



# Multi-resolution 3D-HOG feature learning method for Alzheimer's Disease diagnosis<sup>☆</sup>

Zhiyuan Ding<sup>a</sup>, Yan Liu<sup>b</sup>, Xu Tian<sup>b</sup>, Wenjing Lu<sup>a</sup>, Zheng Wang<sup>c</sup>, Xiangzhu Zeng<sup>c,\*\*</sup>,  
Ling Wang<sup>a,\*</sup>

<sup>a</sup> School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China

<sup>b</sup> School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing, China

<sup>c</sup> Department of Radiology, Peking University Third Hospital, Beijing, China

## ARTICLE INFO

### Article history:

Received 5 July 2021

Revised 25 November 2021

Accepted 1 December 2021

### Keywords:

Feature learning

Multi-resolution

HOG

Alzheimer's Disease

## ABSTRACT

**Background and Objective:** Alzheimer's Disease (AD) is a progressive irreversible neurodegeneration disease and thus timely identification is critical to delay its progression. **Methods:** In this work, we focus on the traditional branch to design discriminative feature extraction and selection strategies to achieve explainable AD identification. Specifically, a spatial pyramid based three-dimensional histogram of oriented gradient (3D-HOG) feature learning method is proposed. Both global and local texture changes are included in spatial pyramid 3D-HOG (SPHOG) features for comprehensive analysis. Then a modified wrapper-based feature selection algorithm is introduced to select the discriminative features for AD identification while reduce feature dimensions. **Results:** Discriminative SPHOG histograms with various resolutions are selected, which can represent the atrophy characteristics of cerebral cortex with promising performance. As subareas corresponding to selected histograms are consistent with clinical experience, explanatory is emphasized and illustrated with Hippocampus. **Conclusion:** Experimental results illustrate the effectiveness of the proposed method on feature learning based on samples obtained from common dataset and a clinical dataset. The proposed method will be useful for further medical analysis as its explanatory on other region-of-interests (ROIs) of the brain for early diagnosis of AD.

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction

Alzheimer's Disease (AD) is a progressive neurodegenerative disease, which reflects anatomical atrophy or functional neurodegeneration of cerebral cortex. In recent years, some machine learning methods have been used to extract useful features from magnetic resonance imaging (MRI) scanned anatomical data to identify AD from Healthy Controls (HC). For feature extraction methods, three-dimension (3D) image based methods can effectively preserve the spatial feature information of MRI data compared with two-dimension (2D) image based methods. These methods directly

or indirectly extract 3D features from MRI data using traditional feature extraction methods or deep learning methods respectively.

Most traditional feature extraction methods directly extract anatomical features of cerebral cortex from MRI images. Among them, the feature-based machine learning methods use the clinical parameters as features, such as the volume of gray matter, the cortical thickness, the mean curvature and area of cortical which are extracted by FreeSurfer image analysis suite [1]. Some image-based feature extraction methods extract features from different transform domain, such as texture-based Gabor transform method [2], multi-resolution-based discrete wavelet transform (DWT) method [3]. And there are other image-based methods, which directly extract image-based feature to describe the atrophy or shape changes of region-of-interests (ROIs) of the brain, such as ROIs-based sparse feature learning method [4,5], local binary pattern (LBP) method [6] and histogram of oriented gradient (HOG) method [7,8]. As an image gradient based feature extraction method, HOG method extracts image gradients within a region to reflect its edge gradient changes [9], which is used for region detection in medical images [10,11]. Considering that there is volume atrophy and shape changes of cortical in AD, HOG method is also

<sup>☆</sup> This work was supported by "National Natural Science Foundation of China (NSFC)" (No. 61971106, 61603077, 41776204), "Fundamental Research Funds for the Central Universities"

\* Corresponding author.

\*\* Corresponding author.

E-mail addresses: [2017020904008@std.uestc.edu.cn](mailto:2017020904008@std.uestc.edu.cn) (Z. Ding), [yanliu@ucas.ac.cn](mailto:yanliu@ucas.ac.cn) (Y. Liu), [tianxu19@mails.ucas.ac.cn](mailto:tianxu19@mails.ucas.ac.cn) (X. Tian), [2017020911022@std.uestc.edu.cn](mailto:2017020911022@std.uestc.edu.cn) (W. Lu), [quendy@sina.com](mailto:quendy@sina.com) (Z. Wang), [xiangzhuzeng@126.com](mailto:xiangzhuzeng@126.com) (X. Zeng), [ewwangling@uestc.edu.cn](mailto:ewwangling@uestc.edu.cn) (L. Wang).

used to detect the local texture shape changes or gradient changes of the image for early diagnosis of AD. For example, Devvi *et al* extracted HOG feature from three orthogonal of planes to describe the dynamic texture changes of MRI brain images [7]. Zhu *et al* proposed a new multi-view learning method to learn the mappings from the HOG feature space to the ROI feature space, which uses 3D-HOG features as local features to reflect small or subtle changes within brain [12]. And in [8], small scale HOG features are extracted from ROIs and used to quantify spatial gradients of 18F-FDG PET images for AD diagnosis. 3D-HOG features represent the local texture changes within a volume statistic of spatial gradient and overcome the information loss generated from 2D-HOG representation. However, there are two main limits for 3D-HOG method: (1) features with same scale only represent local visual features with the same resolution, which cannot represent the characteristics of the image comprehensively. (2) HOG features consist of numerous histograms and thus cover invalid information. Effective characteristic should be obtained in additional step for accurate classification.

Although there are exhaustive extracted features, irrelevant or redundant features may reduce the efficiency of learning algorithms, i.e. not all extracted features are useful for the classification problems. As discussed in [13] and [14], feature extraction usually encounter the so-called 'High Dimension, Low Sample Size (HDLSS)' problem. In order to resolve this problem, subspace learning methods and feature selection methods are used to reduce the feature dimensions to choose discriminating features. Subspace learning methods include linear methods, such as principle component analysis (PCA) [4,15,16], linear discriminant analysis (LDA) [17,18], and non-linear methods, such as multi-kernel methods [19,20]. Feature selection methods generally choose discriminative feature subset for the following classification, which are divided into class-dependent methods and class-independent methods. Class-independent feature selection methods choose potential features while ignoring different classes. Class-dependent feature selection methods utilize different feature subsets to discriminate different classes and obtain better performance than class-independent feature selection [21], which can be further divided into filter approaches and wrapper approaches depending on whether classifier is used or not. Filter approaches utilize various feature importance ranking methods for feature selection and selected features are used for comprehensive classification. For example, Minimal-RedundancyMaximal-Relevancy Measure (mRMR) method [22] selects attributes with maximal relevance and the minimal redundancy based on calculating the mutual information. RELIEF method [23] is a weighted method which tends to minimize intra-class distance and maximize inter-class distance. Class Separability Measure (CSM) method [24] is proceeded by calculating the intra-class and inter-class ratio which is used to evaluate the contribution of each attribute. The wrapper based method [21] finds discriminative feature subsets for each class and then uses the class-dependent subset for final classification. In [25], we proposed a wrapper-based feature selection method to rank the feature importance, which can select most important feature parameters or featured ROIs for AD identification compared with other feature selection methods.

