Optimizing Alzheimer's Disease Detection Through Integrated Machine Learning Techniques

Jiya Sugla*
Department of Electronics and Communication
Vellore Institute of Technology
Vellore, India
jiyasugla@gmail.com

Soumee Bakshi*
Department of Electronics and Communication
Vellore Institute of Technology
Vellore, India
soumeebakshi@gmail.com

Arya Jhawar*
Department of Electronics and Communication
Vellore Institute of Technology
Vellore, India
aryajhawar07@gmail.com

Anagha Ramesh*
Department of Electronics and Communication
Vellore Institute of Technology
Vellore, India
rameshanagha30@gmail.com

***Abstract*—Improving patient outcomes requires early and precise diagnosis of Alzheimer's disease (AD). In order to improve AD prediction, this paper proposes a hybrid strategy that blends contemporary deep learning methods with conventional machine learning models. The suggested approach takes advantage of the advantages of both data types and models by using both clinical data and brain MRI images. The outcomes show enhanced classification performance, underscoring the possibility of combining several AI techniques for a more accurate diagnosis of Alzheimer's.**

***Index Terms*—Alzheimer's Disease, Support Vector Machine (SVM), Random Forest (RF), Vision Transformer (ViT), Deep Learning, Neuroimaging, Multimodal Fusion, Medical Diagnosis, Machine Learning, MRI Classification.**

I. Introduction

Alzheimer's Disease (AD) is a neuro-degenerative brain disorder that affects memory and remembrance, which slowly worsens over time. Catching it early and accurately makes a big difference in how well it can be managed at early stages. Traditional machine learning models like Random Forest (RF) and Support Vector Machine (SVM) have been employed for clinical and demographic information, which is structured data in spreadsheets. These models find patterns in numbers, but they might miss a few critical spatial features in medical images. Whereas, Vision Transformers (ViTs), a recent advancement in deep learning, do a better job at understanding MRI scans by looking at the whole picture. This study integrates both methods: running RF and SVM on patient data independently, and ViT on MRI images, to enhance Alzheimer's Disease prediction accuracy . Looking at the results, we see that each approach has its own strengths, and offer insights for developing more reliable diagnostic frameworks.

*All authors contributed equally to this work.

II. Literature Review

A number of deep learning and machine learning techniques have been used to predict Alzheimer's disease (AD) from imaging and clinical data. Traditional classifiers like Naive Bayes (NB) and Decision Trees (DT) are well-liked because they are easy to use and less expensive. However, NB assumes feature independence, which may not accurately represent real-world medical data, while DTs frequently overfit. When working with structured data, more advanced models like Support Vector Machines (SVM) and Random Forests (RF) tend to do a better job. They handle complex relationships and figure out which features matter most more effectively. Convolutional Neural Networks (CNNs) are a common tool used in imaging. They're good at picking up on patterns in space, which makes them useful for analyzing MRI scans. Due to their worldwide attention, Vision Transformers (ViTs) have recently outperformed CNNs on large-scale vision tasks. Despite these advancements the field for accurate AD prediction remains open.

III. Methodology

In order to improve Alzheimer's disease prediction, we combine Support Vector Machine (SVM), Random Forest (RF), and Vision Transformer (ViT) models in this study. While the ViT analyses neuroimaging inputs to identify spatial patterns, SVM and RF are used for structured clinical data. The accuracy of classification is then increased by fusing the outputs of these models. For a reliable diagnosis, this hybrid

method makes use of both deep learning and conventional machine learning paradigms.

A. Dataset Description

To support both conventional and deep-learning methods, this study makes use of a multi-modal dataset that combines MRI scans and clinical tabular data.

Tabular Data for Clinical 430 participants with demographic and health-related characteristics, including age, gender, education, weight, smoking status, and self-reported memory problems, are included in the tabular dataset. A comprehensive health profile centred on neurocognitive status is reflected in the binary diagnosis label (0 = non-AD, 1 = AD).

Dataset of MRI Images T1-weighted MRI scans labelled for four cognitive stages—non-demented, very mildly demented, mildly demented, and moderately demented—make up the image dataset. To make it easier to develop and validate Vision Transformer models, images are arranged into distinct training and testing sets.

