Alz-SAENet: A Deep Sparse Autoencoder based Model for Alzheimer's Classification

G Nagarjuna Reddy[1]
Research Scholar
JNT University
Ananthapuramu, India

K Nagi Reddy[2]
Professor, Dept. of ECE
N.B.K.R.I.S.T
Vidyanagar, India

***Abstract*—Precise identification of Alzheimer's Disease (AD) is vital in health care, especially at an early stage, since recognizing the likelihood of incidence and progression allows patients to adopt preventive measures before irreparable brain damage occurs. Magnetic Resonance Imaging is an effective and common clinical strategy to diagnose AD due to its structural details. we built an advanced deep sparse autoencoder-based architecture, named Alz-SAENet for the identification of diseased from typical control subjects using MRI volumes. We focused on a novel optimal feature extraction procedure using the combination of a 3D Convolutional Neural Network (CNN) and deep sparse autoencoder (SAE). Optimal features derived from the bottleneck layer of the hyper-tuned SAE network are subsequently passed via a deep neural network (DNN). This approach results in the improved four-way categorization of AD-prone 3D MRI brain images that prove the capability of this network in AD prognosis to adopt preventive measures. This model is further evaluated using ADNI and Kaggle data and achieved 98.9% and 98.215% accuracy and showed a tremendous response in distinguishing the MRI volumes that are in a transitional phase of AD.**

***Keywords*—Alzheimer's disease; MRI; CNN; sparse autoencoder; DNN; mild cognitive impairment**

I. Introduction

Alzheimer's Disease (AD) is the most predominantly reported dementia observed in elder people more than 60 years old. Late-onset effects of AD are most often seen in the mid-60s, whereas early-onset effects appear between the 30s and the mid-60s. Because of the world's aging population, it is anticipated that around 640 million individuals will be impacted by AD by the year 2050. Each year, more than 10 million fresh dementia cases are diagnosed worldwide, implying one fresh patient for every 3.2 seconds [1]. This situation leads to imposing significant impact on patients' and caretakers' daily living routines, physical and emotional states, and the economy. Early diagnosis is the only solution to find suitable medication that prevents additional damage to the cognitive ability of a patient.

Both the Mini Mental State Examination (MMSE) and the Clinical Dementia Rating (CDR) parameters are commonly used to estimate the severity of AD. Magnetic Resonance Images (MRI) are shown to be the most effective imaging biomarker [2, 3] in clinical assessment for analysing and getting a conclusion about the stage of dementia due to their ability in reflecting the structural details of human brain. In practise, several computer-aided diagnostic tools that employ machine learning methodologies [4, 5, 6, and 7] such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and ensemble methods [8], are suggested, developed and widely implemented using MRI to assist the medical practitioners community.

Due to the rapid advancement of Artificial Intelligence algorithms in diagnosis procedures, the deep learning techniques have been able to categorise, extract high level features, and will also aid in the right diagnosis of AD patients in short span of time. The potential for gratifying feedback from using deep learning algorithms in medical imaging prompted several investigators to emphasize the approach when tackling research difficulties and concerns [9, 10, and 11]. Convolutional Neural Network (CNN) [35] changed the complete picture of pattern recognition especially in AD diagnosis [12] with their capability in extraction of latent features from various objects by fine tuning its hyper parameters using optimizers. Sparse autoencoder (SAE), another architecture of deep learning model has shown exceptional performance in a wide range of unsupervised applications due to their ability in utilizing the sparsity in information bottleneck [13, 14, 15, and 16]. It excels in learning useful feature representations in very complex and large datasets, making it a possible solution to handle the difficulties of AD prediction.

In this research paper, a 3-stage neural network model that combines 3D CNN, SAE, and Deep Neural Network (DNN) is presented. Before feeding MRI volumes into this network, the first MRI volumes are pre-processed and converted into 2D slices in a series. In this research work, only 40 medial slices covering the hippocampal portion were considered so that MRI volume consists of AD symptoms. After pre-processing, convolutional layers in CNN are trained to extract the latent features from MRI data. SAE in the next stage reduces the feature dimension so that only dominant features are incorporated into DNN [17] in the third stage to classify the subjects into AD subcategories namely AD, low and stable Mild Cognitive Impairment (ls-MCI), progressive Mild Cognitive Impairment (p-MCI), and Cognitive Normal (CN). In contrast to earlier approaches, the proposed Alz-SAENet exhibits good accuracy, and fast convergence by leveraging the convolutional layer's potential in CNN and the SAE sparsity.

The key contributions of this work are:

- Alz-SAENet is developed for robust 4-way classification for AD diagnosis.
- For early diagnosis, this model can be utilized since classifying MCI stages were more concentrated.

The rest of this research article is organized as follows: literature is studied and challenges were drawn in Section II. Datasets and design of Alz-SAENet are discussed systematically in Section III. Experimental remarks and discussions are provided and Section IV. Finally, Section V summarized the work and future enhancement of this work.

II. Literature Review

A. Related Works

Jha et al. [18] developed a deep architecture that comprises SAE, scale conjugate gradient, stacked autoencoders, and finally one SoftMax layer for effective classification of AD subjects. When compared to other renowned investigations, this approach demonstrated a significant improvement in diagnosis, yielding an adequate and reliable accuracy of 91.6%.

Jabason et al. [19] described a novel technique to handle the missing data patterns in Alzheimer's diagnostic methods by leveraging the benefits of sparse autoencoder. It is a stacked sparse autoencoder that assign the missing data patterns with values and to select the discriminative features for supporting 3-way Alzheimer's disease diagnosis. The proposed model has attained 95.90% diagnostic accuracy over ADNI dataset by imputing the missing data patterns.

Soliman et al. [20] proposed an enhanced sparse autoencoder base CNN for 3-way Alzheimer's prediction. The authors tuned the hyperparameters of CNN and SAE thoroughly using the Adam optimizer and obtained 87.8% diagnostic accuracy over ADNI datasets. This diagnostic accuracy is attained not only due to the structure of the network and also with the pre-processing phase. Further improvement of diagnostic accuracy is a bit challenging matter.

Venugopalan et al. [21] implemented a deep learning (DL) model for analyzing multimodal data for the 3-way classification of AD. They used Stacked Denoising Autoencoders (SDAE) for feature extraction from multimodal data that comprise both clinical and genetic data and 3D-CNNs are employed for the categorization of MR volumetric data. This model outperformed almost all shallow architectures in terms of key evaluation metrics.

