Kernel-Ridge-Regression-Based Randomized Network for Brain Age Classification and Estimation

Raveendra Pilli , Tripti Goel , *Senior Member, IEEE*, R. Murugan , *Member, IEEE*,
M. Tanveer , *Senior Member, IEEE*, and P. N. Suganthan , *Fellow, IEEE*

Abstract—**Accelerated brain aging and abnormalities are associated with variations in brain patterns. Effective and reliable assessment methods are required to accurately classify and estimate brain age. In this study, a brain age classification and estimation framework is proposed using structural magnetic resonance imaging (sMRI) scans, a 3-D convolutional neural network (3-D-CNN), and a kernel ridge regression-based random vector functional link (KRR-RVFL) network. We used 480 brain MRI images from the publicly available IXI database and segmented them into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) images to show age-related associations by region. Features from MRI images are extracted using 3-D-CNN and fed into the wavelet KRR-RVFL network for brain age classification and prediction. The proposed algorithm achieved high classification accuracy, 97.22%, 99.31%, and 95.83% for GM, WM, and CSF regions, respectively. Moreover, the proposed algorithm demonstrated excellent prediction accuracy with a mean absolute error (MAE) of 3.89 years, 3.64 years, and 4.49 years for GM, WM, and CSF regions, confirming that changes in WM volume are significantly associated with normal brain aging. Additionally, voxel-based morphometry (VBM) examines age-related anatomical alterations in different brain regions in GM, WM, and CSF tissue volumes.**

Index Terms—**Cerebrospinal fluid (CSF), gray matter (GM), kernel ridge regression-random vector functional link (KRR-RVFL), magnetic resonance imaging (MRI), white matter (WM).**

I. INTRODUCTION

AGING is a dynamic biological mechanism from birth to old age, during which the cells in the body grow and die. The human brain is known to undergo consistent changes throughout an individual's whole lifetime. The human brain grows old and eventually stops functioning with increasing age. The brain aging process is challenging to understand because of the different accumulated damage rates of the brain, and sometimes brain aging is accelerated and disrupted by internal pathological changes [3]. Various neurodegenerative disorders like Alzheimer's disease (AD), Parkinson's disease (PD), and cognitive decline are difficult to foresee and detect early on. Understanding the risk of age-related degeneration and how it affects people is critical to offer appropriate care and treatment. Brain age estimation entails training a deep learning (DL) or machine learning (ML) model to detect brain aging patterns using many brain magnetic resonance imaging (MRI) images. Once trained, the model can estimate brain age on new test samples. The disparity between the predicted brain age and chronological age (actual age) is called the brain age gap (BAG). It can be used as a biomarker for general health status and early diagnosis of neurodegenerative ailments. A low BAG score implies normal aging, but a high BAG value shows abnormal aging, indicating a higher risk of age-related neurological illnesses. The human brain has three important compartments: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). GM is occupied in the brain cerebrum, cerebellum, brain stem, and spinal cord of the brain regions. GM consists of neuronal cell bodies, dendrites, and unmyelinated axons. GM is essential for various functions such as motor control, memory preservation, and emotion management. Age-related declines in GM volume occur annually at a rate of around 0.5%, and some brain regions experience more deterioration than others [20]. WM is made up of myelinated axons that link different GM regions of the brain and carry nerve impulses between neurons. Axons are long connecting fibers covered in a fatty material called the myelin sheath, which acts like an insulator, allowing electrochemical impulses to be transmitted efficiently across various areas. WM volume and myelinated axons length are declined with aging. WM volume reaches its peak at the age of 40–50, after which it progressively declines [24]. CSF is a clear, colorless body fluid like plasma in the

brain and spinal cord. The brain generates around 500 mL of CSF daily at a rate of 25 mL per hour. CSF provides various functions, such as protecting the brain from injury, removing the brain's metabolic waste, and preventing brain ischemia. The volume of CSF also varies with normal brain aging [10]. On a global level, there is a negative linear relationship between GM and age and a nonlinear relationship between WM, CSF, and age. On a regional level, linear and nonlinear age-related volumetric variations in various regions of the brain with the aging process [2].

With the emergence of neuroimaging, the structural magnetic resonance imaging (sMRI) modality has been extensively utilized to examine structural changes of the brain during aging and provide insight into the aging of the brain and degeneration. As a result, it is essential to classify and estimate an individual's brain age using high-dimensional morphological changes from sMRI. Convolutional neural networks (CNN), which are frequently employed in image classification and prediction challenges, have shown great promise in visual feature extraction. CNNs' excellent learning ability and automated decision-making pipelines make them perfect for brain age categorization and prediction, which may enhance medical consultation, assessment, and decision-making [33]. In the present work, a 3-D-CNN network is employed to extract the robust features from the MRI scans.

The conventional CNN networks use gradient-based learning methods, which result in slow convergence rates and the chance of being stuck in local minima. A randomized version of the single-layer feed-forward neural network (SLFN), called as random vector functional link (RVFL) network, has been implemented to overcome these constraints. In RVFL, hidden layer weights and biases are randomly initialized and remain constant throughout the network training [29]. The traditional RVFL is renowned for its fast learning rate, simple design, and good generalization abilities. It provides direct connections between input and output nodes and nonlinear hidden nodes with random weights. A feed-forward RVFL network overcomes the limitations of conventional DL algorithms and also provides better accuracy with faster speed. The input features are directly fed to the output layer, which generates outputs employing a linear decision function. In this study, a kernel trick is applied to a standard RVFL network, which transforms the input space into higher dimensional feature space while supplying input features to the RVFL output layer through direct links [4], [38]. As a result, the output layer of the kernel ridge regression-based RVFL (KRR-RVFL) network gets complete nonlinearity by using the synergies of nonlinear activation functions like radial basis (RadBas), triangular basis (TriBas), sigmoid, and kernel function. The KRR-RVFL network benefits from both kernel trick and nonlinear activation functions and provides better performance of classification and regression tasks. In this present work, a kernel ridge regression-based RVFL model is used for the brain age classification and estimation of healthy individuals. This study makes the following contributions:

1) All MRI images are segmented into GM, WM, and CSF tissues. Experiments are conducted on preprocessed GM, WM, and CSF datasets to show their regionwise association with age.

2) Features are extracted using a 3-D-CNN deep network from GM, WM, and CSF datasets.

3) The kernel trick is applied to the input features to get the nonlinearity between the input and output of standard RVFL to improve the efficiency of the classification and regression framework.

4) The association between brain tissues and the natural aging process examined by the suggested deep model.

5) The classification and regression performance metrics of the wavelet KRR-RVFL network are compared with state-of-the-art classifiers and regression networks.

6) Voxel-based morphometry (VBM) analysis is used to observe the age-related anatomical variations in GM, WM, and CSF brain compartments.

The rest of the article is organized as follows: In Section II Related works are discussed. Dataset as well as methodology are presented in Section III. In Section IV results and discussion are illustrated, and in Section V the article is concluded.