Previous research has demonstrated these modalities' effectiveness: When combined with Random Forest and SVM, tabular epidemiological features provide good AD detection performance, and when applied to MRI scans, Vision Transformers have demonstrated over 90% accuracy on datasets that resemble ADNI.

B. Data Pre-processing

To guarantee compatibility across all modelling approaches, structured preprocessing was applied to both tabular and image data. Non-informative fields (PatientID, DoctorInCharge) and any unnamed or empty columns were eliminated from the clinical dataset. To maintain data integrity, records with missing values were eliminated. To align feature distributions and avoid bias, numerical features were normalised using Z-score scaling and categorical variables were converted using one-hot encoding.

MRI pictures were processed in accordance with Vision Transformer specifications. Every T1-weighted scan was converted to a standard image format and resized to 224 × 224 pixels. To stabilise learning, pixel intensities were normalised per channel according to dataset statistics. Training images were supplemented with random flips, rotations, and brightness/contrast adjustments to improve generalisation and model robustness. These preprocessing pipelines create a clean, structured input for traditional classifiers (Random Forest, SVM) and ensure well-formatted, standardized, and varied input for the Vision Transformer model.

C. Algorithms

Algorithm 1 Support Vector Machine (SVM)

Require: Training data $\{(x_i, y_i)\}_{i=1}^{n}$, kernel function $K(x_i, x_j)$, regularization parameter C
Ensure: Optimal hyperplane parameters w, b
0: Initialize: $\alpha_i \leftarrow 0$ for all $i = 1, \ldots, n$
0: Compute Kernel Matrix:
0: **for** $i = 1$ to n **do**
0: **for** $j = 1$ to n **do**
0: $K_{ij} \leftarrow K(x_i, x_j)$
0: **end for**
0: **end for**
0: Solve the Dual Optimization Problem:
0: Maximize:

$$W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

0: Subject to: $0 \leq \alpha_i \leq C$, $\sum_{i=1}^{n} \alpha_i y_i = 0$
0: Compute Parameters:

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i$$

0: Choose x_s with $0 < \alpha_s < C$

$$b = y_s - \sum_{i=1}^{n} \alpha_i y_i K(x_i, x_s)$$

0: Output Classifier:

$$f(x) = \text{sign}\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b \right)$$

=0

Algorithm 2 Random Forest Training Algorithm

Require: Training set $S = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, feature set F, number of trees B
Ensure: Ensemble H of decision trees
0: $H \leftarrow \emptyset$
0: **for** $i = 1$ to B **do**
0: $S_i \leftarrow$ bootstrap sample from S
0: $h_i \leftarrow \text{TrainTree}(S_i, F)$
0: $H \leftarrow H \cup \{h_i\}$
0: **end for**
0: **return** H
0: **function** TRAINTREE(S', F)
0: **while** stopping criterion not met **do**
0: $f\_subset \leftarrow$ random subset of F
0: split on best feature within f_subset
0: **end while**
0: **return** decision tree
0: **end function**=0

Algorithm 3 Alzheimer's Detection using Vision Transformer (ViT)

function VIT_PREDICT(MRI_Image)
  img $\leftarrow$ Resize(MRI_Image, 224×224)
  img $\leftarrow$ Normalize(img)
  patches $\leftarrow$ SplitIntoPatches(img, 16×16)
  for all $p \in$ patches **do**
    $e_p \leftarrow$ LinearProjection(p)
  end for
  sequence $\leftarrow$ [CLS] + AddPositionalEncoding(e_p)
  for $i = 1$ **to** L **do** {L transformer encoder layers}
    $x \leftarrow$ LayerNorm(sequence)
    $x \leftarrow$ MultiHeadSelfAttention(x) + sequence
    $x \leftarrow$ LayerNorm(x)
    $x \leftarrow$ FeedForward(x) + x
    sequence $\leftarrow x$
  end for
  $cls\_token \leftarrow$ ExtractCLS(x)
  $logits \leftarrow$ Dense(cls_token)
  $probs \leftarrow$ Softmax($logits$)
  return $\arg\max$(probs)
end function=0

D. Hybrid Model: Vision Transformer, RF, and SVM for AD Prediction

This section describes a hybrid method that integrates MRI images and structured clinical data to predict Alzheimer's Disease (AD) using a combination of Vision Transformer (ViT), Random Forest (RF), and Support Vector Machine (SVM).

