Article

Early Detection of Alzheimer's Disease Using Polar Harmonic Transforms and Optimized Wavelet Neural Network

Shabana Urooj [1], Satya P. Singh [2,*], Areej Malibari [3], Fadwa Alrowais [4] and Shaeen Kalathil [1]

[1] Department of Electrical Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, Riyadh 84428, Saudi Arabia; smurooj@pnu.edu.sa (S.U.); skalathil@pnu.edu.sa (S.K.)
[2] Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 608232, Singapore
[3] College of Engineering, Princess Nourah Bint Abdulrahman University, Riyadh 84428, Saudi Arabia; aamalibari@pnu.edu.sa
[4] Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 84428, Saudi Arabia; fmalrowais@pnu.edu.sa
* Correspondence: satya@ntu.edu.sg

Abstract: Effective and accurate diagnosis of Alzheimer's disease (AD), as well as early-stage detection, has gained more and more attention in recent years. For AD classification, we propose a new hybrid method for early detection of Alzheimer's disease (AD) using Polar Harmonic Transforms (PHT) and Self-adaptive Differential Evolution Wavelet Neural Network (SaDE-WNN). The orthogonal moments are used for feature extraction from the grey matter tissues of structural Magnetic Resonance Imaging (MRI) data. Irrelevant features are removed by the feature selection process through evaluating the in-class and among-class variance. In recent years, WNNs have gained attention in classification tasks; however, they suffer from the problem of initial parameter tuning, parameter setting. We proposed a WNN with the self-adaptation technique for controlling the Differential Evolution (DE) parameters, i.e., the mutation scale factor (F) and the cross-over rate (CR). Experimental results on the Alzheimer's disease Neuroimaging Initiative (ADNI) database indicate that the proposed method yields the best overall classification results between AD and mild cognitive impairment (MCI) (93.7% accuracy, 86.0% sensitivity, 98.0% specificity, and 0.97 area under the curve (AUC)), MCI and healthy control (HC) (92.9% accuracy, 95.2% sensitivity, 88.9% specificity, and 0.98 AUC), and AD and HC (94.4% accuracy, 88.7% sensitivity, 98.9% specificity and 0.99 AUC).

Keywords: Alzheimer's disease; discriminative analysis; healthy control; mild cognitive impairment; polar harmonic transform; wavelet neural networks

1. Introduction

Alzheimer's disease (AD) is a general form of dementia correlated with the pathological amyloid depositions, structural-atrophy, and metabolic changes in the brain [1,2]. AD usually matures when the nerve cells in the brain die or their functioning becomes abnormal [3,4]. It is a major source of dementia among older people with 47 million people worldwide living with dementia in 2016 [5]. Developing countries are the most affected by this growth rate as 59% of dementia people already living there. This figure maybe around 59% by 2050 [6]. Therefore, early detection of Alzheimer's disease can be the key to slowing, preventing, and stopping the occurrence of dementia at its early stage [7].

In medical observations, the diagnosis of AD is often missed or overlooked at the pre-clinical stage by expert radiologists [8]. A computer-aided diagnosis (CAD) system might increase the judgmental precision of the expert radiologist by providing the advantages of having a second eye to diagnose Alzheimer's at the pre-clinical stage [7,9,10]. Features extraction is the key step for the adequate design of any CAD system [8]. The expert radiologist generally focuses on features such as mesial temporal lobe atrophy and temporoparietal cortical atrophy [11]. The mesial temporal lobe could be analyzed

by visual inspection of the state of hippocampal and Parahippocampal or by indirectly examining the magnification of the Parahippocampal fissures [11] The temporoparietal cortical atrophy is more responsive and explicit. It requires precise volumetric estimation and is usually difficult to analyze by 'eyeballing' [12]. However, there is always some useful information (specific features in the image) about the appearance and changes of the abnormalities in the image in contrast to the image of healthy tissues.

Utilizing the above information, several methods have been suggested to diagnose healthy control (HC), MCI, and AD by using a CAD system. Most studies to the date computed features based on volumetric measurement of segmented *region of interest* (ROI) [13, 14], *voxel-based morphometry* (VBM) [15,16], and voxel-wise statistical approaches [17,18]. Among statistical approaches, *linear discriminant analysis* (LDA) [19] and *principal component analysis* (PCA) [20] are the commonly used statistical tools for feature extraction and data analysis. Another frequently used AD bio-marker is *independent component analysis* (ICA). Khedar et al. [21] used ICA for feature extraction and SVM as a classifier for classifying stages in AD. Zhang et al. [22] extracted 13 features using stationary wavelets (SWE) from volumetric images and classified those features by utilizing a shallow feed-forward neural network. They optimized weights and biases by using predator-prey particle swarm optimization (PSO). However, they used only one axial slice form volumetric images for experimental purposes. Ortiz et al. [23] used a combination of deep-learning (DL) and gray matter (GM) for AD classification and reported high-performance indices. They extracted 3D patches from GM and trained a deep belief network. The statistical approaches are sensitive to geometric deformation/perturbation [24]. Therefore, any unexpected perturbation will affect the measurement [25]. Statistical approaches make high correlated descriptors which led to low computational accuracy [8]. It is evident that statistical approaches produce highly correlated descriptors which led to high computational accuracy. On the contrary, the feature measurement by orthogonal moments is invariant to rotation naturally. Furthermore, with a little modification, they can be invariant to other geometric deformations (i.e., scale, illumination, etc.). Therefore, any unexpected perturbation will not affect the measurement confronting thus the inherent uncertainty.

The orthogonal moments and their functions have been utilized in several applications in medical imaging. The feature extraction by *Zernike Moment* (ZM) and *Pseudo-ZM* (PZM) captures global information and does not require any closed boundary such as Fourier-based-descriptors [26]. Gorji and Haddadnia [27] utilized the properties of PZM for the identification of MCI from AD and HC groups using structural fMRI. In this study, they utilize all the PZMs up to the maximum order of 30 to construct the feature vectors. *Polar Cosine Transform* (PCT), *Polar Sine Transform* (PST), and *Polar Complex Exponential Transform* (PCET) were introduced by Yap et al. [28]. Because the kernels are harmonics in nature, these are jointly called as the Polar Harmonic Transforms (PHTs) which have been adopted in many pattern recognition applications including medical imaging [24], invariant feature extraction [29], color object detection [25], and image hashing [30]. The kernel computation of PCTs is much simpler than ZMs and PZMs. Hence, PHTs perform to a great extent higher speed compared to ZMs [31]. Moreover, PHTs are free from any numerical stability issue [32].

On the classification part, we use Wavelet neural networks (WNNs) for feature classification [33]. WNNs are a new type of feedforward neural networks where the discrete wavelet function is used for activation. WNNs use a gradient descent technique for optimization that often suffers from the local minima, long convergence time problems. Further, the activation function must be differentiable. Therefore, in addition to their own good performance, WNNs face problems with initial parameter tuning and parameter settings. To solve these issues, particle swarm optimization (PSO) and differential evolution (DE) approaches are often used to optimize the WNNs parameters. In [34], authors use DE algorithm to train a WNN and named it Differentially optimized WNN (DEWNN). The model was tested on bankruptcy data set and other datasets such as Iris, breast cancers. The results confirm the advantages of DE with WNN compared to WNN only. In the other

work [35], to improve the classification accuracy of WNN in mammograms, the authors use the PSO approach. Compared to PSO, DE optimized WNNs are robust to noise and show consistent performance over many trials. However, it is likely to be premature and a situation may arrive where one cannot give the guarantee of the diversity of the population. To solve these issues, we adopt the self-adaptation technique [36] for parameter control using two parameters of standard DE, i.e., the mutation scale factor(F) and the cross-over rate (CR).

