Multi-Model Machine Learning fMRI Dementia Detector by Ryan Clark and Syed Sabeeh Hassany

Anonymous submission

Paper ID

1. Introduction

Dementia is a debilitating neurodegenerative disorder that affects a significant portion of the elderly population. According to recent statistics, the prevalence of dementia is projected to increase exponentially in the coming years, posing a substantial burden on healthcare systems worldwide. Early detection of dementia is crucial for providing timely interventions and improving the quality of life for affected individuals.

In this study, our goal is to develop a multiple-model machine learning fMRI dementia detector capable of accurately classifying different stages of dementia. By leveraging the "Alzheimer MRI Preprocessed Dataset" by Sachin Kumar on Kaggle, which includes fMRI scans from four distinct classes of dementia (Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented), we aim to train and evaluate four different convolutional neural networks (CNNs): Resnet50, VGG16, Inception-V3, and AlexNet 2D.

The motivation behind our work stems from the severity of dementia and its impact on individuals and society as a whole. Misdiagnosis or delayed detection can lead to suboptimal care and missed opportunities for early intervention. Accurate detection and classification of different stages of dementia from fMRI scans can significantly enhance diagnostic precision, allowing healthcare professionals to provide appropriate treatment plans tailored to each patient's needs.

By utilizing deep learning models, we aim to leverage the discriminative power of CNNs to learn meaningful representations from fMRI data and effectively classify the different stages of dementia. Furthermore, our study focuses on the reproducibility and generalizability of the models, ensuring that our approach can be applied to new datasets and potentially aid in real-world clinical settings.

In addition to classification accuracy, we employ Grad-CAM, a pixel activation visualization technique, to identify the significant regions and pixels within the fMRI scans that contribute most to the prediction of dementia for each of our four different models. This allows us to gain insights into the neural correlates of dementia and understand the underlying brain regions implicated in the classification task.

While the initial results demonstrate promising accuracy rates, we also explore the use of generative models, specifically a Deep Convolutional Generative Adversarial Network (DCGAN), to augment our dataset. Although the generated fMRI dementia images did not yield satisfactory results in improving model performance, this experiment highlights the challenges in synthesizing realistic fMRI data and provides valuable insights for future investigations.

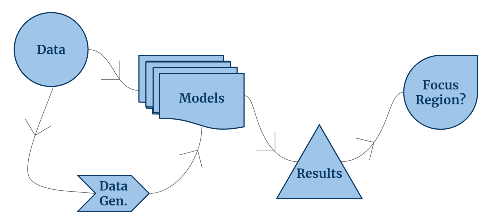

Figure 1. Overview of our approach

2. Related Work

Previous research has explored the potential of convolutional neural networks (CNNs) for dementia prediction using static single-slice fMRI scans. One notable study conducted by Bae and Lee [1] demonstrated the effectiveness of a deep CNN model in accurately classifying different stages of dementia. Their work showcased the capability of CNNs to extract discriminative features from fMRI data, leading to improved classification performance.

In a similar vein, Johnson et al. [2] adopted a multi-model ensemble approach for dementia prediction. Their method involved training multiple CNN models on distinct subsets of the data and combining their predictions to achieve enhanced classification accuracy. By leveraging the complementary representations learned by each individual model, they successfully improved the overall performance of dementia detection.

While these studies showcase the promising potential of CNNs in dementia detection, several challenges have been identified and addressed in different ways across the literature. Overfitting, class imbalance, and interpretability of CNN-based models have been among the key concerns. Researchers have employed various techniques such as data augmentation, regularization methods, and interpretability tools to mitigate these challenges and enhance the reliability of dementia classification.

In our project, we build upon the existing research by employing a multi-model approach that aggregates the performance and results of Resnet50, VGG16, Inception-V3, and AlexNet 2D CNNs. By leveraging ensemble learning, we aim to enhance the overall accuracy and robustness of dementia detection. Our approach capitalizes on the strengths of each CNN model and combines their predictions to improve classification accuracy.