II. RELATED WORKS

Atrophy is the term used to indicate tissue volume loss resulting from cellular loss caused by degeneration or other mechanisms, like pruning. The human brain undergoes atrophy, and brain shrinking has been broadly reported in several studies [34]. The sMRI-derived parameters usually utilized to represent aging can be roughly separated into a measure of GM, WM, and CSF, which together form total brain volume [14]. Resnick et al. [23] conducted research with 92 older adults to find out the distribution and rate of WM and GM loss in elderly persons. They discovered that the tissues in the frontal and parietal regions declined more quickly than tissues in the occipital lobar and temporal areas, and WM tissue is extensively decreased. GM loss in inferior frontal and orbital, cingulate, insular, inferior, and parietal areas, and considerable loss in temporal region. In contrast, [35] examined healthy young people aged 18 to 50. Their findings revealed considerable GM loss in right prefrontal cortex and left cerebellum but no GM loss in the medial temporal area, cingulate gyrus, insular, and temporal neocortex with aging. CSF volume is larger in the elderly than in the young, and CSF-filled ventricles seem to increase quadratically with time, with relative stability until middle adulthood and rapid expansion after that [7]. Liu et al. [13] demonstrated that the volume loss is more apparent in WM than in GM, particularly in the corpus callosum and frontal lobe, and discovered a 26% reduction in WM tissue volume and 14% reduction in GM tissue volume.

Ensemble DL integrates several networks to improve the performance of classification and prediction of the models. Hofmann et al. [8] integrated ensembles of CNN with layer-wise relevance propagation (LRP) algorithm to detect relevant features to predict the brain age. They discovered age-related variations throughout the brain and enlargements of ventricles, WM lesions as well as accelerating aging effect in the frontal lobe. Kuo et al. [12] proposed an ensemble DL brain age prediction model, they first used conventional ML approaches to extract the robust input features, and secondly, the obtained input features fed to ensemble DL model using 6-layer 3-D

residual neural network. The authors achieved higher prediction accuracy: mean absolute error (MAE) = 3.77 years; $R^2 = 0.90$ with multiple concurrent input features. Siar and Teshnehlab [31] implemented an AlexNet-based CNN architecture for brain age classification. The age groups are separated into five categories ranging from 10 to 70 years. Softmax, support vector machine (SVM), and decision trees (DT) classification layers are used in this model, and with the Softmax classification layer the model achieved 79% accuracy. Zhang et al. [39] introduced a novel sparse pretrained RVFL method and they demonstrated the advantage of adopting RVFL across 16 different datasets, and the associated results show that RVFL outperforms another noniterative classifier. In [19] authors implemented the multiclass brain abnormalities classification framework using fast curvelet Tsallis entropy for feature extraction from MRI scans and kernel RVFL as a classifier to classify the five classes: normal, brain tumor, infections, degenerative, and stroke. They used two multiclass datasets, dataset-1 having 200 scans, out of which 40 scans are normal and 160 are abnormal scans, and another dataset-2 having 15 images of each class and a total of 75 scans. The classification accuracy of dataset-1 is 97.33% and for dataset-2 is 93%.

VBM methods have been developed to analyze and quantify volumetric variations in the normal human brain using sMRI. Farokhian et al. [6] used the VBM strategy to investigate regional and global brain GM and WM tissue volume differences between healthy young and older individuals. GM volume declined at the regional level at the frontal, insular, and cingulate cortices regions. WM volume declined at thalamic radiations and increased at occipital and pericentral regions in older subjects compared to young subjects. On a global level, they examined the effects of the aging process on GM, WM, and the total intracranial volume (TIV) of healthy individuals. Spindler et al. [32] detected significant WM variations using VBM in patients with alcohol use disorder (AUD). 416 control normal (CN) and 462 AUD patients are employed for this experiment. This study revealed four significant clusters of WM changes in AUD patients that are anterior and posterior cingulum, corpus callosum, fornix, and internal capsule right posterior limb. In [11] Joy et al. identified the structural GM and WM changes in HIV patients using the VBM technique and surface-based morphometry (SBM) analysis. In this study, 27 HIV-infected individuals and 15 HIV-unaffected CNs are considered to examine the brain's structural changes. Compared to CNs, this study found a link between altered GM and WM volumes in various brain areas of HIV-infected persons, particularly in the middle frontal and parahippocampal regions.

This section overviews important work done in the relevant field to insight the motivation and knowledge.

III. METHODOLOGY

The following section explains the dataset used in the experiments and the proposed brain age classification framework.

A. Dataset

For experiments, the data has been acquired from the publicly available IXI (http://www.brain-development.org/ixi-dataset/)

TABLE I
DEMOGRAPHICAL INFORMATION OF THE DATASET

Class	No. of Samples	Mean Age	SD	TR	TE
Ages (20–40)	160	30.13	5.79	9 Sec	4 Sec
Ages (41–60)	160	50.32	6.14	9 Sec	4 Sec
Ages (61–90)	160	76.24	8.96	9 Sec	4 Sec

Note: SD: standard deviation, TR: repetition time, TE: echo time

database. A total of 480 CN subjects are used in this study. MRI scans are captured using a 3T Phillips scanner with a slice thickness of 1.2 mm. Each MRI image dimension is $256 \times 256 \times 150$. The demographical information of the subjects is illustrated in Table I.

B. Preprocessing

All T1-weighted MRI images are preprocessed by utilizing MATLAB-based statistical parametric mapping-12 toolbox (SPM12) (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), computational anatomy toolbox (CAT)-12 toolbox (http://dbm.neuo.uni-jena.de) and followed by a specific parameter set. First, the MRI scans are segmented into three main components: GM, WM, and CSF. MRI preprocessing includes bias-field distortions correction, removal of nonbrain tissue, and normalization to standard Montreal Neurological Institute (MNI) space by diffeomorphic registration algorithm and then modulated. The pictures of GM, WM, and CSF are smoothed using an 8-mm full-width half maximum (FWHM) Gaussian kernel. The architecture of the proposed KRR-RVFL-based model is illustrated in Fig. 1.

C. Feature Extraction

We used a 3-D-CNN as the backbone to extract deep features from input MRI images. CNN consists of three main layers. The proposed 3-D-CNN contains six blocks. Each block consists of a convolutional layer (CL), batch normalization layer, and rectified linear unit (ReLU) activation layer. A maxpooling layer (MPL) with a kernel size of 2×2 and stride of 2 is applied after every block to reduce the spatial dimensions gradually. The channel numbers used in the blocks are [32, 32, 16, 32, 16, 8].