In this paper, we propose a new hybrid method for early diagnosis of Alzheimer's disease using Polar Harmonic Transforms (PHT) and Self-adaptive Differential Evolution Wavelet Neural Network (SaDE-WNN).

(a)   PHT is utilized for feature extraction. PHT not only has excellent reconstruction properties but also can be constructed without difficulty in an arbitrarily higher order [28]. More importantly, PHT shows robustness to noise, less information redundancy problems, and competent reconstruction ability in image analysis in comparison with ZMs and PZMs.

(b)   SaDE-WNN is an improvement of the *Differential Evolution optimized WNN* (DEWNN) [29] with the self-adaptation technique [30] for parameter control using two parameters of the standard DE. The proposed method is free from any human involvement. The proposed SaDE-WNN improves DE-WNN in terms of parameter tuning.

The rest of the paper are organized as follows. Section 2 presents the material and methods used in this study. Section 3 presents the results and discussion. Section 4 covers the conclusion and future works.

2. Materials and Methods

2.1. Data Collection

Data were downloaded from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (available from the site, http://adni.loni.usc.edu). The ADNI was launched in the year 2004 as a five-year program partnering with the food and drug administration, the national institute of biomedical imaging and bioengineering, and the national institute of aging. Currently, ADNI reaches to ADNI-3 stage through ADNI-1, ADNI-GO, and ADNI-2. We have collected subjects from the ADNI-2 phase. We use 3T baseline T1-weighted images in this study. All the subjects used (T1 MRI) in this study are between 55 and 90 years old with general addition/omission criteria [1,27].

(a)   HC subjects: *mini-mental state examination* (MMSE) scores between 24 and 30, non-depressed, non-MCI, a *clinical dementia rating* (CDR) scale of 0, and non-demented.

(b)   MCI subjects: MMSE scores between 24 and 30; an absence of dementia; CDR of 0.5; strongly retained activities of daily living, and memory objection and objective memory loss measured by education-adjusted scores.

(c)   AD subjects: MMSE scores between 20 and 26 (inclusive); meet NINCDS/ADRDA criteria for portable AD, *Geriatric Depression Scale* (GDS) less than 6 and 5, and CDR of 0.5 or 1.0.

(d)   All the subjects were excluded if they had any other significant neurological disorder other than Alzheimer's disease. In total, 892 subjects (AD = 258, MCI = 304, and HC = 330) were included in the current study. The demographics of the cohort are given in Table 1.

Table 1. Summary of clinical subject demographics and dementia status.

	AD (n = 258, 108F/150M)			MCI (n = 304, 124F/180M)			HC (n = 330, 150F/180M)		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**	**Mean**	**SD**	**Range**
Age [years]	74.8	7.6	68–77	74.7	7.1	61–88	75.3	5.2	58–88
Education [years]	14.9	3.4	8–20	15.9	2.9	12–20	15.8	3.2	12–20
MMSE	23.6	1.9	20–26	27.0	1.7	24–30	28.6	1.3	25–30
CDR	0.7	0.3	0.5–1.0	0.5	0.0	0.5–0.5	0.0	0.0	0.5–1.0

2.2. Pre-Processing

All the subjects were scanned from 3T MRI scanner with a standardized MRI protocol developed for ADNI [37]. All MRI images are preprocessed as follows: (1) Raw Dicom data is transformed to NIfTI using MRIcron software. We use SPM12 (thus MATLAB 2018) for field intensity inhomogeneity correction, and tissue (grey matter (GM)) segmentation. The segmented image is then normalized to MNI space. From the processed GM images, we selected 116 regions of interest (ROI) using Automated anatomical labeling (AAL) atlas. Then from each ROI, we calculated features using the PHT moments as described in the following section. The criterion for selecting the number of features from each ROI will be discussed in the *Feature extraction* section. Further, important features were selected through evaluating the in-class and among-class variance. All these steps are shown in Figure 1, and we will discuss feature extraction, selection, and classification steps in detail in Section 2.4.
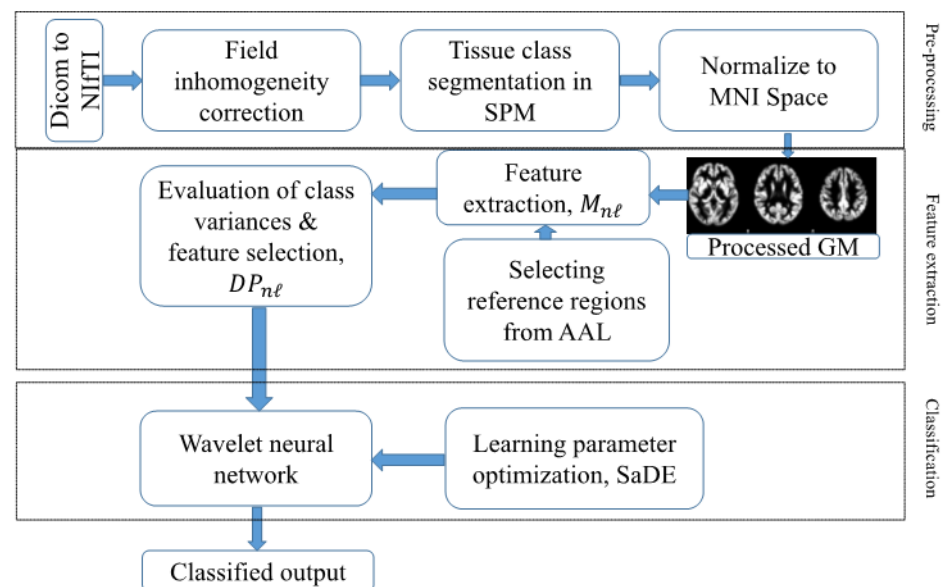
Figure 1. Systematic flow diagram of the proposed approach.

