Texture and Structure Incorporated ScatterNet Hybrid Deep Learning Network (TS-SHDL) For Brain Matter Segmentation

Amarjot Singh
Signal Processing Group
Department of Engineering
University of Cambridge, U.K.
as2436@cam.ac.uk

Devamanyu Hazarika
School of Computing
National University of Singapore
Singapore
devamanyu@comp.nus.edu.sg

Aniruddha Bhattacharya
Department of Computer Science
National Institute of Technology
Warangal, India
baniruddha2@student.nitw.ac.in

Abstract

Automation of brain matter segmentation from MR images is a challenging task due to the irregular boundaries between the grey and white matter regions. In addition, the presence of intensity inhomogeneity in the MR images further complicates the problem. In this paper, we propose a texture and vesselness incorporated version of the ScatterNet Hybrid Deep Learning Network (TS-SHDL) that extracts hierarchical invariant mid-level features, used by fisher vector encoding and a conditional random field (CRF) to perform the desired segmentation. The performance of the proposed network is evaluated by extensive experimentation and comparison with the state-of-the-art methods on several 2D MRI scans taken from the synthetic McGill Brain Web as well as on the MRBrainS dataset of real 3D MRI scans. The advantages of the TS-SHDL network over supervised deep learning networks is also presented in addition to its superior performance over the state-of-the-art.

1. Introduction

Brain matter segmentation is an important task that is essential for the study of various ailments like Alzheimer and Parkinson disease [18] etc. It is typically done by skilled professionals who manually label white and grey matter in the Magnetic Resonance (MR) images of the brain. The images also often contain noise incurred through acquisition defects and errors [37] that further adds to the complexity of the problem. Manual labeling leads to errors and

inconsistencies thus requiring the need for automatic labeling. This task has been performed in three different ways over the years.

The first class of methods utilized hand-engineered features extracted using filters designed to capture edge and texture representations [13, 1, 9, 30]. For example, SIFT [19] features were used in the application of reconstructing tractograms of whole brain [34]. These hand-engineered features are easy to design but are constrained by marginally good performance.

The second class of methods used deep networks for brain matter segmentation as demonstrated in multiple works [4, 11]. Deep networks have achieved the state-of-the-art performance by learning discriminative class-specific features. However, deep networks suffer from overwhelming complexity of trainable parameters and design of optimal configurations. Thus, training with limited sized datasets, common in medical imaging, still remains a difficult task.

The third class of methods used hand-crafted features to first extract low-level edge features which are further used by supervised or unsupervised learning methods to learn higher levels of representations, thus resulting in a hybrid network. For example, Zhang et al. [40] extracted anisotropic features which were used by a CNN to learn higher-level representations. Given the reduced parameter set to be learned while training, these networks generally perform better than normal deep neural networks on small datasets with less training instance, a trait always seen in brain segmentation datasets.