D. KRR-RVFL

RVFL is a randomized neural network, and its graphical representation as shown in Fig. 2. RVFL eliminates the drawbacks of back propagation-based classifiers, such as slow convergence, overfitting, and being trapped in a local minimum. In RVFL randomness occurs between the input layer and hidden layer, weights w and biases b are randomly initialized with uniform distribution from the sets [0, 1] and [−1, 1], respectively, and remain the same during training and testing. From the hidden layer to the output layer, the output weights α are calculated analytically. RVFL's output layer is supplied with X original input features and G hidden layer output features, which can be written as $D = [G\ X]$. X and G are as shown as

$$X = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{N1} & \dots & x_{Nn} \end{bmatrix}, \quad G = \begin{bmatrix} g_1(x_1) & \dots & g_L(x_1) \\ \dots & \dots & \dots \\ g_1(x_N) & \dots & g_L(x_N) \end{bmatrix}$$
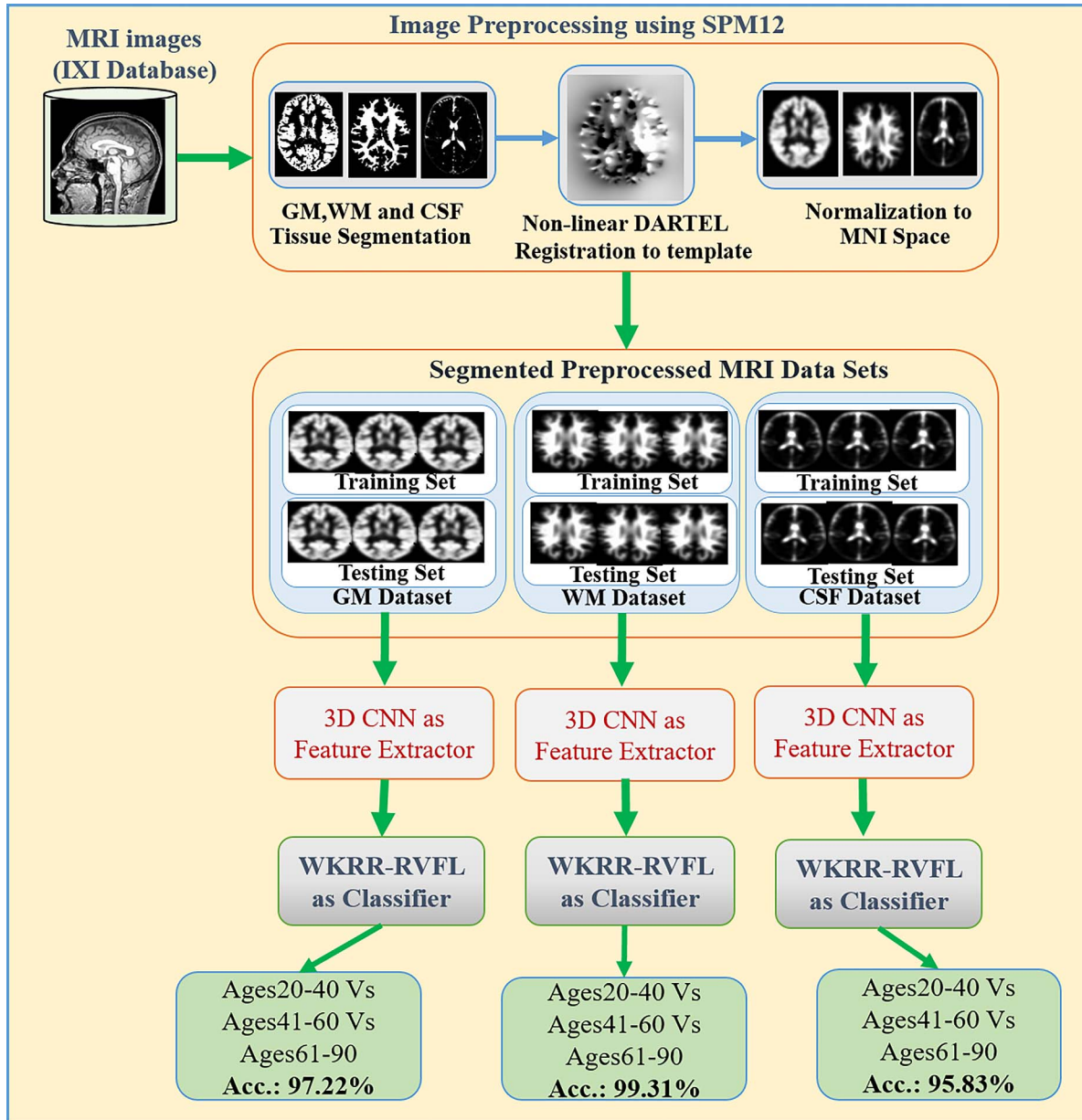
Fig. 1. Architecture of the proposed wavelet KRR-RVFL-based deep model.
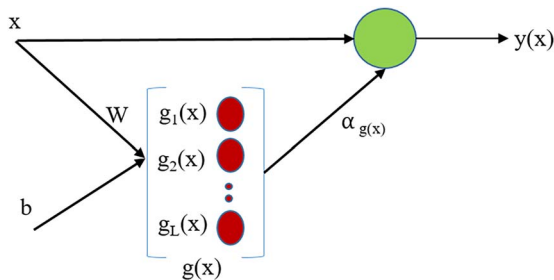
Fig. 2. Graphical illustration of RVFL.

Here, $g_1(x_1) = h(w_1.x_1 + b_1)$ and so on, $h(.)$ is nonlinear activation function. If x are the input features and L are the hidden nodes, then $(x + L)$ features are supplied to output layer. Mathematical expression of the optimization function is

$$\min_{\alpha} \|D\alpha - O\|^2 + \eta \|\alpha\|^2 \tag{1}$$

where $\eta = 1/C$ is regularization parameter and O is output target. The above equation can be computed by using ridge regression (i.e., $\eta \neq 0$) or Moore-Penrose pseudoinverse (i.e., $\eta = 0$). If $\eta = 0$, then $\alpha = D^+O$ and $\eta \neq 0$ then

$$\alpha = (D^T D + \eta I)^{-1} D^T O \tag{2}$$

$$\alpha = D^T \left(DD^T + \frac{I}{C} \right)^{-1} O. \tag{3}$$

The above equations are primal and dual solutions, respectively.

In standard RVFL, input features are directly applied to the output layer, outputs generated using a linear decision bounds,

Fig. 3. Graphical representation of KRR-RVFL.

and there is a possibility of instability in the classification. To overcome this, kernel trick is used in RVFL network, which transforms the input space. The idea of kernelized RVFL is shown in Fig. 3.

The matrix D is written as $D = [G\ \phi(X)]$ and $\phi(X)$ is the mapping function that maps input space X into a higher dimension. The (3) can be represented as

$$\alpha = \begin{bmatrix} G & \phi(X) \end{bmatrix}^T \left(\begin{bmatrix} G & \phi(X) \end{bmatrix} \begin{bmatrix} G^T \\ \phi(X)^T \end{bmatrix} + \frac{I}{C} \right)^{-1} O \quad (4)$$

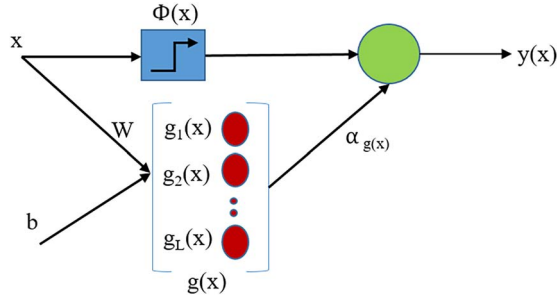$$\alpha = \begin{bmatrix} G & \phi(X) \end{bmatrix}^T \left(GG^T + \phi(X)\phi(X)^T + \frac{I}{C} \right)^{-1} O. \quad (5)$$

Now, applying the kernel trick to (5) and representing K, as follows:

$$K = \phi(X)\phi(X)^T = \begin{bmatrix} k(x_1,x_1) & . & . & k(x_1,x_n) \\ . & . & . & . \\ . & . & . & . \\ k(x_n,x_1) & . & . & k(x_n,x_n) \end{bmatrix}. \quad (6)$$

The Eq. 5 can be expressed as

$$\alpha = \begin{bmatrix} G & \phi(X) \end{bmatrix}^T \left(GG^T + K + \frac{I}{C} \right)^{-1} O. \quad (7)$$

The final output expression can be written as

$$y(x) = [g(x)^T G^T + \phi(x)^T \phi(X)^T] \left[GG^T + K + \frac{I}{C} \right]^{-1} O \quad (8)$$

$$y(x) = [g(x)^T G^T + k(x^T, X^T)] \left[GG^T + K + \frac{I}{C} \right]^{-1} O \quad (9)$$

where $k = x^T X^T$.

