

# Hybrid PCA–ACO–SVM Approach for Multi-Stage Classification of Alzheimer’s Disease using Brain MRI Images

G. Sravani<sup>1</sup>, Chiranjeevi Althi<sup>2</sup>, Dr. T. Ramashri<sup>3</sup>

<sup>1</sup>M. Tech Scholar, Dept of ECE, SVU College of Engineering, S.V. University, Tirupati, AP, India

<sup>2</sup>Academic Consultant, Dept of ECE, SVU College of Engineering, S.V. University, Tirupati, AP, India

<sup>3</sup>Professor, Dept of ECE, SVU College of Engineering, S.V. University, Tirupati, AP, India

[doi.org/10.64643/IJIRT12I6-187851-459](https://doi.org/10.64643/IJIRT12I6-187851-459)

**Abstract**—Alzheimer (AD) is a neurological complaint of the brain that impacts primarily on thinking capacities and makes one forget. This is a common disease in the number of people today and it is important to determine the level of AD. This suggested architecture divides the brain MRI data into four phases using a variety of machine learning techniques. The framework begins with a preprocessing method to improve MRI image clarity, then performs Principal Component Analysis (PCA) to reduce dimensionality, then selects the most informative feature using Ant Colony Optimization (ACO), and finally classifies the selected features with Support Vector Machine (SVM) using the RBF kernel for non-linear classification. To evaluate the system's performance, metrics including recall, accuracy, precision, and F1-score are employed. The confusion matrix and chosen feature mask representations are also included in the evaluation. As a result, the simulation results were 98.52% accurate, outperforming earlier methods based on machine learning.

**Index Terms**— ACO, Four Stages of Alzheimer's disease, machine learning, MRI images, PCA, SVM.

## I. INTRODUCTION.

Deterioration of brain cells occurs over time in Alzheimer's disease [1,2]. The majority of the world's elderly are suffering, yet it is tough to notice. It is mostly affected by brain injuries, aging, stress, sadness, anxiety, and diet [3]. Symptoms such as forgetfulness, disorientation, and inability to put things back where they belong are hallmarks of Alzheimer's disease [1,4]. The most important way to avoid Alzheimer's disease is to detect it in its early stages [5]. There are four main phases of Alzheimer's disease:

Stage 1 is non-demented; it is the first stage of Alzheimer's disease, with no symptoms and a healthy appearance of the human brain. Stage 2 is Very Mild-Demented; this is the second stage, and brain changes begin, but they may be misinterpreted as normal aging symptoms, such as misplacing objects, forgetting names, and having difficulty recalling recent events [2]. Mild-Demented is the third stage; there are apparent symptoms that interfere with daily life, such as forgetting conversations, events, or places (behavioral changes may occur) [3,4]. The final stage is Stage-4, also known as Moderate-Demented; this is the final stage of Alzheimer's disease, and the symptoms are severe, and individuals struggle in their everyday lives. The symptoms include extreme memory loss; they are unable to recollect their personal history or the names of close family members [1,6].

Magnetic Resonance Imaging (MRI) is a highly effective brain scan in the medical industry. It provides high-resolution brain information for people with Alzheimer's disease [7,8]. Images from magnetic resonance imaging (MRI) of the brain may provide particular information about brain tissues, such as whether they are completely normal, mildly abnormal, severely abnormal, or severely demented [9,10]. MRI brain scans can reveal comprehensive information concerning AD, such as hippocampus shrinkage, ventricular enlargement, and cortical thinning [7,8]. Diagnosing the stage of Alzheimer's disease using brain imaging is not always easy, and the process is time-consuming and prone to errors [5,8]. Time savings, improved early detection, and the elimination of human error are all possible outcomes of integrating

and categorizing brain scans using a variety of machine learning approaches [1,3,9].

## II. METHODOLOGY

### 2.1. Principal Component Analysis (PCA)

One phase in unsupervised machine learning that might aid in dimensionality reduction is principal component analysis. Principal component analysis (PCA) is a data reduction technique that keeps the most important information and provides main components while reducing the number of characteristics (dimensions) [1,9]. Principal components are new variables that are used to combine the original features [1,9]. PCA is also used for feature extraction, which involves extracting features from MRI scans [10]. Feature extraction is the process of finding and extracting the most relevant information from raw photos, such as edges, corners, texture, shape, and color.