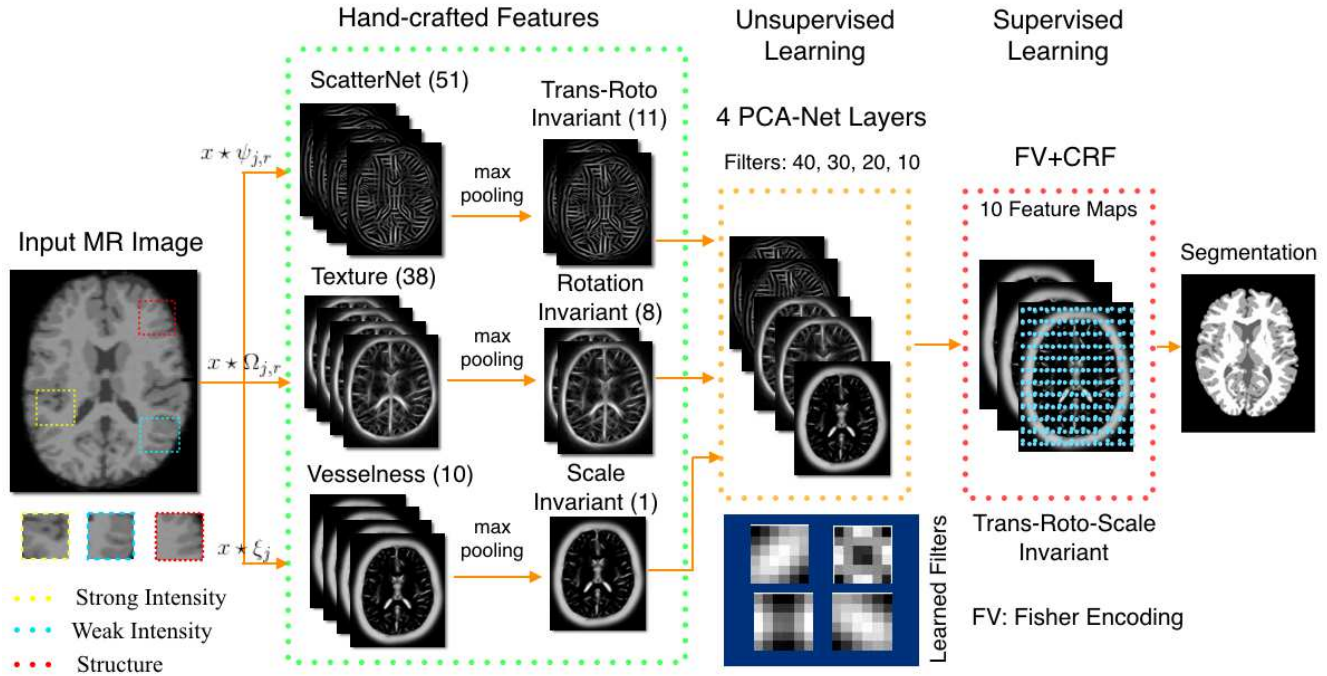Singh et. al. [33] recently proposed a ScatterNet Hybrid

Figure 1. The Illustration shows the proposed TS-SHDL network that captures Intensity (region boundaries and texture) and structural (ridge like structures) properties of the brain matter with handcrafted (ScatterNet, Vesselness, and Texture) features which are concatenated and used by four stacked layer of the PCANet to learn 40, 30, 20 and 10 filters at the first, second, third and fourth layers respectively, of size 5×5. The filters capture a hierarchy of translation, rotation, and scale-invariant mid-level features. The invariant mid-level features are encoded into a dense and compact representation Fisher vector encoding of 500 dimensions using 5 Gaussian mixtures. The encoded features are used by the CRF to produce the desired segmentation.

Deep Learning (SHDL) network that first extracts Scatter-Net handcrafted low-level descriptors which are used by two layers of PCA-Net based unsupervised learning module to learn hierarchical object specific mid-level features. The mid-level features are finally used by a supervised module composed of orthogonal least squares and support vector machine (SVM) to perform classification. Each layer of the SHDL network is designed and optimized automatically using cross-validation that produces a computationally efficient architecture.

This paper proposes an improved texture and vesselness incorporated ScatterNet Hybrid Deep Learning (TS-SHDL) network for brain matter segmentation in the MR images. The improved network incorporates the vesselness and texture hand-crafted features in addition to the two-layer Scat-terNet descriptors to capture the intensity and structural properties of brain matter as shown in Fig. 1. The extracted hand-crafted features are used by four stacked layers of unsupervised learning based PCA-Net to learn translation, rotation and scale invariant mid-level features. The mid-level features are used by the supervised Fisher encoding (FK) to gradually reduce the dimensions of the feature vector. The reduced feature vector is finally used by a conditional random field (CRF) to perform the desired brain matter segmentation. The contributions and reasoning for the choosing the above-mentioned handcrafted features, unsupervised, and supervised learning, module are discussed below.

The main contributions of the paper are stated below:

- *Hand-crafted Module*: Texture and Veselness feature are used jointly with ScatterNet hand-crafted descriptors as they capture the intensity and structural properties of the brain matter. The intensity properties are represented in the brain matter by region boundaries and texture while the ridge-like structures correspond to the structural properties of the matter as shown in Fig. 1. Parametric DTCWT ScatterNet [32] and Texture [28] filters are used to represent the intensity properties of the brain while the structural properties were extracted using a vesselness [16] filter. The intensity and structural properties can appear at any position, orientation and scale in the MR images. Hence, invariance to translation, rotation, and scale is introduced in each sub-feature.