E. Wavelet Analysis and Wavelet Kernel

Wavelet analysis outperforms the Fourier transform (FT) by providing exact time-frequency analysis. Wavelet kernels of Morlet or Mexican Hat types are widely recognized for their generalization and strong function-fitting capabilities, and they can be termed as local kernels. Wavelet analysis approximates a function via dilations and translations of mother wavelet $\phi(x)$

$$\phi_{a,m}(x) = |a|^{-1/2}\ \phi\left(\frac{x-m}{a} \right) \quad (10)$$

where $a, m \in R$, a is dilation factor, and m is the factor of translation. The expression for wavelet transform function $f(x)$ is

$$W_{a,m}(f) = < f(x), \phi_{a,m}(x) >. \quad (11)$$

From (11), $f(x)$ is decomposed on wavelet basis $\phi_{a,m}(x)$. The mother wavelet function $\phi(x)$ must satisfy the following condition

$$W_\phi = \int_0^\infty \frac{|\phi(w)|^2}{|w|} dw < \infty \quad (12)$$

where $\phi(w)$ is $FT(\phi(x))$. The function $f(x)$ can be reconstructed as

$$f(x) = \frac{1}{W_\phi} \int_{-\infty}^\infty \int_0^\infty W_{a,m}(f)\phi_{a,m}(x) da dm. \quad (13)$$

The approximation of $f(x)$ is expressed as

$$\tilde{f}(x) = \sum_{i=1}^l W_i \phi_{a_i,m_i}(x). \quad (14)$$

A basic multidimensional wavelet function can be written as a combination of one dimensional wavelet functions

$$\phi(x) = \prod_{i=1}^N \phi(x_i). \quad (15)$$

If $\phi(x)$ is mother wavelet kernel, and if $x, x' \in R^N$, then the wavelet kernels are

$$K(x,x') = \prod_{i=1}^N \phi\left(\frac{x_i - m_i}{a} \right) \phi\left(\frac{x_i' - m_i'}{a} \right) \quad (16)$$

and the transnational invariant wavelet kernel function is

$$K(x,x') = \prod_{i=1}^N \phi\left(\frac{x_i - x_i'}{a} \right). \quad (17)$$

The mother wavelet of Morlet wavelet function is

$$\phi(x) = \cos(b * x)\exp\left(-\frac{x^2}{2} \right) \quad (18)$$

then Morlet wavelet kernel function is

$$K(x,x') = \prod_{i=1}^N \left(\cos\left(\frac{b\,\|(x_i - x_i')\|}{a} \right) \exp\left(\frac{-\,\|(x_i - x_i')\|^2}{e} \right) \right). \quad (19)$$

Kernel-based classification approaches capture the nonlinearity of data characteristics in the feature space and are often sensitive to the selection of parameters.

Overall, the 3-D-CNN network extracted all levels of features from brain tissue components. After, the extracted features are applied to the wavelet KRR-RVFL (WKRR-RVFL) deep model, and classification is done for brain aging.
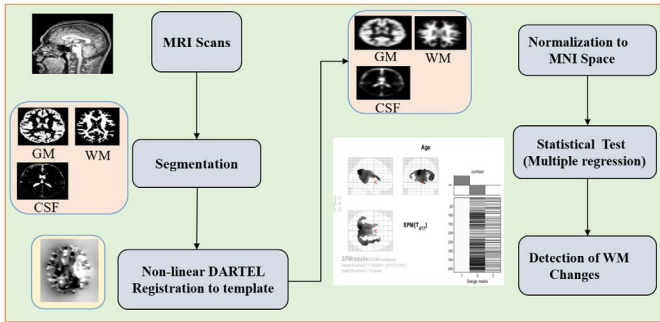
Fig. 4. Architecture of VBM framework.

F. Voxel-Based Morphometry

VBM is a whole-brain unbiased, objective technique developed to assess brain differences using sMRI images. VBM seeks to identify variations in the local composition of brain tissue while overlooking significant scale variances in physical anatomy and location. This is achieved by segmenting the MRI images into GM, WM, and CSF compartments and then spatially normalizing all the MRI images to the MNI space, smoothing the GM, WM, and CSF scans, and finally performing a statistical analysis to identify significant variations in two or more experimental groups. In this study, TIV is used as a global value for estimating the GM, WM, and CSF map volumes, while age and gender are used as covariates. The multiple regression test is then performed using familywise error (FWE) correction and $p < 0.05$ threshold. The 0.2 voxel extent threshold is selected and finally, MATLAB xjview toolbox is utilized to record and observe voxel brain region (visualized with pseudo color), with informative variations, activation volume (cluster), activation intensity (assessed statistically using multiple regression test, and presented as a T-value, where T-value increases as the intensity increases). Fig. 4 shows the VBM analysis processing framework.

IV. RESULTS AND DISCUSSION

A. Implementation Details

The assessments are carried out on a machine equipped with MATLAB R2021a, an Intel(R) i7-8700 CPU-3.20GHz, 16GB RAM, and Windows 10. The preprocessed GM, WM, and CSF datasets are used in this experiment. The samples from each dataset are randomly partitioned into 70:30 ratios. For 3-D-CNN the training parameters are taken as follows: minibatch size (16), learning rate (0.01), optimizer (sgmd), and epochs (10). The hyperparameter values for the experiments are taken as $b = 3$, $a = 2$, $e = 2a^2$, regularization parameter $\eta = 1/C$, C is chosen from the range 2^z, where $z = \{-6, -4, \ldots, 12\}$, and RadBas activation function and wavelet kernel function are incorporated in RVFL network.

B. Experiments

This segment describes the performance measures used to assess the efficacy of classification and regression algorithms.

TABLE II
PERFORMANCE EVALUATION OF PROPOSED KRR-RVFL WITH WAVELET FUNCTION (IN %)

Data	Accuracy	Recall	Specificity	Precision	F1-Score
GM	97.22	100	95.83	92.31	96.00
WM	99.31	100	98.96	97.96	98.97
CSF	95.83	100	93.75	88.89	94.12
Ensemble	99.56	100	99.12	98.23	99.14

Accuracy measures the classifier's ability to differentiate between ages 20–40, ages 41–60, and ages 61–90. Sensitivity or Recall defines the fraction of correctly classified positive occurrences. Precision describes the proportion of true positive predictions out of all positive occurrences. Specificity represents the fraction of correctly classified negative occurrences. F1-score measures accuracy through the weighted harmonic mean of precision and recall. The confusion matrix (CM) evaluates the performance metrics of classification models on a given test data set, and true values are employed to construct CM. The receiver operating characteristic (ROC) curve is a specificity versus sensitivity 2-D plot, illustrating the classification algorithm performance for all values. To assess the accuracy of brain age prediction models, MAE, BAG, coefficient of determination (R^2), and root mean squared error (RMSE) between individual predicted brain ages and real ages are evaluated.

