Deep Learning on SDF for Classifying Brain Biomarkers

Zhangsihao Yang[1], Jianfeng Wu[1], Paul M Thompson[2], and Yalin Wang[3]

Abstract— Biomarkers are one of the primary medical signs to facilitate the early detection of Alzheimer's disease. The small beta-amyloid (Aβ) peptide is an important indicator for the disease. However, current methods to detect Aβ pathology are either invasive (lumbar puncture) or quite costly and not widely available (amyloid PET). Thus a less invasive and cheaper approach is demanded. MRI which has been used widely in preclinical AD has recently shown the capability to predict brain Aβ positivity. This motivates us to develop a method, SDF sparse convolution, taking MRI to predict Aβ positivity. We obtain subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and use our method to discriminate Aβ positivity. Theoretically, we provide analysis towards the understanding of what the network has learned. Empirically, it shows strong performance on par or even better than state of the art.

I. INTRODUCTION

By 2050, 1 out of 85 people worldwide will be affected by Alzheimer's disease (AD) [3]. Degenerative diseases have put a great burden on worldwide healthcare systems in terms of cost and therapies [22]. The small beta-amyloid (Aβ) peptide is one of the major histopathological hallmarks of the disease [10] and found promising for early detection of AD. Current amyloid measures are either invasive or too expensive. Even Blood-based biomarkers (BBBs) are developed for screening Alzheimer's disease, only weak or no correlation is found with the estimate of brain amyloid positivity and [2] BBBs' usefulness for differential diagnosis and prognosis is unclear [11]. On the other hand, sMRI is widely available in clinical practice. Therefore, there is a strong interest to develop new techniques which predict amyloid burden with sMRI analysis.

Recently, prediction of brain amyloid with structural MRI received growing interest because of its wide availability and non-invasive nature. Traditionally, brain image biomarkers [25] use cortical or sub-cortical structure volume. Recent works demonstrate that surface-based brain imaging biomarkers ourperforms such volume measures because it overcomes partial volume effects. A example is the radial distance mapping extracted on biomarkers is applied to segment the hippocampus [1].

Meanwhile, deep model succeeded in many areas [7], including image recognition [24], 3D object classification and segmentation [19], and biomedical segmentation [21].

Fig. 1. **The visual comparisons among different data formats to represent a 3D object.** (b) shows the vertices on the mesh in (a). (c) shows the points sampled from vertices in (b). Details of sampling points from vertices are in section II-B.2. For (d) and (e), the red dots present the points outside of the mesh, and the blue dots present the interior points. (e) shows the cross-sections of (d) along the axis perpendicular to the paper.

However, due to different structure, regular grid based conventional CNN is hard to be directly applied to surface manifold. Most of pioneering work are designed on point cloud [20], [19]. This data format provides efficiency while it loses some intrinsic geometry information which is demanded in discovering subtle brain shape features.

In this work, we propose SDF sparse convolution to address these problems. The inspiration comes from the success of using deep neural network generating SDF and the inherent sparse nature of SDF. The ability to generate high quality 3D objects with SDF demonstrates that the deep neural network has the potential to analyzing SDF. And using sparse convolution to analyze SDF is a feasible approach. This motivates us to introduce this work which is designed to parse SDF instead of other 3D data formats for brain diagnosis. Our work has two contributions. (1) To our knowledge, this is the first paper using deep neural network to classify SDF on brain image biomarkers. (2) We propose and demonstrate an effective and standard schedule for using deep learning for 3D biomarkers.

II. METHOD

We take the hippocampus as a composition of two 2D manifolds (figure 1). Our work is to extract features on the signed distance field (SDF) of the hippocampus with sparse convolution [8], [9] to classify it. We first propose our SDF sparse convolution and the network architecture that could learn features on a field instead of a 2D manifold. Finally, we briefly explain two baselines used to classify the hippocampus, a commonly-used method that leverages the hippocampus' volume, and a state-of-the-art 3D object recognition deep learning approach.