### 2.2. Ant Colony Optimization (ACO)

After extracting the image's features, it is processed for feature selection using metaheuristic optimization. Metaheuristic optimization is simply the process of applying problem-solving approaches inspired by nature, such as ants and birds [3,4]. ACO is an algorithm which is naturally inspired and simulates the foraging behaviour of actual ants. The ants locate the food particles after taking the shortest path possible within the colonies and bring food back to their homes. They do so by emitting pheromones into them

develops environment to motivate other ants. Incidentally, the features recovered are then selected as per the artificial ants. ACO selects the most appropriate details out of the features collected since, PCA can give redundant details. SVM is then used to classify the selected traits into diseased or non-diseased ones.

### 2.3. Support Vector Machine (SVM)

A supervised machine learning technique called Support Vector Machine (SVM) is often used for regression and classification tasks [1,8,9]. In nonlinear classification, this technique is applied to high-dimensional features and a radial basis function (RBF) kernel. The RBF kernel is often used to find common features across the four stages of Alzheimer's disease (AD), including non-dementia, very mild dementia, mild dementia, and moderate dementia. It works well with small set of data and reduces the repetition of data.

### 2.4. Block Diagram

We shall consider the suggested framework that is aimed at classifying the Alzheimer disease into four stages: a state of mild dementia, moderate dementia, severe dementia, and non-dementia. Image capture, preprocessing, feature extraction, dimensional reduction, feature selection, and classification are just a few of the many steps in the proposed system. Figure 1 illustrates a block diagram of the proposed framework.

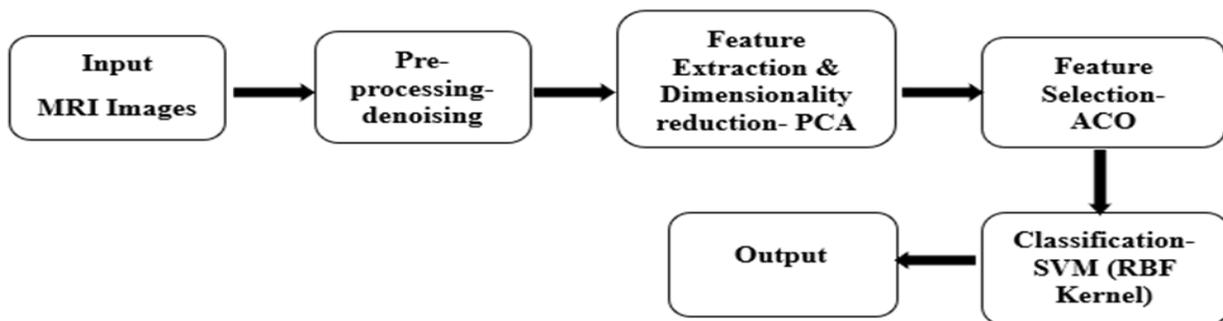


Figure 1. Block Diagram of the Proposed Framework

#### 2.4.1. Dataset description.

The 6,400 brain MRI scans in the Kaggle datasets were classified as no dementia, mild dementia, moderate dementia, or very mild dementia. Classification

testing accounts for 20% of the dataset, whereas training accounts for 80%.

2.4.2. Pre-processing During the pre-processing phase, the original image will be denoised, scaled, sharpened, and normalized.

2.4.2.1 Denoising using Non-Linear Mean Filter (NLM) This is an advanced denoising filter that is better suited for MRI images. It is mostly used for global denoising (finding comparable pixels throughout the entire image), which includes edges, textures, and minor structures.

2.4.2.2 Resize It is used to resize the original MRI pictures to a size of 128x128 pixels.

2.4.2.3 Normalization Ranging pixel intensity levels from 0 to 1.

2.4.2.4 Sharpening Finally, an image sharpening filter is employed to improve clarity in the image's edges and boundaries, allowing for better classification of AD.

2.4.3 Feature extraction and dimensionality reduction with PCA.

We use principal component analysis (PCA) to downsize features and preprocess data after preprocessing. By reducing a high-dimensional feature to a low-dimensional feature for improved classification. It is used to extract the most significant feature following preprocessing.

2.4.4 Feature selection with ACO

ACO picks the most informative features by using PCA's extracted features as input. After a few cycles of altering the number of ants, the pheromone evaporation rate determines the best feature among the extracted features. The best features will be chosen

after iterations to improve AD classification utilizing brain imaging.