Zhu et al. [22] presented a sparse regression approach that utilizes a novel feature selection algorithm that considers task-wise relations amongst the clinical labels and neuroimaging features and 'self-representation' relations. Authors performed both binary and multiclass classification and this procedure outstripped all the conventional methods by improving the accuracy by an average of 4.5%.

Yagis et al. [23] focused on the end-to-end development of an AD detection technique that integrates supervised prediction with unsupervised representation using convolutional autoencoders. To capture hidden representations in structural MRI slices, a 2D convolutional autoencoder is built. Testing the network over OASIS repository data, it is revealed that their model beats several competing classifiers with 74.66% accuracy when employing a single slice.

Lin et al. [24] employed CNN to extract discriminative features from MRI volumetric image data. The authors concentrated on some specific subjects in the transition period from MCI to AD. They attained 79.9% accuracy with their model by a solid balance between specificity and sensitivity.

Basheera et al. [25] demonstrated an innovative method to extract the grey matter from the brain voxels, and CNN is employed for AD classification. The authors resliced 18,017 voxels from 1820 MRI volumes that were retrieved from the ADNI repository. With the support of their enhanced Independent Component Analysis (ICA), they extracted hidden structural features from pre-processed voxels. With this hybrid method, authors achieved 90.47% accuracy in the 3-way classification of AD, MCI, and CN subjects. Slow convergence and moderate diagnostic accuracy are challenges here.

Akramifard et al. [26] designed and developed a hybrid method that combines Autoencoder and SVM for Alzheimer's diagnosis using multimodal datasets including MMSE, MRI, PET, CSF, and personal information. An autoencoder is designed to inputting the missing data. PCA is employed for reducing data dimension. Finally, the SVM algorithm is utilized for classification. During the evaluation, their algorithm yielded 95.57% accuracy for multimodal data whereas for only MRI it produced 84.46% diagnostic accuracy. Although the algorithm performance is superior, processing such limited datasets in all phases is a challenging task.

Almuqhim et al. [27] developed ASD-SAENet, a deep sparse network for distinguishing autism spectrum disorder (ASD) patients from typical control subjects. The authors proposed and implemented an SAE, that proposed an optimized feature extraction procedure for Autism patients' classification. These features are subsequently loaded into a DNN for accurate classification of ASD-prone fMRI brain voxels. Based on both the restored data error and the classifier error, this model is trained to extract optimal hidden details. The model is evaluated over 1,035 Autism Brain Imaging Data Exchange (ABIDE) datasets, and 17 research centers and achieved 70.8% accuracy and 79.1% specificity.

B. Review

Detailed literature analysis is tabulated in Table I. Although the existing literature has many advantages, they are suffering from certain issues viz. number of AD stages, convergence, feature selection, overfitting, and hyperparameters tuning. These challenges to be addressed with productive outcomes by developing a novel network. From overall existing works, it is very clear that 3D CNN will holistically extract MRI latent features and converge quickly. The feature vector obtained from CNN would have huge redundant details that may not affect the classifier overall performance. In second stage, the sparsity of SAE aids in the mapping of low-level features from high-dimensional features while preventing overfitting due to its sample size. Finally, a

DNN is employed for the classification of subjects into four AD categories viz. CN, ls-MCI, p-MCI, and AD. With the hyperparameters tuning, SAE can avoid overfitting by minimizing its cost function and with quick convergence.

TABLE I. LITERATURE DEEP ANALYSIS

Authors	Data	Methodology	Features	Challenges
Jha et al. [18]	OASIS	SAE	Diagnostic accuracy is obtained as 91.6% over very small datasets.	Diagnostic accuracy is low.
Jabason et al. [19]	ADNI	SAE	Imputation of missing data patterrns is employed. 3-way classification is performed.	Accuracy is moderate. Early diagnosis is not precise.
Soliman et al. [20]	ADNI	SAE+CNN	Differentiates healthy controls from AD subjects. 3-way classification is performed.	Suffers from classifying MCI subjects. Less accuracy 87.8%.
Venugopalan et al. [21]	ADNI	3D CNN+ stacked DAE	3-way classification is performed. Multimodal data is used as biomarker for training.	Suffers from limited dataset sizes. Takes long running time. Early diagnosis is a bit tough.
Almuqhim et al. [27]	ADNI	SAE+DNN	Authors designed a network named ASD-SAENet to diagnose Autism Sepctrum Disorder. Evaluated their model on ABIDE dataset and 70.8% accuracy is achieved.	Very low accuracy i.e. 70.8%. Features were not extracted separately in this network.
Basheera et al. [25]	ADNI	CNN+ICA	An enhanced Independent Component Analysis is employed for segmentation. For classification, CNN is utilised.	Multiclass classification accuracy is 90.47%. Diagnostic accuracy is moderate. Converging slowly.
Zhu et al. [22]	ADNI	SVM	A novel feature selection approach is employed using task specific relations. Binary and Multiclass classification is performed.	Deep learning could be used to improve the analysis. Diagnostic Accuracy is less.
Yagis et al. [23]	OASIS	2D CAE	Integrated supervised prediction with unsupervised representations using CAE. Single slice of MRI is used in diagnosis.	3D volumetric latent features were not extracted. Diagnostic accuracy is also less i.e 74.66%.
Akramifard et al. [26]	ADNI	CNN+SAE	2D CNN and 3D CNN both were employed. 3-way classification with 89.47% accuracy is obtained.	Tuned hyper-parameters could further enhance these findings.
Dongren et al [8]	ADNI	Ensemble methods	Authors presented a hierarchical grouping process in feature selection method. Binary and 4-way classification both were performed.	Deep learning architectures may improve accuracy further. Diagnostic accuracy is very poor i.e., 54.375%.
Vu et al [32]	ADNI	CNN+SAE	A novel classification model is presented in a multimodality fusion of MRI and PET. 91.14% accuracy is obtained using MRI-PET.	Instead of random, authors could select a specific patch. Overall, system converge very slowly.
Yang et al.	OASIS	SVM+ICA	Machine learning is employed so it is compulsory to specify ROI. 3-way AD diagnosis is employed.	classification accuracy is still not optimal due to various factors. Age and gender are not considered.
Lin et al. [24]	ADNI	CNN	This model is centric about MCI to AD conversion subjects. This approach achieved an accuracy of 79.9%.	AD progression was not analyzed. Diagnostic accuracy is moderate.