2.3. Feature Extraction

Moment theory had been effectively employed in several engineering applications including medical imaging and pattern recognition problems [25,29,38]. The orthogonal moments are similar to statistical moments such as mean, variance, kurtoses are equal to first, second, and third order of the PHT, and so on. PHT moments can extract an infinite number of features from the images including texture, shape, mean, variance and this list may extend to infinity [39]. The distinct orders of PHT moments signify distinct spatial distribution of image intensity variations. In this context, a collection of these moments can be molded to be a global shape descriptor of a particular image. The features (including AD, MCI, and HC features) are then extracted using the proposed PHT algorithm (disused in the next section) from the segmented (GM) tissue. The regions for feature extraction are

decided using Automated anatomical labeling (AAL) atlas. The details about the feature extraction and selection are as follows:

Feature extraction using PHT. PHTs [28] of order n and repetition ℓ of a piecewise continuous real function, $f(x, y)$ are defined over the unit disk $D = \{(x, y) : x^2 + y^2 \leq 1\}$ as follows:

$$M_{n\ell} = \lambda_N \int_0^{2\pi} \int_0^1 f(r, \theta) [H_{n\ell}(r, \theta)]^* r dr d\theta \tag{1}$$

where $[.]^*$ denotes complex conjugate, r and θ be the polar co-ordinates radius and angle. $H_{n\ell}(r, \theta)$ can be decomposed into radial and angular components:

$$H_{n\ell}(r, \theta) = R_n(r) \exp(i\ell\theta) \tag{2}$$

The radial kernel can be written explicitly as $\exp(i2\pi nr^2)$, $\cos(\pi nr^2)$, and $\sin(\pi nr^2)$ for PCET, PCT, and PST respectively. The normalization constant λ_N is given as:

$$\lambda_N = 1/\pi \; for \; \text{PCET} \tag{3}$$

$$\lambda_N = \begin{cases} 1\pi & n = 0 \\ 2\pi & n \neq 0 \end{cases} \; for \; \text{PCT \& PST} \tag{4}$$

Now, let the input image rotates by an angle of α. PHTs of the original and its rotated sample can be related as:

$$M_{n\ell}^{rotated} = \lambda_N \int_0^{2\pi} \int_0^1 R_{n\ell} \exp(j\ell, \theta) f(r, \theta - \alpha) r dr d\theta \tag{5}$$

By changing the variable $\theta' = \theta - \alpha$, we write:

$$\begin{aligned} M_{n\ell}^{rotated} &= \lambda_N \int_0^{2\pi} \int_0^1 R_{n\ell} \exp(j\ell(\theta' + \alpha)) f(r, \theta') r dr d\theta \\ M_{n\ell}^{rotated} &= M_{n\ell} \exp(j\ell\theta) \end{aligned} \tag{6}$$

Equation (6) expresses the rotational property of PHTs. If we nullify the exponential term of (6), the completely rotational invariant system is obtained. The straightforward way to get the absolute invariant is to take modulo on both sides of (6) which results in:

$$\left| M_{n\ell}^{rotated} \right| = |M_{n\ell}| \tag{7}$$

where, $M_{n\ell}^{rotated}$ and $M_{n\ell}$ are the PHTs of the rotated and the original images, respectively. Thus, the proposed features are the magnitude of the PHTs as texture descriptors.

The total number of features for PCET, and PCT is $(1 + 2n_{max})(1 + 2\ell_{max})$, $(1 + n_{max})(1 + 2\ell_{max})$ and) respectively. For further comparative analysis, we also experimented ZM [27] using the proposed approach. To observe the effect of non-orthogonality, we also utilize the properties of Rotational moments (RMs) [38] as texture features. They are the non-orthogonal moments, and the features are extracted similarly to PHTs. The total number of features for RM is $(1 + n_{max})(1 + 2\ell_{max})$. The extracted features are classified into AD/MCI/HC using the appropriate classification scheme.

2.4. Feature Selection through Evaluating the In-Class and Among-Class Variance

Execution of the classifier highly depends on the discriminative strength of the extracted features [40]. The performance of the classifier could be significantly improved by assessing the in-class and among-class variance of the training dataset. We divide the entire dataset into two groups i.e., training and testing sets. Let the dataset have k classes and N_{tr}^c be the number of training samples for each class c; in total, we create $k \times N_{tr}^c$ training images. Extract the invariant features of PCET, ZM, PCT, and RM at some maximum order of moments M_{max}. Thereafter, construct a training set matrix, $F_{n\ell}^c$, $n, \ell \in 0, 1, \ldots M_{max}$ and

$c \in 0, 1, \ldots k$ for all classes and all training images and arranged in a matrix form. The mean value of the moments of each class, $\mu_{n\ell}^c$ is defined as

$$\mu_{n\ell}^c = \frac{1}{N_{tr}^c} \sum_{t=1}^{N_{tr}^c} F_{n\ell}^c(t) \tag{8}$$

For the sake of discriminative capability of the moment's in-class, we compute the variance of each moment's in-class as follows:

$$V_{n\ell}^c = \frac{1}{k} \sum_{c=1}^{k} \sum_{t=1}^{N_{tr}^c} \left| F_{n\ell}^c(t) - \mu_{n\ell}^c \right|^2 \tag{9}$$

In similar manners, the total variance of all training samples is calculated by

$$V_{n\ell}^T = \sum_{c=1}^{k} \sum_{t=1}^{N_{tr}^c} \left| F_{n\ell}^c(t) - \mu_{n\ell}^T \right|^2 \tag{10}$$

where, $\mu_{n\ell}^T$ is the mean value of the moments of the complete training samples and given as,

$$\mu_{n\ell}^T = \frac{1}{k \times N_{tr}^c} \sum_{c=1}^{k} \sum_{t=1}^{N_{tr}^c} F_{n\ell}^c(t) \tag{11}$$

Finally, we define the discriminative power of PHTs of the order n and repetition ℓ:

$$DP_{n\ell} = \frac{V_{n\ell}^T}{V_{n\ell}^c} \tag{12}$$

It could be observed from Equation (11) that a high value of $DP_{n\ell}$ results in small in-class variations concerning the total variations. This means that the higher value of $DP_{n\ell}$ will lead to higher discrimination ability of the interrelated moments. Therefore, it could be anticipated that feature selection using $DP_{n\ell}$ attain the higher accuracy of the proposed classifier compared to the traditional approaches. All the feature vectors are arranged in descending order, and we select only those moments PHTs and RM which have the highest $DP_{n\ell}$.

2.5. Classification Using the SaDE-WNN Method

In the context of WNN, initial parameter tuning, and parameter setting is an important task, and the performance of WNN depends highly on the values of these parameters. For example, if the learning rate parameter is not set properly, the network can either lead to vibrations or be stuck in an indefinite training time. The momentum rate is used to accelerate the error convergence rate [24]. WNN uses gradient descent learning for training which suffers from some underlying problems such as long convergence-time, entry into local minima, and the need for differentiation of objective function. Furthermore, the determination of the optimal number of neurons in hidden layers is another important task. With no or few hidden neurons, the network may not be able to classify the complex set of problems. By contrast, with too many hidden neurons/hidden layers, the network becomes too complex and the training of the network becomes highly time come consuming. There have been wide uses of Wavelet Neural Networks (WNN) for classification problems [33]. Among all, the Particle Swarm Optimized WNN (PSOWNN) [41], and Differential Evolution optimized WNN (DEWNN) [34] are the most popular ones as they handle the problems that remained in WNN regarding the initial parameter tuning, parameter setting such as the momentum rate, and determining the optimal number of hidden neurons/hidden layer [42]. In the current work, we adopt the self-adaptation technique [36] for parameter control using two parameters of standard DE, i.e., the mutation

scale factor(F) and the cross-over rate (CR). The output is the function of w_{ih} (weight-vector between input and hidden-layer), w_{ho} (weight vector from hidden layer to the output layer), a (dilation parameter), b (translation parameter), and u (input sample). During the training phase, these parameters $T = [w_{ih}, w_{ho}, b, a]$ are predicted and modified by minimizing the Normalized Root Mean Square Error (NRMSE). An initial random population (P) consists of such vectors: $P = [T_1, \ldots, T_{NP}]$ is generated within the boundaries:

$$T_i = T^i_{min} + rand(j) \times \left(T^i_{max} - T^i_{min} \right) \tag{13}$$

where $rand(j)$ represents uniformly distribution of number in the range of $[0, 1]$ *for* $j = 1, 2, 3 \ldots$ which generates a fresh value for every decision parameter. The factor 'F' controls the evolving rate of the population. The CR varies in the range of $[0, 1]$ and indicates the probability with which a trial individual inherits the actual individual's gene. A relatively small CR increases the probability of stagnation and decelerates the search-process [24]. By contrast, if F is relatively high, the population diversity increases and may result in premature convergence [43]. Therefore, these parameters are modified to control these parameters, self-adaptation of two parameters (F and CR) are modified as follows:

$$F_{i,G+1} = \begin{cases} F_\ell + rand1 \times F_u & if \ rand2 < p_1 \\ F_{i,G} & otherwise \end{cases} \tag{14}$$

$$CR_{i,G+1} = \begin{cases} rand3 & if \ rand4 < p_2 \\ CR_{i,G} & otherwise \end{cases} \tag{15}$$

where $i = 1, 2, \ldots, NP$ and NP denotes the number of members in the population. p_1 and p_2 are the probability to adjust the control parameter F and CR. We set $p_1 = p_2 = F_\ell = 0.1$, and $F_u = 0.9$ [43].

Using this method to determine F and CR, a self-adaptive algorithm has been developed. Both tests are vector generation strategy and their corresponding parameter values progressively self-adapted by learning from past events of generating likely solutions. The new parent vector F and CR randomly selects the value from $[0, 1]$. Moreover, $F_{i,G+1}$ and $CR_{i,G+1}$ are accomplished before mutation.

3. Experiments

3.1. Experimental Setting

We have implemented the proposed SaDE-WNN method against the related methods—DEWNN [34] and Ada-DEWNN [8] for comparison in each scenario. The parameters used for the proposed SaDE-WNN classifier are the number of hidden nodes, learning rate, and momentum rate. To determine the optimum number of hidden nodes, we start with a small value of M and estimate the corresponding weight of the output layer, and compute the Normalized Root Mean Square Error (NRMSE). If the NRMSE is less than the predefined threshold value (4–10), we stop. Otherwise, increase M and repeat the same process. The value of n is set to be 50 with 500 training epochs after 100 generations. The maximum number of iterations $G_m = 50$; the new value of Mutation scale factor-F varies between $[0.1, 1]$ and the CR varies between $[0, 1]$. The value of F_{max} and F_{min} have been selected as 0.9 and 0.2 respectively. The $L^{factor}_{min} = 0.00127$; the $M^{factor}_{min} = 0.8927$ and $N^{hidden}_{min} = 12$. The data set is divided into training, validation, and test set by ration 60:20:20. Five-fold cross-validation scheme has been utilized to avoid any over-fitting and/or uncertainty in the training of the classifier [42].

The effectiveness of the proposed CAD system is validated by the following indices: sensitivity, specificity, accuracy, F1-score (F1), and area under the ROC curve (AUC). The accuracy of diagnostic modality in terms of its trade-off between specificity and sensitivity is evaluated by receiver operating characteristics (ROC) analysis. Besides, the difference between the two feature sets is analyzed using the chi-square test. The AUC, the statistical

significance of differences between ROC index and parameters is evaluated and compared by employing the z-test.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{16}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{17}$$

$$Specificity = \frac{TN}{TN + FP} \tag{18}$$

$$F1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{19}$$

3.2. Results

Classification of AD vs. MCI: The first experiment is intended towards the testing and validation of the proposed method for the AD vs. MCI classification. The classification results are shown in Table 2 (AD vs. MCI column). The best results were obtained by the PCT feature sets with the proposed SaDE-WNN are 93.74% accuracy, 86.0% sensitivity, 98.0% specificity, and 88.5 F1. The best results for the RM feature sets are 92.0% accuracy, 84.1% sensitivity, 95.6% specificity, and 86.0 F1. The difference between accuracy, sensitivity, specificity, AUC between the PCT texture and the RM texture features are statistically significant ($p < 0.05$). However, differences are not statistically significant for PCET and PCT ($p > 0.05$).

Table 2. Experimental results for comparative analysis of the Self-adaptive Differential Evolution Wavelet Neural Network (SaDE-WNN) with Differentially optimized WNN (DEWNN) and Ada-DEWNN. Results are compared on the features extracted by non-orthogonal moments (RM) and orthogonal moments (ZM, PCET, and PCT).

Feature Extraction Methods	Classifiers	AD vs. MCI				AD vs. HC				MCI vs. HC			
		Sens.	Spec.	Acc.	F1	Sens.	Spec.	Acc.	F1	Sens.	Spec.	Acc.	F1
RM	DEWNN	70.2	96.5	82.7	81.0	83.7	88.8	90.7	89.2	84.9	87.9	82.5	88.7
ZM		72.0	94.9	87.8	82.8	85.5	92.3	92.7	90.4	92.5	92.8	87.2	89.7
PCET		83.0	95.2	91.6	88.6	86.9	94.2	91.9	90.2	93.4	83.5	91.0	89.9
PCT		84.1	95.6	92.0	86.0	87.6	96.6	93.3	91.7	93.5	86.1	92.5	91.3
RM	Ada-DEWNN	70.5	96.9	83.2	81.2	85.5	92.5	89.9	83.2	88.6	87.3	83.8	89.9
ZM		72.8	98.2	89.4	85.3	86.7	94.8	92.1	91.1	92.5	90.6	90.5	90.1
PCET		80.9	97.9	91.8	86.7	86.9	95.5	92.7	92.7	93.9	84.8	91.7	91.7
PCT		85.9	97.2	92.7	87.2	87.7	97.3	93.0	92.5	94.7	87.7	92.5	92.0
RM	SaDE-WNN	70.7	97.5	85.7	83.7	83.3	95.6	90.8	91.0	89.0	88.6	88.8	87.8
ZM		72.3	99.7	91.2	85.8	85.0	96.7	92.1	91.2	88.0	93.9	90.7	89.5
PCET		81.7	95.9	93.0	87.0	93.1	93.5	93.4	93.2	96.1	85.1	91.6	91.4
PCT		86.0	98.0	93.7	88.5	88.7	98.9	94.4	93.8	95.2	88.9	92.9	93.2

Classification of AD vs. HC: The classification results are shown in Table 2 (AD vs. HC column) in terms of accuracy, sensitivity, specificity for each feature set. The best results obtained by the PCT feature sets with the proposed SaDE-WNN are 94.4% accuracy, 88.7% sensitivity, 98.9% specificity, and 93.8 F1. The best results for the RM feature sets with SaDE-WNN are 93.3% accuracy, 87.6% sensitivity, 96.6% specificity, and 91.7 F1 scores. However, we did not notice any significant improvement in classification results. The difference between accuracy, sensitivity, specificity, between the PCT texture features and the RM texture features are statistically significant, but these differences are not statistically significant for combined and ZM ($p > 0.05$).

