International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 15 (2023) pp. 321-331

Multi-Class Classification of Alzheimer's Disease Stages Using SqueezeNet based Approach for Automated Diagnosis

Anil Kumar Pallikonda[1] & P. Suresh Varma[2]*

[1]Assistant Professor, Department of CSE, PVP Siddhartha Institute of Technology,
Research Scholar in University College of Engineering, Adikavi Nannaya University.
Rajahmundry, Andhra Pradesh 533296.
anilkumarnew857@gmail.com

[2] Professor,Department of CSE, University College of Engineering, Adikavi Nannaya University
Rajahmundry, Andhra Pradesh 533296.

Abstract: **A neurodegenerative disorder is known as Alzheimer's disease (AD). AD causes cognitive impairment and memory loss as a result of brain cell death. Although research on AD has improved dramatically over the years, however, the early detection of this disease is difficult because of the complexity of the brain structure and its functions. This research mainly focuses on multiple stages of AD classification. In this paper, proposed is a SqueezeNet with Harris Hawks Optimization technique (HHO-SqueezeNet) for classifying the stages including MCI, LMCI, EMCI, AD, and CN. Principal component analysis (PCA) is used to minimize the feature dimension after MRI pre-processing, and also the feature set is selected by introducing the CML-ELM approach for each task to consider the intrinsic relevance of several related tasks. The accuracy, sensitivity, specificity, precision, and recall of the models are used to assess their performance. We discovered that our networks were capable of accurately classifying the subjects. Using the proposed model, we enhanced accuracy for all AD stages using the proposed model, with CN, EMCI, MCI, LMCI, and AD achieving 98.5 %, 98.54 %, 98.25 %, 99.02 %, and 99.2 %, respectively. We achieved better classification results with an average accuracy of 98.82% based on overall performance. The MTL algorithm based on SqueezeNet established in this research is an efficient, improved, and practical technique for diagnosing AD.**

Keywords: Alzheimer's Disease (Ad),Multiple Stages, Principal Component Analysis, A Neurodegenerative Disorder, And Feature Dimension.

I. Introduction

Mainly due to AD, dementia affects the majority of senior adults. Alzheimer's disease is a progressive form of dementia that evolves from mild to moderate to severe [1]. In the brain abnormal protein growth, it is occurred, such as tau tangles and amyloid plaques. Changes in the brain occur in the preclinical stage decades before AD is diagnosed [2]. Mild cognitive impairment (MCI) is the next stage, in which memory and cognitive functions are affected at a minor but perceptible level. Dementia is the final stage, which is characterized by memory loss and difficulties with everyday activities. None of the available treatments are more effective in curing the condition. Early detection of the disease is critical for the efficient implementation of preventive measures [3].

Neuroimaging is a significant technique for disease diagnosis in neurodegenerative disorders like Alzheimer's disease, as it can uncover abnormalities that relate to the early stage of the treatment [4]. Positron emission tomography (PET) and magnetic resonance imaging (MRI) have recently been shown to be beneficial in determining the neurophysiological aspects of MCI and AD [5]. Moreover, neuroimaging has revealed that a total of 15 biomarkers can be used to diagnose AD and MCI [6].

The loss of neurons and synapses in the cerebral cortex and certain subcortical regions characterizes Alzheimer's disease, as a result, the affected areas experience severe atrophy, causing parietal and temporal lobes' degeneration, as well as cingulate gyrus and frontal cortex region's degeneration [7]. Various researches have linked specific symptoms of AD to atrophy in different parts of the brain, like the amygdala, temporal lobe, and hippocampus [8]. Moreover, Alzheimer's patients are distinguished from cognitively normal (CN) people by using the identification of major atrophied regions across the entire brain [9].

Alzheimer's disease affects people of all ages which is a multi-class problem but a lot of research is focused on binary classification, which simply shows whether or not an individual has AD [10] [11]. This is insignificant in the diagnosis of AD because the stage of the disease is more important [12]. Multiple clinical assessments are required for the AD diagnosis, resulting in a significant amount of data samples. As a result, manually analyzing data for AD stage detection is not possible. Evidence suggests that people with AD have both structural and functional abnormalities in their brains, including grey matter volume reduction [13]. But these observations are generally based on statistical comparisons at the group level. As a result, they're only useful for diagnosing diseases on an individual basis [14]. MRI and machine learning algorithms are used to solve these problems [15]. The effectiveness of computer-based decision-making in detecting early-stage diseases is demonstrated.

Earlier, AD is diagnosed with the help of clinical measures such as the clinical dementia rating (CDR) [16]. AD is diagnosed early by using multiple machine learning techniques with neuroimaging images like PET, and MRI due to significant improvements in computer and imaging techniques. [17] [18]. A single modality or a combination of

modalities is used for multiple investigations. The single modality is used to simplify the task and save money and time, but it also leads to more inaccurate predictions [19] [20].
Multiclass classification of multiple AD stages is conducted in this study that includes AD, CN, MCI, EMCI, and LMCI. It can be challenging to classify data with similar characteristics across different classes. As deep learning has gained wide popularity in nearly all areas of image processing and computer vision in recent years, this paper's deep learning model is adopted to multiple AD stage classification. By training our models around the ADNI dataset as well as progressive classification of the dataset, we evaluated their performance. The results obtained from the deep feature-based method outperform other methods.
The main contribution of our research is

- First, the pre-processing is performed for removing the noise by using a high boost filter and adaptive histogram equalization.
- After that, the PCA approach is used to extract the features and some of the features are selected by using the CML-ELM algorithm.
- For the classification problem, a novel SqueezeNet-based multitask learning system is examined. For enhancing the model's robustness and generalization ability, specifically the HHO parameter optimization approach is used in DBN
- The ADNI dataset is used for the multiple AD stage classification, the experiments are performed in the Python platform. It's employed to compare and contrast the proposed and existing methods. According to the experimental results, the SqueezeNet-HHO outperforms the previous techniques.
- The research's remaining parts are structured as follows. The various literary works related to this work are explained in section 2. Section 3 describes the proposed method for the classification of the multiple stages of AD. Section 4 provides the experimental details and results. Finally, section 5 has the conclusion.