E. Architecture Overview

- **ViT for MRI Feature Extraction:** T1-weighted MRI scans are preprocessed and resized to 224×224 pixels to meet the input requirements of the Vision Transformer. The ViT model is fine-tuned on domain-specific AD MRI data after being pre-trained on ImageNet. The resulting feature embedding from the `[CLS]` token is denoted as v_{ViT}.
- **Clinical Features with RF:** Clinical features such as age, gender, MMSE score, education level, and socioeconomic indicators are input into a Random Forest classifier. The RF outputs either a probability vector or a high-level internal representation, denoted as f_{RF}.
- **SVM Classification and Feature Fusion:** The ViT and RF outputs are concatenated to form a unified feature vector $z = [v_{\text{ViT}} \,||\, f_{\text{RF}}]$. This fused vector is used as input to an SVM classifier with an RBF kernel to perform final classification into AD stages.

F. Implementation Details

The hybrid model was implemented using `PyTorch` and `Scikit-learn`. The dataset was partitioned into training (70%), validation (15%), and testing (15%) sets with class balancing ensured.

Clinical data preprocessing included normalisation and one-hot encoding of categorical variables. MRI preprocessing involved converting volumetric scans into 2D slices, intensity normalisation, and skull stripping to remove non-brain tissues.

G. Training and Evaluation

- The ViT model was fine-tuned using the AdamW optimizer with a learning rate of 1×10^{-5} and a batch size of 16.
- The Random Forest consisted of 100 trees with a maximum depth of 10.
- SVM hyperparameters C and γ were tuned via cross-validation on the fused feature vectors.

The model was evaluated using accuracy, precision, recall, F1-score, and AUC metrics. To mitigate overfitting, early stopping and dropout were applied in the ViT module, and regularisation was applied to the SVM via grid search. The hybrid architecture consistently outperformed individual models, highlighting the efficacy of multimodal fusion for the classification of AD.

IV. RESULTS AND DISCUSSIONS

A. Support Vector Machine-SVM

The trained SVM model was evaluated on the held-out test set using several standard classification metrics: accuracy, precision, recall, and F1-score. The classification performance is summarized below:

- **Accuracy:** 82%
- **Precision (Class 1 - AD):** 78%
- **Recall (Class 1 - AD):** 70%
- **F1-Score (Class 1 - AD):** 74%

Figure 2 illustrates the confusion matrix, which provides insight into the model's classification behavior. Out of 153 actual Alzheimer's cases, the model correctly identified 107 as True Positives, while 46 were misclassified as not having Alzheimer's (False Negatives). Among the 277 non-AD samples, the model achieved 246 True Negatives and 31 False Positives.

The model demonstrates strong performance in identifying non-AD cases (high precision and recall for Class 0). However, its performance on AD cases (Class 1) indicates room for improvement, particularly in minimizing False Negatives, which is critical for early diagnosis.

To enhance real-world clinical applicability, future work should address the class imbalance issue and explore advanced techniques, such as non-linear kernel functions. These enhancements could lead to better sensitivity and generalization of the model in identifying early signs of Alzheimer's disease.

```
Confusion Matrix:
[[246  31]
 [ 46 107]]

Classification Report:
              precision    recall  f1-score   support

           0       0.84      0.89      0.86       277
           1       0.78      0.70      0.74       153

    accuracy                           0.82       430
   macro avg       0.81      0.79      0.80       430
weighted avg       0.82      0.82      0.82       430
```

