(Augmented-Alzheimer Image Classification)

By:

Name	ID
Omar Mustafa Mustafa	42010262
Bassema Abdel Razek	42010683

Under the Supervision of:

Eng. Ahmed Nousir

Introduction:

In this Jupiter Notebook, we delve into the intricate task of classifying MRI scans associated with Alzheimer's disease. Our main goal is to gauge the effectiveness of different pre-trained models in categorizing scans into four specific groups: Nondemented, VeryMildDemented, MildDemented, and ModerateDemented. Through the utilization of advanced machine learning and deep neural networks, this study aspires to provide meaningful contributions to the intersection of medical imaging and the diagnosis of Alzheimer's disease.

Dataset and Preprocessing:

Overview of the Dataset:

In the context of this exploration, the dataset is segmented into two subsets, one comprising augmented MRI scan images and the other containing original images. These subsets are meticulously arranged within their respective directories. The dataset itself is well-structured, offering a substantial training pool of 33,984 augmented samples and a testing set comprising 6,400 original samples. Notably, the training set demonstrates equilibrium in label distribution between the Nondemented and MildDemented classes, with a slight variance observed in the balance of the ModerateDemented class. It's worth noting that this potential class imbalance may influence the overall performance of the models utilized in this study.

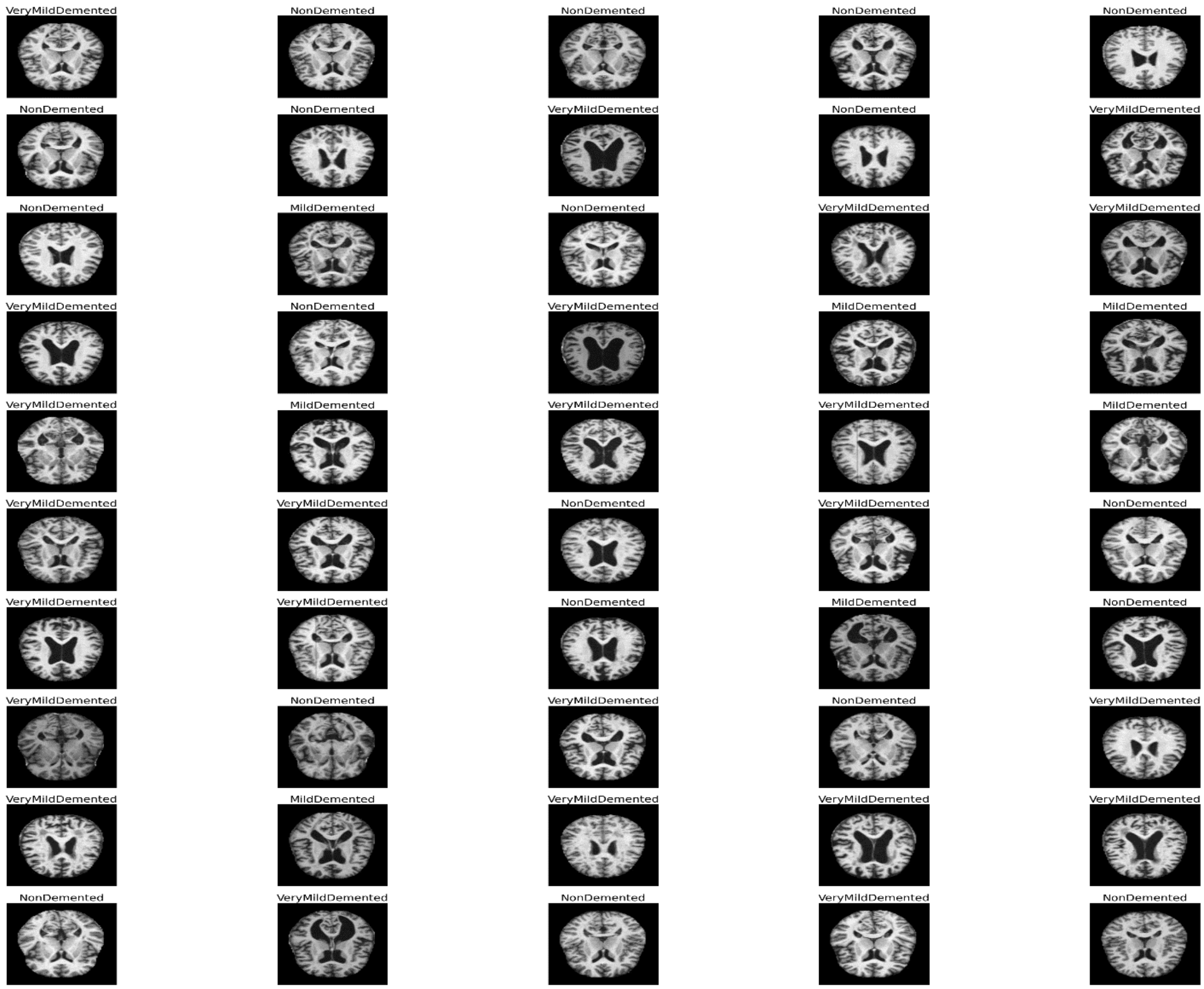

Fig1: The sample of dataset

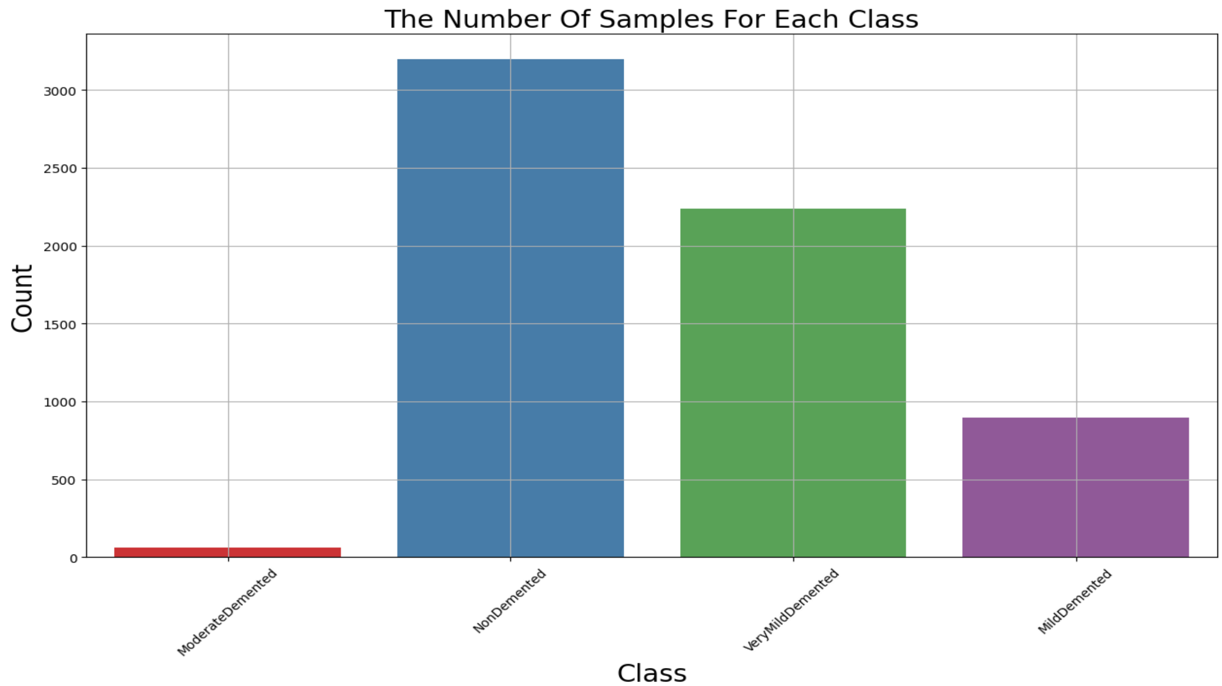

Fig2: The number of samples of each class

Preprocessing includes:

Preparing the Data:

In the initial phase of data preparation, the pixel values of the MRI scan images are rescaled to align with the requirements of neural networks. Following this, the augmented dataset undergoes a partition, creating distinct training and validation sets with a balanced 90/10 split. An essential aspect of this data preparation entails the utilization of the Image Data Generator class, a powerful tool for on-the-fly data augmentation. This strategic approach is pivotal for improving the model's generalization capabilities and effectively addressing concerns related to overfitting during the subsequent training process.

Model Architectures and Training:

Utilizing Advanced Neural Network Architectures:

Within this notebook, the implementation encompasses a suite of cutting-edge Convolutional Neural Network (CNN) architectures. These include VGG16, VGG19, ResNet50, ResNet101, Xception, MobileNet, DenseNet169, DenseNet121, MobileNetV2, and InceptionV3. Employing the transfer learning technique, each model capitalizes on pre-trained weights derived from the ImageNet dataset. To adapt these models for the task at hand, the top layers are removed and replaced with a new dense layer tailored for the classification of the four distinct categories associated with Alzheimer's disease. The models undergo compilation using the Adam optimizer and categorical cross entropy loss. Evaluation metrics primarily include AUC and accuracy, with an early stopping callback strategically implemented to cease training if there is no improvement in validation loss after 8 epochs.

Evaluating Model Performance:

Post the training and evaluation phases, a meticulous examination of each model's performance unfolds, capturing an array of metrics for comprehensive assessment. The specialized function, prepare_for_test, is instrumental in orchestrating the preparation of the test set and predictions for subsequent analysis. Extracted key performance indicators (KPIs) encompass AUC, accuracy, precision, recall, and F1-score, delineated for each specific class. To provide a succinct hierarchy, the models are ranked based on their respective test accuracies, resulting in the following order:

- MobileNet: 95.17%
- VGG16: 94.67%
- VGG19: 93.14%
- DenseNet169: 93.03%
- DenseNet121: 92.64%
- Xception: 91.86%
- InceptionV3: 90.16%
- MobileNetV2: 90.06%
- ResNet101: 84.20%
- ResNet50: 76.66%

Discussion:

In this phase of analysis, the report scrutinizes the potential factors influencing the performance of each model. Considerations such as model depth, complexity, parameter count, and inherent biases present in the dataset are carefully evaluated. The objective is to unravel the nuances that contribute to the observed variations in model performance. Notably, the architectural design of MobileNet emerges as a frontrunner, showcasing a remarkable equilibrium between efficiency and accuracy. This distinctive characteristic positions MobileNet as the leading model in terms of performance.

Conclusion:

To encapsulate the findings, the notebook underscores the overall satisfactory performance of all models, singling out MobileNet for its exceptional accuracy. However, the decision on model selection is not solely guided by performance metrics; practical considerations, including computational resources, play a pivotal role. Models with increased complexity, like ResNet and Dense Net, might impose higher demands on computational power, introducing an additional dimension to the decision-making process.

Recommendations:

In practical deployment scenarios, especially for mobile or web applications, the recommended models would be MobileNet or VGG16, recognized for their remarkable accuracy and efficiency. There exists an opportunity for further refinement through experimentation. Exploring hyperparameter tuning, incorporating diverse data augmentation techniques, or delving into ensemble methods could potentially yield improvements in model performance.

Appendices and References:

To uphold the standards of transparency and reproducibility, the report integrates all pertinent code excerpts, visualizations of output, and a detailed compilation of references to the utilized libraries and frameworks. This meticulous approach aligns with the best practices in ensuring the clarity and replicability of the methodologies applied throughout the study.