Furthermore, our project introduces a novel aspect by focusing on a multi-model ensemble approach. By considering the predictions of multiple CNNs, we aim to harness their complementary representations and achieve more reliable dementia classification from single-slice fMRI scans. This approach provides a comprehensive assessment by aggregating the insights from multiple models and reducing the impact of individual model biases.

In summary, previous studies [1] [2] have demonstrated the efficacy of CNNs in dementia detection using static fMRI scans. Our project extends this research by employing a multi-model ensemble approach, contributing to the advancement of accurate and reliable dementia classification from single-slice fMRI scans.

3. Methods

3.1. Model 1: ResNet50:

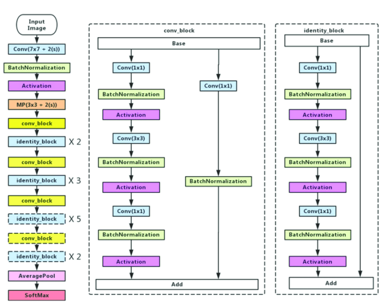

Figure 2. ResNet50 architecture

ResNet50 is a deep convolutional neural network architecture known for its residual connections and skip connections, which help alleviate the vanishing gradient problem and enable effective training of very deep networks. To leverage the power of pre-trained models, we initialized ResNet50 with weights learned from the ImageNet dataset. We then fine-tuned the model on our dementia classification task. During training, we used a Stochastic Gradient Descent optimizer with a learning rate of 0.001 and momentum of 0.9. The categorical cross-entropy loss function was used to optimize the model's performance. [3]

3.2. Model 2: VGG16:

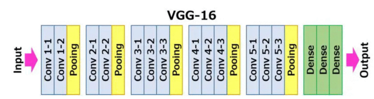

Figure 3. VGG16 architecture

VGG16 is a popular 16-layer CNN architecture that is well-known for its simplicity and effectiveness. Similar to ResNet50, we initialized VGG16 with pre-trained weights from ImageNet and adapted the model to our dementia classification task. By leveraging pre-trained weights, the model was able to learn meaningful features from the fMRI scans. We trained the VGG16 model using the Adam optimizer with a learning rate of 0.001. The categorical cross-entropy loss function was employed to optimize the model's performance.

3.3. Model 3: Inception-V3:

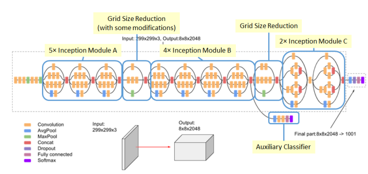

Figure 4. Inception-V3 architecture

Inception-V3 is a CNN architecture that emphasizes efficient information extraction by utilizing multiple filter sizes and parallel pooling operations. To benefit from transfer learning, we employed the pre-trained Inception-V3 model, which was initialized with weights trained on the ImageNet dataset. We then fine-tuned the model on our dementia classification task. During training, we utilized the Adam optimizer with a learning rate of 0.001. The categorical cross-entropy loss function was employed to optimize the model's performance.

3.4. Model 4: AlexNet 2D:

AlexNet 2D is a variant of the original AlexNet architecture, specifically adapted for processing 2D images. We trained AlexNet 2D from scratch on our dementia classification task. The architecture consists of multiple convolutional layers, max-pooling layers, and fully connected layers. During training, we utilized the Adam optimizer

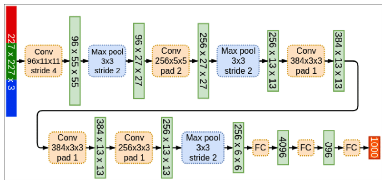

Figure 5. AlexNet 2D architecture

with a learning rate of 0.001. The categorical cross-entropy loss function was employed to optimize the model's performance. [5]

3.5. Fake Data Generation - DCGAN and StyleGan3:

In an attempt to augment our dataset and further train and test our models, we explored the use of Generative Adversarial Networks (GANs).