Classification of MCI vs. HC: In this experiment, the proposed method is tested and validated for the MCI vs. HC classification. The classification results are shown in Table 2 (MCI vs. HC column). The best of accuracy, sensitivity, specificity, and F1 for the PCT with SaDE-WNN are 92.9%, 95.2%, 88.9%, and 93.2%. The overall performance of the proposed

method for the RM feature sets with SaDE-WNN is 92.5% accuracy, 93.5% sensitivity, 86.1% specificity, and 91.3 F1. Again, the results are statistically significant for the RM texture features and PCT. However, as expected, there is no statistical significance is shown between the PCT, PCET, and PST texture feature sets.

3.3. Discussion

AD is the foremost source of dementia in elderly people accounting for 50–60% of all the cases. Correct classification of pre-clinical AD, MCI due to AD, and dementia due to AD remain challenging issues for the researchers. Among the existing literature, much variability has been noticed in the context of sensitivity and specificity [44]. The reported sensitivity has ranged between 41 and 100% with a median of 87% while the reported specificity has ranged between 37 and 100% with a median of 58% [45,46]. Since most of the studies have reported only sensitivity, a general assumption has arisen that the clinical diagnosis is highly accurate. In the current study, we tested the consistency of the proposed method on the following four indices: (1) sensitivity, (2) specificity, (3) accuracy, (4) F1 score, and (5) AUC. Moreover, for a better understanding of the obtained result, we also plotted the ROC curve. Various studies have computed features based on ROI, VBM, and statistical approaches [13,45,46]. These feature extraction techniques are highly sensitive to geometric deformation and/or perturbation of MRI images under consideration. In recent years, orthogonal moments are proven to be excellent feature descriptors in the field of medical imaging [8,13,38,45,46]. These orthogonal moments have the advantage of needing lower precision to represent the difference of the same accuracy as the monomials. Another advantage of the orthogonal moments is that they are rotationally invariant naturally. Therefore, any unexpected perturbation will not affect the measurement confronting thus the inherent uncertainty.

In the first experiment, we test and compare the proposed system for AD vs. MCI. We also plot the ROC curve, as shown in Figure 2a. To understand the effect of orthogonality on feature extraction, we compare RM (non-orthogonal) texture-based features with PHTs (orthogonal) texture-based features to diagnose AD vs. MCI in T-1 MRI images. The classification of these features is done by the proposed SaDE-WNN classifier and for comparative analysis of the proposed classifier. We also classify these features with other similar classifiers available in the literature (DEWNN and Ada-DEWNN). The proposed classifier with PCT texture features obtained a high value of indices (93.7% accuracy, 86.0% sensitivity, 98.0% specificity) followed by PCET, ZM, and RM. It shows more uniform and precise results compared to DEWNN and Ada-DEWNN. The classification results using RM textural features show the lowest results for all cases. A higher value of sensitivity specifies that there are few false-negative results, and thus fewer cases of the disease are missed. On the other hand, high specificity means that there are few false positive results [14,27]. In the current research, false positives are less important than false negatives because the false-negative answer may lead the patient into a serious condition.

In the second experiment, the proposed method shows similar behavior as in the case of MCI vs. HC depicted in Table 2. The classification results based on the non-orthogonal (RM) feature sets show the lowest performance among all other methods. The classification results for PCT with Sade-DEWNN shows higher sensitivity (92.9 %), and AUC (0.98, Figure 2c) compared to ZM, PCET, and RM. The ZM texture-based features show almost the same sensitivity and accuracy as PCT. However, this difference is not statistically significant ($p > 0.05$). It could also be noted that the proposed PCT shows a slight improvement in AUC compared to ZM and PCET (Figure 2).
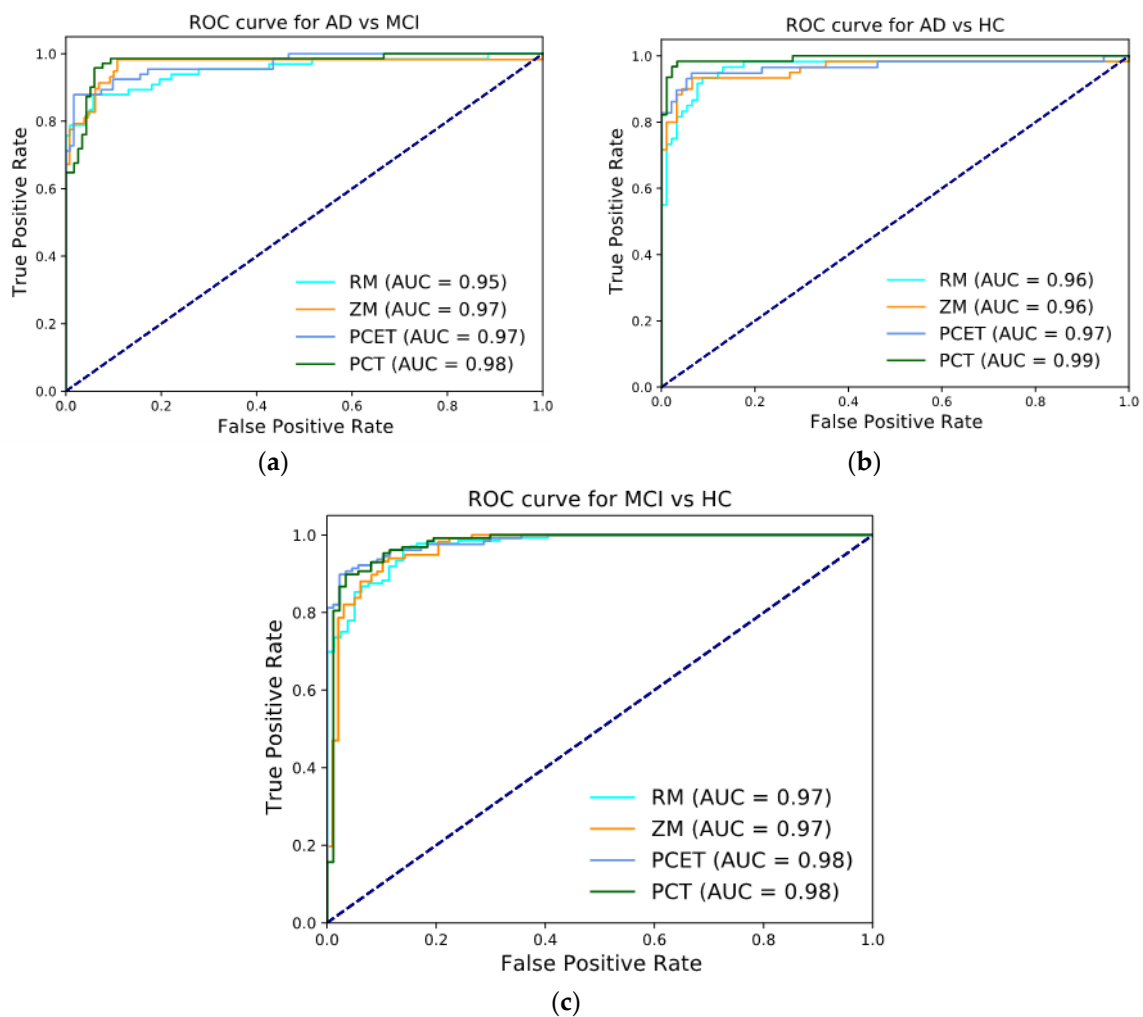
Figure 2. Receiver operating characteristics (ROC) curve for (**a**) Alzheimer's disease (AD) vs. mild cognitive impairments (MCI), (**b**) AD vs. healthy controls (HC), and (**c**) MCI vs. HC.