II. Literature Survey

In a multi-task learning algorithm, a new deep belief network (DBN) is developed by Zeng et al [21] for the problem of classification. The overfitting problem is addressed by using the zero-masking approach and dropout technology and also this method enhances the model's robustness and generalization ability. Then, an AD diagnosis is accurately determined by establishing a novel mechanism depending on the DBN-based multi-task learning. Then, the feature dimension is reduced by developing the principal component analysis, and features are selected by introducing a multi-task feature selection method to entire tasks resulting from considering the internal relevance of several interrelated tasks.

For performing effective feature selection and local similarity learning Zu et al [22] developed a new multi-modality feature selection method in this paper. Multiple imaging modalities are combined specifically by creating a similarity matrix. Simultaneously, a sparse norm constraint is used to conduct feature selection.

In the diagnosis of AD, a transfer learning method is proposed by Zhou et al [23]: As a characteristic parameter, the volume of grey matter (GM) tissue in the anatomical region was calculated after the MR image segmentation. The Spm-Dartel toolbox is used to segment the images and Automatic Anatomical Labeling (AAL) registers the images. For feature selection, the information gained was created. Data from the AD Neuroimaging Initiative (ADNI) database into AD, MCI, and normal controls (NC) are classified by using the TrAdaboost algorithm, while ADNI's "knowledge" was applied to AD samples obtained from a local hospital. Four traditional algorithms were used to calculate classification accuracy, specificity, and sensitivity.

Using consistent metric constraint (MFCC), a different multi-modal neuroimaging feature selection method is proposed by Hao et al [24] for AD diagnosis. Furthermore, pairwise similarity measures are extracted by using the random forest technique for several modalities, and for each modality, the similarity is determined (i.e. VBM-MRI or FDG-PET) separately. The objective function is then constrained using sample similarity constraint regularisation and the group sparsity regularisation to perform feature selection from different modalities. At last for final classification, the selected features are fused by using a multi-kernel support vector machine (MK-SVM) from several models.

Hon et al [25] proposed a new method for solving the difficulties with transfer learning like Inception and VGG are with pre-trained with weights and only a limited number of MRI images are used to retrain the fully connected layer from huge benchmark datasets with natural images. More significant slices are selected by employing the image entropy for training.

III. Proposed Methodology

For AD classification, this study resulted in the development of a more advanced computer-aided system design. Alzheimer's subjects are classified by proposing this methodology into three different stages: image acquisition, pre-processing, feature selection, and classification followed by evaluation. A widely-known database on Alzheimer's disease contains neuroimaging data. In pre-processing, artifacts and noise are removed from data. After Pre-processing, using the expectation-maximization algorithm for segmentation and the features are extracted by PCA. Then some of the features are selected by using the CML-ELM algorithm. After that use SqueezeNet-HHO for classifying the Alzheimer's stages. These processes are graphically represented in Figure 1.

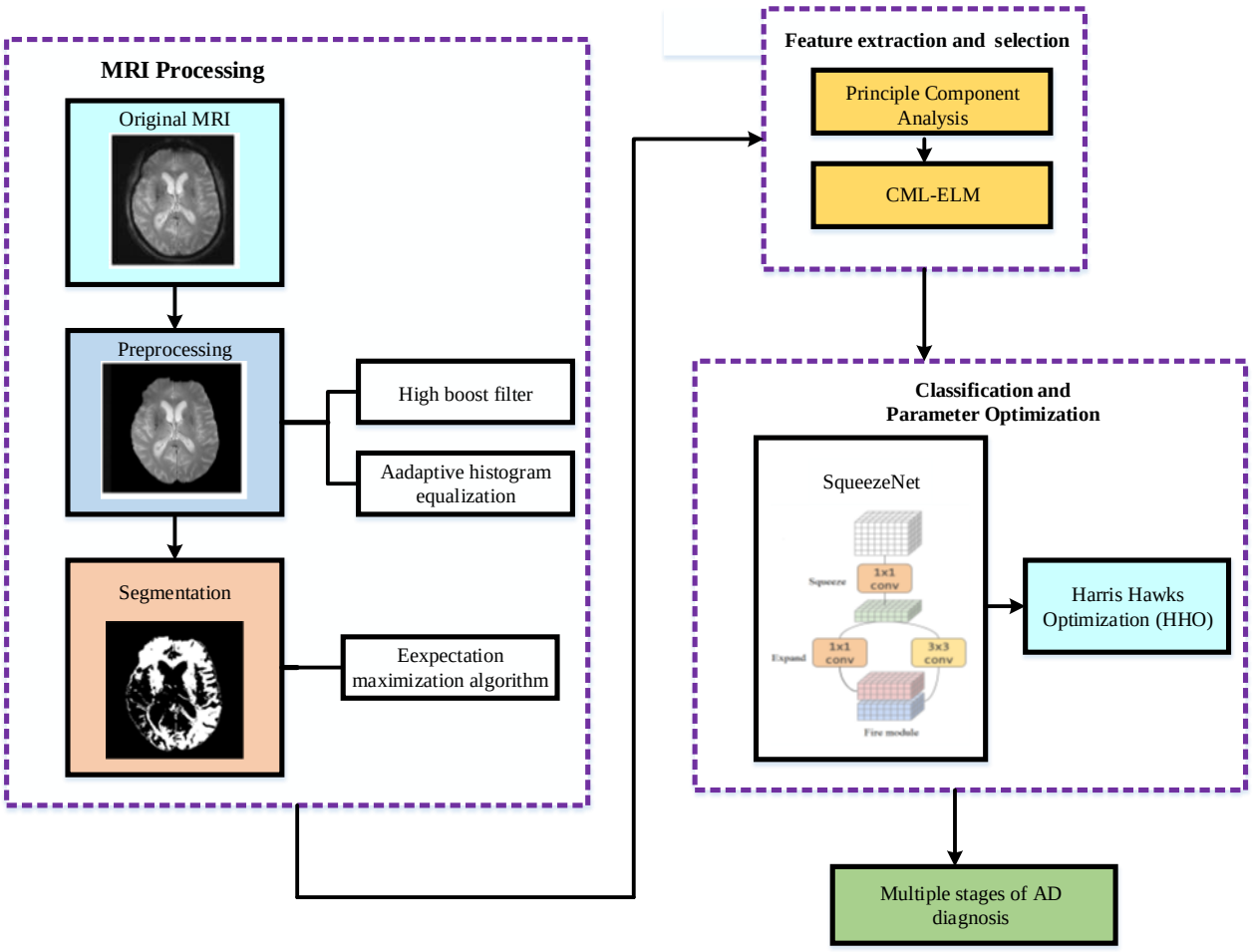


Figure 1. Schematic diagram of the proposed multiple-stage classification of AD diagnosis

A. Problem Statement

Due to the overlapping features of distinct phases of AD, evaluating them is a difficult task. Binary classification is focused on the majority of research in the literature such as the appearance of AD based on neuro-scans. The classification of two or more phases of this disease receives little attention. The goal of this study is to do an AD diagnosis of multiple stages, which comprise CN, EMCI, MCI, LMCI, and AD. It's difficult to data classification with comparable characteristics into various classes. Large datasets with ground truth labels are commonly accessible, which is a difficulty. The SqueezeNet-HHO methodology is proposed to solve this issue, and the performance is improved by using the feature selection method additionally.