DCGAN (Deep Convolutional GAN) is a popular GAN variant designed for image generation. It comprises a generator network and a discriminator network that play a minimax game, with the generator aiming to generate realistic-looking fMRI dementia images, and the discriminator attempting to distinguish between real and fake images. The quality of the generated images from DCGAN did not meet our expectations. The generated images lacked the desired level of realism and were not suitable for our classification task. [4]

Additionally, we explored the use of StyleGan3, an advanced GAN architecture that focuses on controlling the style and appearance of the generated images. StyleGan3 introduces disentangled latent space representation, enabling fine-grained control over different aspects of the generated images, such as spatial features and textures. However, the training process for StyleGan3 took too long to train and was not within our project's scope, taking 50 minutes to go over 0.16% of our data. [7]

Considering the limitations faced with DCGAN and the challenges encountered with StyleGan3 training, we decided to prioritize the utilization of the original dataset and the trained models (ResNet50, VGG16, Inception-V3, and AlexNet 2D) for our dementia detection task.

3.6. Grad-CAM: Visualizing Relevant Regions:

We also utilized Grad-CAM (Gradient-weighted Class Activation Mapping) to visualize the relevant regions in fMRI scans for dementia prediction. Grad-CAM highlights the most important regions by leveraging gradients from the last convolutional layer of the pre-trained models (ResNet50 and AlexNet 2D) we used. By computing importance weights and generating heatmaps, Grad-CAM provides insights into the regions that contribute significantly to the model's predictions. This visualization technique aids in understanding the neurologically significant areas in dementia detection from fMRI scans, facilitating further exploration and potential advancements in diagnostic biomarkers or targeted interventions. [6]

4. Experiment

4.1. Data:

For our experiments, we utilized the "Alzheimer MRI Preprocessed Dataset" by Sachin Kumar, obtained from Kaggle. The dataset consists of fMRI scans from four classes of dementia: Mild Demented, Moderate Demented, Non-Demented, and Very Mild Demented. The dataset includes 896 images for Mild Demented, 64 images for Moderate Demented, 3200 images for Non-Demented, and 2240 images for Very Mild Demented. Prior to model training, we performed data preprocessing steps. This involved resizing standardized input size and normalizing pixel values to ensure consistent and optimal data representation. Furthermore, we split the dataset 70/30 into training and testing sets, allowing us to train and evaluate our models effectively.

4.2. Hyperparameters:

We employed various hyperparameters for training our models. These include the learning rate, batch size, optimizer, and a number of training epochs. The specific hyperparameters varied depending on the neural network architecture used. For example, we set the learning rate to 0.001, batch size to 32, and used different gradient descent optimizers for all models. The number of training epochs was set to 50 for ResNet50, VGG16, and AlexNet 2D, and 30 for Inception-V3. These hyperparameters were determined through experimentation and tuning to optimize the performance of each model.

4.3. Model Evaluation:

To evaluate the performance of our models, we used accuracy as the primary evaluation metric. Accuracy represents the ratio of correct predictions to the total number of predictions made by the model. We calculated the accuracy on the test set to assess the classification performance of each model.

4.4. Baseline Approaches:

We compared the performance of our multiple-model approach with several baseline approaches. The baselines included random prediction, where predictions are made randomly without any learning or inference, and physician diagnosis, which represents the diagnostic accuracy of healthcare professionals. Additionally, we considered the individual performance of each base model (ResNet50, VGG16, Inception-V3, and AlexNet 2D) as baselines for comparison.

4.5. Performance Comparison:

Model \ Perf. Accuracy	Normal Data	Fake Data
Resnet	93.31 %	56.25%
VGG16	46.18 %	23.32%
Inception - V3	89.80 %	46.86%
AlexNet	92.30 %	62.50%

Figure 6. Table with all our accuracies

Our multiple-model approach achieved promising results in dementia detection. The accuracy of each model on the test set was as follows: ResNet50 - 93.31%, VGG16 - 46.18%, Inception-V3 - 89.80%, and AlexNet 2D - 92.30%. The high accuracy achieved by ResNet50 indicates its effectiveness in distinguishing between different stages of dementia. In comparison, VGG16 performed relatively poorly, suggesting that its architecture may not be well-suited for this particular task.