[1]Zhangsihao Yang `zyang195@asu.edu` and Jianfeng Wu `jianfen6@asu.edu` are graduate students of School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, 699 S. Mill Ave. Tempe, AZ 85281, USA

[2]Paul M Thompson is a faculty of USC Institute for Neuroimaging and Informatics, ISI, 4676 Admiralty Way, Ste. 200 Health Sciences Campus, Zonal Avenue, Biggy St, Los Angeles, CA 90033, USA `pthomp@usc.edu`

[3]Yalin Wang is a faculty of School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, 699 S. Mill Ave. Tempe, AZ 85281, USA `ylwang@asu.edu`
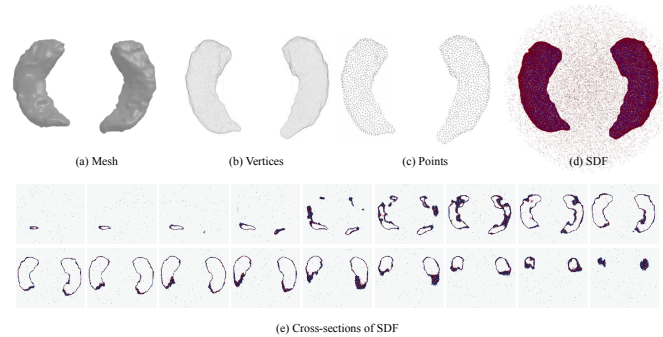
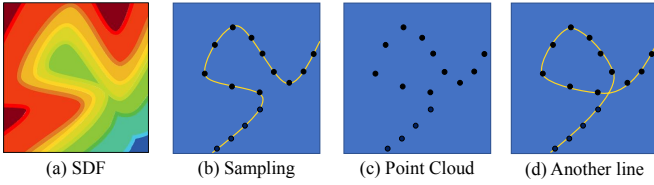(a) SDF (b) Sampling (c) Point Cloud (d) Another line

Fig. 2. **Comparison between SDF and Point Cloud.** (a) is an example of SDF, in which the yellow line represents the surface. The color in (a) shows the distance from the surface. (b) shows sampling a point cloud (c) from the surface. (d) is another possible underlying surface reconstructed from (c).

A. SDF Sparse Convolution

1) Signed Distance Field: A signed distance field is one of the representations to implicitly define the ℓ-level set as a Lipschitz surface $S \in \mathbb{R}^3$.

$$S = \{p \in \mathbb{R}^3 | f_S(p) = \ell\} \tag{1}$$

f_S is a function that maps $\mathbb{R}^3$ to $\mathbb{R}$. There are two popular implicit functions, occupancy function f_S^{OCC}, and signed distance function f_S^{SDF}.

$$f_S^{OCC}(p) = \begin{cases} 1 & p \in S_{int} \\ 0 & p \notin S_{int} \end{cases} \tag{2}$$

$$f_S^{SDF}(p) = (2f_S^{OCC}(p) - 1)dist_S(p) \tag{3}$$

The signed distance function consists of the occupancy function and the distance function $dist_S(p)$ which computes the shortest Euclidean distance from point p to the surface S. Also, using algorithms like ray casting [13] or marching cubes [16] could retrieve the surface from the implicit function.

Some works [5], [17], [14] use deep neural networks to generate 3D shapes via a implicit function. They have achieved competitive performances on the 3D shape generation. However, analyzing SDF remains an open problem. In figure 2, we show the difference between SDF and point cloud in analyzing the underlying surface. Moreover, using a point cloud could lead to a potential misunderstanding of the underlying surface. Therefore, providing an algorithm to probe SDF can be thought of as exploiting the property of SDF to disentangle the ambiguity.

2) SDF Sparse Convolution: As (e) shown in figure 1, the sampled points from the SDF only occupy a small partition of the field. The data represents SDF is inherently sparse. While comparing to voxel data [28], it provides more neighboring information. Sparse convolutional operation is customized for these sparse data and could improve the computational efficiency comparing to standard convolution applied on dense data [9]. Meanwhile, it could also prevent from the dilating observation in every layer.