2.4.5 Classification Using SVM

Because of the non-linear classification of MRI images, after selecting the most informative features from ACO, the RBF kernel is employed to categorize the Alzheimer's stage. Finally, a classification report is built using the following metrics: accuracy, precision, recall/sensitivity, F1 score, and confusion matrix. The final predictions are displayed as heatmaps.

### III. RESULTS AND DISCUSSION

The suggested framework PCA-ACO-SVM has been used to Kaggle 6400 brain MRI pictures from four stages: varying degrees of dementia, from completely non-existent to mild, moderate, or severe. The results are presented and the framework performance is evaluated using metrics such as precision, accuracy, recall, and F1 score on a sample of 1280 samples.

3.1 Confusion matrix

The confusion matrix contrasts the system's projected genuine positive values with test samples from datasets. In Image 2 below, you can see the confusion matrix. The true positive predictions from the Alzheimer's MRI datasets, with the remaining values being false positive predictions. Sample prediction is low in MRI datasets with moderate dementia.

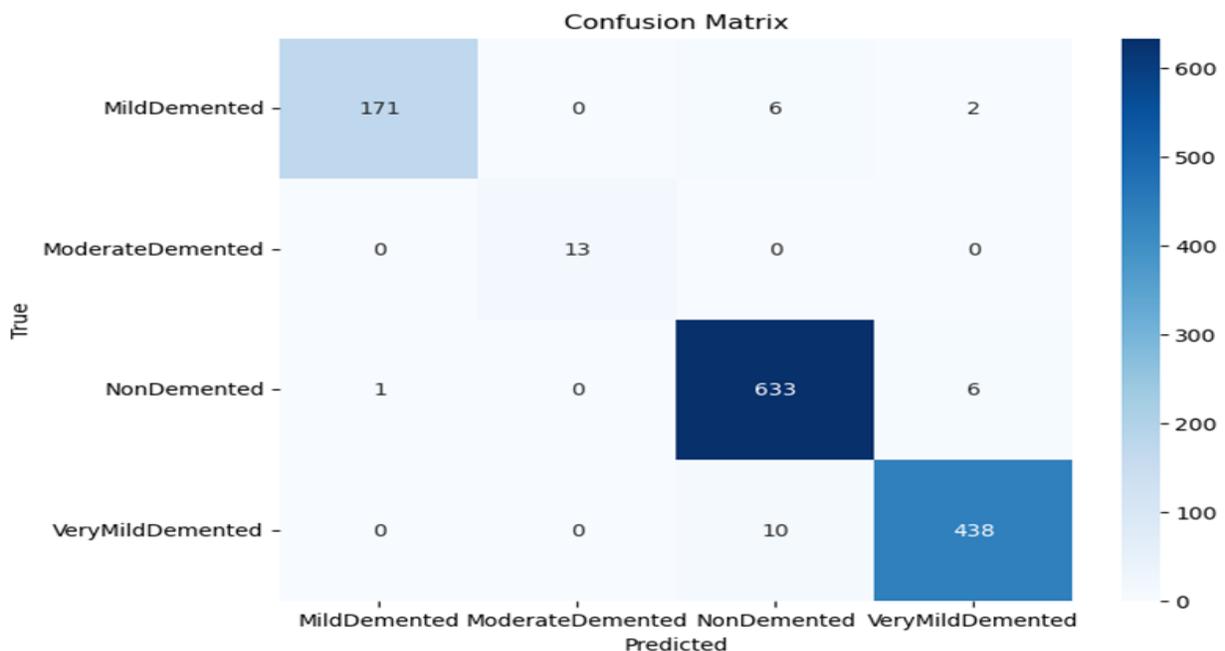


Figure 2: A Framework Confusion Matrix, Proposed.

