IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 March 2024, accepted 18 May 2024, date of publication 3 June 2024, date of current version 14 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3408311

RESEARCH ARTICLE

CirMNet: A Shape-Based Hybrid Feature Extraction Technique Using CNN and CMSMD for Alzheimer's MRI Classification

G. MALU[1], NAYANA UDAY[1], (Member, IEEE), ELIZABETH SHERLY[1], AJITH ABRAHAM[2,3], (Senior Member, IEEE), AND NARENDRA KUBER BODHEY[4]

[1]School of Computer Science and Engineering, Digital University Kerala, Kerala 695317, India
[2]School of Artificial Intelligence, Bennett University, Greater Noida, Uttar Pradesh 201310, India
[3]Innopolis University, Innopolis, 420500 Republic of Tatarstan, Russia
[4]Department of Radiodiagnosis, All India Institute of Medical Sciences Raipur, Chhattisgarh 492099, India

Corresponding authors: G. Malu (malu.g@duk.ac.in) and Ajith Abraham (ajith.abraham@ieee.org)

This work was financially supported by The Analytical Center for the Government of the Russian Federation (Agreement No. 70-2021-00143 dd. 01.11.2021, IGK 000000D730324P540002).

ABSTRACT This research introduces a novel approach called Circular Mesh Network (CirMNet), a shape-based hybrid feature extraction technique. It also proposes innovative Fractal Dimension (FD) and statistical feature extraction techniques for vertically symmetrical images. Convolutional Neural Networks (CNNs) have gained widespread popularity across various domains, including the interpretation of medical images. CNNs excel at extracting prominent features in the initial layers and progressively learn to capture more complex features as they advance. However, the pooling and striding operations involved in CNNs can lead to a loss of spatial and structural details in the image because CNNs require a mechanism to preserve the internal representation and describe the intricate relationships between image components or pixels. Circular Mesh-based Shape and Margin Descriptor (CMSMD) focuses on extracting structural, statistical, and property-based features. However, it does not encompass features such as texture or color. The objective of CirMNet is to leverage the strengths of both CNNs and CMSMD, and to mitigate their respective weaknesses. Structural and texture features were generated from CirMNet, and its performance was evaluated for the diagnosis of neurodegenerative disorders, particularly Alzheimer's Disease (AD). The model can easily identify the permanent shrinkage and destruction of brain cells in MRIs of patients and exhibited a notable accuracy of 97.34% in classifying various stages of AD, encompassing Control, Early Mild Cognitive Impairment (EMCI), and Late Mild Cognitive Impairment (LMCI). This achievement represents a substantial improvement over the existing state-of-the-art methods in the domain.

INDEX TERMS Alzheimer's disease, magnetic resonance imaging, convolutional neural network, structural shapes, feature extraction.

I. INTRODUCTION

HARNESSING the power of Machine Learning (ML) techniques, fields as diverse as bioinformatics, social media, speech recognition, and computer vision have experienced a revolution, propelling advancements and innovation to unprecedented heights [1], [2], [3]. Deep learning has become increasingly prevalent in these domains and plays a crucial role in facilitating intelligent decision-making processes. In numerous medical applications, it is essential to prioritize the examination of the structural characteristics of affected regions [3], [4]. It includes diagnosing neurological disorders

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Giannelli.

and identifying breast and skin cancers [5]. To develop an ML system for object detection or classification, several steps must be identified, such as image acquisition, enhancement, preprocessing, segmentation, pattern recognition, and machine-learning algorithms. Feature extraction and selection are crucial components of ML classification models and are carefully selected based on the image type to achieve high accuracy [6].

Deep learning models incorporate these image-processing and classification tasks into their single architectures. To enhance their performance, these models require extensive training datasets [7]. A complicated hierarchy of layers comprising nonlinear functions and transformations constitutes a Convolutional Neural Network (CNN) architecture [8]. It simulates how humans identify, remember, and analyze features. CNN is one of the best image retrieval, detection, and classification algorithms.

The convolution layer, pooling layer, fully connected layer, and classifier are elements of the CNN architecture [9]. The convolutional layer performs feature extraction and vector generation. Lower-order convolution kernels extract features with a focus on local perception. By contrast, the first few layers extract features that focus on global perception, such as edges, corners, and endpoints. The receptive field can expand this to higher-order convolutions, which also enables the gradual transformation of local characteristics into global features. Throughout the recognition process, the kernel parameters decrease the dimensionality of the calculations using the weight-sharing attribute of each receptive node.

The pooling layer, which follows the convolutional layer(s), downsamples the result by sliding a kernel of a particular size and stride length and taking the maximum or average of the input [8]. The small summaries produced by downsampling yield dimensionality reduction of the feature maps produced by the convolution layer. However, owing to CNN's lack of a mechanism for storing internal representations that would characterize the link between components or pixels, the pooling striding process could result in a loss of spatial and structural features in the image [3]. During sliding window techniques, such as convolution and pooling, the stride parameter controls the kernel displacement in either dimension. For a feature map with dimensions $n_h \times n_w \times n_c$, the dimensions of the output obtained after the pooling layer are $((n_h - f + 1)/s)\ ((n_w - f + 1)/s)\ n_c$ where n_h is the height of the feature map, n_w is the width of the feature map, n_c is the number of channels in the feature map, f is the filter size, and s is the stride length.

The hierarchical, multilayered nature of CNNs makes it possible to extract low, mid, and high-level information. Lower-level features and mid-level features are combined to create high-level features, which represent more abstract information. One of the main reasons CNNs are so popular is their ability to extract features in a hierarchical manner. But in several studies these high-level features are not adequate for good classification results. Possible reasons for the poor performance of CNN models include the absence of a sufficient method to describe the relationship between objects and scenes, scene recognition using object information, the mapping of low-layer to higher layer features, semantic ambiguity in low-layer features, and the loss of detailed object description within the convolutional layers. A linear integration strategy that blends intermediate and high-layer characteristics of a deep CNN at various levels with low-level handcrafted features incorporating object shape, scene depth, and color information can solve these problems to a greater extent [10]. Numerous studies have investigated deep CNNs, including a study that addresses CNN taxonomy based on acceleration methods [10]. In comparison, the intrinsic taxonomy found in current and well-known CNN designs described from 2012 to 2020 is the topic of this survey. The study divides CNN designs into seven major groups: channel boosting, attention-based CNNs, feature-map exploitation, multi-path, width, and spatial exploitation. This survey also sheds light on the fundamental composition of CNNs, their historical context, and the many periods of CNN's growth and accomplishments.

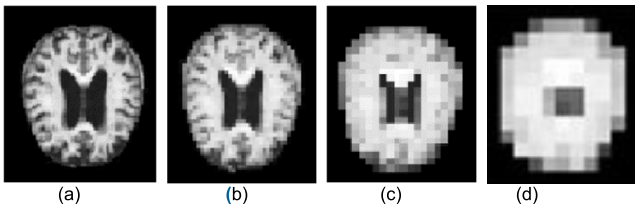

FIGURE 1. Structural loss of images during the pooling operations using 3 × 3 kernel. (a) input image with size 208 × 176 (b) output image of max pool 1 (c) output image of max pool 2 (d) output image of max pool 3.

Here, we used Magnetic Resonance Imaging (MRI) of patients with AD to study the loss of spatial and structural features in CNN operations. AD is the most widespread form of dementia that causes progressive memory loss and hinders daily activities due to damage to the brain cells. We conducted an analysis on a brain MRI shown in Fig. 1, which measured 208 × 176 pixels, and generated images by performing pooling operations. After the third pooling process, the image size decreased to 26 × 22 pixels, and the loss of structural traits was visible in the images. Preserving the original structural features improves the overall performance of intelligent systems [11]. However, a CNN requires a large amount of training data, and the introduction of multiple layers increases runtime. The use of Graphics Processing Units (GPU) can partially resolve this problem.

Several studies have classified AD using ML and deep-learning techniques [12], [13]. Based on segmented gray matter, a hybrid improved independent component analysis study produced 90.47% accuracy [14]. Another MRI study that used a multimodal strategy with a Siamese CNN achieved a performance accuracy of 99.05% [15] in slice-wise datasets. Imaging-based dictionary learning [16], feature learning for multi-class diagnosis [17], structure-based multiview learning [18], encoder-decoder architecture

constructed using ensemble learning [19], and diagnosis utilizing landmark-based features [20] are some of the most recent studies on this subject. Using the CNN architecture, 79.9% of AD diagnoses from mild cognitive impairment were correctly predicted [21].

The CMSMD outperforms CNN (96.72% accuracy) in categorizing AD because it gives more importance to structural characteristics. The features that emerge from this descriptor capture the local and global characteristics of an object. It assesses and classifies items based on minute shapes and marginal detail. An image object of size m × n can be reduced to 2× k using Circular mesh Border Labels (CBL) [2]. The CBL size k is dynamically determined by the size of the circular mesh and the irregularity of the object boundary. The value of k is always less than m and n; thus, the computational complexity is lower. CBL is used to obtain statistical, structural, and property-based features. In addition, it offers an improved accuracy for small datasets. By discarding characteristics other than the structural aspects, the circular mesh reduces the dimensions. However, the CMSMD model has limitations; it primarily deals with structural features and does not handle texture features and occluded objects. In addition, image processing and machine learning techniques require numerous techniques to solve these problems. The benefits and drawbacks of CNN and CMSMD models [22], [23], [24], [25] are united and listed in Table 1.

Considering these limitations and the advantages of both systems, we devised the proposed system.

The motivation behind the design of CirMNet is to:

1) Preserves the high-level structural features such as shape, margin, and edge of the object, which is invariant to translation, rotation, and scaling.

2) Maintain the spatial position of the object during representation, description, and feature extraction.