C. Comparison of Different Segmented Brain Tissue Volumes

Table II displays the results of brain age classification for the KRR-RVFL-based deep model utilizing a wavelet kernel function, with GM, WM, and CSF datasets. The WKRR-RVFL dep network achieved an accuracy rate of 97.22% for GM, 99.31% for WM, and 95.83% for CSF. Notably, WM exhibited the highest accuracy in brain age classification, while CSF provided the lowest performance metrics. Furthermore, the proposed model reached a classification accuracy of 99.56% when considering the combination of GM, WM, and CSF features, which is slightly superior to WM maps. However, it's important to note that the computational complexity of the brain age classification framework is significantly higher when compared to a single modality. The CM and ROC curves of all segmented brain compartments as shown in Fig. 5.

D. Comparison With Different Kernel Functions Introduced in Standard RVFL and State-of-the-Art Classifiers

In this subsection, the efficacy of the proposed KRR-RVFL model with various kernel functions like the polynomial kernel (PK), Gaussian kernel (GK), is compared with standard RVFL, deep RVFL (dRVFL), ensemble dRVFL (edRVFL) and different classifiers like softmax [31], SVM [22], random forest (RF) [27], ensemble bagging (EB) [25], K-nearest neighbor (KNN) [5], Naive Bayes (NB) [37]. The performance matrices of various classifiers with regard to accuracy, recall, specificity, precision, and F1-score are tabulated in Table III. The wavelet KRR-RVFL deep model outperforms standard RVFL, dRVFL, edRVFL as well as other state-of-the-art networks.
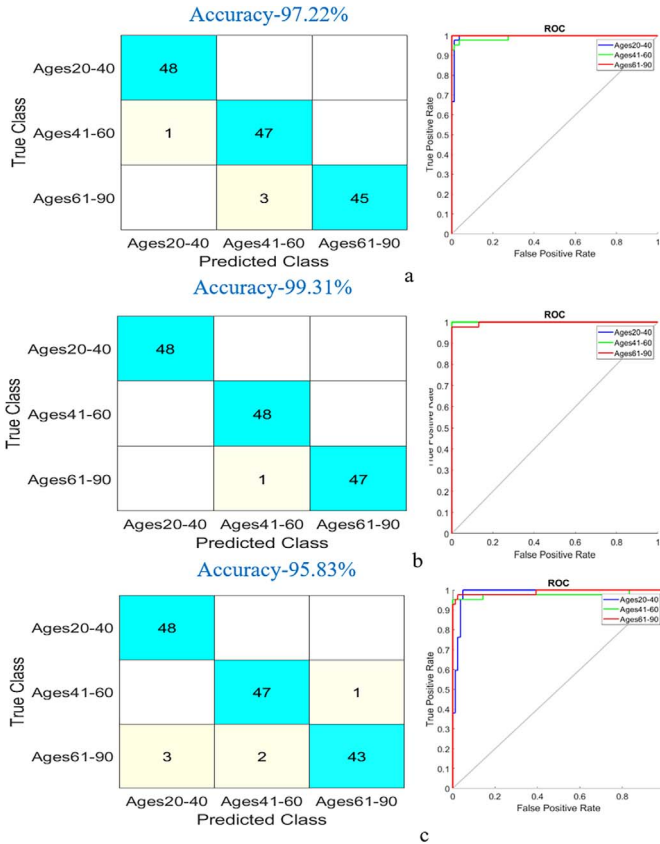
Fig. 5.   CM and ROC curve of a) GM; b) WM; and c) CSF.

TABLE III
COMPARISON BETWEEN THE WKRR-RVFL MODEL AND DIFFERENT
CLASSIFIERS ON WM DATA (IN %)

Model	Accuracy	Recall	Specificity	Precision	F1-Score
NB [37]	90.56	92.86	88.10	79.59	85.71
KNN [5]	91.27	97.62	88.10	80.39	88.17
RF [27]	92.06	100	88.10	80.77	89.36
EB [25]	93.09	87.50	95.83	91.30	89.36
SVM [22]	92.86	95.24	91.67	85.11	89.89
Softmax [31]	93.75	85.42	97.92	95.35	90.11
RVFL [29]	94.44	95.24	94.04	88.89	91.95
dRVFL [28]	95.24	97.62	94.05	89.13	93.18
edRVFL [15]	96.03	97.62	95.24	91.11	94.25
PKRR-RVFL	97.92	100	96.88	94.12	96.97
GKRR-RVFL	98.61	100	97.62	97.96	98.95
WKRR-RVFL	**99.31**	**100**	**98.96**	**97.96**	**98.97**

Note: Bold indicates best result.

E. Comparison With Different Activation Functions

In this subsection, different conventional activation functions like ReLU, scaled exponential linear unit (SELU), sigmoid, sine, hard limit (HardLim), TriBas, and RadBas incorporated individually along with wavelet kernel function to get complete nonlinearity at the output layer of RVFL network. Activation functions are critical in neural networks because they introduce nonlinearity, allow data translation, and substantially influence the network's ability to adapt to input and distinguish between alternative outcomes. The comparative analysis of the

TABLE IV
COMPARISON BETWEEN DIFFERENT ACTIVATION FUNCTIONS ON WM
DATA (IN %)

Activation Function	Accuracy	Recall	Specificity	Precision	F1-Score
ReLU	92.47	96.15	92.19	91.13	91.33
SELU	89.13	92.61	90.32	90.11	88.74
Sigmoid	94.63	95.94	95.10	94.21	92.27
Sine	84.43	89.26	87.21	83.47	83.22
HardLim	97.13	99.75	96.81	95.29	96.49
TriBas	97.83	99.72	96.73	94.27	96.82
RadBas	**99.31**	**100**	**98.96**	**97.96**	**98.97**

Note: Bold indicates best result.

TABLE V
COMPARISON BETWEEN WKRR-RVFL REGRESSION MODEL AND
DIFFERENT REGRESSION MODELS ON GM, WM, CSF TESTING
DATASETS

Data	Regression Model	MAE (Years)	BAG (Years)	RMSE (Years)	R^2 Score
GM	ELM [9]	4.61	0.99	5.24	0.93
	KRR [30]	4.09	0.39	4.88	0.95
	RVFL [29]	4.03	0.36	4.79	0.95
	WKRR-RVFL	3.89	0.22	4.13	0.96
WM	ELM [9]	4.33	0.74	5.07	0.94
	KRR [30]	3.97	0.28	4.81	0.96
	RVFL [29]	3.90	0.26	4.65	0.96
	WKRR-RVFL	**3.64**	**0.13**	**4.04**	**0.97**
CSF	ELM [9]	5.21	1.10	6.74	0.88
	KRR [30]	4.77	0.89	5.40	0.91
	RVFL [29]	4.73	0.84	5.33	0.91
	WKRR-RVFL	4.49	0.63	5.19	0.92

Note: Bold indicates best result.