III. NETWORK DESIGN AND DEVELOPMENT

This research work seeks to properly diagnose AD stages while using less processing power and storage space. To achieve these goals, 3D CNN along with SAE is presented for obtaining high-level dominant features related to multi-class AD from MRI neuroimage intensities. As shown in Fig. 1, unprocessed MRI volumes are thoroughly scaled, resized, and segmented properly before being directed into the hyperparameters tuned CNN followed by SAE to extract latent features. The low dimensional features from the bottleneck layer of SAE are subsequently loaded into a deep neural network (DNN) which results in the enhanced diagnosis of brain voxels more prone to AD.

A. Dataset

From the ADNI repository, 1120 unprocessed MRI volumes of 460 people of various ages and genders were obtained. We employed 1.5-Tesla, T1-weighted, MRI volumes from the patient community [28] that included patients ranging in age from 55 to 70 and of different genders. The configuration of the MRI dataset utilized to implement this research is displayed in Table II. In the entire dataset, 351 CN, 230 ls-MCI, 234 p-MCI, and 305 AD subject volumes were meticulously gathered to orient the study effort toward the AD prognosis. Furthermore, 6400 samples of MRI slices were retrieved from the Kaggle repository and utilized for evaluating the proposed network. In all 6400 samples, 3200 CN subjects, 2240 ls-MCI subjects, 896 p-MCI subjects, and 64 AD subjects MRI slices are selected to validate the network capability in classifying the MCI subcategories.

TABLE II. ADNI DATASET CONFIGURATION [28]

Parameters	CN	p-MCI	ls-MCI	AD
Age	55	60	58	62
Gender	60% male	40% male	50% male	30% male
MMSE	29	24	26	21
CDR	0	0.5	0.3	0.5-1

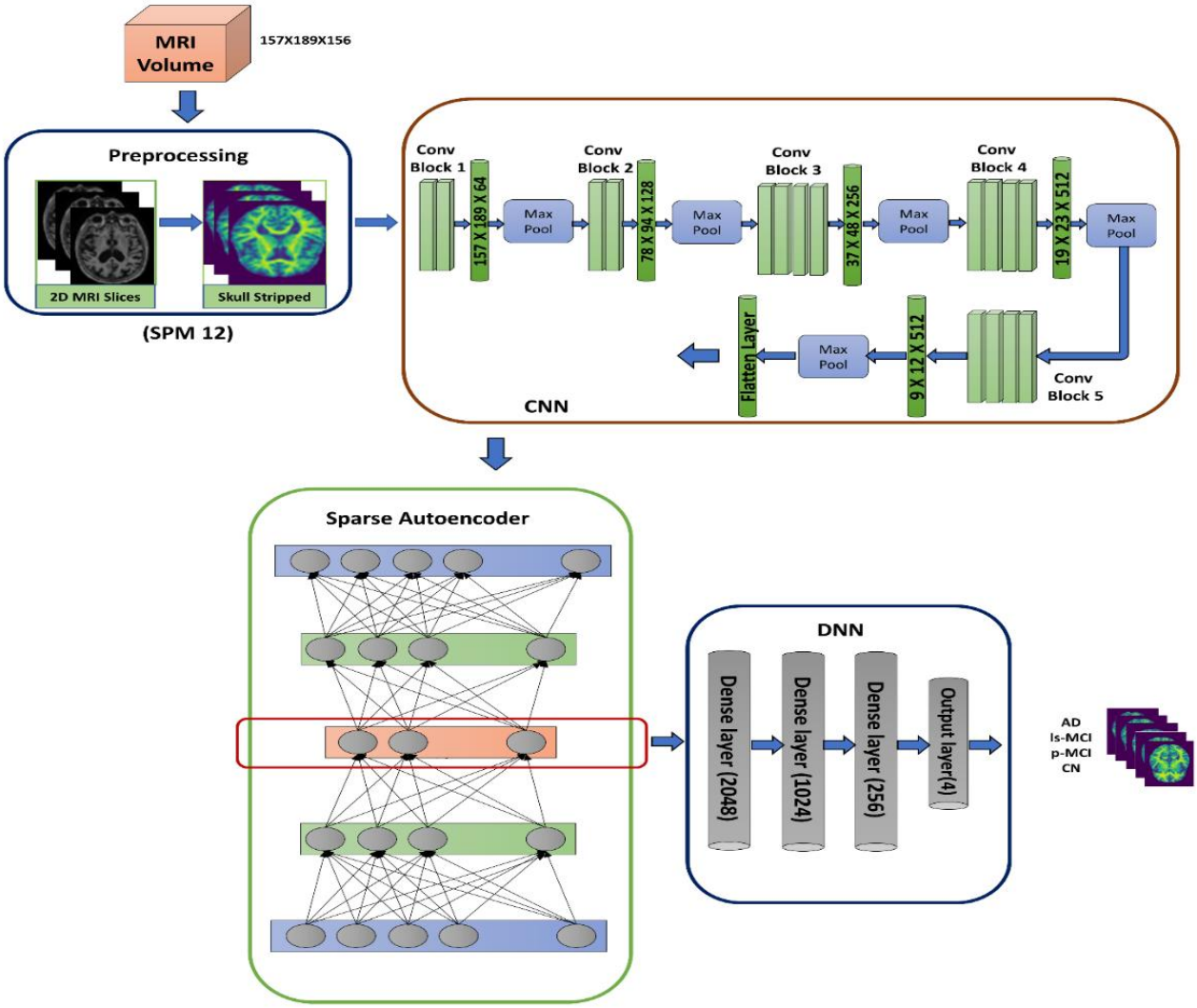

Fig. 1. Alz-SAENet Architecture.

B. Preprocesing

Generally, the raw MRI data are stored in NifTi (.nii) format after retrieving from the ADNI repository. These NifTi files were further pre-processed utilizing the provision of the SPM12 package [29] installed under default settings. This stage pipeline is very much responsible for the removal of non-brain tissues and registering it into Montreal Neurological Institute (MNI) space. After pre-processing, the registered volumes are smoothed and scaled such that all volumes are with the same dimension of 157*189*156 and sliced to generate $1.5*1.5*1.5mm^3$ voxels. Further, the MRI volumetric data is normalized voxel by voxel such that all intensities in the voxel fall in the 0 to 1 range for additional processing by retaining the disease features. For this investigation, only 40 medial MRI slices were considered such that they replicate AD progression clearly via hippocampal lesions to enable the network to distinguish the AD subjects. Sagittal plane views of pre-processed MRI volume for a sample are shown in Fig. 2.

