

Aswathy K Cherian

*Department of Computing Technologies,
SRM Institute of Science and Technology
Chennai
TamilNadu*

Vaidhehi MS

*Department of Computing Technologies,
SRM Institute of Science and Technology
Chennai
TamilNadu*

Arshey M

*Department of CSE
Jain (Deemed-to-be University)
Ernakulam
Kerala*

1. INTRODUCTION

A type of memory loss known as dementia currently affects 50 million individuals worldwide. Amyloid proteins accumulate in brain cells due to Alzheimer's disease (AD), a degenerative illness. As amyloid protein levels rise, cells die, and signals become more difficult to transmit. Public health issues like AD are increasing worse in developed nations. It can be observed in someone's impaired capacity to think clearly, making it challenging to carry out daily tasks or get along with others. Problems interacting with others, memory loss, and difficulty thinking and speaking are all symptoms of the disability. AD is primarily brought on by aging and mainly affects those over 65. Most individuals concur that the brain is one of the body's maximum excellent vital structures. The brain directs and sustains all actions and responses that allow us to reason and trust. It also promotes the retention of trustworthy memories and feelings [1]. A fatal and developing brain condition is AD. Alzheimer's disease slowly worsens and destroys memory cells after it is found, which impairs a person's ability to reason. It is a neurological complaint that deteriorates over time and causes neurons to stop working or even die only four to eight years on the average pass after receiving an Alzheimer's diagnosis. By 2050, AD may affect 1 in every 85 persons, according to new studies [2]. Patients with AD must be identified and treated as soon as feasible. The diagnosis and prognosis of this dementia can be determined using various methods, such as PET, CT (Computed Tomography) scans, and MRI.

This work suggests a mix of deep learning and meta-heuristic optimization as a method for the primary documentation of Alzheimer's disease. There are two stages to the suggested deep learning DCBM framework. Dispersed filtering, an unsupervised two-layer neural network, is used in the first stage to excerpt topographies from the raw data. SoftMax regression is used in the second phase to organize different health statuses founded on the recovered topographies. By using medical data from the National Alzheimer's Coordinating Centre, the proposed framework has a greater classification accuracy when compared to well-known ensemble techniques. The presentation of the collaborative network is better quality by mechanically selecting and optimizing the Deep Convolutional Boltzmann Machine (DCBM) algorithm and extending it to a classification model rather than just a feature extractor with specificity, sensitivity, and accuracy

as a classification of Alzheimer's disease. This paper is organized as following parts: Part 1 presents the background information and research on the suggested algorithm. Part 2 of the proposed method contains a description of its specifics. In instruction to assess the efficiency of the suggested technique, Part 4 provides the experiment and numerical findings. The conclusions and future investigation ideas are obtainable in Section 5.

1.1 a

1.2 a

2. LITERATURE SURVEY

Alickovic et al. [3] 's methodology offered a way to recognize AD using machine learning methods. In the primary stage, histograms convert brain pictures into feature vectors of relevant "brain" features. Then, AD is categorized and recognized using ML techniques. The model mentioned here was developed using the current contribution and performs admirably. The trial's outcomes show that the RF classifier can distinguish between AD and control patients. The proposed method, which employs an RF classifier and a histogram as a feature extractor, has an accuracy of 85.77%. Shankar et al. [4] recommended a method for Alzheimer's disease and brain imaging analysis. (BIA). Unwanted elements of the photos are deleted at this early stage. Transforms, histograms and texture scale-invariants were then derived, among other properties. The detection performance is enhanced using Group Grey Wolf Optimization (GGWO) techniques using CNN, KNN, and decision tree classifiers. They distinguish a more condensed set of traits while preserving performance. This method outperformed other systems in terms of AD detection accuracy, achieving 96.23%. According to Veen et al. [5], FDGPET data automates AD categorization. The sub-scaled profile model and Principal Component Analysis (PCA), which enables the identification of all probable glucose metabolism patterns, are used to recover several characteristics. As a result, Support Vector Machines (SVM) and Generalized Matrix Learning Vector Quantization (GMLVQ), two popular machine learning models, might be used to differentiate between Alzheimer's disease (AD), healthy controls, and partial discharges (PD).

A deep learning-based PET-MRI copy fusion and 3D

convolutional neural network (CNN) for AD multi-classification methods were developed by Kong et al. [6]. All 740 of the 3D photos available in the ADNI database were employed. This research provides preliminary evidence that the A3C-TL-GTO classification system (AD) can be applied to MRI images to diagnose Alzheimer's. The deep learning system for MCI and AD forecast was created and tested by Basaia et al. [7]. 3D T1-weighted pictures were employed in conjunction with the CNN application. CNN was used in the evaluation process to differentiate between c-MCI and s-MCI in AD. Lastly, a higher level of organization precision was attained with ADNI and a combination of ADNI and non-ADNI datasets. CNN was seen as the most effective technique for improved separate patient diagnosis along the AD continuum.

