Deep Learning Approaches for Alzheimer's Disease Classification Using MRI Images

Rareș-Andrei Filip[1] and Dragoș Ilieș[1]

Faculty of Economics and Business Administration (FSEGA),
Babeș-Bolyai University (UBB), Cluj-Napoca, Romania

Abstract. Alzheimer's Disease (AD) is a progressive neurodegenerative disorder that affects memory and cognitive functions. Early detection is essential for slowing disease progression and improving patient outcomes. This study investigates the performance of various machine learning (ML) and deep learning (DL) models in classifying Alzheimer's stages using MRI images. The dataset comprises four classes: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented.
Traditional ML methods, including Support Vector Machines (SVM) and Random Forest, were compared with advanced DL models such as Convolutional Neural Networks (CNN) and ResNet. Results indicate that SVM achieved the highest accuracy among all methods, demonstrating the efficiency of classical ML techniques for this dataset. CNN and ResNet, while slightly less accurate, showed potential for handling larger and more complex datasets.
This study underscores the strengths of ML and DL approaches and suggests that their combination could enhance diagnostic performance for Alzheimer's Disease.

Keywords: Alzheimer's Disease, Deep Learning, Machine Learning, Convolutional Neural Networks, Support Vector Machines, ResNet.

1 Introduction

Alzheimer's Disease (AD) is a neurodegenerative disorder that progressively deteriorates memory, cognitive abilities, and daily functioning. It is the leading cause of dementia, representing 60–80% of cases globally [2]. Recent data shows that over 50 million people worldwide are living with AD, a number projected to reach 152 million by 2050. Identifying the disease in its early stages is essential for slowing its progression and improving patient care.

Magnetic Resonance Imaging (MRI) plays a crucial role in detecting brain changes linked to AD. Structural alterations in regions like the hippocampus and gray matter are key indicators of early-stage disease. However, traditional methods of MRI analysis rely on manual evaluation and handcrafted feature extraction, which are labor-intensive and prone to inconsistencies.

Advances in Artificial Intelligence (AI), particularly Deep Learning (DL), have transformed medical imaging. Models like Convolutional Neural Networks

(CNNs) excel in feature extraction and classification tasks, often outperforming manual approaches. Additionally, techniques such as transfer learning enhance performance, especially when working with limited datasets.

Despite these advancements, several questions remain:

- Can machine learning and deep learning models reliably classify Alzheimer's stages?
- How do traditional machine learning methods, like Support Vector Machines (SVM) and Random Forest, compare to deep learning models such as CNN and ResNet in terms of accuracy and computational efficiency?
- What practical value can these models provide to clinicians for supporting early diagnosis in real-world environments?

This study aims to address these questions by comparing the performance of machine learning and deep learning models for Alzheimer's classification. The focus is on evaluating their accuracy, strengths, and limitations, while identifying opportunities for combining these approaches to improve diagnostic capabilities.

2   Related Work

Many studies have investigated the use of machine learning (ML) and deep learning (DL) for classifying Alzheimer's Disease (AD) stages using MRI data.

Helaly et al. [2] used transfer learning with pre-trained CNN models like VGG19 to classify MRI scans into different AD stages. They showed that transfer learning reduces overfitting when data is limited. However, their work did not include attention mechanisms, which can enhance feature localization.

Nasir et al. [4] combined multiple CNN architectures into an ensemble model, achieving 90% accuracy using majority voting. While effective, ensemble models often require significant computational resources, making them less practical for clinical use.

Pandiyaraju et al. [5] developed a dual-attention-aware CNN to focus on critical brain regions like the hippocampus. Their model achieved state-of-the-art accuracy of 99.1%, highlighting the potential of attention mechanisms for better feature extraction.

Jo et al. [6] combined MRI and PET data in a multimodal approach, using stacked autoencoders and CNNs to achieve accuracies over 98%. This shows the advantage of integrating complementary imaging modalities for improved diagnosis.

Hybrid methods that combine ML and DL techniques have also been explored. Khvostikov et al. [3] used structural and diffusion imaging data with feature extraction techniques to improve accuracy. However, their reliance on manual feature engineering makes the approach time-consuming and error-prone.