In contrast, in our results for MCI vs. HC and AD vs. MCI, we obtained improved results for classification of AD vs. HC for all four indices (sensitivity, specificity, accuracy, and AUC) for all feature sets. The best accuracy, sensitivity, specificity, and AUC of PCT features with SaDE-WNN classifier are 94.4%, 88.7%, 98.9%, and 0.99 respectively (Table 2 and Figure 2b). A significant improvement has been noticed in terms of sensitivity as compared to AD/MCI and MCI/HC experiments for all features, i.e., RM, PCET, ZM, and PCT. Again, RM features show the lowest performance among all the features. PCT-based features give the highest performance for all indices followed by ZM, PCET, and RM. However, no statistical significance is shown between the PCT, PCET, and ZM feature sets. For all cases, RM textural features show poor performance. This is because the orthogonal moments (PHTs) have proven superior feature representation capability and low information redundancy compared to non-orthogonal moments (RM).

We also plotted the regions which are most responsible for distinguishing between normal and disease. Figure 3 shows the most affected regions during dementia through the coronal, sagittal, and axial view. We observed that the temporal middle gyrus right, frontal middle gyrus right, precentral gyrus lefts, left middle occipital gyrus are the most affected regions during the AD progression.
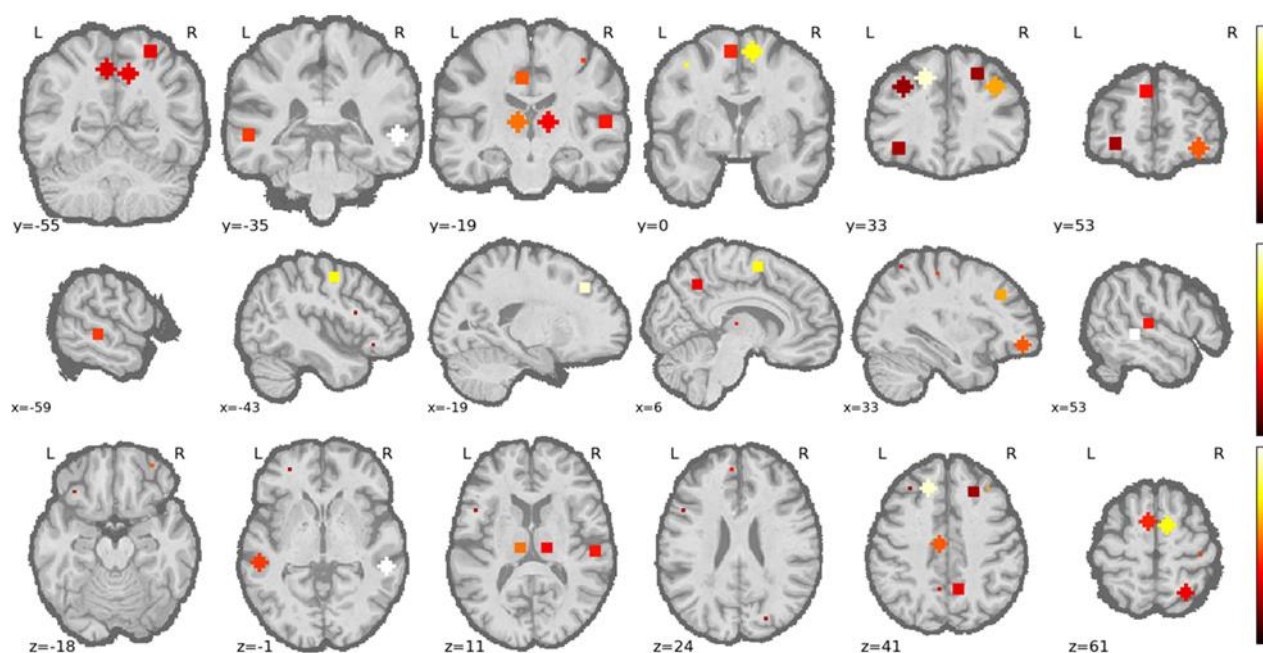
Figure 3. Brain plot for regions that are most affected during the AD progression. From top to bottom are coronal, sagittal, and axial views.

Finally, a brief comparison between the recently developed methods most related to our approach has been presented in Table 2. However, it is quite difficult to compare these methods in a precise manner since different researchers have employed different feature sets and classification methods. According to Table 2, the proposed PHT descriptors represent fair AUC, which is comparable with the other reported systems [46]. The proposed method shows lower sensitivity compared to Challis et al. [46] (100% sensitivity). However, it is superior to Challis et al. (2015) in terms of specificity, accuracy, and AUC for HC vs. MCI cases. Moreover, for MCI vs. AD cases, the proposed method shows the best performance for all indices (sensitivity, specificity, accuracy, and AUC). The feature extraction and data preparation techniques are very similar to [27] except for validation technique, segmentation, and classification scheme. However, the proposed method attains lower sensitivity, accuracy, and specificity for AD/MCI and MCI/HC compared to the results reported in [27]. For MCI/AD, the best results are reported by Khazaee et al. [45]. Our result is concordant with Yang et al. [47] for statistical feature analysis and RM.

In recent years, several research efforts have been made for the classification of AD using machine learning [48–50]. In particular, most of these machine learning approaches looked at the potential of feature extraction using MRI and diffusion tensor imaging [51]. Several other efforts also have been made in multi-modal imaging, which shows that multimodal imaging can enhance the accuracy of classification. In the literature, most of the studies are concerned with gray matter features of the brain. However, progress in diffusion tensor imaging also opening a window to the researchers to explore white matter connectivity features. However, the classification accuracy of the classifier in classifying AD using white matter ranges between 72% and 82% in the literature. Gray matter shows accuracy above 85%. Some studies also show that multimodal features (MRI, FMRI) can increase accuracy above 901% [52–54].

4. Conclusions

In the current research, a framework of computer-aided diagnosis (CAD) system using Polar Harmonic Transforms (PHT) and Self-adaptive Differential Evolution Wavelet Neural Network (SaDE-WNN) has been presented to classify dementia like Alzheimer's disease (AD), mild cognitive impairments (MCI), and healthy controls (HC). The proposed method

works in two phases in which the first one aims to extract the Region-of-Interest (ROI) area using a combination of image processing techniques. The second phase uses feature extraction using PHT and classification with the SaDE-WNN method. The features are firstly extracted by utilizing the properties of the PHT from a structural MRI database followed by a series of image pre-processing steps. The features are then selected by evaluating the in-class and among-class variance of the training database. Finally, the extracted features are classified accordingly using the proposed SaDE-WNN method.

The proposed method has been validated on the Alzheimer's disease Neuroimaging Initiative (ADNI) database. The AUC, the statistical consequence of differences between ROC index and parameters were calculated and compared by employing the z-test. The new method achieved better results than the relevant works; specifically, the overall classification results between AD and MCI, MCI and HC, and for pairs of the AD vs. HC MCI. The classification results were statistically significant between orthogonal moments (RM) and non-orthogonal moments (PHTs). It is finally concluded that the proposed method gives better results specifically in terms of sensitivity which was the main concern of the current research.