- *Unsupervised Learning Module*: This module uses four stacked PCA-Net [10] layers on the concatenated hand-crafted features to learn translation, rota-

tion, and scale invariant robust hierarchical mid-level features. The network is fast and easy to train as opposed to other unsupervised learning modules such autoencoders or RBMs as the minimization of the loss function (Eq. 6) can be obtained in its simplistic form as the eigen decomposition. This makes PCA-Net a good method for learning mid-level features.

- *Supervised Learning Module*: The features obtained from the PCA-Net module are reduced to a more compact and dense representation using Fisher Encoding (FV) [27]. The reduced dimensions results in faster learning of the Conditional Random Field (CRF), used to perform the desired segmentation.

- *Advantages over Supervised Deep Networks*: The proposed network makes use of unsupervised learning to learn a hierarchy of mid-level features that are utilized to perform the desired brain matter segmentation. Similar hierarchical features can be also learned using supervised deep networks. However, supervised deep networks requires large amounts of annotated training data to learn these features which may not available for many applications [29, 12, 25], especially for medical image applications. The proposed TS-SHDL network can be a potential alternative to deep networks for applications with small training datasets.

The proposed framework is used to perform brain matter segmentation on 2D synthetic and 3D real brain MRI datasets. The average segmentation accuracy for each class for both datasets is presented. In addition, an extensive comparison of the proposed pipeline with other deep brain matter segmentation methods is demonstrated.

The paper is divided into the following sections. Section 2 briefly presents the proposed TS-SHDL network. Section 3 presents the experimental results while Section 4 draws conclusions.

2. TS-SHDL Network

This section details the texture and structure feature incorporated ScatterNet Hybrid Deep Learning Network (TS-SHDL) used for brain matter segmentation. The first subsection explains the mathematical formulation of the ScatterNet, texture and structure, sub-features, combined to form the hand-crafted features module. The second subsection details the mathematical formulation of the four layers unsupervised learning based PCA-Net framework that utilizes the hand-crafted features to extract translation, rotation, and scale invariant features. The final sub-section details the supervised Fisher vector encoding that produces a compact and dense feature vector finally used by the Conditional Random Field to produce the desired brain matter

segmentation. The TS-SHDL network is presented in Fig. 1.

2.1. Hand-crafted Features

This section details the hand-crafted features extracted from the MR images that capture the intensity and structure features as shown in Fig. 1.

2.1.1 DTCWT ScatterNet

The parametric log based DTCWT ScatterNet [32] is an improved version (both on classification error and computational efficiency) of the multi-layer Scattering Network [5, 24, 31, 23] that extracts relatively symmetric translation invariant representations from a multi-resolution image using the *dual-tree complex wavelet transform* (DTCWT) [14] and parametric log transformation layer. Below we present the formulation of the parametric DTCWT ScatterNet for a single input image which may then be applied to each of the multi-resolution images.

The first layer is involved in filtering the input signal x using dual-tree complex wavelets $\psi_{j,r}$ at different scales (j). Six pre-defined orientations are fixed for this operation - $15°, 45°, 75°, 105°, 135°$ and $165°$. Point-wise L_2 non-linearity is applied to the filtered signal, to build more translation invariant representation as described below:

$$U[\lambda_{m=1}] = |x \star \psi_{\lambda_1}| = \sqrt{|x \star \psi_{\lambda_1}^a|^2 + |x \star \psi_{\lambda_1}^b|^2} \quad (1)$$

After this a parametric log transformation layer is applied to all the oriented representations ($U[j]$). This extracts representations at a particular scale j with a parameter k_j, and reduces the effect of outliers by introducing relative symmetry. A local average is then computed on the transformation: $|U1[\lambda_{m=1}]|$ that aggregates the coefficients to build the desired translation-invariant representation:

$$S[\lambda_{m=1}] = |U1[\lambda_{m=1}]| \star \phi_{2^J},$$
$$U1[j] = \log(U[j] + k_j), \ U[j] = |x \star \psi_j|, \quad (2)$$

The above transformation leads to loss of high frequency components due to smoothing. These are thus retrieved by cascades wavelet filtering performed in the second layer. However, these retrieved components are not translation invariant, which is restored by first applying the L_2 non-linearity of eq(2) to obtain the regular envelope:

$$U2[\lambda_{m=1}, \lambda_{m=2}] = |U1[\lambda_{m=1}] \star \psi_{\lambda_{m=2}}| \quad (3)$$

A local-smoothing operator is then applied to the regular envelope ($U2[\lambda_{m=1}, \lambda_{m=2}]$) to extract the desired second layer ($m = 2$) translation invariant coefficients:

$$S[\lambda_{m=1}, \lambda_{m=2}] = U2[\lambda_{m=1}, \lambda_{m=2}] \star \phi_{2^J} \quad (4)$$

Figure 2. Demonstrates the pipeline of vesselness features at 3 scales. For each feature map extracted at each scale, a 20×20 patch is selected. Here only 3 scales are shown as opposed to 10 used in our experiments owing to space restrictions. Finally, Max operation is performed across scales to achieve scale invariance.

