent validation accuracy which indicates a better generalization and model stability.

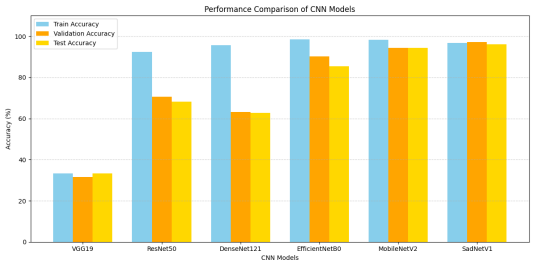

Fig. 9. Comparison of Proposed Model with Other Pre-Trained Model

C. Training Stability

To evaluate training dynamics, accuracy and loss graph of SadNetV1 are presented in Fig. 10 The model exhibits a gentle convergence and is not overfitted as effective regularisation with modules of attention and preprocessing is used.

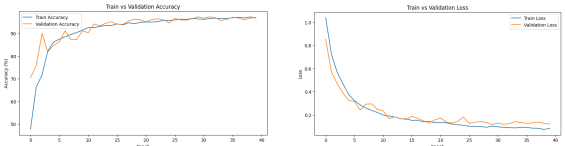

Fig. 10. SadNetV1 Accuracy and Loss

D. Confusion Matrix

The confusion matrix shown in Fig. 11 demonstrates the good per-class separation in validation and test results. Minimal misclassification between AD, PD, and FTD is seen. Rather than that, the marginal confusion was noted between AD and FTD as opposed to our model that did extremely well.

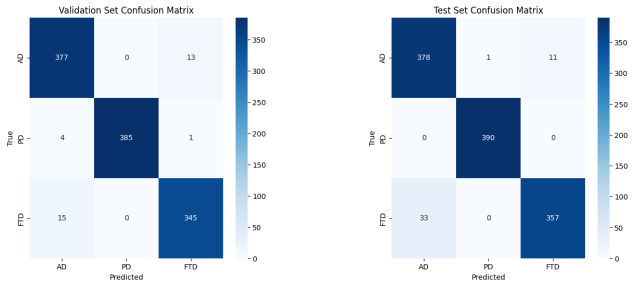

Fig. 11. SadNetV1 Confusion Matrix of Test and Validation

E. Precision, Recall and F1-Score Analysis

To accompany the accuracy-related assessment, precision, recall, and F1-score were also computed across each of the data sets. Such parameters play an important role, especially in the diagnosis of diseases and preventing false negative results and misclassifications.

The comparison of the metric is provided in fig. 12 SadNetV1 demonstrated all sets in the highest values especially 0.8652 (test F1) and 0.8899 (validation F1), which proves its strength and elucidation of differentiating between matching clinical patterns.

	Test			Train			Validate		
Model	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
MobineNetV2	0.7947	0.7673	0.7788	0.9822	0.9840	0.9823	0.7593	0.7315	0.7440
ResNet50	0.5658	0.4106	0.4503	0.9219	0.9279	0.9200	0.5735	0.4298	0.4672
VGG19	0.3328	0.3378	0.3347	0.2156	0.3351	0.2164	0.3142	0.3194	0.3162
EfficientNet B0	0.6979	0.6249	0.6481	0.9824	0.9867	0.9836	0.6426	0.5953	0.6159
DenseNet121	0.6450	0.4757	0.5102	0.9542	0.9624	0.9550	0.6167	0.4519	0.4908
Proposed Model	0.8761	0.8574	0.8652	0.9625	0.9744	0.9635	0.8981	0.8835	0.8899

Fig. 12. Precision, Recall and F1 score of test, train and validation of all models

F. Discussion

The reliable accuracy and F1 results of SadNetV1 show its ability to learn informative spatial aspects of MRI slices. It also improved attention to disease-relevant regions of the brain as the attention modules that were used in combination increased its capacity to supply this attention. Further tasks can involve testing this architecture on bigger datasets and integrating it into the real-time diagnostic applications.

VI. Limitations of The Paper

Despite the fact that the proposed SadNetV1 model portrayed high classification ability, several aspects should be considered. The 3D MRI volumes of the 3D sagittal datasets were trained on 2D slices because they were easy to build the architecture, explaining the reasons why 2D slices were used. Although the major anatomical information is retained in this pseudo-3D method, potential disadvantages are loss in spatial continuity between slices. Furthermore, even though different slice positions led to the axial and coronal slices being explored, the sagittal ones provided the best discriminative performance hence the choice. Generalizability may also be affected by variation in acquisition protocols and amount of data. It is possible that further development of this framework will include 3D or hybrid models and demonstrate it in an even larger multicenter sample.