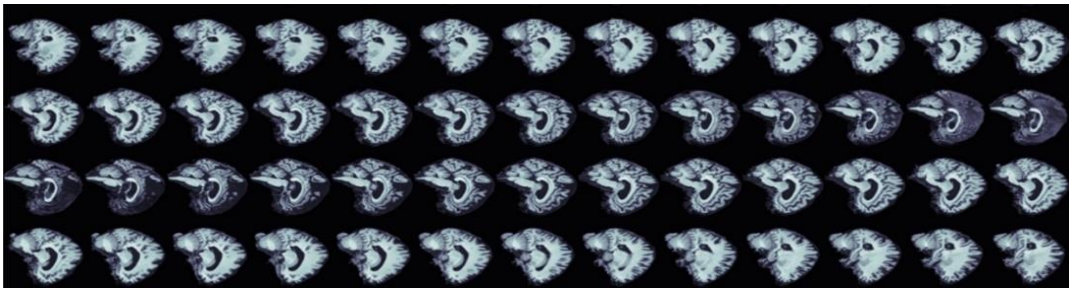

Fig. 2. Pre-processed Medial Sagittal Planes of Brain MRI.

C. 3D CNN

Generally, CNN models consist of convolutional layers, batch normalization layers, pooling layers, and fully connected layers in sequential and backpropagation algorithm is used to learn latent features from MRI slices. Advanced architectures follow more complicated topologies for extracting features in complex applications where highly correlated data is present. In this work, we focused on a sequential model inspired by VGG 19 architecture [30], i.e., one popular version of Visual Geometric Group. VGG19 is a deep CNN that has been trained on millions of images with complex classification tasks.

This architecture consists of five blocks, amongst the very first two blocks comprise two convolutional layers preceded by a pooling layer each, and the last three blocks have four convolutional layers preceded by a pooling layer with a filter size of 2x2x2. In this method, all pretrained layers are frozen, and the output of the last convolution block is considered a discriminative feature of the MRI volume. This feature map is flattened using a flattening layer before being sent to a sparse encoder. All layers, dimensions, and several parameters for training the VGG19 network employed in this work for one MRI slice were tabulated in Table III.

D. Sparse Autoencoder

A sparse autoencoder [31, 32] is merely an autoencoder with a sparsity penalty as a training benchmark. This autoencoder is used to draw more valuable insights from MRI with reasonable dimensionality. In addition, the bias applied at the encoder and decoder forces the AE to restore the input more accurately by avoiding overfitting.

Let s_i is the i^{th} input among the given training samples (s_1, s_2,….s_N). Then, the designed SAE is accomplished to restore the input si maximum similar to the estimated function $h_{w,b}(x^i)$. The cost function of SAE consists of three very important factors, viz., Mean squared error (MSE), Weight Decay (WD) and Sparsity [33]. The mean squared error for all N training samples and weight decay is mathematically expressed as in eq. (1) and eq. (2).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left\|h_{w,b}(x^i) - y^i\right\| \quad (1)$$

$$WD = \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(w_{ji}^{l})^2 \quad (2)$$

This weight decay eq. (2) helps to avoid overfitting. Overfitting may result from a small value of λ, whereas underfitting may result from a big value of λ. To choose lambda for the term's best match, we, therefore, carried out a number of empirical studies.

TABLE III. LAYERS AND FEATURE MAPS OF 3D CNN

Block	Layer	Dimension	Parameters
	Input	[(None, 157, 189, 3)]	0
1	conv1 (Conv)	(None, 157, 189, 64)	1792
	conv2 (Conv)	(None, 157, 189, 64)	36928
	pool (MaxPooling)	(None, 78, 94, 64)	0
2	conv1 (Conv)	(None, 78, 94, 128)	73856
	conv2 (Conv)	(None, 78, 94, 128)	147584
	pool (MaxPooling)	(None, 39, 47, 128)	0
3	conv1 (Conv)	(None, 39, 47, 256)	295168
	conv2 (Conv)	(None, 39, 47, 256)	590080
	conv3 (Conv)	(None, 39, 47, 256)	590080
	conv4 (Conv)	(None, 39, 47, 256)	590080
	pool (MaxPooling)	(None, 19, 23, 256)	0
4	conv1 (Conv)	(None, 19, 23, 512)	1180160
	conv2 (Conv)	(None, 19, 23, 512)	2359808
	conv3 (Conv)	(None, 19, 23, 512)	2359808
	conv4 (Conv)	(None, 19, 23, 512)	2359808
	pool (MaxPooling)	(None, 9, 11, 512)	0
5	conv1 (Conv)	(None, 9, 11, 1024)	2359808
	conv2 (Conv)	(None, 9, 11, 1024)	2359808
	conv3 (Conv)	(None, 9, 11, 1024)	2359808
	conv4 (Conv)	(None, 9, 11, 1024)	2359808
	pool(MaxPooling)	(None, 4, 5, 1024)	0

Sparsity is the last term of the cost function that help to avoid overfitting by generating the activation for hidden neurons of the SAE to avoid overfitting. The eq. (3) is the average activated value of hidden layer, where 'a' is the ReLU activation function.

$$\widehat{p_j} = \frac{1}{N}\sum_{i=1}^{N}[a_j^2(x^i)] \quad (3)$$

The sparsity is calculated so that $\widehat{p_j}$=p constraint is satisfied, where p is the sparsity parameter. The deviation of $\widehat{p_j}$ from p help the network to activate and deactivate neurons in the hidden layer. KL divergence is used to define the parameter as in eq. (4).

$$\sum_{j=1}^{s_l} KL(p||\widehat{p_j}) = \sum_{j=1}^{s_l}[p\log\left(\frac{p}{\widehat{p_j}}\right) + (1-p)\ \log(\frac{1-p}{1-\widehat{p_j}})] \quad (4)$$

By combining the three terms, the cost function of the SAE J_{SAE} is defined as shown in eq. (5).

$$J_{SAE} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left\|h_{w,b}(x^i) - y^i\right\| + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}\left(w_{ji}^{l}\right)^2 + \beta\sum_{j=1}^{s_l} KL(p||\widehat{p_j}) \quad (5)$$

Here β is the sparse penalty.