3) Improve the performance than state-of the-art systems.

In light of these drawbacks and benefits of both models, we developed the proposed systems and the other contributions of this study are as follows:

1) CirMNet, a novel hybrid approach that combines CNN and CMSMD's best features, is proposed.

2) A brand-new statistical and FD-based feature extraction method employing CMSMD is suggested for vertically symmetrical images.

3) The performance of the system is evaluated by implementing it on the classification of neurodegenerative disease.

The dataset must undergo several steps to be ready for CNN and CMSMD to process, which might entail noise reduction, artifact correction, or normalization. Other standard preprocessing techniques for brain MRI scans include skull stripping [26], motion correction, spatial and intensity normalization, filtering, registration [27], and segmentation [6]. Researchers have chosen numerous approaches and tools for each preprocessing step based on the requirements of fixing any errors or noise in the images [6]. Even if the images are not preprocessed and registered, the CNN can still interpret the data; however, the accuracy may be lower. These methods can improve performance by boosting feature recognition and reducing the effects of input image variability. Consequently, the complexities for preprocessing the dataset for CNN, CMSMD, and CirMNet, are the same. However, the algorithmic complexity is higher for the CirMNet model because it combines CNN and CMSMD.

TABLE 1. Characteristics of CNN and CMSMD.

SlNo	CNN	CMSMD
1	Extracts edge, texture, color and structural features.	Extracts shape and margin characters.
2	A variant to rotation and scaling.	Invariant to translation, rotation, and scaling.
3	Recognize the image as clusters of pixels in distinct patterns or components.	Labels the boundary using CBL technique is used to represent the object.
4	No internal representations to describe the relationship between pixels.	Maintain the spatial position of the object with the help of CBL.
5	During the striding process, spatial and structural details of the image may be lost.	Spatial values are maintained during representation, description, and feature extraction.
6	Image matrix used for analysis.	It uses boundary values to represent objects.
7	Boundless performance while classifying images that are very similar to the dataset.	Outstanding performance while classifying objects with well-defined boundaries.
8	Initial layers extract global features such as edges, corners, endpoints, etc.	Can extract global and local shape and margin features.
9	Difficulties in classifying images from different angles.	Face similar problems.
10	Effective processing requires a large amount of training data.	Suitable for smaller datasets.
11	Significantly slower due to operations like pooling and convolution.	Computational complexity depends on the algorithms used in the sequence of tasks.
12	Training takes a long time if there are several layers. A good GPU can resolve the problem.	GPU can resolve the time complexity.

II. CIRCULAR MESH-BASED SHAPE AND MARGIN DESCRIPTOR

The CMSMD is a tool for classifying and recognizing objects [2]. This is because it combines the functions of the structural and global contour-based descriptions. A polar coordinate system that represents a circular mesh structure using track and sector values enables border marking. Equal-sized cells in the circular mesh were labeled using the Circular mesh-based Cell Labeling (CCL) technique. Essential details regarding their forms and margins are transmitted by placing brain border-extracted sections from images on a circular mesh and naming them using Circular mesh Border Labeling (CBL). The borders were then retrieved from the brain images using various procedures, including enhanced Fuzzy C Means Clustering (FCM) segmentation, hole-filling, and boundary extraction, each with a complexity of $O(ndc^2t)$, $O(lm)$, and $O(lm)$. The number of objects (n), number of

dimensions (d), number of clusters (c), number of iterations (t), and image size ($l \times m$) are all given. Subsequently, three procedures— NUMCIR(), KMEANSCAL(), and RADIUSFIND() —with an overall complexity of O($tknd$) are used to build the circular mesh dynamically, where n is the number of objects, k denotes the number of clusters, d denotes the size of each item, and t denotes the total number of iterations. Typically, k, t, and $d < n$ yield the best results when the datasets are distinct or well separated.

Each cell in the dynamic circular mesh was labeled based on the track values (Trk(i)) and sector values (Sec(i)) from the origin and x-axis, respectively. Moreover, the complexity of border labeling is O(nm), where n and m are the row and column sizes of the image, respectively. The border values are stored as sampled or unsampled in the Trk(i) and Sec(i) circular arrays, reducing the complexity to linear values. Border smoothing is achieved by reducing the size of the circular mesh. An object can be recreated using Trk(i) and Sec(i): TrkVar(j) stores the repeating values in adjacent locations of Trk, and TrkOcc(j) stores its occurrences - similarly, SecVar(j) and SecOcc(j) for Sec.

New features were derived from the descriptor using structural, statistical, and CBL property-based methods in linear time to find the peaks and valleys of the embedded contours with linear complexity. The effectiveness of the descriptor was demonstrated by detecting breast lesions using Dynamic Contrast-Enhanced MRI (DCE-MRI), which achieved an accuracy of 94.69% [2].

III. ALZHEIMER'S DISEASE

The basic concept of neurodegenerative disease is simple and self-explanatory. It occurs due to neuronal death or damage, leading to various disorders that affect memory, movement, and behavior. Well-known neurodegenerative disorders include AD, motor neuron disease, Huntington's disease, and Parkinson's Disease (PD). AD exhibits symptoms of mental dysfunction such as memory and behavioral issues, whereas others may show movement symptoms. However, neuronal degeneration begins long before these symptoms appear [28].

Cognitive decline in AD is due to progressive degeneration of neurons, often characterized by the accumulation of abnormal proteins. This leads to intracellular neurofibrillary tangles and extracellular senile plaques [29], [30]. The MRI is a widely used imaging method for AD, providing information on structural changes in the brain, including the hippocampus and cortex. The deposition of beta-amyloid protein leads to brain volume loss, and these volumetric changes can be detected using voxel-based morphometry using MRI. Hippocampal changes may also occur before the onset of AD symptoms and are considered imaging markers for patients with pre-demented AD. Other areas of the brain affected by AD include the entorhinal cortex; amygdala; temporal, orbitofrontal, and parietal cortex, which may show varying degrees of shrinkage or structural changes [31], [32], [33]. These underlying pathological processes are reflected and identified using T1-weighted MRI. Fluoro-deoxyglucose (FDG) and Positron Emission Tomography (PET) using fluorodeoxyglucose are promising imaging methods for detecting neuronal degeneration. Glucose uptake by neurons decreases in AD, but this decrease may not be specific to AD [34], [35], [36] and can also occur in other diseases, resulting in neuronal loss. The role of FDG PET in monitoring AD is limited, and amyloid PET scans for detecting amyloid plaques may be more promising [37], [38]; However, this process may be slow, require monitoring for years, and disease progression may plateau over time. Medical professionals prefer MRI to radionuclide imaging because of its wide availability, and they rely on T1-weighted MRI-based structural assessments as the mainstay of imaging. Other imaging methods include morphometry, functional MRI (fMRI), and optical coherence tomography [39].

The structural MRI (sMRI) displays structural changes in EMCI, LMCI, and AD visible on it. Radiologists can categorize these stages based on volumetric changes in MRI scans. Early detection of AD is crucial to slow or arrest its progression, as irreversible damage leads to poor quality of life. Identifying specific patient groups on MRI scans during mild cognitive impairment or before symptoms appear may help to arrest the disease. Radiologists may detect early changes in deep brain structures and cortical thinning complexes; however, computer-aided algorithms can better assess them. Quantifying the morphometry and deformation in these areas is promising for detecting changes and monitoring patients longitudinally.

TABLE 2. The class-wise distribution of the dataset with training and testing allocation by considering one slice per subject.

Class label	Training	Testing	Total
Control	124	22	146
Early MCI	121	21	142
Late MCI	124	20	144
AD	120	21	141
Total Images	489	84	573

IV. MATERIALS

This study used 573 subject-level instances from the ADNI [40] database (https://adni.loni.usc.edu), in which 1.5 Tesla, T1-weighted MRI scans with 256 × 256 pixels were employed. The Spoiled Gradient Recalled (SPGR) acquisition parameters considered in the study are repetition time (1.6 ms), echo time (5.0 ms and 1.4 ms), flip angle (30^0), inversion time (0 ms), and thickness (5.0 mm). We chose to select one axial slice per subject according to the expert opinion of the radiologist. Based on the degree of neurological degeneration, MRI images are classified as Control, EMCI, LMCI, and AD, and we used 146, 142, 144, and 141 images, respectively. We employed non-overlapping subject-level training and testing sets with sizes of 489 and 84, respectively and displayed in Table 2.

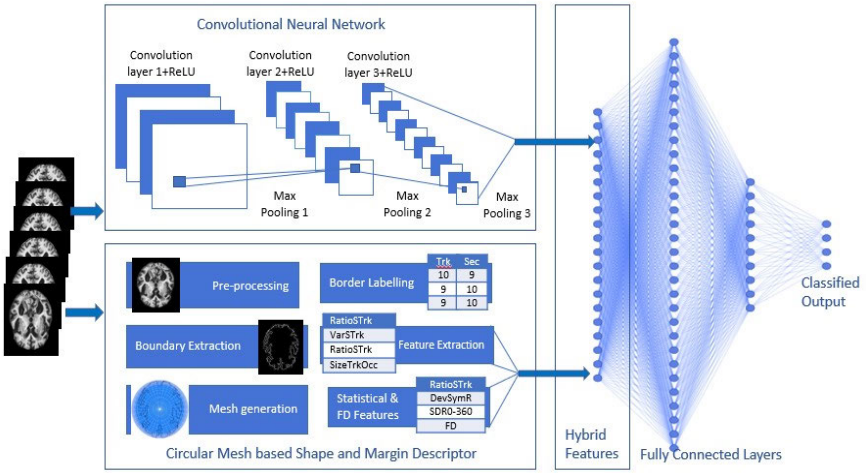

FIGURE 2. CirMNet Architecture with major components: Two parallel layers- Convolutional Neural Network and Circular Mesh based Shape and Margin Descriptor for feature extraction; hybrid feature generation; classification layers.