In figure 3, we compare the basic feature extracting operations in PointNet, sparse convolution, and SDF sparse convolution. The operations we discussed on 2D space could also be applied to a higher dimension. The bottom building block of PointNet++ shares the same structure as PointNet.

$$f(p_1, p_2, ..., p_n) = \gamma \left(\max_{i=1,...,n} \{h(p_i)\} \right) \tag{4}$$
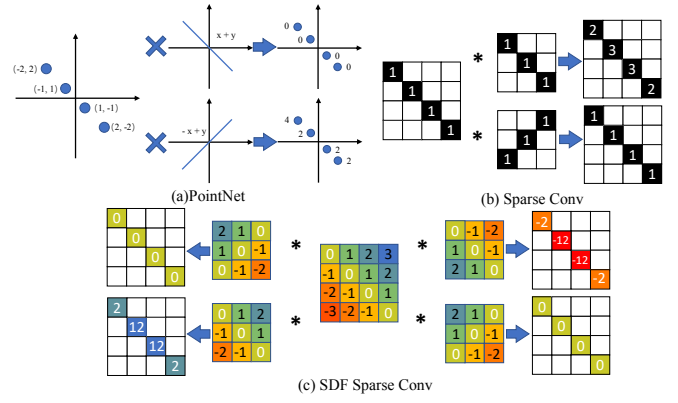

(a)PointNet (b) Sparse Conv

(c) SDF Sparse Conv

Fig. 3. **The computation process of different feature extractors.** (a) is the core computation process for one layer inside PointNet and PointNet++. (b) presents the sparse convolution on a 2D grid. (c) shows the process of computing features on Signed Distance Field using SDF Sparse Convolution. And SDF Sparse Conv has the power of distinguishing the inside and outside of a surface.

The function f in equation 4 maps a unordered set of points with $p_i \in \mathbb{R}^3$ to a vector. γ and h are multi-layer perceptron (MLP) networks. In PointNet, the feature extraction is based on linear projection. The output feature would be smaller if the kernel is similar to the input point cloud and vice versa. Because what PointNet does is to sum the projected distance of the input point cloud onto the kernel ((a) in figure 3). This means that in order to learn variant features the learned kernel should store as much dissimilar kernels compared to the input pattern as possible. For sparse convolution and SDF sparse convolution, the effect is different from PointNet. A similar pattern of the learned kernel relative to the input would generate larger (more significant) output features.

$$[f \circledast g](p) = \sum_{q=q_0}^{q_k} f(q) \cdot g(p - q) \tag{5}$$

The set $q_0, q_1, ..., q_k$ contains the neighboring points of p with a signal signed to it. Not all the neighboring points of p are involved in the computation is the key to sparse convolution. The function f is the kernel function and the function g is the input function. Equation 6 is the function g used in SDF sparse convolution and sparse convolution. Thus the number of patterns to be learned from the network would be potentially smaller comparing to PointNet and PointNet++. As shown in figure 3 (b) and (c), comparing to conventional sparse convolution, SDF sparse convolution absorbs information from neighboring. Therefore, SDF sparse convolution provides more information, such as the location of the surface, the normal direction of the surface, and the connection between sparse points, to enrich feature extraction on 3D objects.

$$g(p) = f_S^{SDF}(p) \quad \text{and} \quad g(p) = 1 \tag{6}$$

3) Data Processing and Neural Network Architecture: We sample 250000 points in a spherical space of the mesh (figure 1 (d)) with 10% uniformly sampled in the sphere and 90% sampled near the surface with perturbation of points on
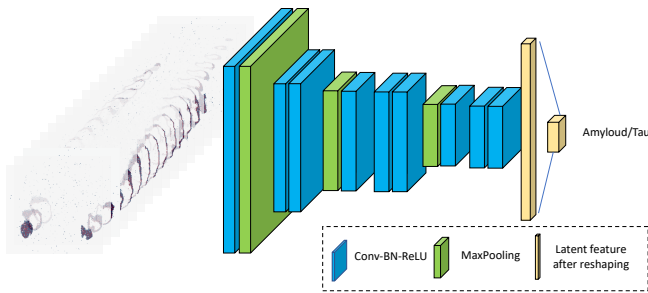
Fig. 4. **The architecture of SDF Sparse Convolutional Network.** The right part represent the SDF. The output of the network is a 2 dimension vector represent the probability of Amyliod and Tau.

the surface by with a Gaussian $\mathcal{N}(0, 1)$ centered at 0 with 1 mm variance.