VII. Conclusion and Future Work

In this paper the lightweight and transparent deep learning architecture called SadNetV1 was proposed to classify neurodegenerative diseases in a multi-class manner using 2-dimensional sagittal magnetic resonance imaging (MRI) slices based on a 3D volume. It demonstrated higher generalization and interpretability due to the combination of MobileNetV2, SE blocks, and spatial attention than by the existing CNN models, according to all evaluation scores. The findings prove that sagittal views comprise discriminative spatial characteristics that are important in early Alzheimer Disease, Parkinson Disease, and Frontotemporal Dementia. Future work of this framework will involve the extension to include volumetric (3D) modeling, cross-domain training using heterogeneous datasets and incorporation of clinical metadata. There will also be an investigation into the real-time deployment and clinical validation with feedback by the radiologists to fill in the gap between the research and diagnosis.

References

[1] C.R. et al. Jack. Nia-aa research framework: Toward a biological definition of alzheimer's disease. *Alzheimer's Dementia*, 14(4):535–562, 2018.

[2] Lorraine V Kalia and Anthony E Lang. Parkinson's disease. *The Lancet*, 386(9996):896–912, 2015.

[3] Jeeyoung Bang, Salvatore Spina, and Bruce L Miller. Frontotemporal dementia. *The Lancet*, 386(10004):1672–1682, 2015.

[4] PC Muhammed Raees and Vinu Thomas. Automated detection of alzheimer's disease using deep learning in mri. In *Journal of Physics: Conference Series*, volume 1921, page 012024. IOP Publishing, 2021.

[5] Suriya Murugan, Chandran Venkatesan, M. G. Sumithra, Xiao-Zhi 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*, 9:90319–90329, 2021.

[6] Ejaz Ul Haq, Qin Yong, Zhou Yuan, Huang Jianjun, and Rizwan Ul Haq. Multimodal fusion diagnosis of the alzheimer's disease via lightweight cnn-lstm model using magnetic resonance imaging (mri). *Available at SSRN 4719918*.

[7] Wu Wang, Junho Lee, Fouzi Harrou, and Ying Sun. Early detection of parkinson's disease using deep learning and machine learning. *IEEE Access*, 8:147635–147646, 2020.

[8] Sura Mahmood Abdullah, Thekra Abbas, Munzir Hubiba Bashir, Ishfaq Ahmad Khaja, Musheer Ahmad, Naglaa F Soliman, and Walid El-Shafai. Deep transfer learning based parkinson's disease detection using optimized feature selection. *IEEE Access*, 11:3511–3524, 2023.

[9] Changqin Quan, Kang Ren, and Zhiwei Luo. A deep learning based method for parkinson's disease detection using dynamic features of speech. *IEEE Access*, 9:10239–10252, 2021.

[10] Liaqat Ali, Ashir Javeed, Adeeb Noor, Hafiz Tayyab Rauf, Seifedine Kadry, and Amir H Gandomi. Parkinson's disease detection based on features refinement through l1 regularized svm and deep neural network. *Scientific Reports*, 14(1):1333, 2024.

[11] Jingjing Hu, Zhao Qing, Renyuan Liu, Xin Zhang, Pin Lv, Maoxue Wang, Yang Wang, Kelei He, Yang Gao, and Bing Zhang. Deep learning-based classification and voxel-based visualization of frontotemporal dementia and alzheimer's disease. *Frontiers in Neuroscience*, Volume 14 - 2020, 2021.

[12] Agnès Pérez-Millan, José Contador, Jordi Juncà-Parella, Beatriz Bosch, Laia Borrell, Adrià Tort-Merino, Neus Falgàs, Sergi Borrego-Écija, Nuria Bargalló, Lorena Rami, Mircea Balasa, Albert Lladó, Raquel Sánchez-Valle, and Roser Sala-Llonch. Classifying alzheimer's disease and frontotemporal dementia using machine learning with cross-sectional and longitudinal magnetic resonance imaging data. *Human Brain Mapping*, 44(6):2234–2244, 2023.

[13] Da Ma, Donghuan Lu, Karteek Popuri, Lei Wang, and Alzheimer's Disease Neuroimaging Initiative Beg, Mirza Faisal and. Differential diagnosis of frontotemporal dementia, alzheimer's disease, and normal aging using a multi-scale multi-type feature generative adversarial deep neural network on structural magnetic resonance images. *Frontiers in Neuroscience*, Volume 14 - 2020, 2020.

[14] Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease neuroimaging initiative (adni), 2025. Accessed: 2024-09-20.
[15] Parkinson's Progression Markers Initiative. Parkinson's progression markers initiative (ppmi), 2025. Accessed: 2024-11-04.
[16] ALLFTD. Artfl-lefftds longitudinal frontotemporal lobar degeneration (allftd) research study, 2025. Accessed: 2024-09-20.