3.2 Visualization of Stepwise Processing.

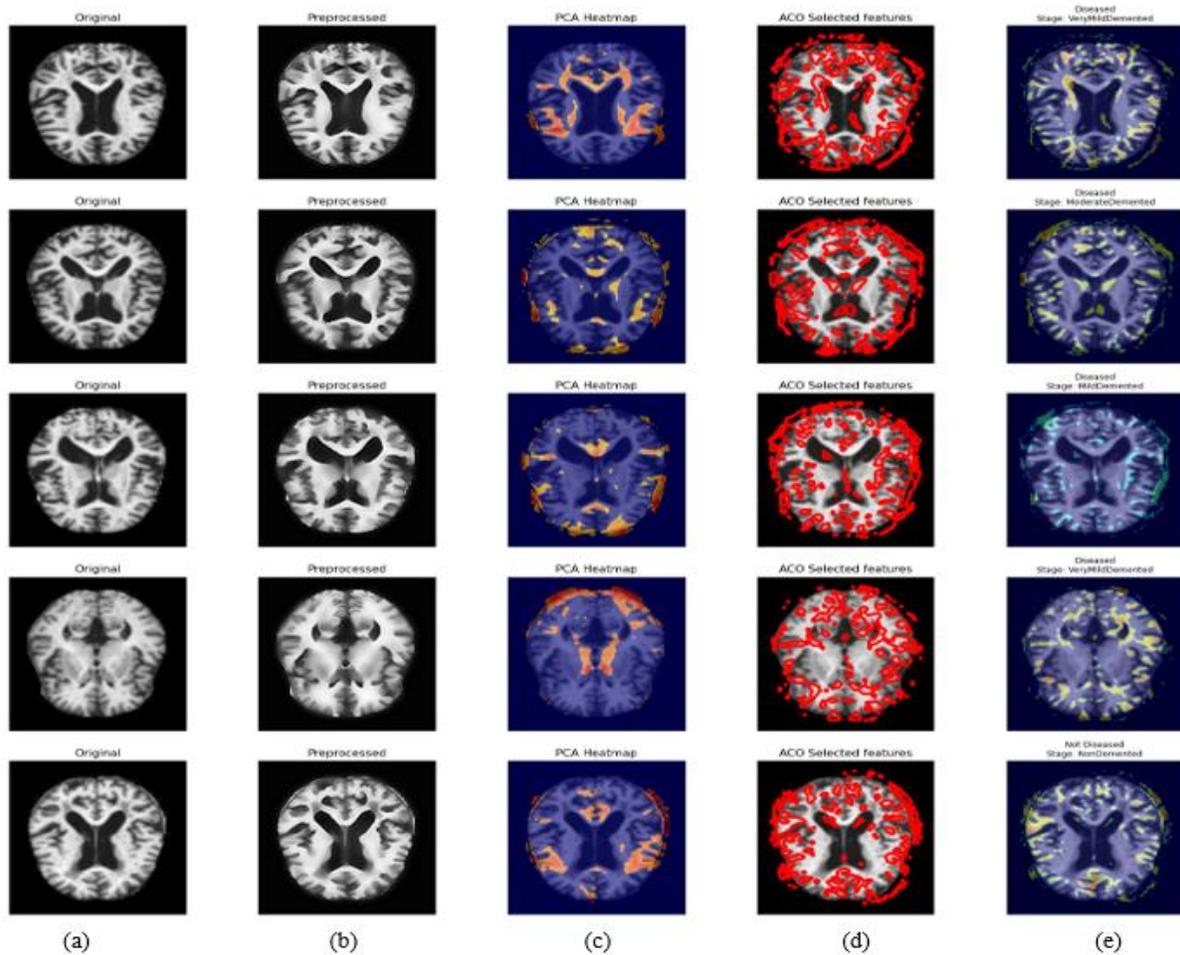


Figure 3: Step-By-Step Depiction of the Suggested Framework

Figure 3 provides a detailed representation of the PCA-ACO-SVM framework that has been developed. Improving Alzheimer's disease classification using MRI data.

- a). The raw MRI image
- b). Pre-processing denoising, scaling, normalising, and sharpening the image
- c). PCA heatmap displays the variance derived from a picture via PCA.
- d). ACO ROIselected image features obtained after PCA extraction with ACO.
- e). Final prediction this is the image's final classification, including infection status.

3.3. Classification Report

The ACO selects 125 of the most informative features from 200 PCA reduced components. The RBF kernel

SVM non-linear classifier achieves 98.52% accuracy, as illustrated in the table- 1 below.

Formulas of the performance metrics:

3.3.1 Accuracy. The proportion of correct predictions relative to all predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3.3.2 Precision. The likelihood of correctly forecasting cons.

$$Precision = \frac{TP}{TP + FP}$$

3.3.3 Recall is the probability that actual drawbacks will be accurately anticipated.

$$Recall = \frac{TP}{TP + FN}$$

3.3.4 F1-score. a harmony between recall and precision. Balances both.

$$F1 = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are the variables in this context.