As the image-based classification methods, deep learning methods use the whole brain or ROIs as input of the network to extract 2D or 3D features from MRI data indirectly. There are different network structures introduced to extract discriminate features from original images, such as convolution neural network (CNN), Residual Network (ResNet) [26], a combination of encoder-decoder network [27], U-Net [28], DenseNet [29], and transfer learning strategy [30]. Although these methods have shown good performance on AD classification, it is difficult to present interpretability on the extracted features or classification results because they incorpo-

**Table 1**  
Demographics of the subjects.

	AD	NC
Number	198	236
Female/Male	110/88	120/116
Age	72.61 ± 8.15	71.58 ± 7.32
Education	11.1 ± 5.78	14.18 ± 3.34
<sup>a</sup> MMSE	17.84 ± 6.11	29.06 ± 1.02
<sup>b</sup> CDR	0.5, 1	0

a: mini-mental state examination b: clinical dementia rating

rate feature extraction and classifier learning into an unified framework [31,32]. Furthermore, the training outcome is unsatisfactory as the lack of samples [33].

In this paper, we proposed a ROIs-based multi-resolution 3D-HOG feature learning method for AD identification. Some preliminary accounts of this study were presented in our early conference papers [25,34]. The main contributions of this paper can be concluded as follows:

- First, we proposed a multi-resolution 3D-HOG feature extraction method to describe local and global texture changes for AD identification, which can represent the characteristics of the image comprehensively compared with previous HOG-based approaches.
- Second, we proposed a histogram based wrapped feature selection method, which can not only select discriminative histograms with promising performance, but also detect distinct subareas of ROIs for AD identification.

The remaining of this paper is organized as follows: Section 2 presents the proposed feature extraction and feature selection method, Section 3 presents the experimental results, followed by the conclusion of this paper in Section 4.

## 2. Methodology

Framework of the proposed method is shown in Fig. 1. First, data preprocessing with MRI T1-weighted input is used to generate ROI-based images for feature extraction and classification. Second, 3D-HOG from various scales is extracted as basic feature unit. Third, spatial pyramid HOG features are constructed with multi-scale 3D-HOG features for informative representation. Finally, feature selection techniques are introduced to search for discriminative features and further efficient classification.

### 2.1. Materials and image preprocessing

There are two dataset used in this study: the first one  $\Phi_1$  is sampled from Peking University Third Hospital of China. The first one  $\Phi_1$  is downloaded from Alzheimer's Disease Neuroimaging Initiative (ADNI)<sup>1</sup>. Dataset  $\Phi_1$  includes 67 AD subjects and 105 NC subjects, which are sampled from different 3T scanners (i.e. GE, SIEMENS, PHILIP). Dataset  $\Phi_2$  includes 131 AD subjects and 131 normal control (NC) subjects, which are sampled on a SIEMENS Trio 3T scanner with the acquisition parameters  $TR/TE = 2350/3.44ms$ , voxel size =  $1 \times 1 \times 1mm^3$ . Both of two dataset are T1 Magnetization Prepared Rapid Gradient Echo (MPRAGE) data.

Details of subjects are shown in Table 1. The general inclusion/exclusion criteria can be described with mini-mental state examination (MMSE) and clinical dementia rating (CDR), i.e. the MMSE score of each NC subjects is in the range of 26 and 30 with CDR of 0, while the MMSE score of each AD subjects is in the range of 10 and 24, with CDR of 0.5 or 1.

<sup>1</sup> <http://www.loni.ucla.edu/> ADNI

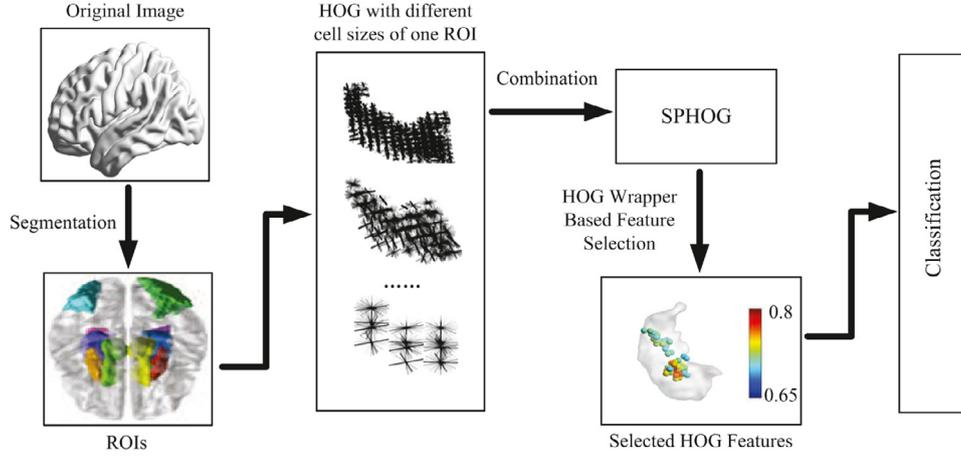


Fig. 1. Framework of the proposed method.

Dataset are examined and preprocessed with SPM8 [35] in the following steps. First, images are denoised and registered on the brain template, and then are segmented into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) maps in the 're-aligned' and 'wrapped' templates by using the 'new segment' toolbox. Second, 90 ROIs masks are obtained by utilizing WFU Pick-Atlas [36] based on automated anatomical labeling (AAL) templates [37] with GM data. Third, each ROI is further cropped with the same size (i.e. the minimal non-zero size of the specified ROI for different samples) in cuboid form  $I^t$ , ( $t = 1, 2, \dots, 90$ ) respectively, which are used for the following feature extraction.