According to Pradhan et al. [8], several stages of AD could be detected. DenseNet169 architecture and VGG19 is the organization technique employed. The data set was obtained from the Kaggle open internet data set repository. Six thousand photos with extremely mild, mild, moderate, and non-demented AD labels are available. The topographies are measured for 80% of the learning and 20% of the testing stages. DenseNet is employed in the classification of images. In this case, VGG19 outperforms DenseNet with an accuracy of 94%. Shankar et al. put forth a strategy for AD and Brain Imaging Analysis [9]. Unwanted areas in the photographs are first deleted in the early step. Several features were retrieved, including transforms, histograms, and texture scale-invariants. Group Grey Wolf Optimization (GGWO) methods improve the detection efficiency of CNN, KNN, and decision tree classifiers. They are utilized to identify a smaller number of features without sacrificing performance. Compared to other competing systems, this method's accuracy in detecting AD was 96.23%. Shaikh et al. [10] have created a unique perfect for involuntary tissue classification and segmentation founded on SVM and PCA. The inputs were the PD, T1, and T2, and the outputs were the WM, CSF, and GM. The atrophy of the segmental tissues was investigated using the texture information from GLCM and the BEP characteristics from DWT. After performing the necessary preprocessing and classification, the SVM classifier extracted the two most important features. The experiment revealed that the proposed approach has a greater tissue categorization accuracy than other compared models.

3. PROPOSED SYSTEM

Structural and molecular/functional biomarkers for AD have been categorized with the help of multimodal neuroimaging data. Enhanced machine learning accuracy has been linked to characteristics like larger capacities or thicker cortical layers in AD-specific brain areas like the entorhinal cortex and hippocampus. Features have been extracted from neuroimaging data using deep learning techniques. Alzheimer's is the most prevalent form of dementia, considered by cell death in the brain's memory regions. Therefore, successful Alzheimer's disease organization and care requires a prompt diagnosis. Between

cognitively normal aging and Alzheimer's disease, amnesic (mild cognitive impairment convertible) MCI is a transitional phase, and AD is developed more than age-matched healthy cognition (HC) by MCI patients. The recommended method introduces the early diagnosis model for AD, built upon a hybrid meta-heuristic optimization. The three processes are feature extraction from the image, image pre-processing, and image classification and detection. Figure 1 depicts the schematic's block diagram.

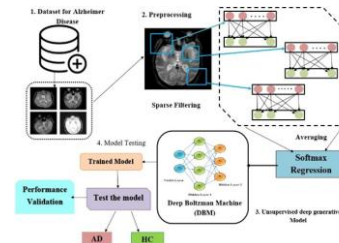


Fig. 3.1: DBM-based categorization approach for Alzheimer's disease

The work developed a two-stage learning technique for processing massive inputs using brain imaging to monitor health issues. The initial part of the process entailed using sparse filtering to learn the expressive features of the images, which was referred to as a two-layer network. A two-layer network SoftMax regression was developed in the second step to differentiate well and sick states automatically. This method relied on a neural network rather than prior knowledge or human labor to learn the features. The superiority of this Deep Boltzmann Machine (DBM) technology was proved using brain scans to classify Alzheimer's illness. Figure 2 depicted samples of classed demented MRI scans ranging from sane illusions to moderate insanity.

3.1 Dataset Used

The National Alzheimer's Coordination Center [11], which originated in 1999 and has preserved a cumulative collection of numerous clinical data kinds, such as clinical evaluations, brain MRI imaging, and neuropathology, is the information's source. Many researchers have visited this site to locate helpful information [12]. The Geriatric Depression Scale (GDS), Unified Parkinson's Disease Rating Scale (UPDRS), Cerebrovascular Disease (CVD), Medical History (MH), Hachinski Ischemic Score (HIS), Neuropsychiatric Inventory Questionnaire (NPIQ), and Functional Activities Questionnaire are the seven groups of measures that were chosen. A sample of 23,165 examples with 100 measurements (qualities) from the NACC UDS (FAQ). Here, CVD and HIS are split into two categories. Some photos for the dataset are shown in Figure 3.