The primary goal of the Sparse Autoencoder [34] in this investigation is to reduce the dimensionality of 55,296 features obtained from 3D CNN. The SAE bottleneck layer provides more dominant insights, i.e., features that can be employed in classification. The bottleneck layer size is 2048 hidden units. Output from bottleneck generally has very limited number of features i.e. 2048 elements in the vector. These features are subsequently passed through DNN in final stage for classification purpose.

E. DNN

DNN is an artificial neural network with several hidden layers for solving complex classification problems. All these layers are dense, and the role of DNN here is in the final phase for classifying the latent features obtained from SAE. This DNN comprises two hidden layers, one input and output layer, with sizes 2048, 1024, 256, and 4, respectively. A SoftMax layer is generally employed as an output layer to estimate the incoming feature vector's four possible classes, i.e., CN, p-MCI, ls-MCI, and AD. A dropout layer is also used between two dense layers to avoid overfitting. Finally, cross-entropy is used for determining the classifier cost function, and the weight decay term is added.

IV. Results and Investigations

Concurrent training of the SAE and the DNN improved feature extraction while optimizing the classifier's decision. The training was completed in 50 epochs with 16 batch size. Sparse penalty, weight decay and sparsity parameter p were initialized 2, 0.0001, and 0.05, respectively. In last 20 epochs, DNN was fine-tuned, and parameters updated to minimize the cost function while the SAE parameters were frozen.

All the experimentations described in this paper were carried out in the Google Colab platform using python scripting with the support of 1X Tesla K80 GPU. To run our deep learning network, the Tensorflow Keras library is employed. To demonstrate the generalizability of the network in AD diagnosis, Alz-SAENet is examined by applying ADNI dataset and the Kaggle dataset in two scenarios.

In the first scenario, the proposed Alz-SAENet architecture is trained over 80% of the whole 1120 MRI volumes of the ADNI dataset that consists of MRI volumes with uniform dimensions of 157*189*156. After rigorous training, the network is tested using the remaining MRI volumes among the dataset acquired. The test results were assessed in terms of diagnostic accuracy, network sensitivity, precision, specificity, Negative Predictive Value (NPV), False Positive Rate (FPR), F1-Score, False Negative Rate (FNR), False Discovery Rate (FDR), and Mathew Correlation Coefficient (MCC). All these parameters are evaluated and listed in Table IV. AD vs CN classification over the ADNI dataset is 99.54% accurate and 100% sensitive, and produce MCC is 89.93%. p-MCI vs ls-MCI classification that facilitates early diagnosis of AD yields 98.56% accuracy, 100% sensitivity and 96.64% MCC. p-MCI vs AD classification is 97.90% accurate, 100% sensitive and 92.10% correlated data. CN vs ls-MCI classification that facilitates early diagnosis of AD yields 97.71% accuracy, 98.15% sensitivity and 95.24% MCC. However, Multiclass diagnosis produced 98.9% accuracy, 97.5% sensitivity and 94.25% MCC. These results demonstrate that Alz-SAENet has significantly classified AD stages, especially for the transitional period between ls-MCI and p-MCI, with an accuracy of 98.56%. The graphical representation is also depicted in Fig. 3.

In the second scenario, the proposed model was evaluated using another repository, i.e. Kaggle, to prove the generalization ability of this network over the unseen data during training. 6400 sample MRI medial slices belonging to four AD classes were acquired from the repository and evaluated in this model. Unlike in the first scenario, the evaluation metrics accuracy, precision, sensitivity, and F1-Score were only obtained and tabulated in Table V. The bar chart for these evaluation metrics is also shown in Fig. 4. AD vs CN classification yields 98.62% accuracy, 92% sensitivity and 98% precision. p-MCI vs ls-MCI diagnosis also attained 98.37% accuracy, 97% sensitivity, and 98% precision. This network produced 98.215% accuracy, 92.25% sensitivity, and 82.25% precision in multiclass classification.

TABLE IV. Alz-SAENet Performance Over ADNI Dataset

Classification	Accuracy	F1-Score	Sensitivity	Specificity	Precision	NPV	MCC
AD vs CN	0.9954	0.8966	1	0.9953	0.8125	1	0.8993
p-MCI vs ls-MCI	0.9856	0.9761	1	0.9797	0.9534	1	0.9664
p-MCI vs AD	0.9790	0.9450	1	0.9821	0.9452	1	0.9210
ls-MCI vs CN	0.9771	0.9807	0.9815	0.9707	0.9799	0.9729	0.9524
Multiclass	0.989		0.975		0.915		0.9425

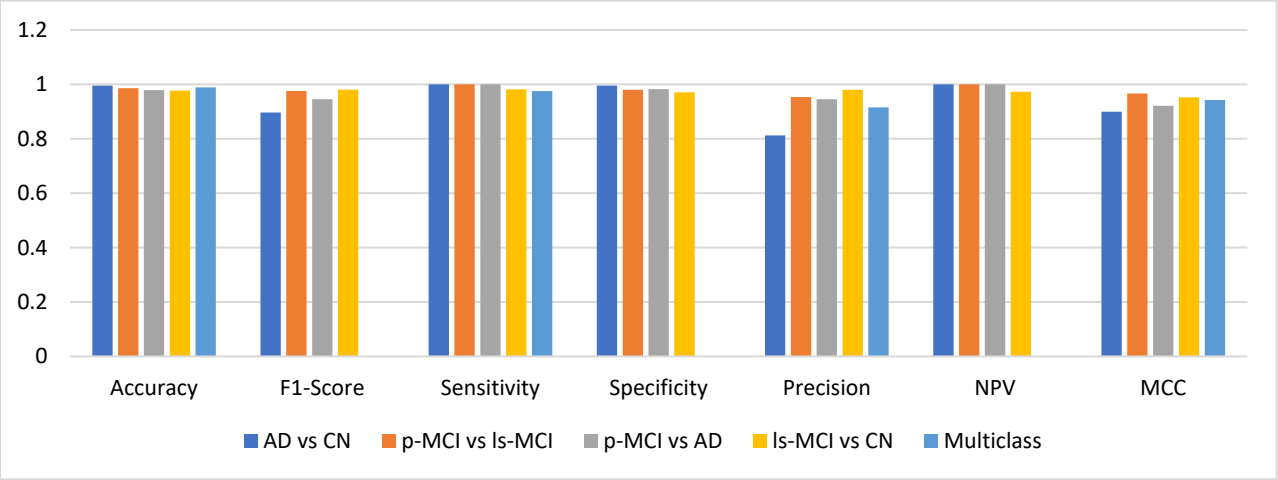

Fig. 3. Graphical Presentation of Alz-SAENet Performance over ADNI Data.