## 2.2. Computation of 3D-HOG

3D-HOG feature represents local texture information within given volume (i.e. cell) and thus is effective to capture volume deformation in MRI. Following is the procedure of 3D-HOG feature extraction. First, spatial gradients at each pixel  $(x, y, z)$  for the  $t$ -th ROI image  $I^t$ , ( $t = 1, 2, \dots, 90$ ) can be approximated as

$$\begin{aligned} \nabla I^t(x, y, z) &= (I_x^t, I_y^t, I_z^t)^T = \left( \frac{\partial I^t}{\partial x}, \frac{\partial I^t}{\partial y}, \frac{\partial I^t}{\partial z} \right)^T \\ &\approx \begin{pmatrix} I^t(x+1, y, z) - I^t(x-1, y, z) \\ I^t(x, y+1, z) - I^t(x, y-1, z) \\ I^t(x, y, z+1) - I^t(x, y, z-1) \end{pmatrix}, \end{aligned} \quad (1)$$

where  $I^t(x, y, z)$  denotes the gray value of pixel  $(x, y, z)$  for  $k$ -th ROIs. And the magnitude and direction of spatial gradient  $\nabla I^t(x, y, z)$  can be computed as

$$\begin{aligned} r &= \sqrt{I_x^2 + I_y^2 + I_z^2}, \\ \theta &= \tan^{-1}(I_y/I_x), \\ \phi &= \cos^{-1}(I_z/r), \end{aligned} \quad (2)$$

where  $r$  is magnitude,  $\theta$  is azimuth, and  $\phi$  is elevation.

Second, histogram for each cell of  $I^t$  is calculated to obtain 3D-HOG descriptors. Spatial regions for each  $I^t$  are referred as a block with same size  $M \times N \times K$ , which is divided into  $m \times n \times k$  cells with size  $S = (l_x, l_y, l_z)^T$ . Each cell represents the small spatial region with length  $l_x$ , width  $l_y$  and height  $l_z$  of the cuboid, in which the statistic is made. In this study, we divide the histogram into  $9 \times 18$  bins for  $\theta$  and  $\phi$ . For each cell, the 2D histogram is created by accumulating the gradient magnitude value  $r$  within different 162 bins for  $\theta$  and  $\phi$ . And we can obtain  $m \times n \times k$  histograms with cell size  $S$ .

Third, histograms from different cells are arranged to construct 3D-HOG features for each ROI. And the dimension of obtained features is of  $162 \times m \times n \times k$  for  $I^k$  with cell size  $S$ . Thus, the 3D-HOG

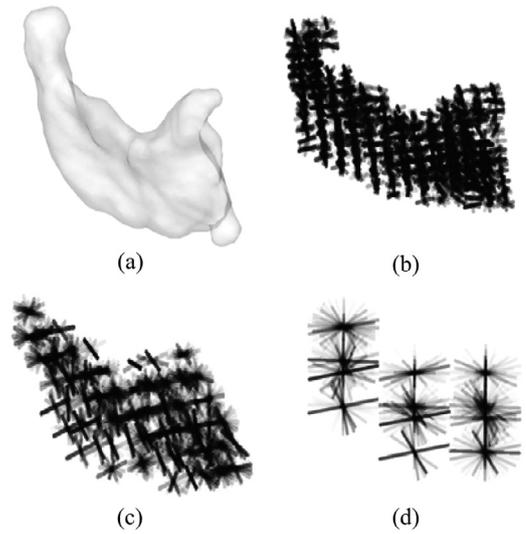


Fig. 2. Left Hippocampus and its HOG feature with different cell sizes: (a) original left Hippocampus; (b) HOG features with cell size  $(2, 2, 2)^T$ ; (c) HOG features with cell size  $(4, 4, 4)^T$ ; (d) HOG features with cell size  $(8, 8, 8)^T$ .

features extracted with cell size  $S$  for  $I^k$  are represented as

$$\begin{aligned} H_S &= (h_1(\theta, \phi), \dots, h_i(\theta, \phi), \dots), \\ S &= (l_x, l_y, l_z)^T, \end{aligned} \quad (3)$$

where  $h_i(\theta, \phi)$  represents the histogram for  $(\theta, \phi)$  of the  $i$ th cell.

## 2.3. Construction of spatial pyramid HOG

As described in [38,39], Spatial Pyramid Representation (SPR) is used as an embedding method for describing both global and local spatial information, which shows good performance on image segmentation and classification. With SPR, discriminative SPR is proposed to improve the effectiveness of feature representation, which is a weighted sum of original features over various pyramid levels [38].

Although extracted 3D-HOG features represent gradient statistical characteristic within a specific range, features with the same cell size represent gradient statistical characteristics with the same resolution, which cannot represent the characteristics of the image comprehensively. As shown in Fig. 2, 3D-HOG features with a small cell size (high resolution) describe detailed spatial information by making statistic within local receptive fields, while the extracted

features with a large cell size (low resolution) describe overall spatial information within global receptive fields. Fig. 2 (a) shows the original left Hippocampus, whose basic shape can be reflected in Fig. 2 (d) with a larger cell size  $S = (8, 8, 8)^T$ . But more detailed information of Hippocampus cannot be found in Fig. 2 (d). If we extract 3D-HOG feature with a smaller cell size  $S = (2, 2, 2)^T$ , we can see more detail information while preserve the basic shape of left Hippocampus in Fig. 2 (b). Fig. 2 (c) shows 3D-HOG features from different resolutions with cell size  $S = (4, 4, 4)^T$ . From Fig. 2, we can see that these HOG features with different cell sizes represent the statistical characteristics of original image in different resolutions.

In this study, we build a multi-resolution 3D-HOG feature in a spatial pyramid form (SPHOG) to capture both local and global texture changes, which can further improve spatial resolution of the extracted 3D-HOG features. The proposed SPHOG feature is a cascading of 3D-HOG features with different cell sizes, which can be described as

$$H_{SP} = \{h|h \in H_S, S = S^1, S^2, \dots, S^t\}. \quad (4)$$

where  $S^k$  represents the  $k$ th cell size,  $H_{SP}$  represents SPHOG feature.