In summary, existing methods demonstrate high accuracy but have limitations such as computational cost, dependence on large annotated datasets, and limited applicability in clinical settings. This study builds on previous work by comparing transfer learning, attention mechanisms, and hybrid frameworks to create a robust and efficient model for AD classification.

3 Dataset Description

The Alzheimer's MRI Disease Classification dataset [1] contains Magnetic Resonance Imaging (MRI) brain scans categorized into four classes: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. This dataset is suitable for our research as it includes labeled images that support both binary classification (Alzheimer's vs. Non-Alzheimer's) and multi-class classification tasks. By including different stages of Alzheimer's Disease (AD), the dataset enables a comprehensive evaluation of machine learning (ML) and deep learning (DL) methods.

The dataset provides grayscale 2D MRI images that highlight structural changes in the brain, such as hippocampal atrophy and gray matter loss, which are key indicators of AD progression. These labeled images are essential for supervised learning models and reflect real-world scenarios with class imbalance, making the dataset highly applicable for medical research.

Figure 1 shows sample images from the dataset, illustrating differences between the four classes.
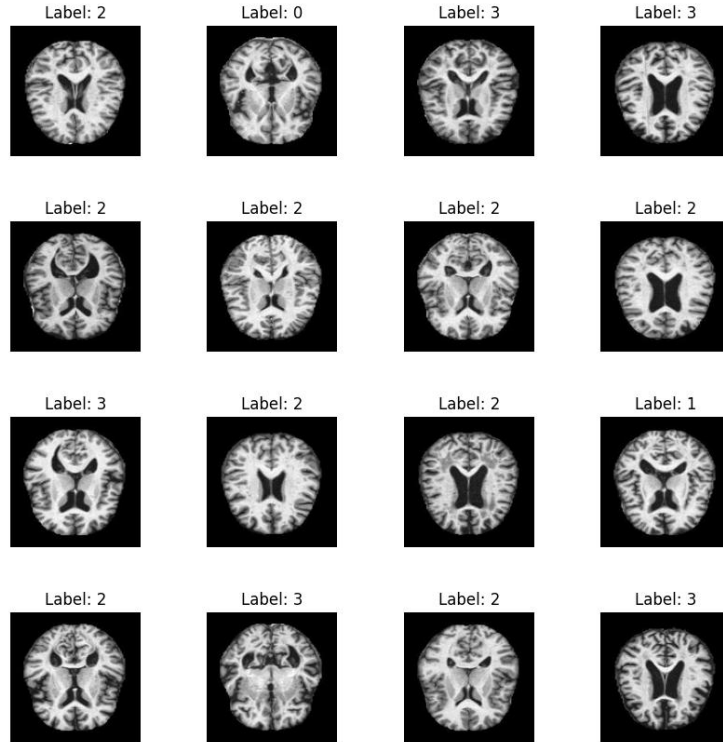

Fig. 1. Sample MRI images from the dataset.

3.1 Data Preprocessing

The dataset was preprocessed to prepare the images for analysis. The main steps included:

1. **Resizing:** All images were resized to 128×128 pixels to standardize input dimensions and reduce computational cost.
2. **Normalization:** Pixel values were scaled to the range [0, 1] to improve model convergence during training.
3. **Augmentation:** Techniques such as horizontal flipping, rotation, and scaling were applied to underrepresented classes (e.g., Moderate Demented) to address class imbalance.