TABLE V. ALZ-SAENET PERFORMANCE OVER KAGGLE DATASET

Classification	**Accuracy**	**F1-Score**	**Sensitivity**	**Precision**
AD vs CN	0.9862	0.95	0.92	0.98
p-MCI vs ls-MCI	0.9837	0.98	0.97	0.98
p-MCI vs AD	0.9759	0.98	0.97	0.98
ls-MCI vs CN	0.9828	0.49	0.81	0.35
Multiclass	0.98215	0.85	0.9225	0.8225

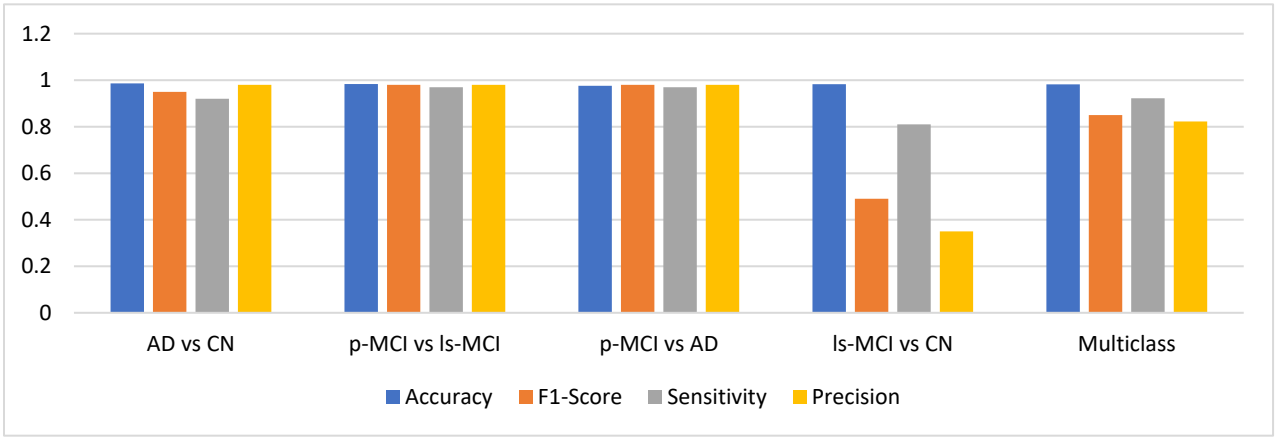

Fig. 4. Graphical view of Alz-SAENet Performance over Kaggle Data.

The results obtained in second scenario reveal that the network has shown good generalizability among the datasets which never seen during training. The model did not overfit due to the unbalanced datasets appeared in training process. This was happened because the SAE hyperparameters are tuned systematically so that its cost function gets minimized as the weight updation is in progress. With this support, the network can reliable outcomes even in non-supportive environments.

Attained results demonstrate that the proposed model, i.e., Alz-SAENet has greater generalizability in classifying the unseen data during the training process. The deep learning techniques from literature [9], [20], [25], and [26] are suffered from different challenges in terms of convergence speed, overfitting, and hyperparameters tuning to obtain optimal performance from their networks. The results of Alz-SAENet have revealed that it has attained almost all those objectives by leveraging the benefits of Sparse autoencoder by minimizing the cost function.

The proposed model exhibited very poor performance over ADNI data in attaining precision, f1-score, and MCC for classifying AD and CN subjects. When applied to Kaggle data, it has not produced inadequate performance in terms of f1-score, sensitivity, precision, etc. for ls-MCI vs CN classification.

TABLE VI. COMPARISON OF ALZ-SAENET WITH LITERATURE REPORTED

Author	Modality	Datasets	Classifier	Subjects	Classification Metrics		
					Accuracy (%)	Precision (%)	Recall (%)
Soliman et al. [20]	MRI	ADNI	SAE+CNN	Multiclass	87.8	91%	88%
Akramifard et al. [26]	MRI	ADNI	Autoencoder + PCA + SVM	AD vs MCI MCI vs CN AD vs CN	66.84 66.97 84.46	-	48.70 78.94 81.87
Basheera et al. [25]	MRI	ADNI	CNN+ICA	Multiclass AD vs MCI MCI vs CN AD vs CN	86.7 96.2 98.0 100	89.6 93.0 96.0 100	86.61 100 100 100
Alz-SAENet	**MRI**	**ADNI, Kaggle**	**3D CNN+SAE +DNN**	**AD vs CN** **CN vs ls-MCI** **AD vs p-MCI** **ls-MCI vs p-MCI** **Multiclass**	**99.54** **97.71** **97.90** **98.56** **98.90**	**81.25** **97.99** **94.52** **95.34** **91.5**	**100** **98.15** **100** **100** **97.5**

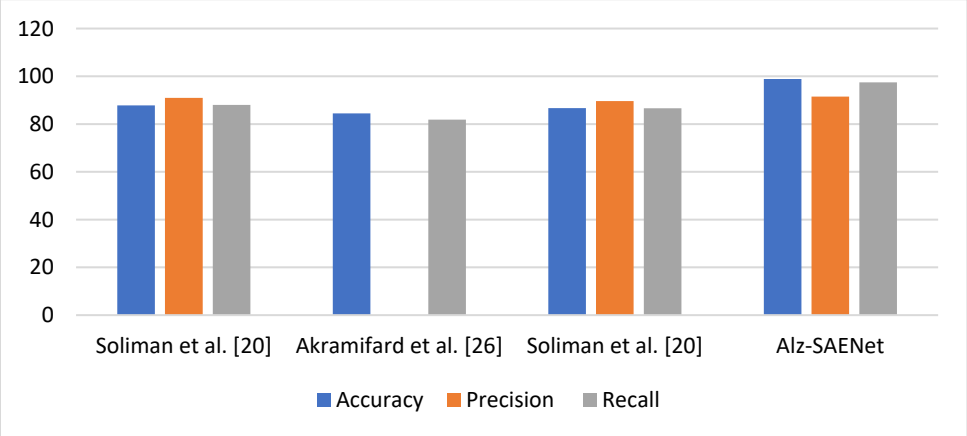

Fig. 5. Comparison of Alz-SAENet with Literature Reported.