Table.1 Comparison Results of Existing and Proposed Methods

S. No	Performance Metrics (%)	Existing method (PCA-SVM) [1]	Proposed method (PCA-ACO-SVM)
1	Accuracy	95	98.52
2	Precision	0.9730	0.9879
3	Recall	0.9319	0.9805
4	F1-score	0.9510	0.9841

### 3.4 Metrics per-class (4-stages of AD)

Table. 2 Shows the Classification Report for PCA-ACO-SVM

S.No	Class	Precision	Recall	F1-Score	Support
1.	Mild-Demented	0.99	0.96	0.97	179
2.	Moderate-Demented	1.00	1.00	1.00	13
3.	Non-Demented	0.98	0.99	0.98	640
4.	Very Mild-Demented	0.98	0.98	0.98	448
Overall Accuracy		98.52%			1280

The F1-score macro accuracy score was 0.9841, while the recall and precision values were 0.9805 and 0.9879, respectively. A total of 98.52% accuracy is achieved when all four types of Alzheimer's disease are considered together. The below formulas are used to calculate classification report. The suggested framework outperforms the existing approaches. The current PCA-SVM hybrid technique has a 95% accuracy rate [1]. The proposed PCA-ACO-SVM, hybrid framework achieved an accuracy of 98.52%. It is improved by including ACO in feature selection and categorization.

The great accuracy was achieved by combining Principal Component Analysis, Ant Colony Optimization, and Support Vector Machine using an RBF kernel. In this approach, the confusion matrix and output step-by-step visualization are highlighted in the process of identifying four phases of Alzheimer's disease using brain MRI data. In future investigations, the low categorization of intermediate dementia could be improved by including additional datasets such as PET brain scans.

## IV. CONCLUSION.

In this study, we developed a strategy to classify Alzheimer's disease cases using brain magnetic resonance imaging (MRI) scans. Integrating dimensionality reduction and feature extraction with ant colony optimization (ACO) and support vector machine (SVM) classification using an RBF kernel, the suggested system is a powerful tool for data science. With a recall, accuracy, and F1-score of 98.52%, the proposed framework was able to simulate successful outcomes. Finally, a step-by-step display of the framework PCA heatmap, and ACO ROI enhanced the model's performance.

## V. FUTURE WORK.

The suggested framework produces good results, although the moderately demented datasets are limited. In future investigations, incorporating multimodal datasets such as PET and fMRI will improve the early diagnosis of Alzheimer's disease.

## REFERENCES

- [1] O.A. Dara, J.M. Lopez-Guede, H.I. Raheem, J. Rahebi, E. Zulueta, and U. Fernandez-Gamiz. Machine Learning for Alzheimer's Disease Diagnosis: A Survey (2023). Applied Science 13, 8298. <https://doi.org/10.3390/app13148298>.
- [2] S. Liu, Y. Liu, and C. Li (2022). "Multi-class Alzheimer's disease classification using machine learning techniques," *Frontiers in Aging Neuroscience*, vol. 14, no. 874973.
- [3] Keles, M.K. & Kilic, U. (2022). Classification of Brain Volumetric Data for Alzheimer's Disease. Using the Artificial Bee Colony Algorithm as a feature selection method. *IEEE Access* 10, 82989–83001.
- [4] R. Abu Zitar, L. Abualigah, J. O. Agushaka, A.E. Ezugwu, and O.O. Akinola (2022). Metaheuristic optimization methods are being utilized to investigate multiclass feature selection. *Neural Computer Applications*, 34: 19751–19790.
- [5] Adeghal K. (2021). Alzheimer's disease can be detected by deep learning-based classification of multimodal MRI images. *Image processing*

- [eess. IV], Université Ibn Zohr, Agadir, Morocco; Bordeaux, France.
- [6] Coretto, P., and L. Coraggio (2023). Determine the overall number of clusters, models, and algorithms. A quadratic discriminant-based unifying method. *Journal of Multivariable Analysis*, 196:105–181.
- [7] Song, X., Liu, L., Liu, S., and Zhang, Y. *Sustainability*, 13: 1224. (2021) A review of deep learning-based techniques for medical image segmentation.
- [8] Hybrid metaheuristic–support vector machine models for biomedical data classification: A review by M. Khishe and M. R. Mosavi, *Applied Soft Computing*, vol. 92, p. 106308.
- [9] R. Gupta, N. Juneja, and S. S. Chauhan (2020) "PCA-based feature reduction and SVM classification for early diagnosis of Alzheimer's disease," in *Procedia Computer Science*, vol. 167, pp. 2410–2249.
- [10] S. Basaia, F. Agosta, M. Wagner, et al. (2019) "A classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," *NeuroImage: Clinical*, volume 21, page 101645.