We implement the feature extractor as a VGG-like [24] structure neural network. The detailed structure of the network is shown in figure 4. The loss functions we use for training PointNet++ and SDF sparse convolution is binary cross-entropy loss.

We train the model from scratch by optimizing the cross-entropy loss function using stochastic gradient descent for 100 epochs. In all our experiments, we use the Adam optimizer [12] with a fixed learning rate of 10^{-4}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$. Besides, the batch size is set to 4 and 10000 sampled points per object.

B. Baselines

1) Volume: After processing the meshes into watertight [26], we compute the volume of meshes. Then we combine Neighborhood Components Analysis (NCA) [6] with k-nearest neighbors classifier [23] to classify the features.

2) PointNet++: Based on a pioneer work PointNet [20], PointNet++ [19] achieves better results on recognition of a point cloud than PointNet. To train PointNet++, we first extract 60,000 vertices, (b) in figure 1, on the hippocampus mesh. During training, we use Furthest Point Sampling (FPS) [15] to sample 2500 points on these vertices to feed into PointNet++.

III. EXPERIMENTS

A. Subjects

We use the ADNI database [18] for testing the performance of baselines and our algorithm. We divide the combined database into five classes according to the stage and the positivity of amyloid biomarkers. The number of each class and the number of training, validation, and test subjects are listed in table I. We split the data set with the ratio of 80% for training and 20% for testing. We further split the training data into 80% to train the network and 20% for validation.

B. Training conditions

The hippocampus meshes have different properties from the 3D objects in ShapeNet [4] or ModelNet [27]. One aspect is the scale. 3D Objects in ShapeNet or ModelNet are scale-invariant. But for hippocampus meshes, the scale is an important criterion to judge whether the brain is suffering atrophy. Another aspect is the data augmentation during

Group	Total number	Training	Validation	Testing
AD Aβ+	151	96	24	31
MCI Aβ+	171	109	27	35
MCI Aβ-	171	109	27	35
CU Aβ+	116	74	18	24
CU Aβ-	232	148	37	47

TABLE I

THE NUMBER OF SUBJECTS CONTAINED IN OUR DATABASE.

Training data processing	Validation	Test
unit-scale & data augmentation (da)	80.3	62.8
unit-scale & no da	85.2	75.6
constant-scale & da	78.7	71.8
constant-scale & no da	85.2	**78.2**

TABLE II

DIFFERENT TRAINING DATA PROCESSING AND THEIR CORRESPONDING ACCURACY. UNIT-SCALE MEANS ALL THE MESH DATA HAS BEEN NORMALIZED INTO A UNIT SPHERE. CONSTANT-SCALE MEANS THAT THE ORIGINAL RELATIVE SCALE IS KEPT.

training. There are three prevalent data augmentation methods, randomly scale, randomly rotate along the gravity axis, and small perturb of the locations. Each of them could increase the robustness of a 3D neural network.

To decide how to utilize these two aspects, we have done some experiments. The results are shown in the table II. We compare the accuracy on validation and test with different data-preprocessing and data augmentation during training. The results show that without having scale information and with data augmentation during the training the accuracy decreases. For a tissue, like a hippocampus, randomly scaling would obstruct the detection of atrophy, the rotation would impede the already-done registration step, and perturbing would destroy the local information for understanding the tissue. From now on we have set all the training with the same setting, keeping the relative scale and training without rotation and perturbation augmentation.

C. Improvement over baselines

In table III, we compare the accuracy over three groups and three methods. We split the data set into 10 folds and take the average test accuracy of each method in each group. Our proposed SDF sparse convolution Network outperforms 19 folds out of 30 folds over the other two methods.