V. CIRMNET

The CNN and CMSMD models were combined in the suggested architecture, CirMNet, to address the shortcomings of each model while maximizing the benefits of the other. A shape-prioritized hybrid deep-learning architecture combines the structural characteristics of a CMSMD with CNN features. The CMSMD features to boost the structural characteristics lost during the striding and pooling operations of the CNN model. This technique can be used for any image classification issue that requires structural specifications to solve the problem. This study classified brain MRI scans into control, EMCI, LMCI, and AD phases. This is because structural alterations in the brain frequently underlie neurodegenerative illnesses, and these structural changes are examined during diagnosis. Two parallel tracks in the CirMNet model are used to process the input images. CNN-based features were retrieved from the first track, whereas the second track extracted CMSMD-based features. Once these features are retrieved, they are aggregated and applied to a fully linked image classification layer. Fig. 2 presents the proposed CirMNet architecture.

A. CONVOLUTIONAL NEURAL NETWORK FEATURES

In the first track of CirMNet, the input image $I_{i,j}$ of size $n_1 \times n_2$ and $I_{i,j} \in R_{i,\mathrm{j}}$ is convolving with kernel $K \in R^{(2h+1)\times(2h2+1)}$ and is given by:

$$(I * K)r, s := \sum_{u=-h_1}^{h_1} \sum_{v=-h_2}^{h_2} K_{u,v} I_{r+u,s+v} \tag{1}$$

where the kernel K is represented as

$$K = \begin{pmatrix} K_{-h_1,-h_2} & \cdots\cdots & K_{-h_1,h_2} \\ \cdot & & \cdot \\ \cdot & K_{0,0} & \cdot \\ \cdot & & \cdot \\ K_{h_1,-h_2} & \cdots\cdots & K_{h_1,h_2} \end{pmatrix} \tag{2}$$

The input layer l of the convolution layer comprises $cnf_1^{(l-1)}$ feature vectors from the former layer, with a size $cnf_2^{(l-1)} \times cnf_3^{(l-1)}$. The value $l = 1$ includes one or more channels for a single input image, I. The output of layer l contains $cnf_1^{(l)}$ feature vectors of size $cnf_2^{(l-1)} \times cnf_3^{(l-1)}$. The i^{th} feature vector in layer l, $Y_i^{(l)}$, is calculated as:

$$Y_i^{(l)} = B_i(l) + \sum_{j=1}^{cnf_1^{(l-1)}} K_{i,j}^{(l)} * Y_j^{(l-1)} \tag{3}$$

where $B_i^{(l)}$ is a bias, and $K_{i,j}^{(l)}$ is a kernel of size $(2h_1^{(l)}+1) \times (2h_2^{(l)}+1)$ connecting the j^{th} feature vector in layer $(l\text{-}1)$ with the i^{th} feature vector in layer l. Subsampling is used to reduce noise and unwanted distortions. Here, subsampling is represented by equation (4).

$$\mathrm{MaxP} = max\left(Y_i^{(l)}\right) \tag{4}$$

All of these CNN features are stored in the feature matrix $\mathrm{CNF} = [cnf_{(1)}, cnf_{(2)}, cnf_{(3)}, \ldots\ldots\ldots, cnf_{(n)}]$.

B. CMSMD FEATURES

The input images were processed in the second track using a number of different techniques, including geometric active contour for skull stripping, a sophisticated diffusion method for noise removal, morphological procedures for hole filling in segmentation, and erosion and subtraction for boundary extraction [6], [26]. Fig. 3 illustrates the sequence of steps applied to the four classes of neurodegenerative conditions: control, EMCI, LMCI, and AD.

The CBL approach, which successfully communicates information about the forms and margins of objects, is represented by a circular mesh framework in a polar coordinate system. According to [2], a dynamic circular mesh was created for this investigation, and its size was set based on the irregularity of the border images. Fig. 4 displays the circular

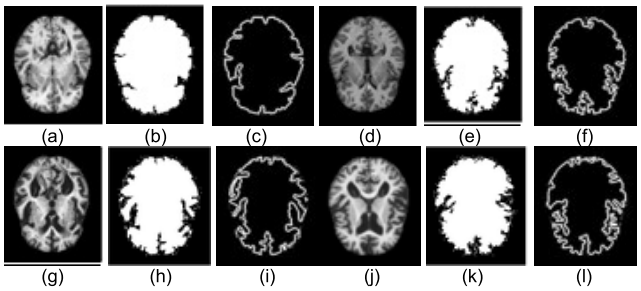

FIGURE 3. Boundary extraction steps of control case: (a) Original Image (b) Segmented hole filled (c) Boundary Extracted; EMCI case: (d) Original Image (e) hole-filled (f) Boundary Extracted; LMCI cases: (g) Original Image (h) Segmented hole-filled (i) Boundary Extracted; AD case: (j) Original Image (k) Segmented hole-filled (l) Boundary Extracted.

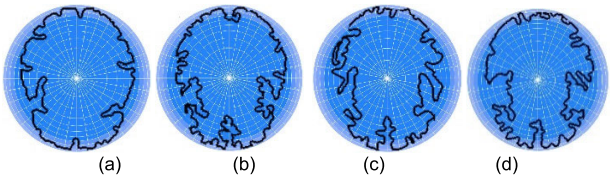

FIGURE 4. Brain Boundary embeddings of circular meshes (a) control, (b) EMCI, (c) LMCI, and (d) AD.

TABLE 3. Trk and Sec values of embedded brain boundaries of Fig. 5 images.

Control		EMCI		LMCI		AD	
Trk	Sec	Trk	Sec	Trk	Sec	Trk	Sec
5	1	5	2	7	2	8	7
7	2	6	2	6	2	8	8
7	2	7	2	10	2	6	8
7	1	5	1	10	13	7	14
6	13	6	13	9	13	6	13
.	.	.		.	.	.	.
.	.	.		.	.	.	.
.	.	.		.	.	.	.
767	767	1151	1151	1203	1203	1246	1246

mesh embeddings, of Fig. 3.c, Fig. 3.f, Fig. 3.i, and Fig. 3.l. The border labels are stored in two circular arrays, Trk and Sec, and displayed in Table 3. The Trk and Sec values for smooth borders remain in one Trk value for a long time, but for uneven borders, the Trk values fluctuate. TrkOcc and SecOcc are variables that count and store the persistence or frequency of Trk and Sec, respectively. The increasing values of Trk, Sec, TrkOcc, and SecOcc for each of the four types of AD indicate the degree of irregularity in the affected brain. The features produced by the CMSMD can be used to analyze and classify objects with subtle shapes and marginal changes. These features include both the local and global attributes. Table 4 lists TrkVar, TrkOcc, SecVar, and SecOcc for the various classes of images shown in Fig. 4.

C. PROPOSED FEATURE EXTRACTION TECHNIQUES FOR BRAIN MRI

In this paper, we propose an FD approach and statistical features to extract the qualities of the structural boundaries. In a

TABLE 4. a. TrkVar, b. TrkOcc, c. SecVar and d. SecOcc of respective classes of embedded brain boundaries of Fig. 4 images.

Control				EMCI				LMCI				AD			
A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
7	1	1	3	7	2	1	0	6	3	1	3	7	2	1	0
6	2	2	1	5	1	2	1	7	1	2	1	5	1	2	1
7	1	1	2	6	2	1	1	4	1	1	2	6	2	1	1
5	2	2	2	7	1	2	3	7	1	2	2	7	1	2	3
8	1	3	2	4	1	1	1	7	9	1	3	2	4	1	1
6	4	9	2	5	2	13	32	6	4	13	2	5	2	13	32
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
39	39	27	27	62	62	39	39	67	67	45	45	72	72	52	52

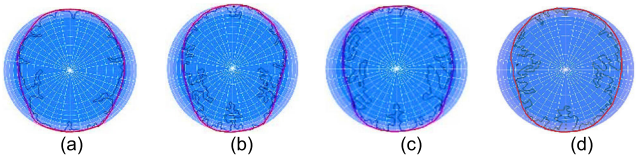

FIGURE 5. Reference boundary for brain contours (a) Control (b) EMCI (c) LMCI (d) AD.

previous study, we developed structural, statistical, and CBL property-based characteristics and applied them to identify breast lesions [2]. In their study, the formulae produced variations from circular or elliptical reference shapes. However, the specific shape of the brain can vary among images, slices, and individuals when using MRI scans, and is neither round nor elliptical. We recommend using a dynamic reference border for each image in this study to precisely examine deviations from this reference and pinpoint the structural abnormalities.

1) REFERENCE BOUNDARY FOR FEATURE GENERATION

The reference image, RefBond, is formed outside the boundary and is similar to the stretching and placement of a rubber band on the border. The track and sector values of the reference boundary are denoted as RTrk and RSec, respectively. Fig. 5 shows the red curve representing the reference boundary produced for the four classes of brain image boundaries. This study suggests the following statistical aspects:

2) STANDARD DEVIATION CONCERNING REFERENCE BOUNDARY

The standard deviation of Trk(i) was one or close to one, indicating a circular boundary. This study considers the reference boundary and saves its track values as RTrk. During feature derivation, we determined the mean Rμ of the reference boundary. Different RTrk values may be utilized in the equation, depending on the circumstances. The complete value of RSec can be used, which corresponds to angles between 0° and 360° in the polar coordinate system. The formula proposed to determine how far the brain boundary deviates from RefBond is as follows:

$$\mathrm{SDR}_{0-360} = \sum_{i=1}^{N} \left((\mathrm{Trk}\,(i) - R\mu)^2\right)^{1/2} \Big/ N \qquad (5)$$

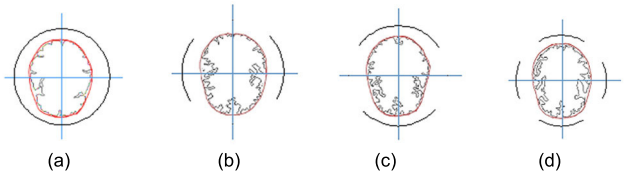

FIGURE 6. Range of angles taken for standard deviation and coefficient of variations to check symmetry concerning the reference boundary RefBond (a) 0°-360° (b) 45°-315°(clockwise), 135°-225° (c) 60°-120°, 240°-300° (d) 30°-330° (clockwise), 150°-210° and 70°-110°, 250°-290°.