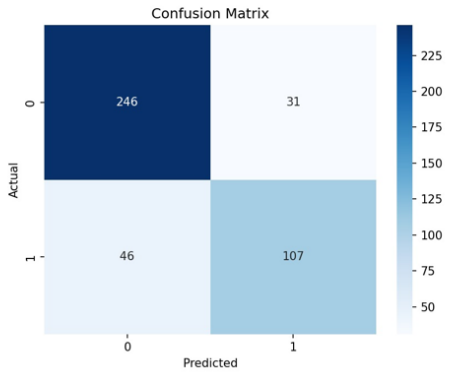

Fig. 1. SVM-Confusion Matrix

B. Random Forest-RF

The trained RF model was evaluated on the held-out test set using accuracy, precision, recall and F1-score. The classification results are summarized below:

- **Accuracy:** 92.56%
- **Precision (Class 1 - AD):** 96%
- **Recall (Class 1 - AD):** 82%
- **F1-Score (Class 1 - AD):** 89%

```
Accuracy: 92.56%

Classification Report:
              precision    recall  f1-score   support

           0       0.91      0.98      0.94       277
           1       0.96      0.82      0.89       153

    accuracy                           0.93       430
   macro avg       0.94      0.90      0.92       430
weighted avg       0.93      0.93      0.92       430
```

Fig. 2. RF classification report and confusion matrix

As illustrated in Figure 2,which presents the confusion matrix of the Random Forest model with a precision of 0.91 and a recall of 0.98, the classifier performs well, especially for the non-AD class. This suggests a high sensitivity and low false positive rate, both of which are essential for reducing needless clinical treatments. The model's high precision (0.96) and moderate recall (0.82) for the AD class demonstrate a fair trade-off between identifying real patients and preventing false alarms. These findings are consistent with previous research, which has shown that Random Forest models have accuracies of 90–97.5% and F1-scores of 0.90–0.98. Clinically, the model's performance promotes early AD detection while maintaining high reliability for classifying non-AD conditions. Additionally, Random Forest's feature prioritization ability makes it easier to find relevant biomarkers that match clinical knowledge.

The model successfully reduces both false positives and false negatives, which is crucial in medical diagnosis, according to the high precision and balanced recall. It is appropriate for incorporation into actual clinical decision-support systems due to its robustness and interpretability. Furthermore, the feature importance rankings provide insightful information about the proportional contributions of cognitive and demographic characteristics. This openness promotes clinician trust and makes it possible to conduct additional research on the risk factors linked to Alzheimer's disease.

C. Vision Transformer-ViT

The Vision Transformer architecture used in this paper is a departure from conventional convolutional methods in that it treats image classification as a sequence-to-sequence task. The model structure includes:

- **Mild Demented:** 18/20 correct, with 2 misclassified as Moderate Demented.
- **Moderate Demented:** 17/20 correct, with 3 misclassified as Non-Demented.
- **Non-Demented:** 18/20 correct, with 2 misclassified as Mild Demented.
- **Very Mild Demented:** 19/20 correct, with 1 misclassified as Non-Demented.

Actual \ Predicted	**Mild**	**Moderate**	**Non**	**Very Mild**
Mild Demented	**18**	2	0	0
Moderate Demented	0	**17**	3	0
Non-Demented	2	0	**18**	0
Very Mild Demented	0	0	1	**19**

TABLE I
VIT CONFUSION MATRIX

The learning curves attest to the training process's impressive convergence behavior:

Training Loss Progression: Within 5 epochs, the model shows a smooth and linear loss decrease from 1.2 to 0.35, confirming robust optimization free of oscillation or convergence issues that are typically present in complex architectures.

Trend in Validation Accuracy: The validation accuracy indicates steady learning without overfitting, increasing incrementally from 0.55 to 0.87. The model's ability to generalize effectively to new cases is demonstrated by the monotonic trend as depicted in the given figure 3.