D. Ablation study

To this end, we utilize group 1 as the benchmark to compare the performance of different choices. We recognize two important aspects of the network design, the input data formats, and network operations. When using PointNet++ to analyze SDF, we randomly sample 2500 points from 250000 points in SDF to feed into the network. When using sparse convolution to analyze a point cloud, we sample 10000 points from vertices of a mesh to feed into the network. In table IV, the results show that PointNet++ could not exploit SDF information comparing to the point cloud even SDF could provide more information than a point cloud. And the performance of sparse convolution drops from 77.2% to 70.9%

Group	AD Aβ+ vs CU Aβ-			MCI Aβ+ vs MCI Aβ-			CU Aβ+ vs CU Aβ-		
Fold	V	P++	SDF SC	V	P++	SDF SC	V	P++	SDF SC
0	76.9	85.2/78.2	83.6/76.9	51.4	72.2/57.1	68.5/48.6	66.2	69.1/60.6	67.3/66.2
1	71.8	85.2/80.8	86.9/71.8	44.3	63.0/61.4	72.2/54.3	54.9	71.0/64.8	69.1/66.2
2	70.5	85.2/69.2	86.9/78.2	52.9	70.9/50.0	61.1/61.4	62.0	67.3/66.2	67.3/66.2
3	71.8	82.0/76.9	78.7/82.1	45.7	75.9/55.7	66.7/57.1	60.6	67.3/66.2	69.1/63.4
4	70.5	86.9/70.5	85.2/75.6	42.9	72.2/52.9	66.7/54.3	64.8	72.7/62.0	67.3/66.2
5	66.7	86.9/80.8	83.6/76.6	52.9	63.0/61.4	61.1/51.4	60.6	67.3/66.2	70.9/66.2
6	67.9	83.6/79.5	78.7/76.9	45.7	79.6/64.3	72.2/54.3	60.6	70.9/63.4	67.3/66.2
7	69.2	86.9/80.8	86.9/79.5	45.7	66.7/61.4	66.7/61.4	57.7	76.4/60.6	69.1/64.8
8	70.5	85.2/71.8	82.0/74.4	40.0	72.2/65.7	68.5/58.6	63.4	69.1/60.6	67.3/66.2
9	66.7	82.0/76.9	77.0/79.5	54.3	61.1/65.7	59.3/62.9	63.4	70.9/57.7	67.3/66.2
m/s	70.3±2.9	76.5±4.2	**77.2±2.7**	47.6±4.7	**59.6±5.2**	56.4±4.4	61.4±3.2	62.8±2.8	**65.8±0.9**

TABLE III

THE COMPARISON AMONG DIFFERENT GROUPS. V REPRESENTS THE METHOD USING VOLUME (DETAIL IN SECTION II-B.1). P++ REPRESENTS POINTNET++ (DETAIL IN SECTION II-B.2) AND SDF SC IS SDF SPARSE CONVOLUTION (DETAIL IN SECTION II-A).

PointNet++	SDF PointNet++	Sparse Convolution	SDF Sparse Convolution
76.5±4.2	69.5±4.9	70.9±5.0	**77.2±2.7**

TABLE IV

ABLATION STUDY ON DIFFERENT SETTINGS. THE EXPERIMENTS ARE DONE ON THE AD Aβ+ VS CU Aβ- GROUP WITH 10 FOLDS.

when only takes a point cloud is the input to the network. The combination between SDF and sparse convolution performs best in all combination settings.

IV. CONCLUSION AND FUTURE WORK

In summary, our work is the first one using SDF-based deep neural network to classify brain amyloid burdens. The combination between signed distance field and sparse convolution provides a new way of analyzing brain images. We propose and demonstrate an effective and standard schedule for using deep learning for 3D manifold data. In the future, with more data being collected, the gap between deep learning methods and traditional methods could become larger and the performance drop between validation accuracy and test accuracy could be smaller.

REFERENCES

[1] Liana G Apostolova, Jonathan H Morra, Amity E Green, Kristy S Hwang, Christina Avedissian, Ellen Woo, Jeffrey L Cummings, Arthur W Toga, Clifford R Jack Jr, Michael W Weiner, et al. Automated 3d mapping of baseline and 12-month associations between three verbal memory measures and hippocampal atrophy in 490 adni subjects. *Neuroimage*, 51(1):488–499, 2010.

[2] Randall J Bateman, Kaj Blennow, Rachelle Doody, Suzanne Hendrix, Simon Lovestone, Stephen Salloway, Rachel Schindler, Michael Weiner, Henrik Zetterberg, P Aisen, et al. Plasma biomarkers of ad emerging as essential tools for drug development: an eu/us ctad task force report. *The journal of prevention of Alzheimer's disease*, 6(3):169–173, 2019.