The outcomes of the Alz-SAENet, i.e. integrated version of 3D CNN, SAE, and DNN illustrate that the performance of AD diagnosis is enhanced with the support of features that has learned and fine-tuning of hyperparameters of SAE during training. The model performs well on the test data, which is an important indicator of its generalizability, given that the model has never seen the data before. In addition, the proposed approach is also compared with recent literature in deep learning domain on ADNI data. This comparison is evidently shown in Table VI, and the proposed network outperformed all those methods focused by the literature deep analysis in Section II in terms of key parameters accuracy, precision, recall, etc. The same comparison is also depicted in Fig. 5.

V. CONCLUSION

In this paper, we built a novel deep sparse autoencoder-based learning model named Alz-SAENet for classifying brain MRI volumes that exhibit severe AD symptoms to cognitive normal. This network utilized 3D CNN, SAE, and DNN components that assisted in understanding the neurobiological foundations of the AD brain in a better way. 3D CNN extracted 55,296 latent features from MRI volume. These features were flattened and their dimension was reduced through SAE via bottleneck layer and fed to DNN for classifying them into four stages of AD. During testing, this network gave an accuracy of 98.9% over the ADNI dataset and 98.215% over the Kaggle data. These results demonstrated that Alz-SAENet outscored all the state-of-the-art approaches that worked on AD classification. In the future, we test the architecture over real-world data along with metadata to improve the generalizability, efficiency, and efficacy of early diagnosis of AD. Furthermore, the results obtained for ls-MCI and CN classification over the Kaggle dataset are needed to be enhanced.

REFERENCES

[1] Gauthier S, Rosa-Neto P, Morais JA, & Webster C. 2021. World Alzheimer Report 2021: Journey through the diagnosis of dementia. London, England: Alzheimer's Disease International.

[2] Johnson KA, Fox NC, Sperling RA, Klunk WE. Brain imaging in Alzheimer disease. Cold Spring Harb Perspect Med. 2012 Apr;2(4):a006213. doi: 10.1101/cshperspect.a006213. PMID: 22474610; PMCID: PMC3312396.

[3] Saad, S.H.S., Alashwah, M.M.A., Alsafa, A.A. et al. The role of brain structural magnetic resonance imaging in the assessment of hippocampal subfields in Alzheimer's disease. Egypt J Radiol Nucl Med 51, 53 (2020). https://doi.org/10.1186/s43055-020-00164-8.

[4] Vichianin Y, Khummongkol A, Chiewvit P, Raksthaput A, Chaichanettee S, Aoonkaew N, Senanarong V. Accuracy of Support-Vector Machines for Diagnosis of Alzheimer's Disease, Using Volume of Brain Obtained by Structural MRI at Siriraj Hospital. Front Neurol. 2021 May 10;12:640696. doi: 10.3389/fneur.2021.640696. PMID: 34040575; PMCID: PMC8141708.

[5] B. Rabeh, F. Benzarti and H. Amiri, "Diagnosis of Alzheimer Diseases in Early Step Using SVM (Support Vector Machine)," 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV), 2016, pp. 364-367, doi: 10.1109/CGiV.2016.76.

[6] S. Aruchamy, V. Mounya and A. Verma, "Alzheimer's Disease Classification in Brain MRI using Modified kNN Algorithm," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), 2020, pp. 1-6, doi: 10.1109/iSSSC50941.2020.9358867.

[7] Reddy, G & K, Nagireddy. (2021). A Robust Machine Learning Approach for Multiclass Alzheimer's Disease Detection using 3D Brain Magnetic Resonance Images. Journal of Engineering Research. 10.36909/jer.10511.

[8] Yao, Dongren; Calhoun, Vince D.; Fu, Zening; Du, Yuhui; Sui, Jing (2018). An Ensemble Learning System for a 4-Way Classification of Alzheimer's Disease and Mild Cognitive Impairment. Journal of Neuroscience Methods, (), S0165027018300797–. doi:10.1016/j.jneumeth.2018.03.008.

[9] Jo T, Nho K, Saykin AJ. Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data. Front Aging Neurosci. 2019 Aug 20;11:220. doi: 10.3389/fnagi.2019.00220. PMID: 31481890; PMCID: PMC6710444.

[10] S. Al-Shoukry, T. H. Rassem and N. M. Makbol, "Alzheimer's Diseases Detection by Using Deep Learning Algorithms: A Mini-Review," in IEEE Access, vol. 8, pp. 77131-77141, 2020, doi: 10.1109/ACCESS.2020.2989396.

[11] S. Gao and D. Lima, A review of the application of deep learning in the detection of Alzheimer's disease, International Journal of Cognitive Computing in Engineering (2022) 1–8.

[12] P. Rath, P. K. Mallick, R. Siddavatam, and G. S. Chae, "An empirical development of hyper-tuned CNN using spotted hyena optimizer for bio-medical image classification," Journal of Natural Science, Biology and Medicine, vol. 12, no. 3, 2021.

[13] Alkabawi, Elham M.; Hilal, Allaa R.; Basir, Otman A. (2017). [IEEE 2017 IEEE International Conference on Systems, Man and Cybernetics (SMC) - Banff, AB (2017.10.5-2017.10.8)] 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) - Feature abstraction for early detection of multi-type of dementia with sparse auto-encoder., (), 3471–3476. doi:10.1109/smc.2017.8123168.

[14] H. B. Baydargil, J. -S. Park and D. -Y. Kang, "Classification of Alzheimer's Disease Using Stacked Sparse Convolutional Autoencoder," 2019 19th International Conference on Control, Automation and Systems (ICCAS), 2019, pp. 891-895, doi: 10.23919/ICCAS47443.2019.8971696.

[15] Liu, J., Wu, G., Luo, Y., Qiu, S., Yang, S., Li, W., & Bi, Y. (2020). EEG-Based Emotion Classification Using a Deep Neural Network and Sparse Autoencoder. Frontiers in Systems Neuroscience, 14. https://doi.org/10.3389/fnsys.2020.00043.