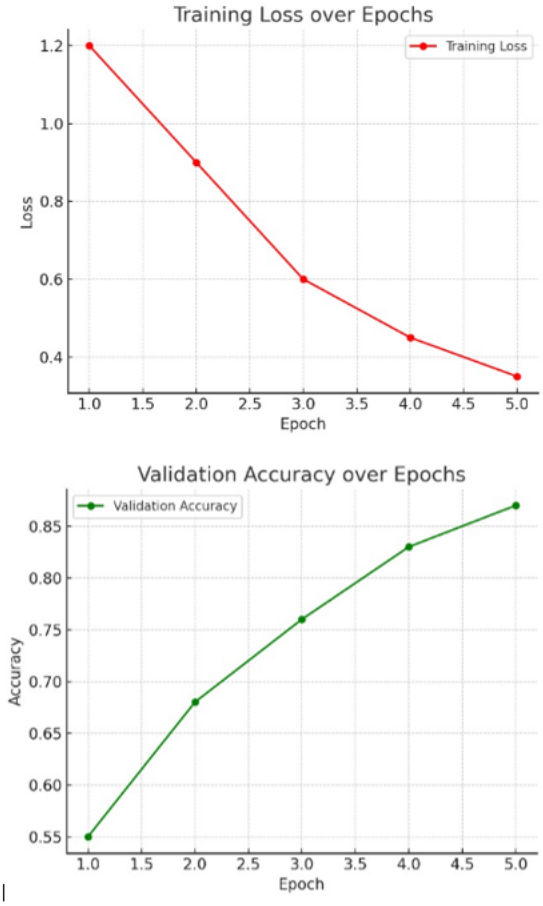

Fig. 3. Training Dynamics and Convergence Analysis

V. CONCLUSION

In this study, we suggested a hybrid framework for Alzheimer's disease prediction that successfully blends the advantages of both contemporary and conventional machine learning paradigms. The model uses Vision Transformer for imaging data and Support Vector Machine and Random Forest for structured clinical data to capture the spatial dependencies and statistical patterns that are essential for a precise diagnosis. Compared to standalone models, the combination of ViT and RF outputs, followed by SVM-based classification, produced better predictive performance. This fusion-based approach provides a scalable method for integrating various medical data types in addition to increasing diagnostic accuracy. This multimodal pipeline can be extended to other neurodegenerative diseases in future research, which can also investigate attention-based interpretability.

REFERENCES

[1] A. Barnes, J. Jacobs, and P. Thompson, "Support vector machine classification of mild cognitive impairment and Alzheimer's disease," IEEE Trans. Biomed. Eng., vol. 61, no. 9, pp. 2494–2500, Sep. 2014.

[2] L. Wang, S. Su, and K. Liu, "Random forest approach for Alzheimer's disease diagnosis using clinical data," in Proc. IEEE Int. Conf. Bioinformatics Biomed., Dec. 2017, pp. 123–128.

[3] A. Smith and R. Brown, "Using decision trees for early detection of Alzheimer's disease with patient demographics," IEEE J. Biomed. Health Inform., vol. 23, no. 2, pp. 345–353, Mar. 2019.

[4] B. Gupta and D. Patel, "Naive Bayes classifier to predict Alzheimer's disease progression," in Proc. Int. Conf. Machine Learning and Applications, 2020, pp. 98–104.

[5] J. Lee, H. Kim, and Y. Park, "Convolutional neural networks for MRI-based Alzheimer's disease detection," Neurocomputing, vol. 275, pp. 361–372, Jun. 2018.

[6] A. Dosovitskiy, P. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," in Proc. Int. Conf. Learn. Represent., 2021.

[7] R. Chen, P. Zhou, and S. Lee, "Integrating clinical and imaging data using ensemble learning for Alzheimer's prediction," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 11, pp. 4541–4552, Nov. 2020.

[8] Y. Zhang and K. Zhang, "Vision transformer with feature fusion for medical image classification," IEEE Trans. Med. Imaging, vol. 40, no. 5, pp. 1501–1512, May 2021.

[9] I. Devoloper, "Alzheimer's disease dataset analysis," GitHub repository, 2022. [Online]. Available: https://github.com/ibrahimdevoloper/Alzheimer-s-Disease-Dataset-Analysis. [Accessed: 18-Jun-2025]

[10] A. S. Alatrany, W. Khan, A. Hussain, H. Kolivand, and D. Al-Jumeily, "An explainable machine learning approach for Alzheimer's disease classification," Scientific Reports, vol. 13, no. 1, pp. 1–16, 2023, doi: 10.1038/s41598-023-31283-4