B. Data acquisition

The image collection is the first step toward detecting Alzheimer's disease (AD). The ADNI website (http://www.loni.ucla.edu/ADNI) is used to collect the data, which is available publicly. ADNI was designed to measure MCI and early AD progression by combining other biological markers, Positron Emission Tomography (PET), serial MRI, and MCI, and early AD can be evaluated using clinical and neuropsychological assessments.

C. Pre-processing

Because it's tough to see the anatomical structures due to the low contrast of the input images, the noise is removed and image contrast is enhanced by performing smoothing of the image and enhancement in the pre-processing stage. Noise removal and contrast enhancement are done by using the adaptive histogram equalization and high boost filter. The high-frequency components are emphasized by using a high boost filter.

$$\text{Highboost} = (A) * (\text{original}) - \text{lowpass}$$

The amplification factor is represented by A. The Alzheimer's disease region is hard to segment effectively. Adaptive histogram equalization is used for improvement to represent the impacted region contrast. The medical image contrast is improved by using this method commonly. The image is segmented by using the expectation-maximization algorithm after pre-processing. Moreover, the regionally normalized brain image is segmented into three main parts by employing spatial normalization and tissue segmentation. Depending on the prior probability distribution of the brain and the image grey value, the bias-corrected normalized total brain, the modulated normalized white matter (WM), and grey matter (GM) are segmented. The GM images' signal-to-noise ratio is improved by implementing the smoothing process finally.

D. Feature Extraction and Feature selection

Smoothed GM images of five classes should be highlighted since they are so identical that they are tough to predict visually. Every fMRI can yield 90 characteristics. In this research, resulting in 32,490 totally features, although many redundant ones are extremely prevalent. As a result, obtaining highly discriminating characteristics through dimension reduction is critical for a better diagnosis.

1) Principal component analysis

Principal component analysis (PCA) is a well-known dimension reduction method that can considerably enhance computational performance. The primitive $X^{M\times N}$ features are mapped towards the novel k features (k\n) by using the technique of PCA. The original data matrix is denoted by, where the obtained features and number of subjects are described by n and m. The cth aspect of the dth subject is denoted by X_i^j :

$$\tilde{X} = X - \frac{\sum_{c=1}^{m} x_c^d}{m} \qquad (1 \le d \le n) \qquad (1)$$

$$COV = \frac{\tilde{X}.\tilde{X}^T}{n-1} \qquad (2)$$

Where the centralization matrix is represented by $\tilde{X}$, this is identical to the original matrix X in terms of information. The ith aspect of the dth subject is denoted by X_c^d

The eigenvalues $D = \{d_1, d_S, \ldots d_n\}$ and the $V^{m\times n}$ of eigenvector matrix are then obtained via a linear transformation. The highest k eigenvectors of the matrix D_K are then set to V_k . Then the newly created data matrix DM is represented by

$$DM = V_k^T.\tilde{X} \qquad (3)$$

In particular, each principal component's contribution rate is represented by the histogram (namely variance), and the sum of the contribution rates is represented by the curve (namely histogram value). In this research, if the rate of cumulative contribution is adjusted to 95 percent, then reducing each fMRI's 90 features using the PCA method.

2) Feature Selection

Overlap between features is eliminated by the feature selection process, it is the significant objective of the research and choose only those features that have been strong enough to be classified effectively. Reducing the number of predictors is the second goal of this step, which aided in the testing process's speed. Through ELM and mutual learning (CML-ELM), we used a process termed correntropy to inspire these two crucial characteristics. Algorithm 1 demonstrates how this method works.

Algorithm 1
CML-ELM feature selection method

Input: $\eta c, \eta c \in \{\eta_1, \eta_2, \eta_3, \ldots, \eta_c\}$

Output: $Q_w(c), Q_w(d) \in \{Q_w(1), Q_w(2), \ldots., Q_w(c)\}$

Start

Step 1: Parameters Initialization

$$Q_w(1) = \eta_c,\ c = 0,1,2,3,\ldots n$$

$$\alpha - 1 = 0$$

$$\alpha_0 = 1$$

$$RL = RL_0$$

Step 2: For c = 1 **to K do**

$$b_c = \frac{\alpha_{c-2} - 1}{\alpha_c - 1}$$

$$A^c = Q_w(c) + b_c\,(Q_w(c) - Q_w(c-1))$$

Step 3: Update $Q_w(c+1)$

Step 4: Determine the RL minimum value among $(RL_{(c-1)},\ 2RL_{(c-1)}, 3RL_{(c-1)}, \ldots)$

:- $f(Q_w(c)) \le g(RL_c, u^c)\ Q_w(c+1)$

Step 5: In the ELM classifier, send the calculated LR values

Step 6: For the ELM classifier, calculate the MSER

Step 7: If $MSER \ge 0.1$

Update RL_{c+1}

Step 8: $\alpha_{c+1} = \frac{1+\sqrt{1+4_c^2}}{2}$

End For

End

The PCA feature extraction algorithm's original feature vector is indicated by the symbol η_c . $Q_w(c)$ signifies a feature vector that has been selected, regularization parameter is denoted by the RL. b_c is a parameter that has been selected. An appropriate pair of $Q_w(c)$ and $Q_w(c-1)$ is A^c . MSER stands for mean squared error, which is calculated using Equation (4).

$$MSER = \frac{1}{n}\sum_{c=1}^{K} (RL - \overline{RL_c})^2 \qquad (4)$$

Where the observed features are denoted by LR_c , and the predicted features are denoted by $\overline{RL_c}$. The MSER was determined each time, and the features were updated if it was more than or equal to 0.1, and 1000 times this technique was carried out. If the goal was not met, the categorization was based on the features from the previous iteration. Thus, a stable vector was created, where $X_1 \times K$ represents the vector dimension is and was represented by $\eta_{Q_w}(1)$, where the total feature's range is specified by K that had been selected, and the total number of images was specified by X1. The PCA feature vector η_d was likewise subjected to this feature selection method and the number of selected features is represented by K.