WKRR-RVFL model with different activation functions on the WM dataset is shown in Table IV.

F. Performance Evaluation of Different Regression Models

The BAG is used as the biomarker to understand better the healthy aging process and the prognosis of severe neuronal ailments. In this study, we predicted the brain age of CN subjects and obtained better prediction outcomes on WM images. The extreme machine learning (ELM) regression network on GM data maps got an MAE value of 4.61 years more than the kernel ridge regression (KRR) and RVFL networks. The BAG is 0.22 years for the WKRR-RVFL model and 0.36 years for the RVFL network. The ELM achieved an MAE of 4.33 years, RMSE of 5.07 years, and R^2 score of 0.94, and for the RVFL network, an MAE of 3.90 years on WM maps, whereas the KRR obtained a BAG of 0.28 years, an RMSE of 4.81 years. The proposed regression model achieved the lowest error outcomes, such as MAE of 3.64 years, BAG of 0.13 years, and RMSE of 4.04 years, than the comparative regression models. On CSF data maps, ELM provided more error outcomes as MAE is 5.21 years and the BAG is 1.10 years. The BAG and error output values are higher for CSF images, indicating that the fluid maps are less associated with aging. The WM tissue is significantly associated with the normal aging process than the GM tissue and the summary of regression models assessment is tabulated in Table V.

TABLE VI
COMPARISON BETWEEN WKRR-RVFL REGRESSION MODEL
AND CONVENTIONAL REGRESSION MODELS

Data	Regression Model	MAE (Years)	BAG (Years)	RMSE (Years)	R^2 Score
GM	SVR [16]	5.66	1.69	6.43	0.85
	GPR [36]	5.42	1.27	6.32	0.86
	RR [17]	5.28	1.33	6.22	0.86
	RF [1]	5.39	1.21	6.59	0.86
	LR [1]	5.45	1.43	6.41	0.86
	WKRR-RVFL	3.89	0.22	4.13	0.96
WM	SVR [16]	5.47	1.14	6.28	0.86
	GPR [36]	5.39	1.11	6.37	0.88
	RR [17]	5.16	1.09	6.19	0.87
	RF [1]	5.23	1.17	6.24	0.87
	LR [1]	5.36	1.23	6.32	0.87
	WKRR-RVFL	**3.64**	**0.13**	**4.04**	**0.97**
CSF	SVR [16]	6.03	1.82	7.10	0.82
	GPR [36]	6.22	1.93	7.09	0.82
	RR [17]	6.45	2.01	7.25	0.81
	RF [1]	6.26	1.97	7.12	0.82
	LR [1]	5.98	1.76	7.01	0.83
	WKRR-RVFL	4.49	0.63	5.19	0.92

Note: Bold indicates best result.

TABLE VII
VOXELWISE MULTIPLE REGRESSION OF AGE ON BRAIN TISSUES
WITH $p < 0.05$

Tissue	Anatomical Area	Voxels	T-Value
WM	Left cerebrum temporal lobe	2645	18.10
	Right cerebrum thalamus	205	11.98
	Right cerebrum extra nuclear	24	9.70
	Right cerebrum internal ventricle	36	6.83
GM	Right cerebrum caudate	590	14.09
	Right cerebrum extra nuclear	15	9.54
	Left cerebrum extra nuclear	9	8.39
CSF	Left cerebrum limbic lobe	80	9.87
	Interhemispheric fissure	10	7.70

G. Comparison Between WKRR-RVFL Regression Network and State-of-the-Art Regression Networks

Table VI shows the comparison between the WKRR-RVFL regression network and conventional regression models like support vector regression (SVR), Gaussian regression process (GPR), ridge regression (RR), RF, and lasso regression (LR) networks for GM. WM, and CSF datasets in terms of prediction accuracy measures.

On the WM dataset, the SVR network achieved an MAE of 5.47 years, a BAG of 1.14 years, and the RF regression network obtained a RMSE value of 6.24 years. The LR network got 5.45 years of MAE and 6.41 years of RMSE values on the GM dataset, whereas, for the CSF dataset, the RR network obtained a BAG of 1.93 years and the RF network achieved the prediction accuracy of a MAE is 6.26 years. The proposed deep model achieved better prediction outcomes than conventional regression models.

H. VBM Analysis

With the voxelwise multiple regression of age on the volumes of brain tissues found significant changes in various brain regions of healthy aging. All regions described here as important survived a statistical threshold of $p < 0.05$ (FWE corrected). Substantial age-related changes in GM volume have been identified in the Right cerebrum caudate, Right cerebrum extranuclear, and Left cerebrum extranuclear. WM volume alterations in Left cerebrum temporal lobe, Right cerebrum thalamus, Right cerebrum extranuclear, Right cerebrum internal ventricle. Areas of accelerated enlargement of the CSF are seen at the Interhemispheric fissure and Left cerebrum limbic lobe. Table VII demonstrated voxelwise multiple regression of age on the WM, GM, and CSF tissues with $p < 0.05$ and the clusters in Fig. 6 depict the GM, WM, and CSF tissues alteration of brain regions.
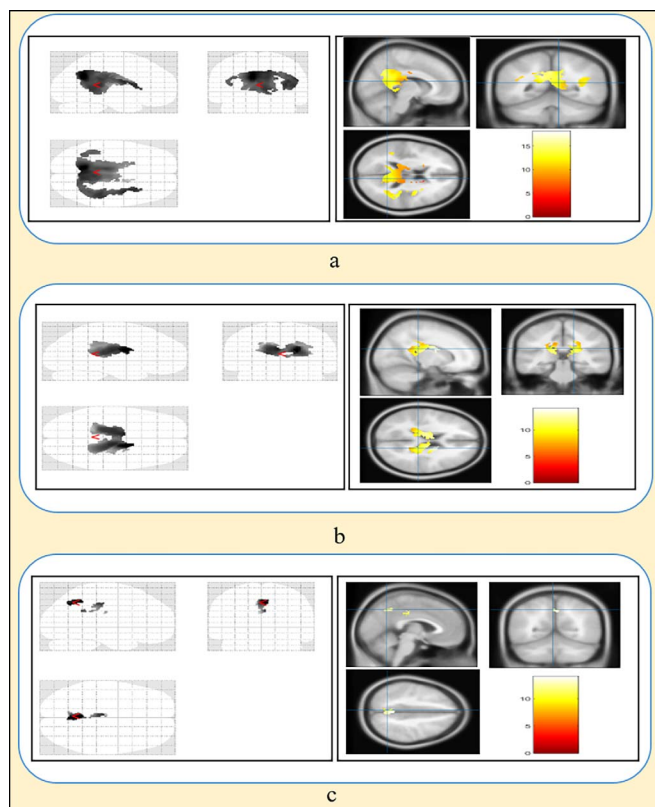

Fig. 6.   VBM analysis for a) WM alterations; b) GM alterations; and c) CSF alterations.