$R\mu$ is the mean of RTrk(i), i varies from 1 to N, and is the size of RTrk.

3) STANDARD DEVIATION OF SYMMETRIC POINTS

To compare the symmetry of the standard deviation of the Trk values with respect to the angular points on the y-axis, we quantified this symmetry using a feature. To perform the symmetry check, we took the Trk values of the border from symmetric angles such as 45°-315° and 135°-225°, as shown in Fig. 6. We considered the brain contour symmetrical if the standard deviations of the Trk values at these symmetric angles were approximately equal. The calculations involved using the mean of the reference boundary denoted as $R\mu$.

By adjusting the range of i in the standard deviation equation (5) and utilizing this as a feature value, we independently tested the upper and lower quadrant symmetric values.

4) COEFFICIENT OF VARIATION

This feature calculates the deviation from symmetry concerning the reference boundary, and the calculation formula is:

$$\text{CoeffVarR} = SD_{i-j}/R\mu \tag{6}$$

where i and j are the angles or Sec value ranges and $R\mu$ is the mean of the reference boundary. This statistical formula helps compare the variations between the two data series.

5) DEVIATION FROM SYMMETRY OF REFERENCE BOUNDARY

To calculate the track symmetry of RefBond, we compare the symmetric track values for a particular angle corresponding to a sector. The brain boundary analysis calculates only the sector position at $\varphi = 90°$ to check the y-axis symmetry. When $\text{Sec}(\varphi) = 90$, θ is the index position in the circular array, and $\text{RTrk}(\theta + i)$ and $\text{RTrk}(\theta - i)$ provide symmetric points. When conducting the symmetry check for $\varphi = 90$, we checked the borders in the first and fourth quadrants, with the borders in the second and third quadrants. The deviation from symmetry can be calculated using the following formula:

$$\text{DevSymR} = \sum_{i=1}^{n/2} ((\text{RTrk}\,(\theta + i) - \text{RTrk}(\theta - i))\Big/ n \tag{7}$$

When $\varphi = 90°$, n denotes the number of elements from Sec (0) to Sec (90).

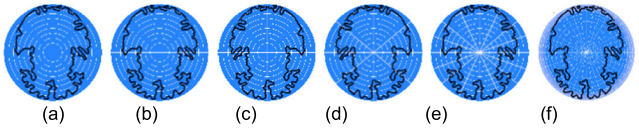

FIGURE 7. Brain boundary embedded on circular mesh with varying sizes of cells (a) Nc = 10 × 20 Ac = 40000, (b) Nc = 10 × 21 Ac = 20000, (c) Nc = 10 × 22 Ac = 10000, (d) Nc = 10 × 23 Ac = 5000, (e) Nc = 10 × 24 Ac = 2500 and (f) Nc = 10 × 25 Ac = 1250.

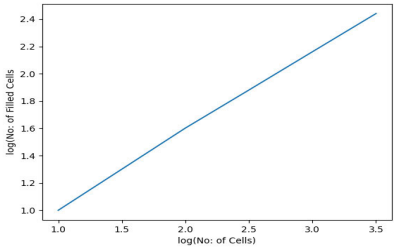

FIGURE 8. The log-log plot of the number of cells against the number of filled cells. The fractal dimension is the slope of the solid line.

6) FRACTAL DIMENSION

This study proposes a new FD technique that uses a circular mesh to measure self-similarity at different scales. Fractal geometry is used in structural analysis and is a better method for detecting complex brain patterns than traditional Euclidean geometry. The FD formula was derived based on the number of cells compared with the number of cells embedded in the border.

$$\text{FD} = \text{Log}(\textit{No. cells})/\text{Log}(\textit{No. filled cells}) \tag{8}$$

where No. filled cells represent cells embedded in the brain regions, and No. cells represent the total number of cells in the circular mesh (Fig. 7). The logarithm of the number of filled cells was plotted against the logarithm of the number of cells (Fig. 8) at different magnifications. The circular mesh is divided into 10, 20, 40, 80, 160, and 320 parts, as shown in Fig. 7. represents the magnification of the image. The FD was calculated as the slope of the best-fit line, as shown in Fig. 8.

D. HYBRID FEATURES

The newly derived statistical and FD features and the features generated using the CMSMD are stored as CMF = [*cmf(1), cmf(2), cmf(3), ….., cmf(m)*]. The CNN features were retained in the CNF = [*cnf(1), cnf(2), cnf(3), ……., cnf(n)*]. CNF and CMF were combined to form a hybrid feature set HYB.

$$\text{HYB} = \textit{Concatenate}\,(\text{CNF, CMF}) \tag{9}$$

These nineteen CMF features, along with 1,479,200 CNF features, are integrated to form 1,479,219 CirMNet features in this study. The CirMNet architecture passes these hybrid feature sets to a fully connected layer for classification.

Algorithm 1 CirMNet (Img)

Require: Input Image
Ensure: CNF = [cnf(1), cnf(2), cnf(3),. . ., cnf(n)]
CNN feature set
CMF = [cmf(1), cmf(2), cmf(3),. . ., cmf(m)]
Circular mesh feature set
HYB = Concatenated Hybrid feature set
CirMFea = Selected hybrid feature used in CirMNet

```
Read Img
CNF = Conv(Img)
for j = 1 to n do
  for k = 1 to m do
    DCNF(j,k) = w(j,k)* cnf(k) + b(k)
  end for
end for
CMF = CMSMD(Img)
for j = 1 to n do
  for k = 1 to m do
    DCMF (j,k) = w(j,k) * cmf(k) + b(k)
  end for
end for
HYB = Concatenate (DCNF, DCMF)
CirMFea = CorHYB(HYB)
for l = 1 to the number of layers do
  if (l-1 == 0) then
    Z(1) = FullyConnected(CirMFea)
  else then
    Z(l) = FullyConnected (Z(l-1))
  end if
end for
return
```

E. CLASSIFICATION LAYER

Input vector HYB goes to the fully connected layer, and the equation for the l^{th} layer in the fully connected layer is

$$Z_i^{(l)} = B_i^l + \sum_{j=0}^{n(l-1)} w_{i,j}^{(l)} A_j \tag{10}$$

When $\mathrm{A_j} = \begin{cases} \mathrm{HYB} & if\ l = 1 \\ Z_i^{l-1} & otherwise \end{cases}$

$w_{i,j}^{(l)}$, the strength of the connection from the $\mathrm{j^{th}}$ unit in the previous layer to the $\mathrm{i^{th}}$ unit in the current layer l. These weights, along with the bias B_i^l, determine the output $Z_i^{(l)}$ by modulating the contributions of the previous layer's activations A_j.

VI. RESULT

The main focus of this study is on techniques for extracting structural features in image processing. State-of-the-art ML techniques work well with small datasets but can be more efficient. Table 5 presents a comparison with state-of-the-art deep learning methods. However, in most cases, a slice-wise dataset was employed [7], [15], [42] which exhibited a higher accuracy rate. It is important to note that using a slice-wise dataset can lead to overfitting. To mitigate this issue, subject-wise studies were proposed, and as such, we employed a subject-wise dataset in this study. Recent advances in deep learning, such as CNNs, have demonstrated outstanding accuracy in various image classification challenges; however, owing to the pooling and striding procedures, they fail to retain the spatial and structural properties of images as they advance to deeper layers. By fusing CNNs with CMSMD, CirMNet design solves this problem while maintaining the structural properties of the objects. The performance of the model is evaluated by measuring the structural changes of brain MRIs. The structural changes contribute for the classification of Neurodegenerative diseases. The assessment of CirMNet's performance involved the classification of brain MRI of 573 subjects into four stages, including control, EMCI, LMCI, and AD. The training dataset consisted of 489 images, while the validation test set contained 84 images. Table 2 displays the class-wise distribution of the dataset, including training and testing allocations. The minimum system requirements for the CirMNet model include a quad-core or higher CPU like Intel Core i7, a mid-range or higher NVIDIA GPU with CUDA support like NVIDIA GeForce GTX 1060 or higher, and 16 GB of RAM.

TABLE 5. Performance comparison of state-of-the-art deep learning techniques for ad classification and diagnosis.

References	Methods	Accuracy
Farooq. et al. [7]	Deep CNN (layers - 22)	98.8%
Khan et al. [41]	Multi-model ML	87%
Basheera et al. [14]	CNN	90.47
Mehmood et al. [15]	Siamese CNN (20)	99.05%
Lin et al. [16]	Dictionary Learning	79.9%
Yildirim et al. [8]	CNN	90%
Madhusudan et al. [42]	CNN with KNN (slicewise)	99.58%
Bae et al. [43]	CNN	83%
Proposed Method	CirMNet (Subject wise dataset)	97.34%

The dataset used in this study went through a series of steps such as noise removal using the Complex Diffusion Method, registration using Statistical Parametric Mapping (SPM), skull-stripping using thresholding, artifacts removal using the diffusion method, and motion correction using Interpolation Methods [44]. The preprocessing techniques mentioned earlier would yield complexities ranging from $\mathrm{O}(n^2)$ to $\mathrm{O}(n^5)$ based on the algorithm employed for image correction. This helps to improve the quality of the data, reduce variability, and extract meaningful information from brain images, thereby increasing the accuracy of AD diagnosis.