[16] An, N., Ding, H., Yang, J., Au, R., and Ang, T. F. A., "Deep ensemble learning for Alzheimers disease classification", arXiv e-prints, 2019.

[17] Gouse, Mohammed. (2021). Deep Neural Network Concepts for Classification using Convolutional Neural Network: A Systematic Review and Evaluation. Technium Romanian Journal of Applied Sciences and Technology. 3. 58-70. 10.47577/technium.v3i8.4554.

[18] Jha, Debesh & Kwon, Goo-Rak. (2017). Alzheimer's Disease Detection Using Sparse Autoencoder, Scale Conjugate Gradient and Softmax Output Layer with Fine Tuning. 7. 13-17. 10.18178/ijmlc.2017.7.1.612.

[19] E. Jabason, M. O. Ahmad and M. N. S. Swamy, "Missing Structural and Clinical Features Imputation for Semi-supervised Alzheimer's Disease Classification using Stacked Sparse Autoencoder," 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS), 2018, pp. 1-4, doi: 10.1109/BIOCAS.2018.8584844.

[20] Soliman, Sarah & El-Dahshan, El-Sayed & Salem, Abdel-Badeeh. (2021). Diagnosis of Alzheimer's Disease by Three-Dimensional Convolutional Neural Network Using Unsupervised Feature Learning Method. International Journal of Intelligent Computing and Information Sciences. 1-15. 10.21608/ijicis.2021.80596.1103.

[21] Venugopalan, J., Tong, L., Hassanzadeh, H.R. et al. Multimodal deep learning models for early detection of Alzheimer's disease stage. Sci Rep 11, 3254 (2021). https://doi.org/10.1038/s41598-020-74399-w.

[22] Zhu X, Suk HI, Lee SW, Shen D. Discriminative self-representation sparse regression for neuroimaging-based alzheimer's disease diagnosis. Brain Imaging Behav. 2019 Feb;13(1):27-40. doi: 10.1007/s11682-017-9731-x. PMID: 28624881; PMCID: PMC5811409.

[23] E. Yagis, A. G. S. de Herrera and L. Citi, "Convolutional Autoencoder based Deep Learning Approach for Alzheimer's Disease Diagnosis using Brain MRI," 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS), 2021, pp. 486-491, doi: 10.1109/CBMS52027.2021.00097.

[24] Lin W, Tong T, Gao Q, Guo D, Du X, Yang Y, Guo G, Xiao M, Du M, Qu X and The Alzheimer's Disease Neuroimaging Initiative (2018) Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer's Disease Prediction From Mild Cognitive Impairment. Front. Neurosci. 12:777. doi: 10.3389/fnins.2018.0077.

[25] Basheera, Shaik & Ram, M. (2020). A novel CNN Based Alzheimer's Disease Classification using Hybrid Enhanced ICA Segmented Gray Matter of MRI. Computerized Medical Imaging and Graphics. 81. 101713. 10.1016/j.compmedimag.2020.101713.

[26] Akramifard H, Balafar MA, Razavi SN, Ramli AR. Early Detection of Alzheimer's Disease Based on Clinical Trials, Three-Dimensional Imaging Data, and Personal Information Using Autoencoders. J Med Signals Sens. 2021 May 24;11(2):120-130. doi: 10.4103/jmss.JMSS_11_20. PMID: 34268100; PMCID: PMC8253314.

[27] Almuqhim F and Saeed F (2021) ASD-SAENet: A Sparse Autoencoder, and Deep-Neural Network Model for Detecting Autism Spectrum Disorder (ASD) Using fMRI Data. Front. Comput. Neurosci. 15:654315. doi: 10.3389/fncom.2021.654315.

[28] G. N. Reddy and K. N. Reddy, "Boosting based Deep hybrid Framework for Alzheimer's Disease classification using 3D MRI," 2022 6th International Conference on Devices, Circuits and Systems (ICDCS), 2022, pp. 100-106, doi: 10.1109/ICDCS54290.2022.9780736.

[29] Ashburner, John & Barnes, Gareth & Chen, Chun-Chuan & Daunizeau, Jean & Flandin, Guillaume & Friston, Karl & Gitelman, Darren & Glauche, Volkmar & Henson, Rik & Hutton, Chloe & Jafarian, Amirhossein & Kiebel, Stefan & Kilner, James & Litvak, Vladimir & Mattout, Jérémie & Moran, Rosalyn & Penny, Will & Phillips, Christophe & Razi, Adeel & Zeidman, Peter. (2021). SPM12 Manual.

[30] Bhatele K. R, Bhadauria S. S. Classification of Neurodegenerative Diseases Based on VGG 19 Deep Transfer Learning Architecture: A Deep Learning Approach. Biosc.Biotech.Res.Comm. 2020;13(4).

[31] Mahmoud A., Karamti H. and Alrowais F. (2020). An Effective Sparse Autoencoders based Deep Learning Framework for fMRI Scans Classification. In Proceedings of the 22nd International Conference on Enterprise Information Systems - Volume 1: ICEIS, ISBN 978-989-758-423-7, pages 540-547. DOI: 10.5220/0009397605400547.

[32] T. Vu, H. Yang, V. Nguyen, A. Oh and M. Kim, "Multimodal learning using convolution neural network and Sparse Autoencoder," in 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, 2017 pp. 309-312. doi: 10.1109/BIGCOMP.2017.7881683.

[33] Antelmi, L., Ayache, N., Robert, P. & Lorenzi, M.. (2019). Sparse Multi-Channel Variational Autoencoder for the Joint Analysis of Heterogeneous Data., Proceedings of the 36th International Conference on Machine Learning in Proceedings of Machine Learning Research, 97:302-311.

[34] P. Bhatkoti and M. Paul, "Early diagnosis of Alzheimer's disease: A multi-class deep learning framework with modified k-sparse autoencoder classification," 2016 International Conference on Image and Vision Computing New Zealand (IVCNZ), 2016, pp. 1-5, doi: 10.1109/IVCNZ.2016.7804459.

[35] Gurukumar Lokku, G Harinatha Reddy, M N Giri Prasad, Optimized Scale-Invariant Feature Transform with Local Tri-directional Patterns for Facial Expression Recognition with Deep Learning Model, The Computer Journal, Volume 65, Issue 9, September 2022, Pages 2506–2527, https://doi.org/10.1093/comjnl/bxab088.