E. Harris Hawks-SqueezeNet based AD Classification

The HHO-SqueezeNet-based rapid diagnostic system's overall architecture is developed. Three main stages compose the proposed system a raw dataset augmented offline, SqueezeNet model training based on Whale optimization, and network decision-making during the testing phase. The proposed method classifies MRI brain scans into five categories: CN, EMCI, LMCI, MCI, and AD. Because of the uneven sample distributions in the raw input images, the

offline augmentation method is used in the first stage. When there are fewer sample numbers in a smaller class then this strategy is appropriate. There are three subsets in the augmented dataset after augmentation: train, validation, and test sets. The SqueezeNet convolutional network is used throughout the training and optimization phases. The input for testing is the test set. The input for testing is the test set. The SqueezeNet convolution network is used in the training and optimization stages, and a more effective and simpler CNN architecture is constructed by squeezing and expanding the layers of the fire modules during training. The Harris Hawks optimization technique is used to optimize the CNN network for acquiring the optimal decision-making model during training, this is an example of a sequential design strategy. The optimization process is updated using a validation error. Lastly, the most effective SqueezeNet model is identified and implemented in the test set's decision-making process. The most effective network model is used to determine the classification performance of the infection classes.

1) SqueezeNet architecture

The fifteen layers are presented in the SqueezeNet structure, including three max-pooling layers, one global average pooling layer, two convolution layers, one softmax output layer, and eight fire layers. The filter's receptive field size is represented by K×K notation, the stride size is denoted by s and l stands for the feature map length, respectively. The network's input has RGB channels and is 227 × 227 in size. To generalize the input images and apply the maximum pooling, convolution is used. The small regions and weights in the input volumes are convoluted by the convolution layer, which uses 3 × 3 kernels. As the positive component of its argument, and element-wise activation function is performed by each convolution layer. The fire is used by SqueezeNet between the convolution layers, which is made up of squeeze and expansion phases. The fire has a consistent input and output tensor scale. The 3 × 3 and 1 × 1 size filters are used in the expansion phase, while the 1 × 1size filter is used in the squeeze phase. At first, the squeeze transmits the H × W × C input tensors. The ReLu units are involved in both the squeezing and extension phases. The depth is compressed by the squeeze operation and also the depth is increased by the expansion while maintaining the size of the features are the same. The concatenate action is used to layering the expansion outputs in the input tensor's depth dimension. Assume the channels and feature maps are defined by C and FM, The squeeze operation's output layer f{y}, w can be represented as with the kernel i.e.,

$$f\{y\} = \sum_{fm1=1}^{FM} \sum_{c=1}^{c} w_c^f \, x_c^{fm1} \qquad (5)$$

Here, $f\{y\} \in R^N$ and $w \in R^{c\times 1\times FM\,2}$. The different tensors' feature maps combined with weights are defined by the squeeze outputs. A down-sampling operation is executed by the max pool layers in the network along the spatial dimensions and the class maps with features are converted into one value by the global average pool. The multiclass probability distributions are generated by the softmax activation function towards the end of the network.

2) Harris Hawks optimization algorithm

Both deep learning and machine learning techniques depend on hyperparameters because the actions of the training algorithms are tightly controlled by this and have a significant impact on model performance. In medical image processing, hyperparameter optimization is critical, specifically in the area of deep learning. However, one optimization challenge is hyperparameter tuning, to handle this, we can develop the HHO. Such optimization problems are solved effectively by using the Harris Hawks optimization algorithm. In this research, the HHO algorithm is used to update the validation error.

Harris hawks are one of the most intelligent birds in the environment. These birds are adept at managing a group and working together to find a certain rabbit. In this phase, various attacking and escaping actions are performed. The optimization problems are solved by developing a method termed HHO in this research by mathematically modeling these interactions. We have a set of hawks in HHO which perform the role of candidate solutions with rabbit serving as the best-obtained solution HHO is a continuous method that has been tested on a set of issues that do not have a binary variable.

The proposed HHO-SqueezeNet model's parameters optimization

Step 1: Set the HHO input parameters to their default values. The space dimension, numbers of generations, boundaries, population, and population size are the parameters in the HHO process.

Step 2: Encode l, n, and feature subset.

Step 3: Manage and calculate population fitness across borders.

Step 4: Update parameters H

$$H = 2H_0(1-\frac{t}{T}) \qquad (6)$$

Where H_o represents a [-1, 1] random number.

Step 5: Using (l, n) and selected features, calculate each swarm member's fitness value according to the following rule

$$\left\{ \begin{array}{c} q_1 = \dfrac{\sum_{i=1}^{k} acc_i}{K} \\ q_2 = 1 - \sum_{j=1}^{n} bin_j \\ q = \alpha \times f_1 + \beta \times f_2 \end{array} \right\} \qquad (7)$$

The objective function must be minimized at this stage. Using K-fold CV, the average accuracy degree is denoted by q_1's the first sub-objective function through the SqueezeNet, the cth fold CV's accuracy is denoted by acci. The dth feature's binary value is displayed in binj for other sub-objective functions q_2 , and the total number of features is

represented by n. We have two scaling factors in the final objective formulation, which is depicted using f. One of these is α , which indicates the user-determined accuracy term's weight. Likewise, the other, marked by the β , specify the value of the selected features on a scale.

Step 6: The first best hawk (solution) is selected $R_{rabbit}(t)$ by using the highest fitness value

Step 7: The position is updated by using the three areas of "hard besiege", "soft besiege", and "soft besiege with rapid progressive divides". During the "soft besiege" phase, hawks update their rule to catch rabbits, which is represented as,

$$R(t+1) = \Delta R(t) - H \mid TR_{rabbit}(t) - R(t) \mid \tag{8}$$

$$\Delta R(t) = R_{rabbit(t)} - R(t) \tag{9}$$

Step 8: Calculate the population's objective.