I. Discussion

In this work, we aimed to classify and estimate brain age based on changes in the brain's tissues. The appropriate detection of brain imaging biomarkers, that have a major influence on brain age, using DL networks can enhance medical services. Monitoring healthy aging using DL models may give new methods for predicting health and the risk of neurodegenerative disorders. Hence, this study aims to implement an accurate brain age prediction framework that can be useful to observe brain age in the clinical context. The significant finding of our study is that age can be more highly correlated with WM than GM and CSF. During young, middle, and old age, WM experiences some of the most significant changes. The brain develops during young adulthood and reaches maturity around

the twenties. WM gradually increases, and specific brain regions are responsible for bodily functioning during this period. In middle age, brain atrophy begins slightly, and some areas are affected, leading to cognitive function decline. In older age, the brain shrinks more rapidly, and there is a significant reduction in WM volume. The substantial decrease in WM volume is consistent with the concept that the aging brain experiences a decrease in neural connectivity. Multiple studies [18], [21], [26] have demonstrated a relatively more significant decline in WM volume with increasing age. In individuals aged 30–90, there is a 26% reduction in WM volume compared to a 14% reduction in GM tissue volume. WM volume loss may be characteristic of later stages of aging, whereas GM loss may initiate earlier and progress gradually. Brain aging is characterized by the deterioration of WM, including features such as myelin discoloration, loss of myelinated fibers, and deformation of myelin sheaths. The proposed brain age classification framework is implemented on 480 brain MRI scans collected from the IXI database and arranged into three groups (ages 20–40, ages 41–60, and ages 61–90). All MRI images are segmented and preprocessed using SPM12 toolbox, and GM, WM, and CSF datasets are prepared for doing experiments on the proposed KRR-RVFL-based deep model. The features of input images are extracted using 3-D-CNN, and classification has been done by the KRR-RVFL network with different kernel functions. The experiments are conducted on GM, WM, and CSF segmented data and achieved 97.22%, 99.31%, and 95.83% classification accuracy, respectively, and for ensemble of three tissues the model provided better accuracy, but the complexity of the model rise up hence we considered only single modality to reduce the complexity. The summary of assessment of wavelet KRR-RVFL model as depicted in Table II. The proposed KRR-RVFL-based deep model with different kernel functions is compared with various classifiers like NB, KNN, EB, RF, and SVM classifiers, as well as with standard RVFL, dRVFL, and edRVFL. WM dataset is considered as input to the model. The KRR-RVFL model with wavelet kernel function yields better performance metrics than other state-of-the-art networks as illustrated in Table III. The utilization of the RadBas activation function and the wavelet kernel function in the RVFL network yielded notable classification results for WM scans when compared with various other activation functions, as summarized in Table IV. The summary of the assessment of different regression models used to predict brain age is shown in Table V. The proposed regression model outperforms other comparative regression models on the WM dataset in terms of MAE, BAG, and RMSE values. The WKRR-RVFL model achieved RMSE of 4.13 years, 4.04 years, and 5.19 years and a BAG of 0.22 years, 0.13 years, and 0.63 years for GM, WM, and CSF datasets, respectively. Conventional regression models like SVR, GPR, RR, RF, and LR networks are used to predict the brain age of CN subjects, and the prediction accuracy measures are compared with the proposed model outcomes are shown in Table VI. Furthermore, we also investigated the age-related anatomical variations in brain tissue volumes using VBM analysis. VBM helps to detect the anatomical variations of brain regions in neurodegenerative ailments like dementia,

PD, schizophrenia (SZ), Huntington's disease, and multiple sclerosis (MS). VBM automatically analyzes each voxel size of segmented tissue to identify the GM, WM, and CSF volume variations. VBM is not biased toward anatomical changes in a specific brain location and prevents subjective differences induced by artificially showing the region of interest. In this work GM volume alterations have been identified in three regions, WM volume variations are detected in five regions, and for CSF two regions are identified. For WM maps T-value (18.10) and for GM, CSF maps T-values (14.09), (9.87), respectively. It indicates that the WM more correlates with age compared with GM and CSF.

V. CONCLUSION

In this study, we classified and estimated the brain age of 480 CN individuals using a kernel ridge regression-based randomized deep model with structural magnetic imaging (MRI) images. We employed the DARTEL preprocessing pipeline to segment the MRI images into GM, WM, and CSF tissues. All images are then normalized to a standard MNI template using the statistical parametric mapping-12 (SPM12) toolbox. The features of the preprocessed MRI scans are extracted using a 3-D convolutional neural network (3-D-CNN), and these extracted features are classified using a wavelet kernel ridge regression (WKRR)-RVFL network. To address the potential linearity issues at the output layer due to the direct link of input features, a kernel trick is applied to the input features before transferring them to the output layer. The WKRR-RVFL-based deep model demonstrated excellent performance metrics, including F1-score, Accuracy, Sensitivity, Specificity, and Precision, Particularly in relation to WM images. Additionally, the model effectively estimated brain age, yielding prominent outcomes in terms of BAG, MAE, root mean square error (RMSE), and coefficient of determination (R^2). When compared to various classifiers and regression networks, the KRR-RVFL network with the wavelet kernel function outperformed others. The experimental results underscore the feasibility and validity of the wavelet KRR-RVFL network in brain age classification and estimation. Furthermore, we conducted a VBM analysis to observe age-related anatomical variations in GM, WM, and CSF brain tissues of CN subjects. Notably, our findings suggest that WM volumes are more significantly associated with the normal aging process than GM and CSF volumes.

The current study's limitations lie in its exclusive use of structural MRI images of CN subjects for brain age classification and estimation. The WKRR-RVFL model requires a significant amount of memory to store the kernel matrix, which makes solving linear systems involving kernel techniques computationally expensive, as well as confronting challenges in terms of hyperparameter fine-tuning, interpretability, and scalability. Future research can expand on this by incorporating functional MRI scans to capture functional changes, positron emission tomography (PET) scans to observe metabolic changes and magnetic resonance spectroscopy (MRS) scans to detect chemical alterations in the brain, facilitating a comprehensive analysis of

age-related variations in the brain. Additionally, future studies may explore the connection between brain aging and neurological disorders.

REFERENCES

[1] N. Amoroso et al., "Deep learning and multiplex networks for accurate modeling of brain age," *Front. Aging Neurosci.*, vol. 11, 2019, Art. no. 115.

[2] I. Beheshti, N. Maikusa, and H. Matsuda, "Effects of aging on brain volumes in healthy individuals across adulthood," *Neurol. Sci.*, vol. 40, no. 6, pp. 1191–1198, 2019.

[3] J. Bijsterbosch, "How old is your brain?" *Nat. Neurosci.*, vol. 22, no. 10, pp. 1611–1612, 2019.

[4] P. Borah, D. Gupta, and S. S. S. Mishra, "Kernelized random vector functional-link network," in *Proc. Pattern Recognit. Data Anal. Appl.*, New York, NY, USA: Springer-Verlag, 2022, pp. 743–750.

[5] P. Cunningham and S. J. Delany, "k-Nearest neighbour classifiers—A tutorial," *ACM Comput. Surv. (CSUR)*, vol. 54, no. 6, pp. 1–25, 2021.