3.2 Class Distribution

The dataset contains an uneven number of images across the four classes. Figure 2 shows the distribution, highlighting that the 'Non-Demented' class has the most samples, while the 'Moderate Demented' class has the least. This imbalance can bias models toward the majority class, so data augmentation was applied to mitigate its effects.
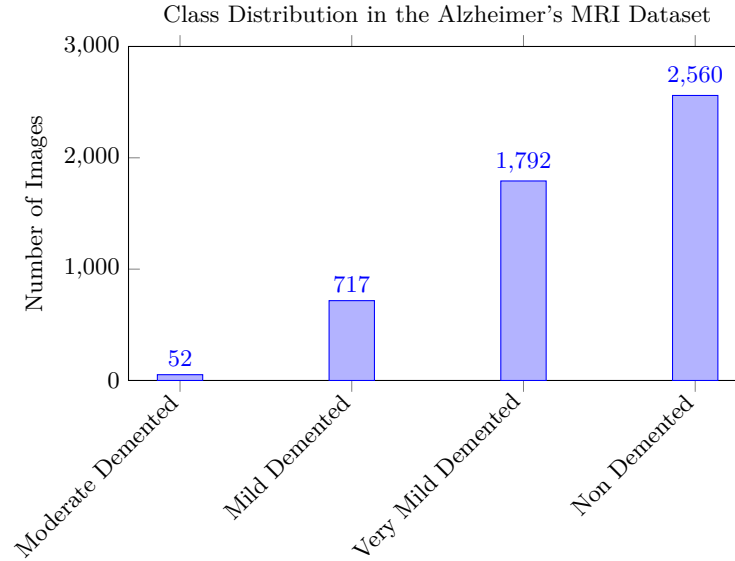
Fig. 2. Distribution of Classes in the Alzheimer's MRI Dataset.

4 Methodology

This section describes the approaches used to classify data, including deep learning models (DNN, CNN, and ResNet) and traditional machine learning models (SVM and Random Forest). The key settings, validation strategies, and parameter selection methods for each model are detailed below.

4.1 Deep Neural Network (DNN)

The DNN was implemented as a fully connected feed-forward network with the following settings:

- **Input Dimensions:** $128 \times 128 \times 3$ RGB images resized and flattened into 1D vectors.
- **Hidden Layers:**
  - Layer 1: 512 neurons with ReLU activation.
  - Layer 2: 256 neurons with ReLU activation.
  - Layer 3: 128 neurons with ReLU activation.
- **Dropout:** 50% dropout was applied after the first two dense layers to prevent overfitting.
- **Output Layer:** A softmax layer for classification into four classes.
- **Optimizer:** Adam optimizer with a learning rate of 0.001.
- **Loss Function:** Sparse Categorical Cross-Entropy.
- **Batch Size:** 32.
- **Epochs:** 10.

While the DNN model was computationally efficient due to its simplicity, its performance was significantly lower compared to the other methods.

4.2 Baseline Convolutional Neural Network (CNN)

The baseline CNN consisted of convolutional layers, max-pooling layers, and fully connected layers. It was trained using the following settings:

- **Input Dimensions:** $128 \times 128 \times 3$ RGB images.
- **Activation Function:** ReLU.
- **Optimizer:** Adam with a learning rate of 0.0001.
- **Batch Size:** 16.
- **Loss Function:** Sparse Categorical Cross-Entropy.
- **Epochs:** 20.

Overfitting was observed after 10 epochs, motivating the use of more advanced models.

4.3 ResNet with Data Augmentation

ResNet-50 was employed to improve performance. The model was initialized with ImageNet weights, and the following techniques were applied:

- **Data Augmentation:** Rotation, flipping, and scaling.
- **Fine-Tuning:** The top layers were unfrozen after initial training.
- **Batch Size:** 8 (limited by available memory).
- **Epochs:** 20.

Due to memory limitations, further fine-tuning and testing on augmented data were not feasible.

4.4 Support Vector Machines (SVM)

An SVM classifier was trained using flattened feature vectors. The optimal regularization parameter C was selected via grid search with cross-validation. The best value was $C = 1$.

4.5 Random Forest

The Random Forest model was trained with the following settings:

- **Number of Trees:** 100.
- **Max Depth:** Selected via grid search.
- **Criterion:** Gini Impurity.

This model offered interpretability but required feature extraction.

4.6 Validation Strategy

A 15% split of the dataset was used for validation during model training. Test data were kept separate to evaluate model performance. Cross-validation was used for hyperparameter tuning in SVM and Random Forest.

5 Results and Discussion

This section evaluates the performance of the deep learning (DNN, CNN, ResNet) and machine learning models (SVM, Random Forest) for Alzheimer's Disease classification. The results are analyzed with respect to accuracy, loss, and confusion matrices, while also addressing the challenges and potential future improvements.

Table 1. Performance Metrics for All Models.