To generate the architecture of CirMNet, a range of hyperparameters were selected based on the expertise and literature review that we conducted and then we applied an induction method for the selection. We constructed different CNN architectures with varying numbers of layers and kernels, as shown in Table 6, to give insight into the performance of CNN on this dataset.

A CNN with three convolutions, three pooling, four fully connected layers, and with node sizes 100, 50, 40, and 10 produced better accuracy with minimum time and space complexities. The same configuration completed the task in 48 minutes, while different architectures with three convolutions, three pooling, and five fully connected layers took around two hours to complete (Table 6). The structural features of the local and global characteristics of the CNN were

TABLE 6. The performance evaluation of CNN for subject-level images. Parameters used in the study: kernel size = 3 × 3, number of epochs = 100, layer order = (convolution, pooling, fully connected), number of kernels = (16,32,32).

Layers	Node size	Elapsed Time	Accuracy
(3,3,4)	(40,30,10,4)	48 min.	0.9366
(3,3,4)	(100,50,40,10)	55min.	0.9489
(3,3,5)	(500,300,200,100,50)	1 hrs. 10 min.	0.9700
(3,3,5)	(700,500,200,100,50)	1 hrs. 25 min.	0.9900
(3,3,9)	(500,300,133,100,30, 12,10,5,4)	1 hrs. 36 min.	0.9600

TABLE 7. Performance evaluation of CMSMD at subject level.

Size of Circular mesh	Training time in Minutes	Testing time in Minutes	Accuracy
10×16	20	6	0.9672
10×32	27	9	0.9463
20×16	22	12	0.9602
20×32	24	13	0.9685
30×16	26	23	0.9610
30×32	23	25	0.9545

analyzed using MRI boundary images of the brain as the input, and tested it with kernels of varying sizes, like 3 × 3, 5 × 5, 15 × 15. It was impossible to preserve spatial symmetry with even-sized kernels, but this odd-sized kernel helped to extract the center pixel and do so. The directionality and symmetry of feature identification were also established with the use of odd kernels. Larger kernels allowed for more complicated feature extraction but it was not good for structural features. The 3 × 3 kernel was more effective at capturing spatial information from the input data and had a bigger receptive field than other kernels, which was important for identifying structural changes. To take the structural summary in our investigation, we, therefore, selected the 3 × 3 kernel.

We conducted a thorough evaluation to assess the performance of the CNN at different layers, and carefully examined the feature maps generated at these layers. Fig. 9 presents the features generated at various layers of the three convolution and max-pooling operations. After performing max pooling 1, the images lose their spatial and structural details, and the resulting feature map represents an input that is spatially invariant to the next layer. Table 6 presents the performance evaluation results of the CNN for various kernel values. This shows that the accuracy increases with the number of layers and kernels; however, the elapsed time also increases noticeably and must be decreased.

The preprocessed MRI underwent a series of steps, such as enhanced FCM segmentation, hole filling, boundary extraction, CCL, and CBL, to extract the CMSMD features. Three procedures, namely NUMCIR(), KMEANSCAL(), and RADIUSFIND(), were used to generate the circular mesh. This process yields an overall complexity of O(*tknd*).

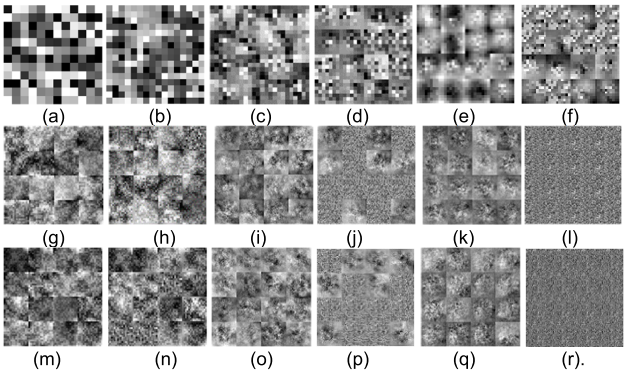

FIGURE 9. Features generated at different layers of three convolutions and Max pooling operations (a) Convolution 1 with kernel 3 × 3 (b) Max pool 1 with kernel 3 × 3 (c) Convolution 2 with kernel 3 × 3 (d) Max pool 2 with kernels 3 × 3 (e) Convolution 3 with kernel 3 × 3 (f) Max pool 3 with kernel 3 × 3 (g) Convolution 1 with kernel 5 × 5 (h) Max pool 1 with kernel 5 × 5 (i) Convolution 2 with kernel 5 × 5 (j) Max pool 2 with kernel 5 × 5 (k) Convolution 3 with kernel 5 × 5 (l) Max pool 3 with kernel 5 × 5 (m) Convolution 1 with kernel 15 × 15 (n) Max pool 1 with kernel 15 × 15 (o) Convolution 2 with kernel 15 × 15 (p) Max pool 2 with kernels 15 × 15 (q) Convolution 3 with kernel 15 × 15 (r) Max pool 3 with kernels 15 × 15.

TABLE 8. Rank order of 19 relevant feature values, labeled as $cmf_{(1)}$ to $cmf_{(19)}$.

Rank	Features	Rank	Features
1	RatioCirPerObjPer	11	CfVarTrkA0-360
2	RatioSTrk	12	CVarTrkA40-140
3	VarSTrk	13	CoeffVArR140-220
4	SizeTrkFD	14	OneSecSSec
5	SD0-360	15	Size of sector
6	SDPT195-265-285-345	16	RatioSSecVarSSec
7	SDR0-360	17	RMTrkOccSTrkVar
8	SDR320-40	18	AvgZeroTrki
9	SDR40-140	19	AvgNegSeci
10	SDPT0-45-135-180		

We evaluated the performance of the CMSMD with various track and sector sizes, namely 10 × 16, 10 × 32, 20 × 16, 20 × 32, 30 × 16, and 30 × 32. The evaluation yielded results of 0.9672, 0.9463, 0.9602, 0.9685, 0.9610, and 0.9545 for subject-level images, and circular mesh with size 10 × 16 is found apt for the study. Table 7 lists the training and testing times, as well as the accuracies. It is clear from the table that the training time for CMSMD is shorter than that for CNN. The space requirement was also lower than that of CNN. CirMNet aims to maintain the structural characteristics of image objects using hybrid feature-extraction methods. The training was conducted on an 11th Gen Intel Core(TM) i7-1165G7@280GHz workstation CPUx64-based processor with 16 GB of memory and a moderate graphics processing unit (GPU) card. The algorithms were implemented in MS Windows 10, Pro 21H2, and MATLAB R2022b.

We constructed the CirMNet model using two parallel tracks: the first track for extracting CNN features (CNF), and the second for CMSMD features (CMF). The input MRI

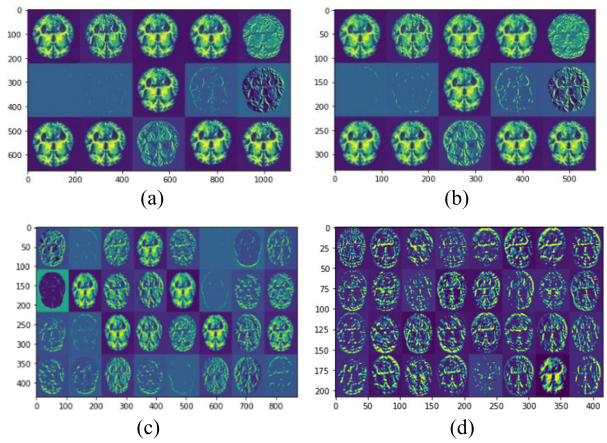

FIGURE 10. Features generated during convolution and pooling (a) Convolution 1 (b) Max pool 1 (c) Convolution 2 (d) Convolution 3.

images used in this study were converted into 224 × 224 in order to analyse the structural anomalies. We conducted experiments by resizing and then examining the images to mark the structural alterations. We found that images with a dimension of 224 × 224 are best suited for this purpose among different sizes like 200 × 200, 256 × 256, and 512 × 512. The CNN track consisted of three convolution layers: ReLU activation and max-pooling operations. The convolutional layers used 16 channels with 3 × 3 kernels. Convolution 1 and max pooling1 produced 788,544 and 781,456 features, respectively. Convolution 2 and max pooling 2 generated 1,534,752 and 1,520,768 features, respectively, whereas convolution 3 and max pooling 3 generated 1,492,992 and 1,479,200 features. In between convolution and pooling operations, batch normalization was performed to improve training stability and speed up convergence. Fig. 10 shows the features generated during the convolution and pooling operations.

In the second track, the images underwent skull stripping, preprocessing, hole-filling, and boundary extraction before generating a circular mesh. The extracted boundary was incorporated in the circular mesh, which was produced dynamically depending on the irregularity of the boundary, with the extreme border point lying on the y-axis. The CMSMD's size is fixed dynamically so that it grows as the borders change from being smooth to being irregular. Smooth borders do not require precise details, thus a descriptor with fewer components is produced; on the other hand, irregular borders do demand fine details, so a descriptor with more elements is produced. If the descriptor size is appropriate, this dynamic creation enables the CMSMD to reflect the shape of an item, and object reconstruction is achievable without sacrificing any features. Because only the pertinent information is retained, this feature makes it easier to accommodate any type of contour, whether it be vast or small [2]. For this study, a circular mesh size of 10 × 16 was shown to be ideal.