Future works in this research field will investigate the integration of Deep Learning to train the SaDE-WNN and enhance the computational cost of the proposed method in this study. In this paper, we used a single modality of images only. A combination of different modalities can increase the effectiveness of the proposed method such as T1, T2. We could also test the prosed method using GM and white matter or using a combination of MRI and fMRI.

References

1. Shen, D.; Wee, C.Y.; Zhang, D.; Zhou, L.; Yap, P.T. Machine learning techniques for AD/MCI diagnosis and prognosis. *Intell. Syst. Ref. Libr.* **2014**, *56*, 147–179. [CrossRef]
2. Chaddad, A.; Desrosiers, C.; Niazi, T. Deep Radiomic Analysis of MRI Related to Alzheimer's Disease. *IEEE Access* **2018**, *6*, 58213–58221. [CrossRef]
3. Costanza, R.; Daly, L.; Fioramonti, L.; Giovannini, E.; Kubiszewski, I.; Mortensen, L.F.; Pickett, K.E.; Ragnarsdottir, K.V.; De Vogli, R.; Wilkinson, R. Modelling and measuring sustainable wellbeing in connection with the UN Sustainable Development Goals. *Ecol. Econ.* **2016**, *130*, 350–355. [CrossRef]
4. Curado, M.; Escolano, F.; Lozano, M.A.; Hancock, E.R. Early detection of Alzheimer's disease: Detecting asymmetries with a return random walk link predictor. *Entropy* **2020**, *22*, 465. [CrossRef] [PubMed]
5. Plocharski, M.; Østergaard, L.R. Extraction of sulcal medial surface and classification of Alzheimer's disease using sulcal features. *Comput. Methods Programs Biomed.* **2016**, *133*, 35–44. [CrossRef]
6. Miao, Y.; Jiang, H.; Liu, H.; Yao, Y. dong An Alzheimers disease related genes identification method based on multiple classifier integration. *Comput. Methods Programs Biomed.* **2017**, *150*, 107–115. [CrossRef]
7. Previtali, F.; Bertolazzi, P.; Felici, G.; Weitschek, E. A novel method and software for automatically classifying Alzheimer's disease patients by magnetic resonance imaging analysis. *Comput. Methods Programs Biomed.* **2017**, *143*, 89–95. [CrossRef]

8. Singh, S.P.; Urooj, S.; Lay-Ekuakille, A. Breast Cancer Detection Using PCPCET and ADEWNN: A Geometric Invariant Approach to Medical X-Ray Image Sensors. *IEEE Sens. J.* **2016**, *16*, 4847–4855. [CrossRef]

9. Kim, J.; Kim, Y.-K. Crosstalk between Depression and Dementia with Resting-State fMRI Studies and Its Relationship with Cognitive Functioning. *Biomedicines* **2021**, *9*, 82. [CrossRef]

10. Dachena, C.; Casu, S.; Fanti, A.; Lodi, M.B.; Mazzarella, G. Combined use of MRI, fMRIand cognitive data for Alzheimer's Disease: Preliminary results. *Appl. Sci.* **2019**, *9*, 3156. [CrossRef]

11. Frisoni, G.B.; Fox, N.C.; Jack, C.R.; Scheltens, P.; Thompson, P.M. The clinical use of structural MRI in Alzheimer disease. *Nat. Rev. Neurol.* **2010**, *6*, 67–77. [CrossRef]

12. Singh, S.P.; Urooj, S. Wavelet packets based spectral estimation of textured images. In Proceedings of the 2015 IEEE International Conference on Computational Intelligence and Communication Technology, Ghaziabad, India, 13–14 February 2015; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2015; pp. 651–654.

13. Eskildsen, S.F.; Coupé, P.; Fonov, V.S.; Pruessner, J.C.; Collins, D.L. Structural imaging biomarkers of Alzheimer's disease: Predicting disease progression. *Neurobiol. Aging* **2015**, *36*, S23–S31. [CrossRef]

14. Jiji, G.W.; Suji, G.E.; Rangini, M. An intelligent technique for detecting Alzheimer's disease based on brain structural changes and hippocampal shape. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2014**, *2*, 121–128. [CrossRef]

15. Ji, X.; Wang, H.; Zhu, M.; He, Y.; Zhang, H.; Chen, X.; Gao, W.; Fu, Y. Brainstem atrophy in the early stage of Alzheimer's disease: A voxel-based morphometry study. *Brain Imaging Behav.* **2020**, 1–11. [CrossRef]

16. Raczek, M.; Cercignani, M.; Banerjee, S. Voxel-based morphometry and resting state fMRI as predictors of neuropsychiatric symptoms in Alzheimer's disease. *Alzheimer's Dement.* **2020**, *16*. [CrossRef]

17. Wu, J.; Simon, M.A.; Brigham, J.C. A comparative analysis of global shape analysis methods for the assessment of the human right ventricle. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2016**, *4*, 327–343. [CrossRef]

18. López, M.; Ramírez, J.; Górriz, J.M.; Álvarez, I.; Salas-Gonzalez, D.; Segovia, F.; Chaves, R.; Padilla, P.; Gómez-Río, M. Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease. *Neurocomputing* **2011**, *74*, 1260–1271. [CrossRef]

19. Nemmi, F.; Sabatini, U.; Rascol, O.; Péran, P. Parkinson's disease and local atrophy in subcortical nuclei: Insight from shape analysis. *Neurobiol. Aging* **2015**, *36*, 424–433. [CrossRef] [PubMed]

20. Cury, C.; Glaunès, J.; Chupin, M.; Colliot, O. Analysis of anatomical variability using diffeomorphic iterative centroid in patients with Alzheimer's disease. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2017**, *5*, 350–358. [CrossRef]

21. Khedher, L.; Illán, I.A.; Górriz, J.M.; Ramírez, J.; Brahim, A.; Meyer-Baese, A. Independent Component Analysis-Support Vector Machine-Based Computer-Aided Diagnosis System for Alzheimer's with Visual Support. *Int. J. Neural Syst.* **2016**, *27*, 1650050. [CrossRef] [PubMed]

22. Zhang, Y.; Wang, S.; Sui, Y.; Yang, M.; Liu, B.; Cheng, H.; Sun, J.; Jia, W.; Phillips, P.; Gorriz, J.M. Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. *J. Alzheimer's Dis.* **2018**, *65*, 855–869. [CrossRef] [PubMed]

23. Ortiz, A.; Munilla, J.; Górriz, J.M.; Ramírez, J. Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer's Disease. *Int. J. Neural Syst.* **2016**, *26*, 1650025. [CrossRef] [PubMed]

24. Singh, S.P.; Urooj, S. An Improved CAD System for Breast Cancer Diagnosis Based on Generalized Pseudo-Zernike Moment and Ada-DEWNN Classifier. *J. Med. Syst.* **2016**, *40*, 1–13. [CrossRef]

25. Singh, S.P.; Urooj, S. Three Types of Moment Invariants for Color Object Recognition Based on Radon and Polar Harmonic Transform in Cℓ (0, 2) Space. *Arab. J. Sci. Eng.* **2016**, *41*, 3051–3060. [CrossRef]