[6] F. Farokhian, C. Yang, I. Beheshti, H. Matsuda, and S. Wu, "Age-related gray and white matter changes in normal adult brains," *Aging Disease*, vol. 8, no. 6, p. 899–909, 2017.

[7] A. M. Fjell et al., "Critical ages in the life course of the adult brain: Nonlinear subcortical aging," *Neurobiol. Aging*, vol. 34, no. 10, pp. 2239–2247, 2013.

[8] S. M. Hofmann et al., "Towards the interpretability of deep learning models for multi-modal neuroimaging: Finding structural changes of the ageing brain," *NeuroImage*, vol. 261, 2022, Art. no. 119504.

[9] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern., B (Cybern.)*, vol. 42, no. 2, pp. 513–529, 2011.

[10] N. Hwangbo et al., "An aging clock using metabolomic CSF," *J. Gerontology - Ser. A Biol. Sci. Med. Sci.*, pp. 744–754, 2021.

[11] A. Joy et al., "Alterations of gray and white matter volumes and cortical thickness in treated HIV-positive patients," *Magn. Reson. Imaging*, vol. 95, pp. 27–38, 2023.

[12] C.-Y. Kuo et al., "Improving individual brain age prediction using an ensemble deep learning framework," *Front. Psychiatry*, vol. 12, 2021, Art. no. 626677.

[13] H. Liu et al., "A voxel-based morphometric study of age-and sex-related changes in white matter volume in the normal aging brain," *Neuropsychiatr. Dis. Treat.*, vol. 12, pp. 453–465, 2016.

[14] M. E. MacDonald and G. B. Pike, "MRI of healthy brain aging: A review," *NMR Biomed.*, vol. 34, no. 9, 2021, Art. no. e4564.

[15] A. K. Malik and M. Tanveer, "Graph embedded ensemble deep randomized network for diagnosis of Alzheimer's disease," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, early access, Sep. 16, 2022.

[16] O. L. Mangasarian and D. R. Musicant, "Robust linear and support vector regression," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 9, pp. 950–955, Sep. 2000.

[17] R. J. Massett et al., "Regional neuroanatomic effects on brain age inferred using magnetic resonance imaging and ridge regression," *J. Gerontol. Ser. A*, vol. 78, no. 6, pp. 872–881, 2023.

[18] C. Morrison, M. Dadar, S. Villeneuve, and D. L. Collins, "White matter lesions may be an early marker for age-related cognitive decline," *NeuroImage, Clin.*, vol. 35, 2022, Art. no. 103096.

[19] D. R. Nayak, R. Dash, B. Majhi, and U. R. Acharya, "Application of fast curvelet Tsallis entropy and kernel random vector functional link network for automated detection of multiclass brain abnormalities," *Computerized Med. Imag. Graph.*, vol. 77, 2019, Art. no. 101656.

[20] N. Neufeld et al., "Longitudinal changes in grey matter and cognitive performance over four years of healthy aging," *Neuroimage, Rep.*, vol. 2, no. 4, 2022, Art. no. 100140.

[21] R. Pilli, T. Goel, R. Murugan, and M. Tanveer, "Association of white matter volume with brain age classification using deep learning network and region wise analysis," *Eng. Appl. Artif. Intell.*, vol. 125, 2023, Art. no. 106596.

[22] D. A. Pisner and D. M. Schnyer, "Support vector machine," in *Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 101–121.

[23] S. M. Resnick, D. L. Pham, M. A. Kraut, A. B. Zonderman, and C. Davatzikos, "Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain," *J. Neurosci.*, vol. 23, no. 8, pp. 3295–3301, 2003.

[24] S. Safaiyan et al., "White matter aging drives microglial diversity," *Neuron*, vol. 109, no. 7, pp. 1100–1117, 2021.

[25] R. Saini and S. K. Ghosh, "Ensemble classifiers in remote sensing: A review," in *Proc. Int. Conf. Comput., Commun. Automat. (ICCCA)*, Piscataway, NJ, USA: IEEE Press, 2017, pp. 1148–1152.

[26] K. G. Schilling et al., "Aging and white matter microstructure and macrostructure: A longitudinal multi-site diffusion MRI study of 1218 participants," *Brain Struct. Funct.*, vol. 227, no. 6, pp. 2111–2125, 2022.

[27] A. B. Shaik and S. Srinivasan, "A brief survey on Random Forest ensembles in classification model," in *Proc. Int. Conf. Innovative Comput. Commun.*, New York, NY, USA: Springer-Verlag, 2019, pp. 253–260.

[28] R. Sharma, T. Goel, M. Tanveer, S. Dwivedi, and R. Murugan, "FAF-DRVFL: Fuzzy activation function based deep random vector functional links network for early diagnosis of Alzheimer disease," *Appl. Soft Comput.*, vol. 106, 2021, Art. no. 107371.

[29] Q. Shi, R. Katuwal, P. N. Suganthan, and M. Tanveer, "Random vector functional link neural network based ensemble deep learning," *Pattern Recognit.*, vol. 117, 2021, Art. no. 107978.

[30] C. Saunders, A. Gammerman, and V. Vovk, "Ridge regression learning algorithm in dual variables," in *Proc. 15th Int. Conf. Mach. Learn.*, 1998, pp. 515–521.

[31] M. Siar and M. Teshnehlab, "Age detection from brain MRI images using the deep learning," in *Proc. 9th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Piscataway, NJ, USA: IEEE Press, 2019, pp. 369–374.

[32] C. Spindler, L. Mallien, S. Trautmann, N. Alexander, and M. Muehlhan, "A coordinate-based meta-analysis of white matter alterations in patients with alcohol use disorder," *Transl. Psychiatry*, vol. 12, no. 1, pp. 1–9, 2022.

[33] J. Steffener, C. Habeck, D. O'Shea, Q. Razlighi, L. Bherer, and Y. Stern, "Differences between chronological and brain age are related to education and self-reported physical activity," *Neurobiol. Aging*, vol. 40, pp. 138–144, Apr. 2016.

[34] M. Tanveer et al., "Deep learning for brain age estimation: A systematic review," *Inf. Fusion,* vol. 96, pp. 130–143, 2023.

[35] D. Terribilli et al., "Age-related gray matter volume changes in the brain during non-elderly adulthood," *Neurobiol. Aging*, vol. 32, no. 2, pp. 354–368, 2011.

[36] A. G. Wilson, D. A. Knowles, and Z. Ghahramani, "Gaussian process regression networks," in *29th Int. Conf. Mach. Learn.*, 2012, *arXiv: 1110.4411.*

[37] F.-J. Yang, "An implementation of Naive Bayes classifier," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Piscataway, NJ, USA: IEEE Press, 2018, pp. 301–306.

[38] P.-B. Zhang and Z.-X. Yang, "A new learning paradigm for random vector functional-link network: RVFL+," *Neural Netw.*, vol. 122, no. Issue C, pp. 94–105, 2020.

[39] Y. Zhang, J. Wu, Z. Cai, B. Du, and S. Y. Philip, "An unsupervised parameter learning model for RVFL neural network," *Neural Netw.*, vol. 112, pp. 85–97, 2019.