The CBL mechanism labels the boundary regions, reducing the space requirement from 5.93 KB to 144 bytes for a single image with a circular mesh size of 10 × 16. New

TABLE 9. Layer description of CirMNet architecture.

Layer	No of Channels	Size of Kernel	Feature Map
CNN			
Input	224×224	-	224×224
Conv1+ReLU	224×224	3×3×16	111×111×16
Maxpool1	111×111×16	2×2×32	109×109×32
Conv2+ReLU	109×109×32	3×3×32	54×54×32
Maxpool2	54×54×32	2×2×32	52×52× 32
Conv3+ReLU	52×52× 32	3×3×32	26×26×32
Maxpool3	26×26×32	2×2×32	24×24×32
CMSMD		Mesh	
Circular Mesh		10×26	19
CirMNet	1479200+19	-	1479219
Fully Connected		Nodes	
Layer 1	40x1	40	59168800
Layer 2	30×1	30	1230
Layer 3	10×1	10	310
Layer 4	4×1	4	44
Classification			Class

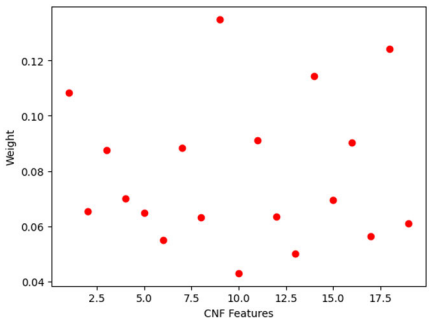

FIGURE 11. A plot of CNF features against weight.

statistical and FD features were derived for vertically symmetrical images and used along with statistical, structural, and label property-based features. The 19 highest priority features and their rank order based on the ReliefF algorithm are listed in Table 8. These nineteen CNF features, along with 14,79,200 CMF features, are integrated to form 14,79,219 CirMNet features.

The CNF's weight against features and CMF weights against features are also plotted to get a better understanding and are shown in Fig 11 and 12. In Fig. 13 and 14, CNF and CMF features are combined, and the rank versus features and weights versus features are plotted. The blue color represents the CMF items, whereas the red color represents the CNF items. From this analysis, it is clear that CMF has higher rank positions and weights than CNF.

A fully connected layer, consisting of three layers, was used to process the HYD features for image classification. Table 9 lists the architecture of the proposed model. In this study, we used different learning rates (0.1, 0.01, 0.001, 0.0001, 0.00001), and it converges at 20, 25, 28, 40, and 48 epochs, respectively with, 92.86, 94.65, 96.21, 97.34, 97.32 accuracies. The binary cross-entropy losses, in conjunction with the Adam optimizer, were used to train the CirMNet architecture. By comparing the results for each

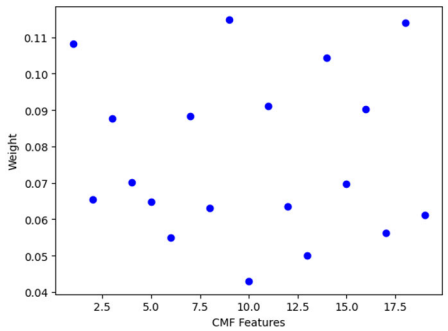

FIGURE 12. A plot of CMF features against weight.

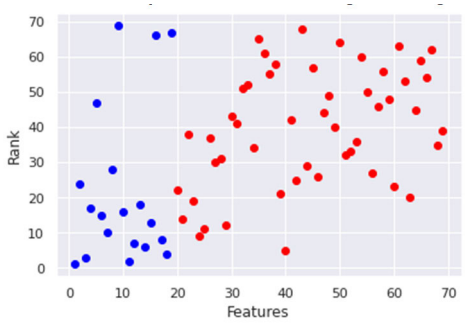

FIGURE 13. A plot of CMFand CNF against rank order.

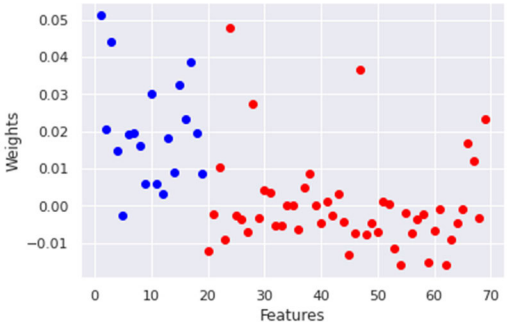

FIGURE 14. A plot of CMF and CNF against weight.

learning rate, we found that the best accuracy and loss performance was achieved using a learning rate of 0.0001. Fig. 15 shows the training and validation loss and validation accuracy of the CirMNet model, and the results indicate that there is no overfitting or overshooting for this model. The performance evaluation measures accuracy, precision, recall, F1 score, and loss values, are listed in Table 10. In addition to this, we employed the SHAP technique to interpret the important areas that contribute more to the disease classification purpose and plotted in Fig. 16. From this analysis, it is evident that the boundaryareas are contributing more to the classification activity.

In the CirMNet, we conducted a 5-fold cross-validation. The 2D CNN models analyzed form 3D brain image data, so we performed a subject-level cross-validation to prevent

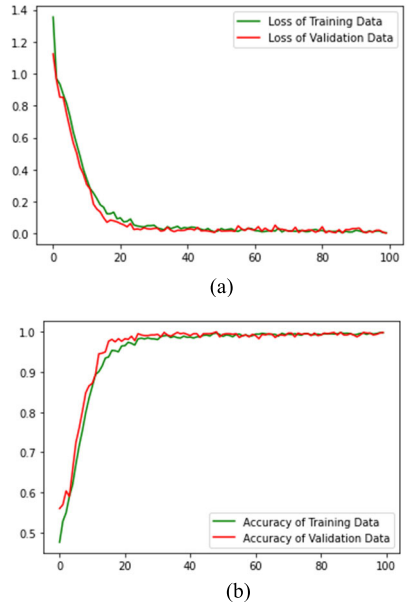

FIGURE 15. (a) CirMNet training and validation loss (b) CirMNet training and validation accuracy.

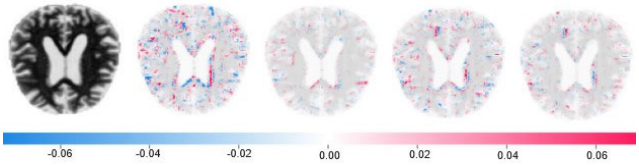

FIGURE 16. Plots the boundary areas that contribute more to the classification using the CirMNet Model using the SHAP interpretation technique.

TABLE 10. Performance evaluation of CNN, CMSMD, and CIRMNet: accuracy, precision, recall, F1 score, sensitivity, loss.

Method		Acc	Pre	Rec	F1	Sens	Loss
CNN		0.9366	0.9376	0.9368	0.9365	0.9368	0.0634
CMSMD		0.9672	0.9673	0.9671	0.9530	0.9671	0.0328
CirMNet	Fold 1	0.9734	0.9734	0.9734	0.9734	0.9734	0.0266
	Fold 2	0.9754	0.9755	0.9754	0.9754	0.9754	0.0246
	Fold 3	0.9815	0.9819	0.9816	0.9816	0.9816	0.0185
	Fold 4	0.9775	0.9777	0.9774	0.9775	0.9774	0.0225
	Fold 5	0.9595	0.9601	0.9595	0.9595	0.9595	0.0405
		0.9734	0.9737	0.9734	0.9735	0.9734	0.0265

data leakage [45]. We performed the training and testing of the model ten times, using a different fold as the test set and the remaining four folds as the training set in each iteration. We obtained the accuracy measurement for each fold, and displayed in Table 10. Fig. 17 shows the confusion matrix for four classes using the proposed method. This indicates that CirMNet provides the best classification results for the dataset, with an accuracy of 0. 9734, a precision of 0. 9737, a recall of 0.9734, an F1 Score of 0.9735, and a sensitivity of 0.9734. The training loss was only 2.66%. This study will help the radiologists for easy and accurate diagnosis by using a system parallel to the existing Picture Archiving and

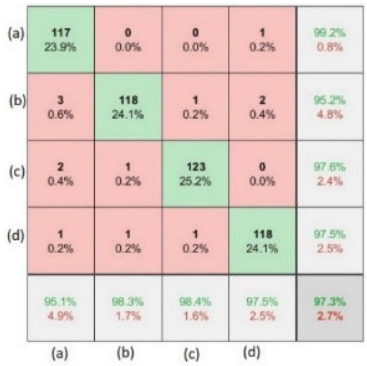

FIGURE 17. Confusion matrix for four classes: (a) LMCI (b) AD (c) Control (d) EMCI for the subject.

Communication System (PACS) server. The generalization of the model with an independent test set and different datasets will also be carried out in the future to make the model more adaptable to various other studies. We are planning to develop a CAD system by adapting the CirMNet Model to work in different MRI scanners.

VII. CONCLUSION

In this paper, we propose a hybrid feature extraction technique called CirMNet to address the limitations of CNN by incorporating additional shape features from CMSMD. We also present new statistical and FD-based feature extraction methods using the CMSMD. The CirMNet architecture consists of two parallel tracks: one utilizing a CNN and the other utilizing a CMSMD. From the CNN track, 1,479,200 features were extracted using three convolution layers, three pooling layers, and five fully connected layers and combined with the 19 most relevant features from the CMSMD track. The CMSMD features had the highest priority, and were among the top 70 ranked out of 1,479,219 features. The performance accuracy of CNN is 93.66%, CMSMD is 96.72%, and that of the CirMNet model is 97.34%, with times of 48 minutes, 20 minutes, and 63 minutes, respectively. The investigation showed that the suggested model performed better in classifying brain MRI scans into Control, EMCI, LMCI, and AD by integrating CMSMD features as well as CNN features. Considering the performance, CMSMD appears to be superior to CNN, and CirMNet outperforms CNN, while the complexities of the models followed the increasing order of CMSMD, CNN, and CirMNet. While considering the evaluation time, CMSMD outperforms better than the other two models, but CMSMD does not extract the texture and color features, which are also relevant for object classification studies. The development of a deep learning technique, which incorporates texture, color, and other essential features along with shape and margin for classification, is in the future scope of the study. We're also working to create a deep learning-enabled brain atlas for automated Alzheimer's disease (AD) detection, and we intend to provide a computer-aided detection (CAD) system in real-time clinical situations.