3.2 Pre-processing using Sparse Filtering

Unsupervised feature learning techniques have been working hard to approximate a fair distribution of the obtained data. Examples include dispersed coding, constrained Boltzmann machines, scattered self-encryptions, etc. They often necessitate fine-tuning different stage

features, which is a challenging undertaking. For instance, it is necessary to determine the characteristics, dispersion correction, learning rate, weight degradation, and movement for a limited Boltzmann machine. If these limitations are wrongly adjusted, the developed topographies could lead to low diagnostic accuracy. Nigam et al. [13] created the sparse filtering approach of unsupervised feature learning. The number of features is the sole parameter required for light filtering. So, for sparse filtering, adjusting the parameters is essential. Sparse filtering produces a discrete circulation of the characteristics calculated by the information composed as an unmanaged two-layer network, as opposed to mimicking the data distribution.

The sparse filtering of examples is translated to the topographies of using the matrix of \mathbf{X} , using the training set of \mathbf{X} where \mathbf{x} are a example and A is the amount of examples.

The normalized feature of \mathbf{l}_2 is used by sparse filtering to optimize a function. It is essential to keep in mind that the \mathbf{l}_p norm of \mathbf{t} is written as $\|\mathbf{t}\|_p$. where \mathbf{t} The feature matrix is made by \mathbf{X} . The first step is to apply the \mathbf{l}_2 norm to every row of the feature matrix across all training data.

The features are mapped onto the \mathbf{l}_2 -ball by standardizing each column with the \mathbf{l}_2 norm.

To solve the weight requirements in Eq. 1, the \mathbf{l}_1 norm of the cost purpose constraint is optimized as stated below.

The \mathbf{L}_2 norm is typically employed to quantify component dispersal, indicating that maximum of the modules are zero. The word in equation (4) processes the sample's characteristics. Since the \mathbf{l}_2 -ball unit is the only location where the normalized feature of \mathbf{l}_2 of \mathbf{x} can exist, the cost purpose in Eq. (4) is reduced when these topographies are discrete. Sparse filtering can be rehabilitated into nonlinear filtering with the use of an activation function, and the activation function is recommended. The enlarged version of Therefore, Eq. (5) is

The input data can have acceptable unreliability when nonlinear information is extracted using the trained features by improving the objective functions of the characteristics in Eq. 5.

3.3 Softmax Regression

Softmax regression is commonly used in the top layer of neural networks for multiclass classification [14]. It's straightforward, and quick calculations are possible. Assume we have a labeled set of \mathbf{Y} and a training set of \mathbf{X} , where \mathbf{y} and \mathbf{x} . The perfect attempts to predict the probability of \mathbf{y} for each of the labels of \mathbf{Y} for each input sample \mathbf{x} . As a result, the SoftMax regression's underlying premise yields a vector containing the N-estimated probability that each label belongs to the input sample of \mathbf{x} . The premise of is as follows.

Where \mathbf{T} . The Softmax regression model's T parameter. It should be noticed that the formula normalizes the distribution \mathbf{p} , which makes the assumption's parts add up to 1.

The perfect is completed by minimalizing the cost purpose of according to the assumption.

Where \mathbf{p} is the purpose of the marker, which, if true, gives condition one and denotes the length of weight reduction. While some Softmax regression variables are forced to take values close to zero throughout the weight reduction phase, other restrictions can maintain their relatively high standards, increasing generality. The Softmax regression model makes sure that there is only one theoretical solution and the cost function of \mathbf{J} strictly grows convex with the term of \mathbf{w} weight reduction (for every \mathbf{w}). Assuming that the combination of sample features may be used to determine the likelihood that a sample would match a given health status label, a unique approach to tackling the classification task known as Softmax regression also makes this assumption. For instance, softmax regression offers a probable classification.

3.4 Deep Generative Model

The DBM is a single-directional deep generative classical. Each layer's specific unique variable is present in it. In DBM, a bipartition structure is used to model the organization of images. The suggested partitioned creation of the DBM is shown in Figure 4.

Examine the shared probability distribution of the energy function E for the DBM, with \mathbf{h}_1 and \mathbf{h}_2 acting as the two hidden layers in addition to the visible layer \mathbf{v} . The probability distribution appears beneath P :

DBM offers a bipartition architecture for classifying spectral-spatial images, unlike earlier generative models like the limited Boltzmann machine. The dispersion of each unit is active because each adjacent layer uses the Bernoulli variable to generate the DBM situation distribution. For categorization, deep learning models primarily extract features from autoencoders. Its architecture includes both an encoder and a decode. The encoder and decoder both make use of simple matrix multiplication. As a normalizing function, the encoder's gradient function is employed. The network training is finished by changing the autoencoder's weight (\mathbf{w}) and biases (\mathbf{b}).