[3] Ron Brookmeyer, Elizabeth Johnson, Kathryn Ziegler-Graham, and H Michael Arrighi. Forecasting the global burden of alzheimer's disease. *Alzheimer's & dementia*, 3(3):186–191, 2007.

[4] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*, 2015.

[5] Zhiqin Chen and Hao Zhang. Learning implicit fields for generative shape modeling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5939–5948, 2019.

[6] Jacob Goldberger, Geoffrey E Hinton, Sam Roweis, and Russ R Salakhutdinov. Neighbourhood components analysis. *Advances in neural information processing systems*, 17:513–520, 2004.

[7] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*, volume 1. 2016.

[8] Benjamin Graham, Martin Engelcke, and Laurens van der Maaten. 3d semantic segmentation with submanifold sparse convolutional networks. *CVPR*, 2018.

[9] Benjamin Graham and Laurens van der Maaten. Submanifold sparse convolutional networks. *arXiv preprint arXiv:1706.01307*, 2017.

[10] Bradley T Hyman. Amyloid-dependent and amyloid-independent stages of alzheimer disease. *Archives of neurology*, 68(8):1062–1064, 2011.

[11] Shorena Janelidze, Niklas Mattsson, Sebastian Palmqvist, Ruben Smith, Thomas G Beach, Geidy E Serrano, Xiyun Chai, Nicholas K Proctor, Udo Eichenlaub, Henrik Zetterberg, et al. Plasma p-tau181 in alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to alzheimer's dementia. *Nature medicine*, 26(3):379–386, 2020.

[12] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[13] Aaron Knoll, Younis Hijazi, Andrew Kensler, Mathias Schott, Charles Hansen, and Hans Hagen. Fast ray tracing of arbitrary implicit surfaces with interval and affine arithmetic. In *Computer Graphics Forum*, volume 28, pages 26–40. Wiley Online Library, 2009.

[14] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4460–4470, 2019.

[15] Carsten Moenning and Neil A Dodgson. Fast marching farthest point sampling. Technical report, University of Cambridge, Computer Laboratory, 2003.

[16] Timothy S Newman and Hong Yi. A survey of the marching cubes algorithm. *Computers & Graphics*, 30(5):854–879, 2006.

[17] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 165–174, 2019.

[18] Ronald Carl Petersen, PS Aisen, Laurel A Beckett, MC Donohue, AC Gamst, Danielle J Harvey, CR Jack, WJ Jagust, LM Shaw, AW Toga, et al. Alzheimer's disease neuroimaging initiative (adni): clinical characterization. *Neurology*, 74(3):201–209, 2010.

[19] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*, 2017.

[20] Charles Ruizhongtai Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. *CoRR*, abs/1612.00593, 2016.

[21] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[22] Christian Salvatore, Antonio Cerasa, Petronilla Battista, Maria Carla Gilardi, Aldo Quattrone, and Isabella Castiglioni. Magnetic resonance imaging biomarkers for the early diagnosis of alzheimer's disease: a machine learning approach. *Frontiers in neuroscience*, 9:307, 2015.

[23] Bernard W Silverman and M Christopher Jones. E. fix and jl hodges (1951): An important contribution to nonparametric discriminant analysis and density estimation: Commentary on fix and hodges (1951). *International Statistical Review/Revue Internationale de Statistique*, pages 233–238, 1989.

[24] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[25] Kyle Strimbu and Jorge A Tavel. What are biomarkers? *Current Opinion in HIV and AIDS*, 5(6):463, 2010.

[26] Jianfeng Wu, Qunxi Dong, Jie Gui, Jie Zhang, Yi Su, Kewei Chen, Paul M Thompson, Richard J Caselli, Eric M Reiman, Jieping Ye, et al. Predicting brain amyloid using multivariate morphometry statistics, sparse coding, and correntropy: Validation in 1,125 individuals from the adni and oasis databases. *bioRxiv*, 2020.

[27] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1912–1920, 2015.

[28] Yu Xiang, Wongun Choi, Yuanqing Lin, and Silvio Savarese. Data-driven 3d voxel patterns for object category recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1903–1911, 2015.