Step 9: Find the best pair of a parameter (l, n) and use the test set to input the best feature subset (k, m) into SqueezeNet.

Along with the momentum values, the proposed deep HHO-SqueezeNet model optimizes the deep network architecture's most significant hyperparameter, called "initial learning rate."

IV. Results and Discussion

This section discusses the proposed medical image classification as well as the implementation results. It is executed on python for the implementation on NP5SHR9 workstation Intel(R) Core(TM) CPU @ 2.60GHz with 8 GB RAM.

A. Dataset Description

Neuroimaging data is acquired from Alzheimer's disease Neuroimaging Initiative (ADNI) (http://adni.loni.usc.edu/about/ADNI) database that was used in various research for AD classification. It aims to identify genetic, biochemical, neuroimaging, and clinical biomarkers that could serve as diagnostic, prognosticate and track AD. Neuroimages are available in various modes in DNI including PET, DTI, fMRI, and MRI. Our paper used MRI brain scans from ADNI. Over 800 subjects are available in the ADNI database including 493 CN, 204 EMCI, 61 LMCI, 198 MCI, and 145 AD. The ADNI obtains MRI brain images by conducting experiments at the University of Southern California's Laboratory of Neuro-Imaging (LONI).

B. Performance metrics

On MRI images, experiments on deep learning models were conducted to detect AD. HHO-SqueezeNet has been trained for 100 epochs. The comparison results of the performance evaluations of the previous and proposed techniques are represented in terms of accuracy, specificity, sensitivity, precision, and recall. MRI images were classified using the neural network. Evaluation parameters include:

$$Accuracy(AC) = \frac{TP + TN}{TP + FP + TN + FN} \tag{10}$$

$$Precision\ (PR) = \frac{TP}{TP + FP} \tag{11}$$

$$Recall\ (RC) = \frac{TP}{TP + FN} \tag{12}$$

$$Specificity\ (SP) = \frac{TN}{TN + FP} \tag{13}$$

$$Sensitivity\ (SN) = \frac{TP}{TP + FN} \tag{14}$$

The False Negative, True Negative, False Positive, and True Positive are represented by FN, TN, FP, and TP, correspondingly.

C. Experimental Results

We develop HHO-SqueezeNet based multimodal classification framework in this research. Image pre-processing, extracting and selecting the effective features, and disease classification are the three processes in our proposed method. Noise is removed and contrast is improved by performing the pre-processing with a adaptive histogram equalization and high boost filter. For image segmentation, the algorithm of expectation-maximization is then used. After segmenting the image, some of the features are extracted from the MRI brain images by using PCA, then features are selected by CML-ELM, which will produce the final segmented image. Objects detected in segmented images would then be classified to enhance the adaptability of the network.

Experiments on the ADNI dataset's 830 patients are used for evaluating the effectiveness of our proposed methodology. The proposed method results revealed that the results can not only use complementary information in several sources for the AD stage classification and also help in the finding of disease biomarkers to better understand the pathological mechanism.

The classification performance of 830 ADNI individuals was assessed using MRI modalities. There are many classification problems in this study, including MCI, CN, AD, EMCI, and LMCI. The tenfold cross-validation is used for evaluating the proposed SqueezeNet-based method's performance. The training samples are selected at random from 70% of the associated data for each specific task, while the remaining 30% is used as the testing samples and all trials maintain the same training and testing set. It's important to note that the two types of sets never intersect. Each of the five trials is performed ten times, with the average findings provided. The classification accuracy, sensitivity, specificity, precision, and recall are all used to assess performance. Among all subjects, a total of 210 subjects were correctly classified and it is used to calculate classification accuracy (ACC). The proportion of correctly predicted patients is represented by the sensitivity (SEN), while the proportion of correctly diagnosed normal controls is represented by the specificity (SPE). The input images are shown in Figure.2

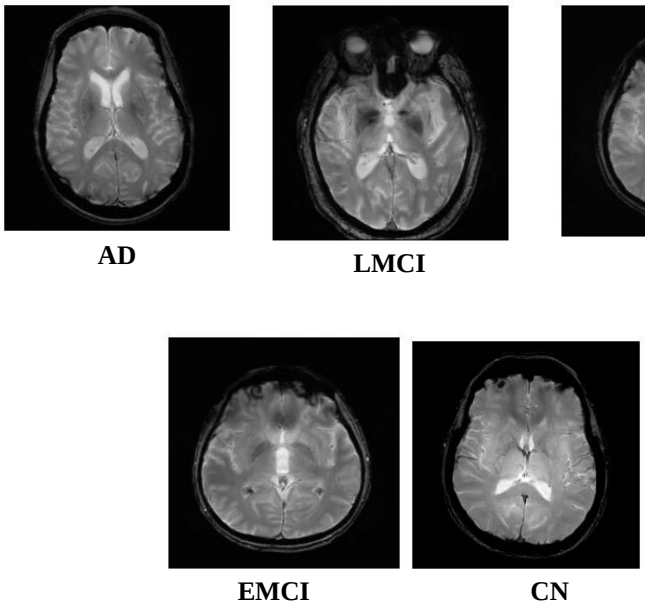


Figure 2. The multiple AD stages of sample images collected from the ADNI dataset

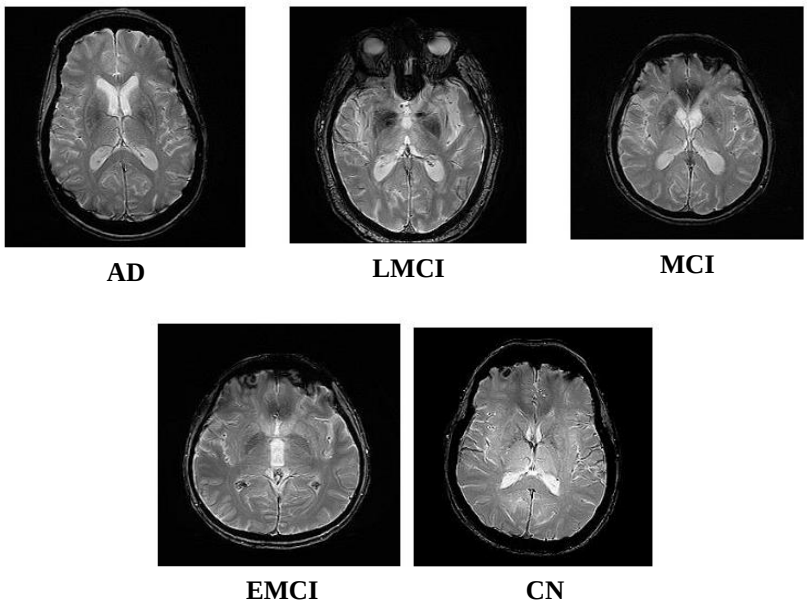
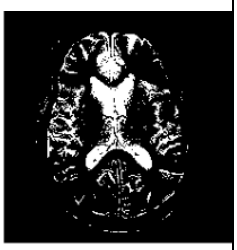
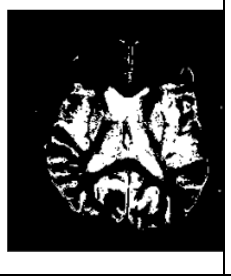