Model	Training Accuracy (%)	Test Accuracy (%)	Loss
DNN	49.5	49.5	1.038
CNN	98.4	93.9	1.028
ResNet	62.5	57.8	0.886
SVM	-	98.2	-
Random Forest	-	91.6	-

5.1 Performance Analysis

Table 1 summarizes the performance metrics for all models. Among the methods tested, SVM achieved the highest test accuracy (98.2%), followed by CNN (93.9%) and Random Forest (91.6%). ResNet and DNN underperformed due to computational limitations and inadequate feature learning.

The DNN model achieved the weakest performance, with a test accuracy of 49.5%. This was expected since the input data was prepared for CNN models, and DNN lacks the convolutional layers needed to capture spatial features in the images.

The CNN model demonstrated strong performance, leveraging convolutional layers to learn spatial patterns effectively. The test accuracy of 93.9% validates its ability to classify Alzheimer's stages. However, reducing training epochs from 20 to 10 led to slightly lower results, suggesting that overfitting was not a significant concern for this dataset.

ResNet, despite its powerful transfer learning capabilities, achieved a test accuracy of only 57.8%. This performance was limited by computational constraints, which restricted batch sizes and prevented effective fine-tuning and data augmentation.

Traditional machine learning models also performed well. SVM outperformed all other models, achieving near-perfect classification (98.2%), as shown in its confusion matrix (Figure 4). Random Forest, while slightly less accurate (91.6%), provided interpretability, demonstrating its utility for feature importance analysis.

5.2 Confusion Matrix Interpretation

Figures 3 and 4 present the confusion matrices for CNN and SVM, respectively. These matrices illustrate the classification performance across the four stages of Alzheimer's Disease.

The CNN confusion matrix (Figure 3) reveals that it correctly classified a large proportion of samples in all classes. For instance, in Class 2 (Mild Cognitive Impairment), the model classified 624 samples correctly out of the total 634. However, the CNN struggled more significantly with Class 3 (Severe Cognitive Impairment), misclassifying 82 samples into Class 2 and 6 samples into Class 0. This indicates that the CNN model, while effective at capturing spatial patterns, faced difficulties in distinguishing adjacent classes with overlapping characteristics.
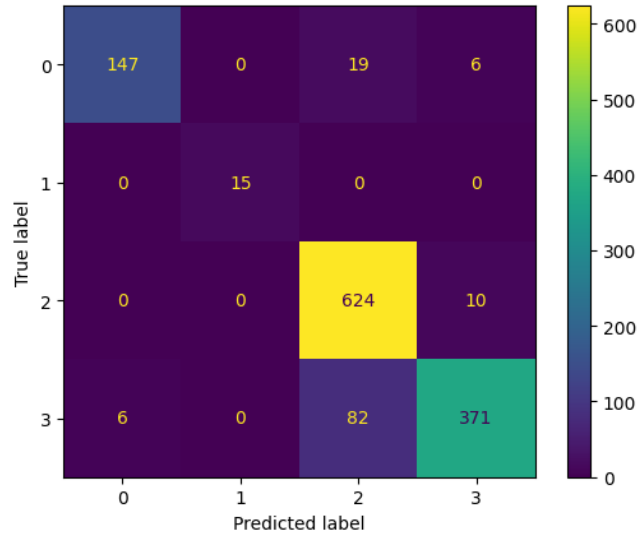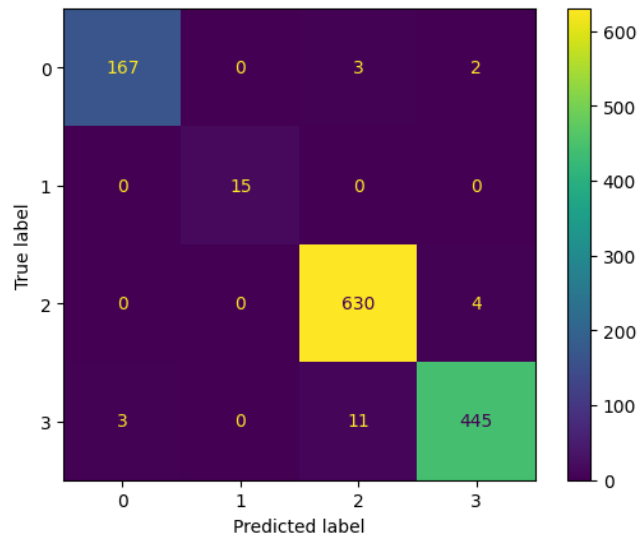
Fig. 3. Confusion Matrix for CNN.