ACKNOWLEDGMENT

The Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) provided the information for this article. The ADNI investigators made design and implementation contributions, but they did not take part in the analysis or writing of this publication. https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf contains a complete list of ADNI researchers.

REFERENCES

[1] E. Miranda, M. Aryuni, and E. Irwansyah, "A survey of medical image classification techniques," in *Proc. Int. Conf. Inf. Manag. Technol. (ICIMTech)*, Nov. 2016, doi: 10.1109/icimtech.2016.7930302.
[2] G. Malu, S. Elizabeth, and S. M. Koshy, "Circular mesh-based shape and margin descriptor for object detection," *Pattern Recognit.*, vol. 84, pp. 97–111, Dec. 2018, doi: 10.1016/j.patcog.2018.07.004.
[3] M. K. Patrick, A. F. Adekoya, A. A. Mighty, and B. Y. Edward, "Capsule networks—A survey," *J. King Saud Univ., Comput. Inf. Sci.*, Sep. 2019, doi: 10.1016/j.jksuci.2019.09.014.
[4] X. Chen, X. Wang, K. Zhang, K.-M. Fung, T. C. Thai, K. Moore, R. S. Mannel, H. Liu, B. Zheng, and Y. Qiu, "Recent advances and clinical applications of deep learning in medical image analysis," *Med. Image Anal.*, vol. 79, Jul. 2022, Art. no. 102444, doi: 10.1016/j.media.2022.102444.
[5] M. A. Khan et al., "Multiclass skin lesion classification using deep learning networks optimal information fusion," *Deleted J.*, vol. 6, no. 6, May 2024, doi: 10.1007/s42452-024-05998-9.
[6] R. C. G. Lez and R. E. Woods, *Digital Image Processing*, 4th ed. Noida, India: Pearson, 2018.
[7] A. Farooq, S. Anwar, M. Awais, and S. Rehman, "A deep CNN based multi-class classification of Alzheimer's disease using MRI," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2017, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8261460
[8] M. Yildirim and A. Cinar, "Classification of Alzheimer's disease MRI images with CNN based hybrid method," *Ingénierie des Systèmes d Inf.*, vol. 25, no. 4, pp. 413–418, Sep. 2020, doi: 10.18280/isi.250402.
[9] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6, doi: 10.1109/ICEngTechnol.2017.8308186.
[10] Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, and B. Yu, "Recent advances in convolutional neural network acceleration," *Neurocomputing*, vol. 323, pp. 37–51, Jan. 2019, doi: 10.1016/j.neucom.2018.09.038.
[11] G. Douaud, H. Refsum, C. A. de Jager, R. Jacoby, T. E. Nichols, S. M. Smith, and A. D. Smith, "Preventing Alzheimer's disease-related gray matter atrophy by B-vitamin treatment," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 23, pp. 9523–9528, May 2013, doi: 10.1073/pnas.1301816110.
[12] B. Khagi and G.-R. Kwon, "3D CNN design for the classification of Alzheimer's disease using brain MRI and PET," *IEEE Access*, vol. 8, pp. 217830–217847, 2020, doi: 10.1109/ACCESS.2020.3040486.
[13] M. Tanveer, B. Richhariya, R. U. Khan, A. H. Rashid, P. Khanna, M. Prasad, and C. T. Lin, "Machine learning techniques for the diagnosis of Alzheimer's disease," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 16, no. 1s, pp. 1–35, Apr. 2020, doi: 10.1145/3344998.
[14] S. Basheera and M. S. Ram, "Convolution neural network–based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation," *Alzheimer's Dementia, Transl. Res. Clin. Intervent.*, vol. 5, no. 1, pp. 974–986, Jan. 2019, doi: 10.1016/j.trci.2019.10.001.
[15] A. Mehmood, M. Maqsood, M. Bashir, and Y. Shuyuan, "A deep Siamese convolution neural network for multi-class classification of Alzheimer disease," *Brain Sci.*, vol. 10, no. 2, p. 84, Feb. 2020, doi: 10.3390/brainsci10020084.

[16] Y. Lin, K. Huang, H. Xu, Z. Qiao, S. Cai, Y. Wang, and L. Huang, "Predicting the progression of mild cognitive impairment to Alzheimer's disease by longitudinal magnetic resonance imaging-based dictionary learning," *Clin. Neurophysiol.*, vol. 131, no. 10, pp. 2429–2439, Oct. 2020, doi: 10.1016/j.clinph.2020.07.016.

[17] S. Liu, S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, and M. J. Fulham, "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1132–1140, Apr. 2015, doi: 10.1109/TBME.2014.2372011.

[18] M. Liu, D. Zhang, E. Adeli, and D. Shen, "Inherent structure-based multiview learning with multitemplate feature representation for Alzheimer's disease diagnosis," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1473–1482, Jul. 2016, doi: 10.1109/TBME.2015.2496233.

[19] S. Sreelakshmi, G. Malu, E. Sherly, and R. Mathew, "M-Net: An encoder–decoder architecture for medical image analysis using ensemble learning," *Results Eng.*, vol. 17, Mar. 2023, Art. no. 100927, doi: 10.1016/j.rineng.2023.100927.

[20] J. Zhang, M. Liu, L. An, Y. Gao, and D. Shen, "Alzheimer's disease diagnosis using landmark-based features from longitudinal structural MR images," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 6, pp. 1607–1616, Nov. 2017, doi: 10.1109/JBHI.2017.2704614.

[21] V. Valotassiou, J. Malamitsi, J. Papatriantafyllou, E. Dardiotis, I. Tsougos, D. Psimadas, S. Alexiou, G. Hadjigeorgiou, and P. Georgoulias, "SPECT and PET imaging in Alzheimer's disease," *Ann. Nucl. Med.*, vol. 32, no. 9, pp. 583–593, Aug. 2018, doi: 10.1007/s12149-018-1292-6.

[22] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, Apr. 2020, doi: 10.1007/s10462-020-09825-6.

[23] X. Yao, X. Wang, S.-H. Wang, and Y.-D. Zhang, "A comprehensive survey on convolutional neural network in medical image analysis," *Multimedia Tools Appl.*, vol. 81, no. 29, pp. 41361–41405, Dec. 2022, doi: 10.1007/s11042-020-09634-7.

[24] D. Bhatt, C. Patel, H. Talsania, J. Patel, R. Vaghela, S. Pandya, K. Modi, and H. Ghayvat, "CNN variants for computer vision: History, architecture, application, challenges and future scope," *Electronics*, vol. 10, no. 20, p. 2470, Oct. 2021, doi: 10.3390/electronics10202470.

[25] N. Akhtar and U. Ragavendran, "Interpretation of intelligence in CNN-pooling processes: A methodological survey," *Neural Comput. Appl.*, vol. 32, no. 3, pp. 879–898, Feb. 2020, doi: 10.1007/s00521-019-04296-5.

[26] H. Zhang, J. Liu, Z. Zhu, and H. Li, "An automated and simple method for brain MR image extraction," *Biomed. Eng. OnLine*, vol. 10, no. 1, Sep. 2011, Art. no. 81, doi: 10.1186/1475-925x-10-81.

[27] P. J. Kostelec and S. Periaswamy, "Image registration for MRI," Modern Signal Process., MSRI Publications, Cambridge Univ. Press, pp. 161–184, 2003, vol. 46. [Online]. Available: http://library.msri.org/books/Book46/files/07kostelec.pdf

[28] H. Braak and E. Braak, "Neuropathological stageing of Alzheimer-related changes," *Acta Neuropathologica*, vol. 82, pp. 239–259, Sep. 1991. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/1759558

[29] L. Harper, F. Bouwman, E. J. Burton, F. Barkhof, P. Scheltens, J. T. O'Brien, N. C. Fox, G. R. Ridgway, and J. M. Schott, "Patterns of atrophy in pathologically confirmed dementias: A voxelwise analysis," *J. Neurol., Neurosurgery Psychiatry*, vol. 88, no. 11, pp. 908–916, Nov. 2017, doi: 10.1136/jnnp-2016-314978.

[30] C. R. Jack, D. W. Dickson, J. E. Parisi, Y. C. Xu, R. H. Cha, P. C. O'Brien, S. D. Edland, G. E. Smith, B. F. Boeve, E. G. Tangalos, E. Kokmen, and R. C. Petersen, "Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia," *Neurology*, vol. 58, no. 5, pp. 750–757, Mar. 2002, doi: 10.1212/wnl.58.5.750.

[31] B. D. C. Boon, P. J. W. Pouwels, L. E. Jonkman, M. J. Keijzer, P. Preziosa, W. D. J. van de Berg, J. J. G. Geurts, P. Scheltens, F. Barkhof, A. J. M. Rozemuller, F. H. Bouwman, and M. D. Steenwijk, "Can post-mortem MRI be used as a proxy for in vivo? A case study," *Brain Commun.*, vol. 1, no. 1, Jan. 2019, Art. no. fcz030, doi: 10.1093/braincomms/fcz030.

[32] D. L. G. Hill et al., "Coalition against major diseases/European medicines agency biomarker qualification of hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's disease," *Alzheimer's Dementia*, vol. 10, no. 4, p. 421, Jul. 2014, doi: 10.1016/j.jalz.2013.07.003.