#### 2.4. Histogram based wrapped feature selection

Dimension of SPHOG is much higher than 3D-HOG. Although SPHOG feature describes the gradient changes in various pyramid levels comprehensively, direct classification with SPHOG is non-efficient as the existence of some underlying features. Redundant features which occur repeatedly in various scales would submerge informative feature and larger feature size brings more irrelevant features into potential feature set. Therefore, it is necessary to select discriminative features using some effective feature selection methods.

In this study, a modified wrapper based feature selection algorithm is proposed to select the discriminative feature subsets from extracted SPHOG features.

First, discriminative feature subsets for each class (i.e. AD or HC in this study) are selected based on forward search algorithm. For the original HOG descriptor, it concatenates feature of every bin all into one vector, in which each bin is used as a feature. Considering cells divided with different sizes within one block (i.e. one ROI) can reflect structural and statistical characteristics of ROI in different resolutions, histograms for a cell (i.e. 3D-HOG descriptor) are used as one feature to represent the characteristic, i.e. a histogram is selected or not selected unitarily. In this way, spatial integrality is preserved for further analysis and thus strengthen reliability and interpretability of result. Same as general wrapper approaches, an importance ranking is processed on histogram basis in advance to evaluate the performance of each feature. This is an open step and various measures can be applied, such as accuracy or sensitivity for classification.

During forward selection process, 3D-HOG histograms are added into the feature subset one by one in the order of importance ranking to form a new feature subset. In this way, we can obtain an incremental feature selection curve with the increment of sensitivity or specificity, in which the feature selection point can be determined until a stopping condition is met. Stopping condition makes sure that the forward search process ends when the classification performance is decreasing. In this work, single histograms with classification accuracy higher than 0.7 are selected for incremental selection. Then, these selected histograms are sorted in specific order for the following steps.

Second, a direct incremental selection method is utilized in this histogram-based classification process. Different from the weighted method which generates final probability estimation based on spe-

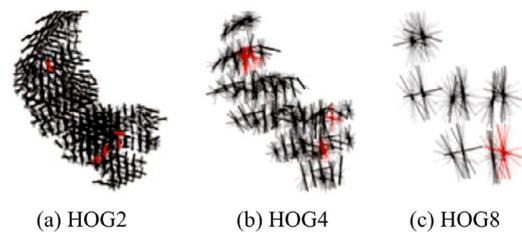


Fig. 3. The selected SPHOG and 3D-HOG of Hippocampus-L. Some selected SPHOG features are shown in red. (a) HOG features with cell size  $(2, 2, 2)^T$ ; (b) HOG features with cell size  $(4, 4, 4)^T$ ; (c) HOG features with cell size  $(8, 8, 8)^T$ . (The black lines are the HOGs in different positions, and the red lines are selected HOGs.)

cific feature subsets, as described in [40], direct method constructs the classification feature set according to classification performance of single histogram. The direct method is applied as we find that utilizing weighted method with proposed SPHOG selects similar histograms for each class and cannot generate distinct result each class. The direct method selects histograms for feature combination just concentrating on single histogram classification accuracy and in this way the selected histograms distinguish for both classes simultaneously.

With the proposed feature selection method, the form of histogram, i.e. SPHOG, is used as one feature entirely during feature selection, in which the importance of each histogram for each class is measured. However, in general wrapper-based feature selection, the general method usually uses feature in one bin (i.e. in one direction of the gradient) as the input of feature selection. It cannot reflect the global or statistical information of the cell and also increase the dimension of features. Such processing methods cannot easily focus on the global or statistical information in unit of cell as the integrity of bins within a histogram is disorganized when single bin ranking proceeds. And the general operation on bins in this way impedes the intact acquiring of local texture information. Compared these two methods, the proposed feature selection method shows the structural and statistical characteristics of the image while saves about 80% of time consumption.

### 3. Experimental result and discussion

In the following experiments, 20 times Monte-Carlo simulations are carried out to: (a) illustrate the effectiveness of SPHOG; (b) illustrate the effectiveness of the proposed feature selection method; (c) compare with other traditional machine learning methods and deep learning methods; (d) illustrate the effectiveness of the selected features. (e) illustrate the effectiveness of our proposed methods on multi-region based identification.

65 AD subjects and 65 HC subjects (totally 130 samples) are randomly selected as test sets and the rest samples serve as train sets. Classification accuracy, true positive rate (TPR), false positive rate (FPR), precision, F1-score and Kappa coefficient are used as indicators for classification results. According to the diagnostic experience and previous studies, Hippocampus is one of the most discriminative ROIs for AD identification. And thus we use left Hippocampus (i.e. Hippocampus-L) to illustrate the effectiveness of the proposed method in the first three experiments.

#### 3.1. Effectiveness of SPHOG

In this work, SPHOG is built based on 3D-HOG with 5 cell sizes  $l = 4, 5, 6, 7, 8$  with  $l_x = l_y = l_z = l$ . It is used to combine local and global features to represent the atrophy of the brain comprehensively. Some 3D-HOG features are shown in Fig. 3, in which some histograms are selected as feature subsets during feature selection (shown in red). Based on 5 sizes of 3D-HOG features, we select

**Table 2**  
Classification results based on 3D-HOG and SPHOG for Hippocampus-L .

Name	Accuracy	TPR	FPR	Precision	F1	Kappa
HOG2	0.777	0.784	0.207	0.762	0.770	0.572
HOG3	0.794	0.828	0.214	0.767	0.794	0.607
HOG4	0.839	0.828	0.133	0.840	0.832	0.694
HOG5	0.843	0.817	0.115	0.857	<b>0.835</b>	0.701
HOG6	0.846	0.798	<b>0.094</b>	<b>0.875</b>	0.833	0.706
HOG7	0.821	0.777	0.122	0.845	0.807	0.655
HOG8	0.836	0.806	0.118	0.850	0.825	0.687
<b>SPHOG</b>	<b>0.859</b>	<b>0.875</b>	0.100	0.867	0.833	<b>0.708</b>