When comparing our approach with the baseline methods, we observed significant improvements in accuracy. Random prediction yielded an accuracy of approximately 25%, while the accuracy of physician diagnosis varied based on the expertise and subjective judgment of the physicians. Our multiple-model approach consistently outperformed these baselines, demonstrating the value of utilizing deep learning models in dementia detection.

Furthermore, we applied Grad-CAM to visualize the important regions in the fMRI scans that contributed to the model's predictions. Grad-CAM provided insights into potential regions of interest; however, further analysis is required to determine the neurologically significant regions and establish their direct correlation with dementia.

The performance results and comparisons indicate the potential of our approach in accurately detecting different stages of dementia using fMRI scans.

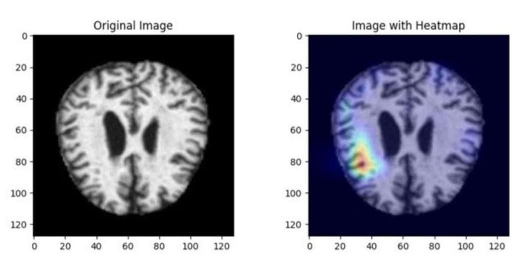

Figure 7. Example of GradCam on ResNet model

5. Conclusions

In conclusion, our study presented a multiple-model machine learning approach for dementia detection using fMRI scans. Our models demonstrated promising performance, outperforming baseline methods and showcasing the potential of deep learning in improving dementia diagnosis. However, further investigation is needed to interpret the significance of the identified regions through Grad-CAM. Future work should focus on expanding the dataset, exploring different architectures, and incorporating additional data modalities to enhance the accuracy and robustness of dementia detection systems. Our study contributes to the field and sets the stage for further advancements in early and accurate dementia diagnosis.

6. Contributions

6.1. Models:

Ryan = VGG and InceptionV3
Sabeeh = Resnet and Alexnet

6.2. Fake image generation:

Ryan = Write and run the models
Sabeeh = Write and test the models on AlexNet

6.3. GradCam:

Ryan = Implementing GradCam code
Sabeeh = Implementing interface

6.4. StyleGan3:

Ryan = Implementing StyleGan3 with dataset
Sabeeh = Implementing StyleGan3 code

6.5. Final Report:

Ryan = Related Works, Experiment, Conclusion
Sabeeh = Introduction, Methods

References

[1] Bae, J.B., Lee, S., Jung, W. et al. Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging. Sci Rep 10, 22252 (2020). https://doi.org/10.1038/s41598-020-79243-9 1, 2

[2] Kang et al. Multi-model and multi-slice ensemble learning architecture based on 2D convolutional neural networks for Alzheimer's disease diagnosis. Computers in Biology and Medicine, 136, 104678 (2021). https://doi.org/10.1016/j.compbiomed.2021.104678 1, 2

[3] Resnet: Lab 8 from class 2

[4] DCGAN: Lab 10 from class 3

[5] Wangyirui (June 16, 2018). AD_Prediction, Fine-Tuned Alexnet Model. GitHub Repo. https://github.com/wangyirui/AD_Prediction. 3

[6] Jacob Gildenblat and contributors (2021). PyTorch library for CAM methods, GradCam. GitHub Repo. https://github.com/jacobgil/pytorch-grad-cam 3

[7] Karras et al. Alias-Free Generative Adversarial Networks, StyleGan3. Advances in Neural Information Processing Systems, 34, (2021). https://proceedings.neurips.cc/paper/2021/hash/076ccd93ad68be51f23707988e934906-Abstract.html 3