Take the training of an MRI database with two hidden layers as an example.

Here is the mean-field value displayed:

Where the Gibbs energy G listed in (16) is as follows:

Equations (17) and (18) provide the new weight value and display the weight shift, respectively. Bias $\mathbf{b} = 0$ is applied to each layer.

The proposed DBM, instead of another structure, uses Gibbs sampling, which updates only one variable at a time. This makes it possible to classify spatial-spectral images precisely. Using Gibbs sampling, an image patch can be split into two update blocks. All image patches for layer \mathbf{l} must have maximum boundaries of $\mathbf{r} = \mathbf{l} + 1$. Two hidden layers are employed for processing. Hence, each

layer was concurrently and separately sampled.

3.4.1 Evaluation of DBM Model

A typical MRI database produced by individuals with Alzheimer's disease is used to apply the suggested DBM model [26]. The image patch is used to process the database for categorization. The 850 nm image was chosen for this study because it can be pre-processed more precisely and offers more spectral and spatial information than photos collected at other wavelengths. With the data set in a fully-connected network and weight and bias expressed as $z = w \cdot x + b$, where w stands for the layer's weight and b for the layer's bias, this database is shown in an input vector for classification. The expression $\sigma(z)$, which determines the change in the reply weight and stands for the non-linear learning rate.

The following represents the texture feature's most excellent likelihood estimate in the MRI database:

Where r signifies the amount of practice examples, the picture patches are separated using this dataset.

The generated picture patching is sent to unconstrained DBM for categorization, as seen in Fig. 4. Instead of using the whole database. The DBM model uses a discriminative small dataset. This is done by training and extracting features from a volumetric database using multi-instance learning. The source dataset for the specified training set is where the pre-processed training set inherits its properties, this is a list of the loss in the inherited database:

The probability of successfully classifying using the DBM

4. RESULT AND DISCUSSION

4.1 Experiment setup

In this experiment, a skull stripping was performed, picture segmentation, and cutting using MATLAB R2023a. The classifier was trained and validated using Google Colab and a compute 3.7, GPU (1xTesla K80 with 2496 CUDA cores, 24GB (23.439GB Usable) GDDR5 VRAM) available to the public as open source software. Parallel processing is obligatory to train deep neural network information. The Keras library was chosen over TensorFlow components to generate our optional mode.

4.2 Precision Analysis

Fig. 5 and Tab. 2 compare the DBM approach's precision evaluation with other existing methods. The graph demonstrates how the deep learning approach performs more precisely. For example, the precision of the DBM model for 1000 data is 97.175%, whereas the precision values of the GMLVQ-SVM, TrAdaboost, GWO, LPBM, and RNN models are 88.094%, 89.668%, 91.675%, 93.456%, and 95.234%, respectively. Similarly, the DBM model has performed best with various data sizes. Likewise, under 3000 data, the precision value of the DBM is 97.098%, compared to GMLVQ-SVM, TrAdaboost, GWO, LPBM, and RNN models, which have precision

values of 88.536%, 90.854%, 92.345%, 94.234%, and 95.098%, respectively. The DBM model has, nevertheless, performed best with various data sizes. The GMLVQ-SVM, TrAdaboost, GWO, LPBM, and RNN models' precision values under 6000 data are 89.738%, 90.654%, 93.098%, 94.675%, and 96.578%, respectively while the DBM method has 97.675% of precision.

5. CONCLUSION

The Deep Boltzmann Machine (DBM) method is improved in this paper by transforming it from a feature extractor to a classification model. The proposed method also utilize a multi-task learning framework based on identification and verification tasks to reduce overfitting. This is crucial for early AD diagnosis because it increases inter-subject differences and minimizes intra-subject variances.