26. Farokhi, S.; Sheikh, U.U.; Flusser, J.; Yang, B. Near infrared face recognition using Zernike moments and Hermite kernels. *Inf. Sci. (Ny).* **2015**, *316*, 234–245. [CrossRef]

27. Gorji, H.T.; Haddadnia, J. A novel method for early diagnosis of Alzheimer's disease based on pseudo Zernike moment from structural MRI. *Neuroscience* **2015**, *305*, 361–371. [CrossRef]

28. Yap, P.T.; Jiang, X.; Chichung Kot, A. Two-dimensional polar harmonic transforms for invariant image representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1259–1270. [CrossRef]

29. Li, L.; Li, S.; Abraham, A.; Pan, J.S. Geometrically invariant image watermarking using Polar Harmonic Transforms. *Inf. Sci. (Ny).* **2012**, *199*, 1–19. [CrossRef]

30. Li, Y. Image copy-move forgery detection based on polar cosine transform and approximate nearest neighbor searching. *Forensic Sci. Int.* **2013**, *224*, 59–67. [CrossRef]

31. Singh, S.P.; Urooj, S. Localized Radon Polar Harmonic Transform (LRPHT) Based Rotation Invariant Analysis of Textured Images. *Int. J. Syst. Dyn. Appl.* **2015**, *4*, 21–41. [CrossRef]

32. Urooj, S.; Singh, S.P. Rotation invariant detection of benign and malignant masses using PHT. In Proceedings of the 2015 International Conference on Computing for Sustainable Global Development, INDIACom 2015, New Delhi, India, 11–13 March 2015; pp. 1627–1632.

33. Lv, J.; Jiang, X.; Li, X.; Zhu, D.; Chen, H.; Zhang, T.; Zhang, S.; Hu, X.; Han, J.; Huang, H.; et al. Sparse representation of whole-brain fMRI signals for identification of functional networks. *Med. Image Anal.* **2015**, *20*, 112–134. [CrossRef]

34. Chauhan, N.; Ravi, V.; Karthik Chandra, D. Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks. *Expert Syst. Appl.* **2009**, *36*, 7659–7665. [CrossRef]

35. Dheeba, J.; Albert Singh, N.; Tamil Selvi, S. Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. *J. Biomed. Inform.* **2014**, *49*, 45–52. [CrossRef]

36. Omran, M.G.H.; Salman, A.; Engelbrecht, A.P. Self-adaptive differential evolution. In *Lecture Notes in Computer Science*; including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3801 LNAI, pp. 192–199.

37. Jack, C.R., Jr.; Bernstein, M.A.; Fox, N.C.; Thompson, P.; Alexander, G.; Harvey, D.; Borowski, B.; Britson, P.J.; Whitwell, J.L.; Ward, C.; et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* **2008**, *27*, 685–691. [CrossRef] [PubMed]

38. Teh, C.H.; Chin, R.T. On image analysis by the method of moments. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *10*, 496–513. [CrossRef]

39. Urooj, S.; Singh, S.P. Geometric invariant feature extraction of medical images using Hu's invariants. In Proceedings of the 3rd International Conference on Computing for Sustainable Global Development, INDIACom 2016, New Delhi, India, 16–18 March 2016; pp. 1560–1562.

40. Dabbaghchian, S.; Ghaemmaghami, M.P.; Aghagolzadeh, A. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology. *Pattern Recognit.* **2010**, *43*, 1431–1440. [CrossRef]

41. Dheeba, J.; Selvi, S.T. A swarm optimized neural network system for classification of microcalcification in mammograms. *J. Med. Syst.* **2012**, *36*, 3051–3061. [CrossRef]

42. Zhang, Q.; Benveniste, A. Wavelet Networks. *IEEE Trans. Neural Networks* **1992**, *3*, 889–898. [CrossRef] [PubMed]

43. Brest, J.; Greiner, S.; Bošković, B.; Mernik, M.; Zumer, V. Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems. *IEEE Trans. Evol. Comput.* **2006**, *10*, 646–657. [CrossRef]

44. Beach, T.G.; Monsell, S.E.; Phillips, L.E.; Kukull, W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. *J. Neuropathol. Exp. Neurol.* **2012**, *71*, 266–273. [CrossRef] [PubMed]

45. Khazaee, A.; Ebrahimzadeh, A.; Babajani-Feremi, A. Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease. *Brain Imaging Behav.* **2016**, *10*, 799–817. [CrossRef] [PubMed]

46. Challis, E.; Hurley, P.; Serra, L.; Bozzali, M.; Oliver, S.; Cercignani, M. Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI. *Neuroimage* **2015**, *112*, 232–243. [CrossRef] [PubMed]

47. Yang, S.-T.; Lee, J.-D.; Chang, T.-C.; Huang, C.-H.; Wang, J.-J.; Hsu, W.-C.; Chan, H.-L.; Wai, Y.-Y.; Li, K.-Y. Discrimination between Alzheimer's Disease and Mild Cognitive Impairment Using SOM and PSO-SVM. *Comput. Math. Methods Med.* **2013**, *2013*, 1–10. [CrossRef]

48. Nahata, H.; Singh, S.P. Deep Learning Solutions for Skin Cancer Detection and Diagnosis. In *Machine Learning with Health Care Perspective*; Springer: Cham, Switzerland, 2020; pp. 159–182.

49. Singh, S.P.; Wang, L.; Gupta, S.; Goli, H.; Padmanabhan, P.; Gulyás, B. 3d deep learning on medical images: A review. *Sensors* **2020**, *20*, 5097. [CrossRef]

50. Singh, S.P.; Wang, L.; Gupta, S.; Gulyas, B.; Padmanabhan, P. Shallow 3D CNN for Detecting Acute Brain Hemorrhage from Medical Imaging Sensors. *IEEE Sens. J.* **2020**. [CrossRef]

51. Tucholka, A.; Grau-Rivera, O.; Falcon, C.; Rami, L.; Sánchez-Valle, R.; Lladó, A.; Gispert, J.D.; Molinuevo, J.L. Structural connectivity alterations along the Alzheimer's disease continuum: Reproducibility across two independent samples and correlation with cerebrospinal fluid amyloid-β and tau. *J. Alzheimer's Dis.* **2018**, *61*, 1575–1587. [CrossRef]

52. Nemy, M.; Grothe, M.; Barroso, J.; Teipel, S.J.; Westman, E.; Ferreira, D. Multimodal imaging reveals human cholinergic system functional and structural integrity in vivo. *Alzheimer's Dement.* **2020**, *16*. [CrossRef]

53. Herdick, M.; Dyrba, M.; Fritz, H.C.J.; Altenstein, S.; Ballarini, T.; Brosseron, F.; Buerger, K.; Can Cetindag, A.; Dechent, P.; Dobisch, L.; et al. Multimodal MRI analysis of basal forebrain structure and function across the Alzheimer's disease spectrum. *NeuroImage Clin.* **2020**, *28*, 102495. [CrossRef] [PubMed]

54. Bi, X.A.; Hu, X.; Wu, H.; Wang, Y. Multimodal Data Analysis of Alzheimer's Disease Based on Clustering Evolutionary Random Forest. *IEEE J. Biomed. Heal. Inform.* **2020**, *24*, 2973–2983. [CrossRef]