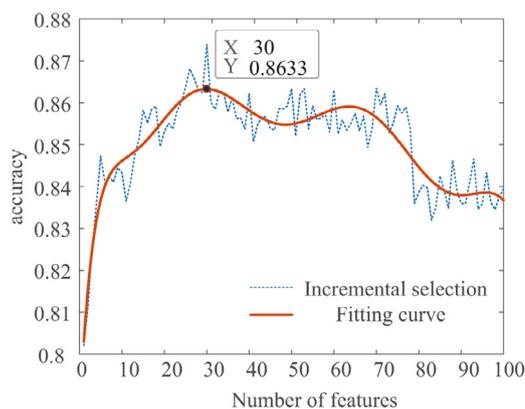
3D-HOG histograms from SPHOG for left Hippocampus. HOG4 feature extracted with cell size  $(4, 4, 4)^T$  can represent the local texture of the structural MRI image, and HOG8 with cell size  $(8, 8, 8)^T$  can reflect the global characteristic of left Hippocampus. In this way, the most discriminative features from different spatial resolutions are selected, which can achieve performance better than single resolution representation. From Table 2, we can see that 3D-HOG features with cell size smaller than 4 generally performs unstable as common characteristic cannot be analyzed in such over-detailed texture. And thus in the process of SPHOG construction, we do not use 3D-HOG with cell size 2. Furthermore, considering the time-consumption and performance about classification utilizing 3D-HOG with cell size 3, we also give it up in the combination of SPHOG. The result of 3D-HOG with cell size of 5 and 6 also illustrates that this medial cell sizes match texture information of left Hippocampus appropriately. The classification performance of HOG7 and HOG8 descends slightly shows that these large size options causes information lost in these options. SPHOG shows the best performance compared with other 3D-HOG with different cell sizes from  $S = (2, 2, 2)^T$  (i.e. HOG2) to  $S = (8, 8, 8)^T$  (i.e. HOG8).

### 3.2. Effectiveness of histogram based wrappered feature selection

With extracted SPHOG features, we compare the performance of feature selection between the proposed histogram based wrapper feature selection method and other feature selection methods. General feature selection methods mRMR, RELIEF, CSM and general wrapper-based method are utilized. And furthermore XGBoost [41], Random Forest [42] group-manner methods are also used for comparison. Two extra methods proposed more recently, a supervised method based on Gini distance statistics in [43] and an unsupervised method SOCFs [44], are utilized for comparison.

First, all of 3D-HOG features are preprocessed. Some bins with smaller variance (such as  $10^{-5}$ ) within a 3D-HOG are removed from this 3D-HOG feature, which compress bins with zeros or nearly the same for all samples. Meanwhile, the preprocess procedure can also be seen as a feature denoising part to exclude the irrelevant bins and make extracted features more robust.

Second, based on the preprocessed 3D-HOG features, direct method is used to generate final classification feature set according to incremental feature selection process. As shown in Fig. 4, this is an incremental figure of feature selection curve. Blue curve represents the classification result on incremental selection process and the red curve is the fitting result of 9th order polynomial fitting. With the fitting curve we obtain the optimal amount of histograms. In this study, there are more than  $10^3$  histograms before feature selection and approximate 30 histograms left after feature selection for histogram based wrappered feature selection method with SPHOG features. For the other feature selection methods, every bin within a 3D-HOG is used as a feature, i.e. there are lots of bins for SPHOG. And the parameters in each method is tuned with a grid search before formal experiment. In order to reduce time consumption, for general wrapper-based feature selection method,



**Fig. 4.** Incremental curve of selected features.

**Table 3**  
Performance comparison between different feature selection methods with the direct SVM classification (i.e. the top line) and the weighted classification method (i.e. the bottom line) respectively.

Name	Accuracy	TPR	FPR	Precision	F1	Kappa
Wrapper	0.852	0.813	0.116	0.767	0.788	0.685
mRMR	0.852	0.813	0.116	0.766	0.788	0.685
RELIEF	0.861	0.808	0.140	0.792	0.799	0.704
CSM	0.856	0.813	0.110	0.761	0.793	0.694
XGBoost	0.846	0.806	0.121	0.848	0.825	0.687
RF	0.837	0.859	0.163	0.839	<b>0.848</b>	0.682
Gini_Cov	0.861	0.798	0.142	0.809	0.831	0.713
Gini_Cor	<b>0.871</b>	0.800	<b>0.066</b>	0.858	<b>0.848</b>	<b>0.737</b>
SOCFS	0.807	0.829	0.235	0.796	0.774	0.653
<b>Our</b>	0.859	<b>0.875</b>	0.100	<b>0.867</b>	0.833	0.708
<b>Method</b>	0.844	0.843	0.171	0.815	0.835	0.687

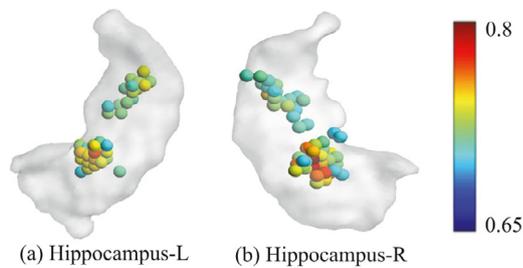
10 bins are added into the feature subset at a time for the pre-selected 2000 bins during incremental feature selection.

Finally, the direct method is used as classifiers for the selected features and we also make a comparison with the weighted method which generate AD and HC subsets respectively. Compared with general feature selection methods, histogram based wrapper method maintains the histogram structure and thus preserves the spatial information in features. As shown in Table 3, the direct method works better compared with weighted method. The proposed histogram based wrapper feature selection method performs better than the general wrapper-based method, specially, 10% higher in precision and 2.3% in Kappa coefficient. In particular, the Recall (TPR) rate increases from 0.843 to 0.875, which means the proposed method can select features with high recognition sensitivity to AD. Though the RELIEFF method obtains higher result in classification accuracy compared with histogram based wrapper method, we prefer the latter as its better performance in other indicators, which illustrates more reliable and comprehensive classification ability.

Compared with other methods, our method is still competitive. Gini distance based methods perform better especially in low FPR, same as Zhang *et al.* found in original work, but TPR is not acceptable as this means a lower identification rate in our scenes. And group based methods XGBoost and RF do not perform quite well as the dimension of extracted feature.

Moreover, bins within a histogram contains complementary information about texture information and histogram based selection contributes to the final feature in an integrated manner without fragmented bins. In this way, the risk of neglecting some informative bins is minimized.

To illustrate the effectiveness of region selection, the locations of the selected 3D-HOG in Hippocampus are illustrated in Fig. 5.



**Fig. 5.** The locations of the selected features in Hippocampus. (a) Left Hippocampus. (b) Right Hippocampus. The locations of the selected features are dotted with different color to reflect the difference of classification accuracy. The color denotes the classification accuracy from low (0.65) to high (0.811).