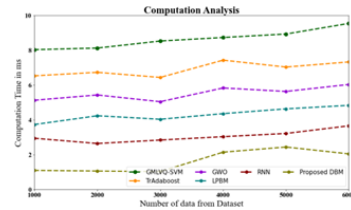


Fig. 5.1: Analysis of the DBM method's computation time using current systems

Two steps comprise the proposed deep learning DCBM framework: the first stage, an unsupervised two-layer neural network, employs dispersed filtering to extract topographies from the basic information, and the second phase uses SoftMax regression to classify various health positions founded on the extracted topographies. The proposed method was evaluated using a clinical dataset from the National Alzheimer's Coordinating Center and establish that it has a greater classification accuracy than well-known ensemble approaches. Given the advancement of deep learning algorithms, it is possible to use these methods to detect AD using multimodal neuroimaging data in the future. Upcoming deep learning-based research in AD may improve performance by using more such as omics data, increase transparency and hybrid data sources. by implementing explainable methods to understand specific disease-related features and mechanisms better. The proposed approach can increase false negative rate, precision, false positive rate, f1-measure, accuracy, recall, and computation time by 97.675%, 90.738%, 94.637%, 99.213%, 0.0232%, 0.0473%, and 2.039sec, when compared with Support vector machine (SVM), Generalized Matrix Learning Vector Quantization (GMLVQ), Grey Wolf Optimization (GWO), and TrAdaboost, linear program boosting method (LPBM), Recurrent Neural Network (RNN), respectively. We also intend to use sagittal MRI pictures and integrate information from sagittal MRI capacities. Because the entire brain can access sagittal data, forecasts are improved.

REFERENCES

- [1] Orouskhani, M.; Rostamian, S.; Zadeh, F.S.; Shafiei, M.; Orouskhani, Y. Alzheimer's Disease Detection from Structural MRI Using Conditional Deep Triplet Network. *Neurosci. Inform.* 2022, 100066.
- [2] Cui, R.; Liu, M. RNN-based longitudinal analysis for diagnosis of Alzheimer's disease. *Comput. Med. Imaging Graph.* 2019, 73, 1–10.
- [3] Alickovic, E.; Subasi, A. Automatic detection of Alzheimer's disease based on histogram and random forest. In *Proceedings of the International Conference on Medical and Biological Engineering*, Banja Luka, Bosnia and Herzegovina, 16–18 May 2019; Almir, B., Ranko, A., Lejla, G.P., Eds.; Springer: Cham, Switzerland, 2020.
- [4] Shankar, K.; Lakshmanprabu, S.K.; Khanna, A.; Tanwar, S.; Rodrigues, J.J.; Roy, N.R. Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. *Comput. Electr. Eng.* 2019, 77, 230–243.
- [5] Venugopalan J, Tong L, Hassanzadeh HR. Multimodal deep learning models for Early Detection of Alzheimer's disease stage. *Sci Rep.* 2021. <https://doi.org/10.1038/s41598-020-74399-w>
- [6] Odusami M, Maskeliunas R, Damaševičius R, Krilavičius T. Detection of early stage from functional brain changes in magnetic resonance images using a Finetuned ResNet18 network. *Diagnostics.* 2021;11(1071):1–16. <https://doi.org/10.3390/diagnostics11061071>
- [7] Kong, Z.; Zhang, M.; Zhu, W.; Yi, Y.; Wang, T.; Zhang, B. Multi-modal data Alzheimer's disease detection based on 3D convolution. *Biomed. Signal Process. Control* 2022, 75, 103565
- [8] Pradhan A, Gige J, Eliazar M. Detection of Alzheimer's Disease (AD) in MRI Images using deep learning. *Int J Eng Res.* 2021;10(3):580–5. <https://doi.org/10.17577/IJERTV10IS030310>
- [9] Lim, H.; Dewaraja, Y.K. Y-90 Patients PET/CT & SPECT/CT and Corresponding Contours Dataset [Data Set]; University of Michigan's Deep Blue Data; University of Michigan: Ann Arbor, MI, USA, 2019
- [10] Shaikh TA, Ali R. Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images. *Magn Reson Imaging.* 2019;62:167-173
- [11] Mohan,P.; Veerappampalayam Easwaramoorthy, S.;Subramanian, M.; Meckanzi, S. Handcrafted Deep-Feature-Based Brain Tumor Detection and Classification Using MRI Images. *Electronics* 2022, 11, 4178. <https://doi.org/10.3390/electronics11244178>.
- [12] S.Raghavendar,A.Hardhavardhan,R.Partheepan,Ranj an Walia and V.Chandra Shekar Rao, Multilayer Stacked Probabilistic Belief Network-Based Brain Tumor Segmentation and Classification –International Journal of Foundations of Computer Science–, <https://doi.org/10.1142/S0129054122420047>.
- [13] Chiai AI-Atroshi, J. Rene Beulah, Kranthi Kumar Singamaneni, C. Pretty Diana Cyril, & S. Velmurugan (2022) Automated speech based evaluation of mild cognitive impairment and Alzheimer's disease detection using with deep belief network model, *International Journal of Healthcare Management*, DOI: 10.1080/20479700.2022.2097764.
- [14] Razavi, F., M. J. Tarokh, and M. Alborzi. "An intelligent Alzheimer's disease diagnosis method using unsupervised feature learning. *J. Big Data.* 6, 32 (2019).