Figure 3. The result of pre-processed images of multiple AD stages

Stages	WM	GM	CSF	Detected AD Stages
AD				

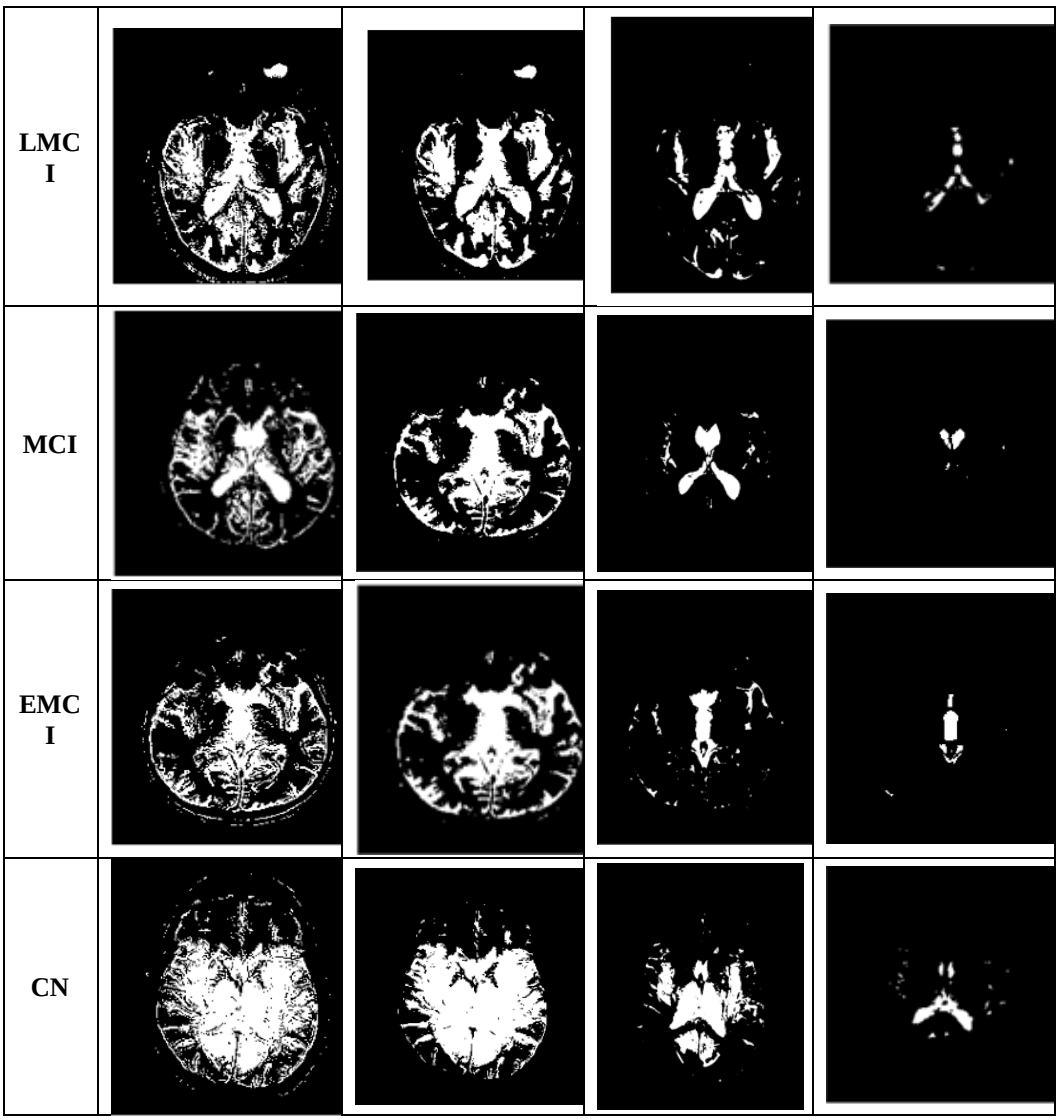


Figure 4. The performance result of segmented and detected images of multiple AD stages

Figure 3 demonstrates pre-processed image results from the original MR images with a high boost filter and adaptive histogram equalization. Figure 4 reveals the segmentation and classification results of the proposed HHO-SqueezeNet using the MRI images of the ADNI dataset. Even though our model employs a better classification, training the network with an improved classification remains critical to achieving a high level of global segmentation accuracy. In addition, our contribution does ensure a high level of efficiency by mainly segmenting brain images, the system detects abnormal objects, classifies them according to their potential classes, and correctly diagnoses the subject under research.

D. Performance Results

The experiment was analyzed with five distinct measures for assessing the proposed method's performance. Specificity, sensitivity, recall, precision, and accuracy are used for the analysis. Figure 5 reveals the multiple AD stage performance comparison graph and the classification findings are presented in Table 1.