**Table 4**  
Performance comparison between different feature extraction methods .

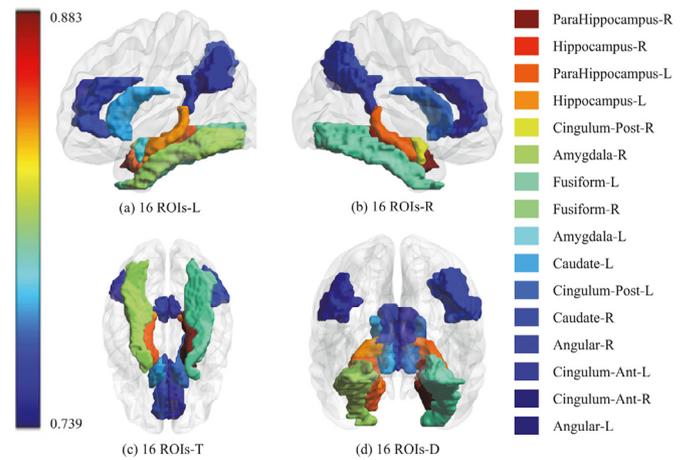
Name	Accuracy	TPR	FPR	Precision	F1	Kappa
DWT	0.836	0.867	0.188	0.801	0.829	0.672
LBP	0.838	0.867	0.184	0.802	0.829	0.675
SIFT	0.853	0.875	0.206	0.821	0.863	0.705
ELP	0.842	0.859	0.172	0.839	0.845	0.696
3D U-Net	0.842	0.847	0.163	0.856	0.840	0.683
3D ResNet	0.837	0.807	0.121	0.860	0.844	0.673
3D DenseNet	0.827	0.796	0.229	0.824	0.803	0.652
MedicalNet3D	0.846	0.835	0.130	0.849	<b>0.856</b>	0.693
<b>Our Method</b>	<b>0.859</b>	<b>0.875</b>	<b>0.100</b>	<b>0.867</b>	0.833	<b>0.708</b>

From Fig. 5, we can see that most of the selected features are located in cornu ammonis 1(CA1) and Subiculum of Hippocampus. It is noticeable that some distinct Hippocampal Subfields (such as CA1) are selected with the top 20% identification accuracy of AD, which is consistent with the clinical experience.

### 3.3. Comparison with different feature extraction methods

In this section, we compare the performance of feature extraction between the proposed SPHOG method and other feature extraction methods. For the traditional feature extraction methods, the extracted features based on 3D Bi-orthogonal Discrete Wavelet Transform (DWT) and LBP are selected using RELIEF method and classified using SVM. SIFT with automatic key point detection and scale determination [45] and Encoded Local Projection(ELP) [46] are used with wrapper selection method for comparison. For the deep learning methods, some general 3D deep learning methods are compared, such as 3D U-net [47], 3D ResNet (i.e. ResNet18 is used in this study), DenseNet [48] and 3D ResNet based transfer learning method (MedicalNet3D) [30]. The network structure of 3D U-Net is the same as described in [47]. For MedicalNet3D, the transfer learning work is similar to that described in [30], in which ResNet18 is used and fine-tuned for classification. Parameters during training are set as follows: the batch size is 32; the learning rate is 0.001 and decays to 0.5 times every 50 epoch. The stopping condition for training is set as 500 epochs or early stopping when the classification accuracy is higher than 0.97 and not increases for successive 5 epochs. We use Adaptive Movement Estimation (Adam) as optimization algorithm and Mean Square Error (MSE) as loss function during training. And a grid search is used in norm regularization to prevent over-fitting.

From the Table 4, we can see that the traditional feature extraction methods and deep learning methods achieve similar classification performance. Limited by sample size, the performance of deep learning methods is worse than the traditional methods. Compared with traditional methods, the proposed method outperforms in various indicators and in this way proves SPHOG's effectiveness in feature construction. As classifier is not the central problem in our



**Fig. 6.** The identification accuracy for different ROIs. The color denotes the classification accuracy from low (0.739) to high (0.883). (a)-(d) shows 16 ROIs of the brain with different classification accuracy from different views (i.e. left, right, top and down) respectively.

algorithm, SVM is employed for every traditional method in this experiment. Compared with deep learning methods, the proposed method achieves better classification performance and meanwhile consumes much less time and computation resources in operating process.

### 3.4. Classification results for other ROIs

According to diagnostic experience and previous studies [34,49,50], 16 distinct ROIs are selected for AD identification to show experimental results. Experimental settings are consistent with the general settings in Left Hippocampus.

As shown in Table 5, proposed method achieves excellent performance on these clinical discriminative ROIs. From Table 5 and Fig. 6, we can see that the average classification accuracy on ParaHippocampus and Hippocampus all are higher than 0.85. Furthermore, other brain regions related to cognition, such as Amygdala and Fusiform, also get distinct performance for AD diagnosis. This result is accord closely with clinic analysis that these ROIs are associated with cognitive function and illustrates the effectiveness of our algorithm.

### 3.5. Classification results on multi ROIs

With the classification performance of various brain regions in Section 3.4, a more comprehensive identification can be obtained across different brain regions based on the histogram based wrapper strategy. In particular, selected histograms from various brain region combinations are utilized for incremental feature selection and final classification. In this part, three different region combinations are used based on the result in Table 6. AHP includes Hippocampus, Amygdala and ParaHippocampus regions and has strong connection with AD. Furthermore, these three brain regions also gets best classification performance compared with other brain regions as in Table 4. Considering special performance of right Hippocampus, left and right ParaHippocampus, we also employ them as TOP3 combination.

Experiment result is shown in Table. Compared with single region classification performance, multi-region performs better and this illustrates complementary information in different brain regions. AHP-R also outperforms AHP-L, which illustrates the right laterality of brain. Moreover, TOP3 combination get best performance among three combination options, which illustrates the texture distinction property of ParaHippocampus and Hippocampus.