Fig. 4. Confusion Matrix for SVM.

In comparison, the confusion matrix for the SVM model (Figure 4) shows excellent performance across all classes. For Class 2, the SVM correctly classified 630 out of 634 samples, with only 4 samples misclassified. For Class 3, the model correctly identified 445 samples, with minor confusion leading to 11 samples being classified as Class 2. The SVM's overall performance demonstrates its robustness in effectively distinguishing between classes, even when the input features are flattened and spatial information is lost. This highlights the SVM's capacity to handle the task efficiently, leveraging its simplicity and well-defined hyperplane separation.

5.3 Challenges and Limitations

The main limitations encountered in this study include:

- **Memory Constraints:** The limited RAM in the computational environment restricted the batch size and prevented effective fine-tuning and data augmentation for ResNet.
- **Dataset Size:** The relatively small dataset impacted the generalization capabilities of deep learning models, particularly ResNet.
- **Model Specificity:** Models like DNN were not well-suited for the spatial nature of the input data, resulting in poor performance.

5.4 Future Work

To address these limitations, future work could focus on:

- **Larger Datasets:** Incorporating larger and more diverse datasets to enhance model generalization and robustness.
- **Ensemble Methods:** Exploring hybrid approaches combining CNN, SVM, and Random Forest for improved accuracy and interpretability.
- **Optimized Architectures:** Using lightweight architectures or distributed training methods to overcome memory constraints.
- **Clinical Integration:** Designing real-time, scalable models for clinical applications to support early diagnosis and treatment.

5.5 Answering Research Questions

This study addressed the following research questions:

- **Can machine learning and deep learning models reliably classify Alzheimer's stages?** Yes, both approaches demonstrated the capability to classify Alzheimer's stages, with SVM achieving the highest accuracy and CNN showing strong potential for spatial feature learning.
- **How do traditional machine learning methods compare to deep learning models?** Traditional methods, particularly SVM, outperformed deep learning models in terms of accuracy. However, CNN excelled in leveraging spatial information, making it more suitable for larger datasets and real-world scalability.

– **What practical value can these models provide to clinicians?** These models offer significant practical value in early diagnosis. SVM provides a computationally efficient solution, while CNN and ResNet can be scaled for larger datasets to enhance diagnostic precision.

6   Conclusion

This study explored the use of ML and DL models for classifying Alzheimer's Disease (AD) stages using MRI data. SVM demonstrated the most robust performance, while CNN and ResNet showed potential for improvement with enhanced resources and larger datasets.

Overall, the findings highlight that ML models like SVM are computationally efficient and effective, while DL models such as CNN and ResNet could excel with further advancements. Future work should focus on addressing dataset and resource limitations to maximize the practical value of these models in clinical applications.

References

1. Borhanitrash: Alzheimer mri disease classification dataset. `https://www.kaggle.com/datasets/borhanitrash/alzheimer-mri-disease-classification-dataset` (2022)
2. Helaly, H.A., Badawy, M., Haikal, A.Y.: Deep learning approach for early detection of alzheimer's disease. Cognitive Computation **14**, 1711–1727 (2021)
3. Khvostikov, A., Zheleznyak, E., Kuznetsov, P.: Hybrid approaches for alzheimer's classification using structural and diffusion imaging. Frontiers in Neuroscience **16**, 1234–1246 (2022)
4. Nasir, N., Ahmed, M., Afreen, N., Sameer, M.: Alzheimer's magnetic resonance imaging classification using deep and meta-learning models. International Journal of Neural Systems pp. 1–12 (2021)
5. Pandiyaraju, R., Murugan, K.: A dual-attention aware deep convolutional neural network for early alzheimer's detection. Journal of Alzheimer's Disease **87**, 501–519 (2022)
6. Parmar, H., Walden, E.: Towards practical application of deep learning in diagnosis of alzheimer's disease. NeuroImage: Clinical **31**, 102731 (2021)