METRICS	CN (%)	EMCI (%)	MCI (%)	LMCI (%)	AD (%)
Accuracy	98.8	99	98.2	99.02	99.2
Precision	99.01	98.89	99.2	98.9	99.8
Recall	96.02	98.54	97.45	96.25	98.21
Sensitivity	97.2	98.2	99.25	97.51	97.08
Specificity	96.8	95.9	97.45	96.39	97.36

Table 1. Performance results for our proposed models

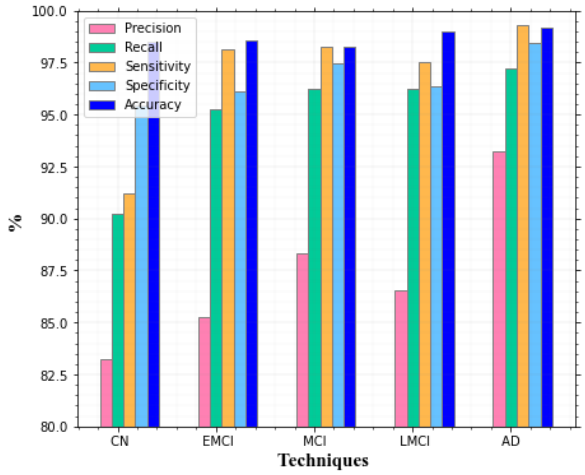


Figure 5. Multiple stage performance metrics comparison graphs

Table 2 compares the proposed HHO-SqueezeNet-based method to five other state-of-the-art machine learning methods in terms of performance. As shown in Table 2, the proposed technique can produce impressive outcomes in many phases of AD. It's also important to note that when applied to the five stages of AD, the developed method produces the best results (98.82 %).

Methods	Accuracy (%)	Precision (%)	Recall (%)	Sensitivity (%)	Specificity (%)
KNN+SVM+MLP	96.31	96.05	86.8	91.27	89.90
AlexNet	97.63	95.54	87.6	93.2	92.45
SCCA	82.0	93.45	89.87	83.7	80.3
ResNet-18	97.6	98.13	97.92	96.23	94.6
DBN	97.62	96.67	92.31	91.89	87.78
Proposed method	98.82	99.2	98	97.92	96.3

Table 2. The classification performance comparison

All of the preceding algorithms should be examined under the same conditions as our method in the strict understanding. However, we want to make sure that the developed structure is a reasonable solution within the same database, and the findings in Table 2 show that the framework that has been proposed is effective, demonstrating that the proposed technique is entirely capable of diagnosing Alzheimer's disease.

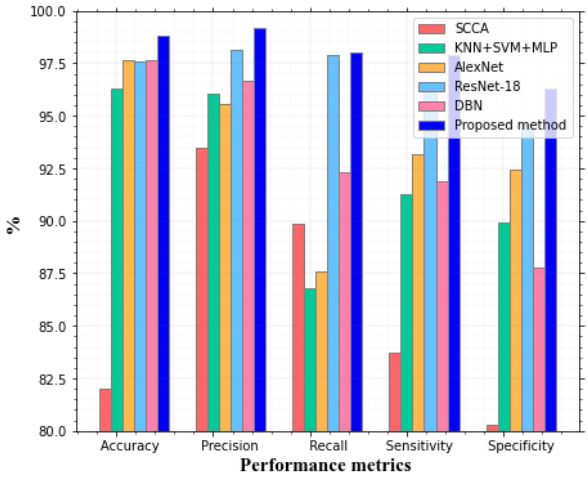


Figure 6. Performance comparison graph

The performance comparison graph for proposed and existing methods is shown in Figure 6. Using our enhanced model, we improved accuracy for all AD stages, with CN, EMCI, MCI, LMCI, and AD achieving 98.5 %, 98.54 %, 98.25 %, 99.02 %, and 99.2 %, respectively. In comparing the results of KNN+SVM+MLP, AlexNet, SCCA, ResNet-18, and DBN, our proposed method achieves a high classifying accuracy is obtained with each AD stage of the proposed model on comparing with the overall average of other existing techniques. The proposed approach achieved better performance by having 98.82 (%) for accuracy, 99.2 (%) for precision, 98 (%) for recall, 97.92 (%) for sensitivity, and 96.3 (%) for specificity respectively in detecting the AD stage. However, DBN performed better than KNN and AlexNet having achieved 98.2(%) and followed by DBN.

The sensitivity (SEN), specificity (SPE), precision and recall measures also improved the existing methods. While compared to the existing methods, the AD diagnosis problem is effectively performed by the proposed HHO-SqueezeNet based multi-task learning framework with strong generalization ability, high accuracy, and fast operation.

V. Conclusion

AD is a severe neurological disease that affects a large portion of the global population. With the help of brain MRI data analysis, we proposed an efficient method of AD diagnosis. An HHO-SqueezeNet-based algorithm is established for multiple stages of AD classifications. Specifically, the feature selection is done by using PCA and CML-ELM. Our method's accuracy in classification tasks for CN, EMCI, MCI, LMCI, and AD is 98.5 %, 98.54 %, 98.25 %, 91.89 percent, 99.02 %, and 99.2 %, respectively, based on MRI data received from the ADNI database. It demonstrates its utility and effectiveness in the identification of AD at various stages. And Our SqueezeNet obtained the best average accuracy of 98.82% for multiple AD stages. Various metrics were calculated for the model evaluation and compared with existing methods to illustrate its accuracy. We aim to include PET or another neuroimaging modality besides MRI in our future work to fully comprehend the interactions between multiple modalities and provide more accurate expert assistance in the biomedical area.

Conflict of interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

Funding: Not applicable

Availability of data and material: Not applicable

Code availability: Not applicable

Authors' contributions: The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Ethics approval: This material is the authors' own original work, which has not been previously published Elsewhere. The paper reflects the authors' own research and analysis in a truthful and complete manner.

References

[1] DeTure, M.A. and Dickson, D.W., 2019. The neuropathological diagnosis of Alzheimer's disease. *Molecular neurodegeneration*, 14(1), pp.1-18.

[2] Zhang, M., Yang, Y., Shen, F., Zhang, H. and Wang, Y., 2017. Multi-view feature selection and classification for Alzheimer's disease diagnosis. *Multimedia Tools and Applications*, 76(8), pp.10761-10775.

[3] Porsteinsson, A.P., Isaacson, R.S., Knox, S., Sabbagh, M.N. and Rubino, I., 2021. Diagnosis of early Alzheimer's disease: clinical practice in 2021. *The journal of prevention of Alzheimer's disease*, vol. 8(3), pp.371-386.

[4] Raj, R.J.S., Shobana, S.J., Pustokhina, I.V., Pustokhin, D.A., Gupta, D. and Shankar, K., 2020. Optimal feature selection-based medical image classification using deep learning model in internet of medical things. *IEEE Access*, vol. 8, pp.58006-58017.

[5] Trambaiolli, L.R., Spolaôr, N., Lorena, A.C., Anghinah, R. and Sato, J.R., 2017. Feature selection before EEG classification supports the diagnosis of Alzheimer's disease. *Clinical Neurophysiology*, vol. 128(10), pp.2058-2067.