**Table 5**  
Classification results for different ROIs .

Name	Accuracy	TPR	FPR	Precision	F1	Kappa
Hippocampus-L	0.859	0.875	0.100	0.867	0.833	0.708
Hippocampus-R	0.880	0.850	0.094	0.884	<b>0.866</b>	0.757
Amygdala-L	0.821	0.785	0.146	0.823	0.801	0.638
Amygdala-R	0.840	0.785	0.113	0.853	0.815	0.674
ParaHipp-L	0.874	<b>0.859</b>	0.113	0.866	0.860	0.745
<b>ParaHipp-R</b>	<b>0.883</b>	0.823	<b>0.066</b>	<b>0.912</b>	0.864	<b>0.762</b>
Caudate-L	0.813	0.781	0.159	0.809	0.793	0.621
Caudate-R	0.802	0.759	0.162	0.796	0.775	0.598
Cingulum-Ant-L	0.745	0.605	0.134	0.791	0.681	0.477
Cingulum-Ant-R	0.739	0.739	0.257	0.711	0.721	0.476
Cingulum-Post-L	0.807	0.879	0.265	0.780	0.826	0.608
Cingulum-Post-R	0.827	0.867	0.221	0.816	0.841	0.650
Fusiform-L	0.823	0.718	0.091	0.864	0.782	0.635
Fusiform-R	0.822	0.789	0.148	0.822	0.802	0.640
Angular-L	0.739	0.551	0.104	0.811	0.650	0.651
Angular-R	0.786	0.634	0.089	0.852	0.723	0.556

**Table 6**  
Classification results based on multi brain regions combination .

Name	Accuracy	TPR	FPR	Precision	F1	Kappa
AHP	0.898	0.870	0.077	0.904	0.885	0.792
AHP-L	0.873	0.861	0.116	0.861	0.859	0.743
AHP-R	0.895	0.860	<b>0.075</b>	<b>0.905</b>	0.880	0.787
<b>TOP3</b>	<b>0.904</b>	<b>0.880</b>	0.086	0.897	<b>0.887</b>	<b>0.792</b>

## 4. Conclusions

In this paper, we proposed a novel feature learning method for AD identification. From spatial pyramid representation, the multi-resolution SPHOG features are constructed to distinguish the deformation characteristics of cerebral cortex comprehensively. With the proposed histogram based wrapped feature selection algorithm, the discriminative SPHOG features are selected and the feature dimensions are reduced. Experimental results show that the selected SPHOG features outperforms other 3D-HOG features with single resolution. Furthermore, some distinct Hippocampal Subfields and some ROIs associated with cognitive function show high identification accuracy of AD, which is consistent with the clinical analysis.

Considering for future improvement, our method mainly focus on a general explainable framework for AD identification and in each step more specific work can be done. More informative feature extraction methods and feature selection process can strengthen identification performance. Moreover, with current detected volumes more exploration can be done in both CAD and clinical fields to analyze these sensitive regions. That could be useful for further medical analysis to find distinct subfields of ROIs of the brain for early diagnosis of AD.

## Declaration of Competing Interest

The authors declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work. There is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

## References

- [1] B. Fischl, A.V.D. Kouve, C. Destrieux, E. Halgren, A.M. Dale, Automatically parcellating the human cerebral cortex, *Cerebral Cortex* 14 (1) (2004) 11–22.
- [2] H. Kilian, T.A. Vinh-Thong, M.J. V. C. Pierrick, Adaptive fusion of texture-based grading for Alzheimer's disease classification, *Computerized Medical Imaging and Graphics* 70 (2018) 8–16.
- [3] S. Chaplot, L.M. Patnaik, N.R. Jagannathan, Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network, *Biomed Signal Process Control* 1 (1) (2006) 86–92.
- [4] Y. Liu, L. Wang, X. Zeng, Compartmental sparse feature selection method for Alzheimer's disease identification, *Conf Proc IEEE Eng Med Biol Soc* 2017 (2017) 3073–3076.
- [5] L. Wang, Y. Liu, X. Zeng, H. Cheng, Q. Wang, Region-of-interest based sparse feature learning method for Alzheimer's disease identification, *Comput Methods Programs Biomed* 187 (2019) 1–10.
- [6] H. Bhasin, R.K. Agrawal, A combination of 3-d discrete wavelet transform and 3-d local binary pattern for classification of mild cognitive impairment, *BMC Med Inform Decis Mak* 20 (1) (2020).
- [7] D. Sarwinda, A. Bustamam, 3D-hog features-based classification using mri images to early diagnosis of Alzheimer's disease, in: 17th IEEE/ACIS International Conference on Computer and Information Science, ICIS, 2018, pp. 457–462.
- [8] X. Pan, M. Adel, C. Fossati, T. Gaidon, J. Wojak, E. Guedj, Multiscale spatial gradient features for 18f-fdg pet image-guided diagnosis of Alzheimer's disease, *Comput Methods Programs Biomed* 180 (2019).
- [9] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [10] A. Albarak, F. Coenen, Y. Zheng, Volumetric image classification using homogeneous decomposition and dictionary learning: a study using retinal optical coherence tomography for detecting age-related macular degeneration, *Computerized Medical Imaging and Graphics* 55 (2017) 113–123.
- [11] M. Erdt, O. Knapp, K. Drechsler, S. Wesarg, Region detection in medical images using hog classifiers and a body landmark network, volume 8670, 2013, pp. 1–7. USA
- [12] X. Zhu, H.-I. Suk, Y. Zhu, K.-H. Thung, G. Wu, D. Shen, Multi-view classification for identification of alzheimer's disease, in: L. Zhou, L. Wang, Q. Wang, Y. Shi (Eds.), *Machine Learning in Medical Imaging*, Springer International Publishing, Cham, 2015, pp. 255–262.
- [13] C. Chu, A.L. Hsu, K.H. Chou, P. Bandettini, C.P. Lin, Does feature selection improve classification accuracy? impact of sample size and feature selection on classification using anatomical magnetic resonance images, *Neuroimage* 60 (1) (2012) 59–70.
- [14] Y. Fan, H. Rao, H. Hurt, J. Giannetta, M. Korczykowski, D. Shera, B.B. Avants, J.C. Gee, J. Wang, D. Shen, Multivariate examination of brain abnormality using both structural and functional mri, *Neuroimage* 36 (4) (2007) 1189–1199.
- [15] Y. Zhang, S. Wang, Detection of Alzheimer's disease by displacement field and machine learning, *PeerJ* 3 (Supplement 1) (2015) 1–29.
- [16] Y.D. Zhang, S.H. Wang, Z.C. Dong, Classification of Alzheimer's disease based on structural magnetic resonance imaging by kernel support vector machine decision tree, *PROG ELECTROMAGN RES* (2014) 171–184.
- [17] E. Adeli-Mosabbeh, K.-H. Thung, L. An, F. Shi, D. Shen, Robust feature-sample linear discriminant analysis for brain disorders diagnosis, volume 2015-January, 2015, pp. 658–666.
- [18] Z. Wang, Y. Zheng, D. Zhu, A. Bozoki, T. Li, Classification of Alzheimer's disease, mild cognitive impairment and normal control subjects using resting-state fmri based network connectivity analysis, *IEEE J Transl Eng Health Med* 6 (2018) 1–9.
- [19] S. Alam, G.-R. Kwon, Alzheimer's disease classification using kpca, lda, and multi-kernel learning svm, *Int J Imaging Syst Technol* 27 (2) (2017) 133–143.
- [20] J. Peng, X. Zhu, Y. Wang, L. An, D. Shen, Structured sparsity regularized multiple kernel learning for Alzheimer's disease diagnosis, *Pattern Recognit* 88 (2019) 370–382.
- [21] L. Wang, N. Zhou, F. Chu, A general wrapper approach to selection of class-dependent features, *IEEE Trans. Neural Networks* 19 (7) (2008) 1267–1278.
- [22] H. Peng, F. Long, C. Ding, Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy, *IEEE Trans Pattern Anal Mach Intell* (2005).