[33] H. Hampel, K. Bürger, S. J. Teipel, A. L. W. Bokde, H. Zetterberg, and K. Blennow, "Core candidate neurochemical and imaging biomarkers of Alzheimer's disease," *Alzheimer's Dementia*, vol. 4, no. 1, pp. 38–48, Jan. 2008, doi: 10.1016/j.jalz.2007.08.006.

[34] S. J. Teipel, M. Grothe, S. Lista, N. Toschi, F. G. Garaci, and H. Hampel, "Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease," *Med. Clinics North Amer.*, vol. 97, no. 3, pp. 399–424, May 2013, doi: 10.1016/j.mcna.2012.12.013.

[35] J. P. Lerch, J. Pruessner, A. P. Zijdenbos, D. L. Collins, S. J. Teipel, H. Hampel, and A. C. Evans, "Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls," *Neurobiol. Aging*, vol. 29, no. 1, pp. 23–30, Jan. 2008, doi: 10.1016/j.neurobiolaging.2006.09.013.

[36] C. R. Jack, H. J. Wiste, T. G. Lesnick, S. D. Weigand, D. S. Knopman, P. Vemuri, V. S. Pankratz, M. L. Senjem, J. L. Gunter, M. M. Mielke, V. J. Lowe, B. F. Boeve, and R. C. Petersen, "Brain β-amyloid load approaches a Plateau," *Neurology*, vol. 80, no. 10, pp. 890–896, Mar. 2013, doi: 10.1212/wnl.0b013e3182840bbe.

[37] V. L. Villemagne, V. Doré, P. Bourgeat, S. C. Burnham, S. Laws, O. Salvado, C. L. Masters, and C. C. Rowe, "A β-amyloid and tau imaging in dementia," *Seminars Nucl. Med.*, vol. 47, no. 1, pp. 75–88, Jan. 2017, doi: 10.1053/j.semnuclmed.2016.09.006.

[38] W. M. van Oostveen and E. C. M. de Lange, "Imaging techniques in Alzheimer's disease: A review of applications in early diagnosis and longitudinal monitoring," *Int. J. Mol. Sci.*, vol. 22, no. 4, p. 2110, Feb. 2021, doi: 10.3390/ijms22042110.

[39] R. Indrakumari, T. Poongodi, and K. Singh, "Introduction to deep learning," in *Advanced Deep Learning for Engineers and Scientists* (EAI/Springer Innovations in Communication and Computing). Cham, Switzerland: Springer, 2021, pp. 1–22, doi: 10.1007/978-3-030-66519-7_1.

[40] *ADNI Dataset*. Accessed: May 25, 2021. [Online]. Available: http://adni.loni.usc.edu/methods/mri-too

[41] A. Khan and S. Zubair, "An improved multi-modal based machine learning approach for the prognosis of Alzheimer's disease," *J. King Saud Univ., Comput. Inf. Sci.*, vol. 34, no. 6, pp. 2688–2706, Jun. 2022, doi: 10.1016/j.jksuci.2020.04.004.

[42] M. G. Lanjewar, J. S. Parab, and A. Y. Shaikh, "Development of framework by combining CNN with KNN to detect Alzheimer's disease using MRI images," *Multimedia Tools Appl.*, vol. 82, no. 8, pp. 12699–12717, Sep. 2022, doi: 10.1007/s11042-022-13935-4.

[43] J. B. Bae, S. Lee, W. Jung, S. Park, W. Kim, H. Oh, J. W. Han, G. E. Kim, J. S. Kim, J. H. Kim, and K. W. Kim, "Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging," *Sci. Rep.*, vol. 10, no. 1, p. 22252, Dec. 2020.

[44] A. Nachmani, R. Schurr, L. Joskowicz, and A. A. Mezer, "The effect of motion correction interpolation on quantitative T1 mapping with MRI," *Med. Image Anal.*, vol. 52, pp. 119–127, Feb. 2019, doi: 10.1016/j.media.2018.11.012.

[45] E. Yagis, S. W. Atnafu, A. G. S. de Herrera, C. Marzi, R. Scheda, M. Giannelli, C. Tessa, L. Citi, and S. Diciotti, "Effect of data leakage in brain MRI classification using 2D convolutional neural networks," *Sci. Rep.*, vol. 11, no. 1, Nov. 2021, Art. no. 22544, doi: 10.1038/s41598-021-01681-w.

G. MALU was born in Kollam, Kerala, India. She received the B.Sc., M.C.A., and Ph.D. degrees from the University of Kerala, in 2000, 2004, and 2017, respectively. From 2005 to 2012, she was an Assistant Professor with the Department of Computer Applications, SNIT, University of Kerala. From 2012 to 2016, she was a Research Scholar with the University of Kerala. From 2017 to 2021, she was a Senior Research Fellow and a Postdoctoral Fellow with the Indian Institute of Information Technology and Management–Kerala. In 2022, she joined the Kerala University of Digital Sciences, Innovation, and Technology, as a Research

Officer, where she is currently an Assistant Professor. In more than 19 years of teaching and research, she has gathered expertise in pattern recognition, medical image processing, shape and margin descriptors, and graph algorithms. She received an Indian patent titled "An Automated Lesion Detection System for DCE-MRI Using Circular Mesh-Based Shape and Margin Descriptor." She was a recipient of the SPEED-IT Research Fellowship and is part of several central government and state government projects. Her academic portfolio includes numerous papers in international and national journals and conferences, along with editorial responsibilities in various scholarly publications.

NAYANA UDAY (Member, IEEE) was born in Kollam, Kerala, India, in 1989. She received the B.Sc. and M.Sc. degrees from the University of Kerala, in 2010 and 2012, respectively, and the M.Phil. degree in computer science from Cochin University of Science and Technology, in 2018. She is currently pursuing the Ph.D. degree in computer science with Kerala University of Digital Sciences, Innovation, and Technology. From 2013 to 2015, she was a programmer in a leading software firm in Kerala. Her research interests include pattern recognition, medical image processing, graph algorithms, brain connectome models, and deep learning algorithms.

ELIZABETH SHERLY was born in Thiruvalla, Kerala, India, in 1962. She received the Ph.D. degree in computer science from the University of Kerala, in 1995, focusing on artificial neural networks in certain control systems. She began her career as a Research Assistant with the University of Kerala, contributing to a Renal Autoregulation Project. In 2001, she joined the Indian Institute of Information Technology and Management–Kerala (IIITM-K), an Assistant Professor until 2020. She is currently holding the position of a Senior Professor with Kerala University of Digital Sciences, Innovation, and Technology. Her research interests include pattern recognition, digital image processing, and natural language processing, resulting in over 100 publications across international, national, and conference platforms. Additionally, she has supervised 12 Ph.D.'s and led projects sponsored by the Ministry of Electronics and IT of the Government of India. She received an accolades, such as the Eureka Fellowship in the U.K., and two awards from the MHRD for Cloud Initiative and e-Governance Initiatives, in 2009 and 2011, respectively. She served as the Director of IIITM-K and the International Center for Free and Open Software in Thiruvananthapuram, Kerala.

AJITH ABRAHAM (Senior Member, IEEE) received the B.Tech. degree in electrical and electronic engineering from the University of Calicut, in 1990, the M.Sc. degree from Nanyang Technological University, Singapore, in 1998, and the Ph.D. degree in computer science from Monash University, Melbourne, Australia, in 2001. He is the Vice Chancellor of Bennett University, New Delhi. Prior to this, he was the Dean of the Faculty of Computing and Mathematical Sciences, FLAME University, Pune, and the Founding Director of the Machine Intelligence Research Labs (MIR Labs), USA, a not-for-profit scientific network for innovation and research excellence connecting industry and academia. He also held two international university professorial appointments, such as, a Professor of artificial intelligence with Innopolis University, Russia, and the Yayasan Tun Ismail Mohamed Ali Professorial Chair of artificial intelligence with UCSI, Malaysia. He is involved in a multidisciplinary environment. He has authored/coauthored more than 1 500 research publications out of which there are more than 100 books covering various aspects of computer science. One of his books was translated into Japanese and a few other articles were translated into Russian and Chinese. He has more than 60 000 academic citations (H-Index of more than 115 as per Google Scholar). He has given more than 250 plenary lectures and conference tutorials (in more than 20 countries). He was the Chair of the IEEE Systems Man and Cybernetics Society Technical Committee on Soft Computing (which has over more than 200 members), from 2008 to 2021, and served as a Distinguished Lecturer for the IEEE Computer Society Representing Europe, from 2011 to 2013. He was the Editor-in-Chief of Engineering Applications of Artificial Intelligence (EAAI), from 2016 to 2021, and serves/served on the Editorial Board for over 15 international journals indexed by Thomson ISI.

NARENDRA KUBER BODHEY was born in Nagpur, Maharashtra, India. He received the M.B.B.S. and M.D. degrees in radiology from the University of Nagpur, Maharashtra, India, in 1995 and 1999, respectively, and the D.N.B. degree in radiodiagnosis from the National Board of Examinations, New Delhi, India, in 2002. From 2000 to 2013, he served as a Radiology Faculty at the Sri Chitra Tirunal Institute of Medical Sciences, Thiruvananthapuram, Kerala, and currently as a Professor and the Head of Radiodiagnosis at the All India Institute of Medical Sciences Raipur, Chhattisgarh, India. Specializing in interventional radiology, neuroimaging, cardiovascular and thoracic imaging, he has approximately 100 publications in various national and international peer-reviewed journals, along with extensive mentoring experience. He completed numerous research projects and was awarded the BOYSCAST fellowship by the Department of Science and Technology, Government of India.