[6] Shao, W., Xiang, S., Zhang, Z., Huang, K. and Zhang, J., 2021. Hyper-graph based sparse canonical correlation analysis for the diagnosis of Alzheimer's disease from multi-dimensional genomic data. *Methods*, 189, pp.86-94.

[7] Shao, W., Peng, Y., Zu, C., Wang, M., Zhang, D. and Alzheimer's Disease Neuroimaging Initiative, 2020. Hypergraph based multi-task feature selection for multimodal classification of Alzheimer's disease. *Computerized Medical Imaging and Graphics*, 80, p.101663.

[8] Weller, J. and Budson, A., 2018. Current understanding of Alzheimer's disease diagnosis and treatment. F1000Research, 7.

[9] Altaf, T., Anwar, S.M., Gul, N., Majeed, M.N. and Majid, M., 2018. Multi-class Alzheimer's disease classification using image and clinical features. Biomedical Signal Processing and Control, 43, pp.64-74.

[10] Böhle, M., Eitel, F., Weygandt, M. and Ritter, K., 2019. Layer-wise relevance propagation for explaining deep neural network decisions in MRI-based Alzheimer's disease classification. Frontiers in aging neuroscience, 11, p.194.

[11] Zhu, X., Suk, H.I., Wang, L., Lee, S.W., Shen, D. and Alzheimer's Disease Neuroimaging Initiative, 2017. A novel relational regularization feature selection method for joint regression and classification in AD diagnosis. *Medical image analysis*, vol. 38, pp.205-214.

[12] Feng, C., Elazab, A., Yang, P., Wang, T., Zhou, F., Hu, H., ...& Lei, B. (2019). Deep learning framework for Alzheimer's disease diagnosis via 3D-CNN and FSBi-LSTM. *IEEE Access*, vol. 7, pp. 63605-63618.

[13] Divya, R. and Shantha Selva Kumari, R., 2021. Genetic algorithm with logistic regression feature selection for Alzheimer's disease classification. *Neural Computing and Applications*, vol. 33(14), pp.8435-8444.

[14] Luo, P., Kang, G. and Xu, X., 2020, January. A novel feature selection and classification method of Alzheimer's disease based on multi-features in MRI. In Proceedings of the 2020 10th International Conference on Bioscience, *Biochemistry and Bioinformatics* (pp. 114-119).

[15] Afshar, P., Mohammadi, A. and Plataniotis, K.N., 2020. BayesCap: a Bayesian approach to brain tumor classification using capsule networks. IEEE Signal Processing Letters, 27, pp.2024-2028..

[16] Nawaz, H., Maqsood, M., Afzal, S., Aadil, F., Mehmood, I. and Rho, S., 2021. A deep feature-based real-time system for Alzheimer disease stage detection. *Multimedia Tools and Applications*, vol. 80(28), pp.35789-35807.

[17] Yildirim, M., &Cinar, A. C. (2020). Classification of Alzheimer's Disease MRI Images with CNN Based Hybrid Method. *Ingénierie des Systèmes d Inf.*, vol. 25(4), 413-418.

[18] Tuan, T. A., Pham, T. B., Kim, J. Y., & Tavares, J. M. R. (2020). Alzheimer's diagnosis using deep learning in segmenting and classifying 3D brain MR images. *International Journal of Neuroscience*, 1-10.

[19] Jain, R., Jain, N., Aggarwal, A., &Hemanth, D. J. (2019). Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research*, vol. 57, pp. 147-159.

[20] Ayadi, W., Elhamzi, W., Charfi, I. and Atri, M., 2021. Deep CNN for brain tumor classification. *Neural Processing Letters*, 53(1), pp.671-700.

[21] Zeng, N., Li, H. and Peng, Y., 2021. A new deep belief network-based multi-task learning for diagnosis of Alzheimer's disease. *Neural Computing and Applications*, pp.1-12.

[22] Zu 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 1542-1545). IEEE. , C., Wang, Y., Zhou, L., Wang, L. and Zhang, D., 2018,

April. Multi-modality feature selection with adaptive similarity learning for classification of Alzheimer's disease. In 2018 IEEE.

[23] Zhou, K., He, W., Xu, Y., Xiong, G. and Cai, J., 2018. Feature selection and transfer learning for Alzheimer's disease clinical diagnosis. *Applied Sciences*, vol. 8(8), p.1372.

[24] Hao, X., Bao, Y., Guo, Y., Yu, M., Zhang, D., Risacher, S.L., Saykin, A.J., Yao, X., Shen, L. and Alzheimer's Disease Neuroimaging Initiative, 2020. Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease. *Medical image analysis*, vol. 60, p.101625.

[25] Hon, M. and Khan, N.M., 2017, November. Towards Alzheimer's disease classification through transfer learning. In 2017 IEEE International conference on bioinformatics and biomedicine (BIBM) (pp. 1166-1169). IEEE.

Author Biographies

Mr P. Anil Kumar Working as Asst.Professor in Computer Science and Engineering Department from 2012 to till date in Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada. He is a Research scholar Pursuing Ph.D in Computer Science Engineering (2020 – till) in Adikavi Nannya University, Rajamahendravaram. He has more than 15 years of experience in Teaching and his area of interest includes Machine Learning, Deep Learning and Image Processing.

Prof P. Suresh Varma presently serving as Dean Academic Affairs and he served as Vice Chancellor(FAC) and Rector, Aadikavi Nannaya University. A senior most Professor in the University, currently holds the position of Dean Faculty of Engineering and Technology. Apart from being an eminent scholar in Computer science and Engineering, he has also served the university in other capacities, which include, Executive Council Member, Senate Member, Dean College Development Council, Dean IT, Dean Faculty of Engineering and Technology and Chairman Board of Studies in Computer science & Engineering Department, Aadikavi Nannaya university.He also served the University as Principal, College of Engineering from 31-Dec-2014 to 12 July, 2017, Principal University College from 4-Dec-2009 to 13-June-2012, Dean College Development Council, twice during 14-June-2012 to 13-June-2015 and 14-June-2018 to 25-April-2019.Prior to joining at Aadikavi Nannaya university, in October, 2008, Prof.Varma worked with various organizations like GITAM, GMRIT, VIGNAN IIT and Dr. L.B.college in between 1994 and 2008.