- [23] K. Kira, L.A. Rendell, The feature selection problem: traditional methods and a new algorithm, in: Tenth National Conference on Artificial Intelligence, 1992.
- [24] X. Fu, L. Wang, Data dimensionality reduction with application to simplifying RBF network structure and improving classification performance, *Systems Man and Cybernetics Part B Cybernetics IEEE Transactions on* 33 (3) (2003) 399–409.
- [25] Y. Liu, X. Tian, X. Zeng, L. Wang, Z. Wang, Q. Wang, A wrapper-based feature learning method used for Alzheimer's disease identification, in: The 15th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 18–22 Dec., 2020.
- [26] D. Jin, J. Xu, K. Zhao, F. Hu, Z. Yang, B. Liu, T. Jiang, Y. Liu, Attention-based 3D convolutional network for Alzheimer's disease diagnosis and biomarkers exploration, 2019, pp. 1047–1051. Piscataway, NJ, USA
- [27] F. Gao, T. Wu, X. Chu, H. Yoon, Y. Xu, B. Patel, Deep residual inception encoder-decoder network for medical imaging synthesis, *IEEE J Biomed Health Inform* 24 (1) (2020) 39–49.
- [28] A. Sikka, S.V. Peri, D.R. Bathula, Mri to FDG-PET: Cross-modal synthesis using 3D U-net for multi-modal Alzheimer's classification, volume 11037 LNCS, 2018, pp. 80–89.
- [29] R. Cui, M. Liu, Hippocampus analysis by combination of 3-d densenet and shapes for Alzheimer's disease diagnosis, *IEEE J Biomed Health Inform* 23 (5) (2019) 2099–2107.
- [30] S. Chen, K. Ma, Y. Zheng, Med3D: Transfer learning for 3D medical image analysis (2019).
- [31] M. Tanveer, B. Richhariya, R. Khan, A. Rashid, P. Khanna, M. Prasad, C. Lin, Machine learning techniques for the diagnosis of Alzheimer's disease: a review, *ACM Transactions on Multimedia Computing, Communications and Applications* 16(1s)
- [32] J.M. Mateos-Pérez, M. Dadar, M. Lacalle-Aurioles, Y. Iturria-Medina, Y. Zeighami, A.C. Evans, Structural neuroimaging as clinical predictor: a review of machine learning applications, *NeuroImage: Clinical* (2018) 506–522.
- [33] C. Perone, J. Cohen-Adad, Promises and limitations of deep learning for medical image segmentation, *Journal of Medical Artificial Intelligence* 2 (2019). 1–1
- [34] Z. Ding, L. Wang, Y. Liu, X. Zeng, Z. Wang, Spatial pyramid based 3d hog feature extraction for Alzheimer's disease identification, in: The 15th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 18–22 Dec., 2020.
- [35] K. Smith Friston, Statistical parametric maps in functional imaging: a general linear approach, *Hum Brain Map* 2 (1995).
- [36] J. Maldjian, P. Laurienti, R. Kraft, J. Burdette, An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of FMRI data sets, *Neuroimage* 19 (2003) 1–9, doi:10.1016/S1053-8119(03)00169-1.
- [37] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot, Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain, *Neuroimage* 15 (1) (2002) 273–289.
- [38] T. Harada, Y. Ushiku, Y. Yamashita, Y. Kuniyoshi, Discriminative spatial pyramid, in: The 24th IEEE Conference on Computer Vision and Pattern Recognition, CVPR, Colorado Springs, CO, USA, 20–25 June, 2011.
- [39] G. Sharma, F. Jurie, C. Schmid, Discriminative spatial saliency for image classification, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 3506–3513.
- [40] M.P. Baggenstoss, The pdf projection theorem and the class-specific method, *IEEE Trans. Signal Process.* 51 (3) (2003) 672–685.
- [41] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
- [42] L. Breiman, Random forests, *Mach Learn* 45 (1) (2001) 5–32.
- [43] S. Zhang, X. Dang, D. Nguyen, D. Wilkins, Y. Chen, Estimating feature-label dependence using Gini distance statistics, *IEEE Trans Pattern Anal Mach Intell* (2019).
- [44] D. Han, J. Kim, Unsupervised simultaneous orthogonal basis clustering feature selection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 5016–5023.
- [45] B. Rister, bbrister/sift3d, 2021 <https://github.com/bbrister/SIFT3D>.
- [46] H.R. Tizhoosh, M. Babaie, Representing medical images with encoded local projections, *IEEE Trans. Biomed. Eng.* 65 (10) (2018) 2267–2277.
- [47] O. Cicek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-net: Learning dense volumetric segmentation from sparse annotation, volume 9901 LNCS, 2016, pp. 424–432.
- [48] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
- [49] A.J. Bastos Leite, P. Scheltens, F. Barkhof, Pathological aging of the brain: an overview, *Topics in Magnetic Resonance Imaging TMRI* 15 (6) (2004) 369–389.
- [50] G. Chetelat, B. Desgranges, L.S.V. De, F. Viader, F. Eustache, J.C. Baron, Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment, *Neuroreport* 13 (15) (2002) 1939–1943.