Finally, the scattering coefficients obtained at each layer are:

$$S = \left(x \star \phi_{2^J}, U1[\lambda_{m=1}] \star \phi_{2^J}, U2[\lambda_{m=1}, \lambda_{m=2}] \star \phi_{2^J}\right) \tag{5}$$

2.1.2 Vesselness

Traits of ridge-like and tubular structures depicting structural properties can enhance segmentation tasks. The brain matter contains these structures that can be exploited for superior segmentation. In the recent past, these structures have been extracted using eigen-decomposition of the Hessian calculated at each pixel [3, 16]. These filters ξ_j, for 10 different scales (j), were used to identify narrow vessels in low contrast. The Vesselness filter uses the Hessian matrix and decomposes its eigenvalues into pre-defined (3) orthonormal directions. Scale invariance is obtained using a filter patch of size 10×10 at different scales j, followed by a max operator across the scales to obtain the final estimate of vesselness. This process is presented in Figure 2.

2.1.3 Texture

Brain Matter typically demonstrates different texture properties for grey and white matter. Grey matter generally present a rougher texture as compared to white matter. In this paper, we use MR8 filter [28] to achieve rotations and scale invariance while extracting texture information. The filter consists of a Gaussian and a Laplacian of Gaussian filter, an edge filter $\Omega_{j,r}^1$ at 3 scales (j) and a bar filter $\Omega_{j,r}^2$ at the same 3 scales (j). The later two filters are oriented at 6 directions (r). We take the maximum response across orientations to achieve rotational invariance. This reduces the number of filter responses from 38 (6 orientations at 3 scales for 2 oriented filters, plus 2 isotropic) to 8 (3 scales for 2 filters, plus 2 isotropic). Following this, max operation is performed on these features across scales to get scale invariance (similar to invariance in the previous section).

2.2. Unsupervised Learning: PCANet

The features extracted using ScatterNet, vesselness, and texture, as described in the previous section, are concatenated and fed to the stacked four-layer PCANet [10]. The PCANet at each layer learns a translation, rotation, and scale invariant mid-level representation using unsupervised learning from the features of the previous layer. The invariant mid-level features are learned by cropping patches of size 5×5 from all the feature as $x_1, x_2, ..., x_{mn}$. After this, patch mean is subtracted from each patch to obtain $\bar{X} = [\bar{x}_1, \bar{x}_2, ..., \bar{x}_{mn}]$, where $\bar{x}_j$ is a mean-removed patch.

The objective of the layer PCANet is to minimize the reconstruction error within a family of orthonormal filters. This can be seen in the following equation,

$$\min_{V \in R^{z_1 z_2 \times K}} \|X - VV^T X\|_F^2, \text{ s.t. } V^T V = I_K, \tag{6}$$

where,

$$X = [\bar{X}_1, \bar{X}_2, ..., \bar{X}_N] \in R^{z_1 z_2 \times N}. \tag{7}$$

N is the number of training images, I_K is identity matrix of size $K \times K$. Here K is the number of filters comprising of the leading K principal eigenvectors capture the main variance amongst the patches for each layer.

2.3. Supervised Fisher Encoding

The Fisher Kernel [27], is used to transform the PCA features obtained from the unsupervised learning module into a compact and dense Fisher vector encoding using supervised learning. The output from the PCA layer is unrolled to get a vector $X = \{x_1, x_2, ...x_T\}$. Assuming a probability density function u_λ which models the generation process of X. The kernel is the described as:

$$K(X, Y) = G_\lambda^{X'} F_\lambda^{-1} G_\lambda^Y \tag{8}$$

here F_λ is the Fisher information matrix, K is the Fisher Kernel on the gradients of log-likelihood. The dimensionality of the Fisher Vector depends on the number of pa-
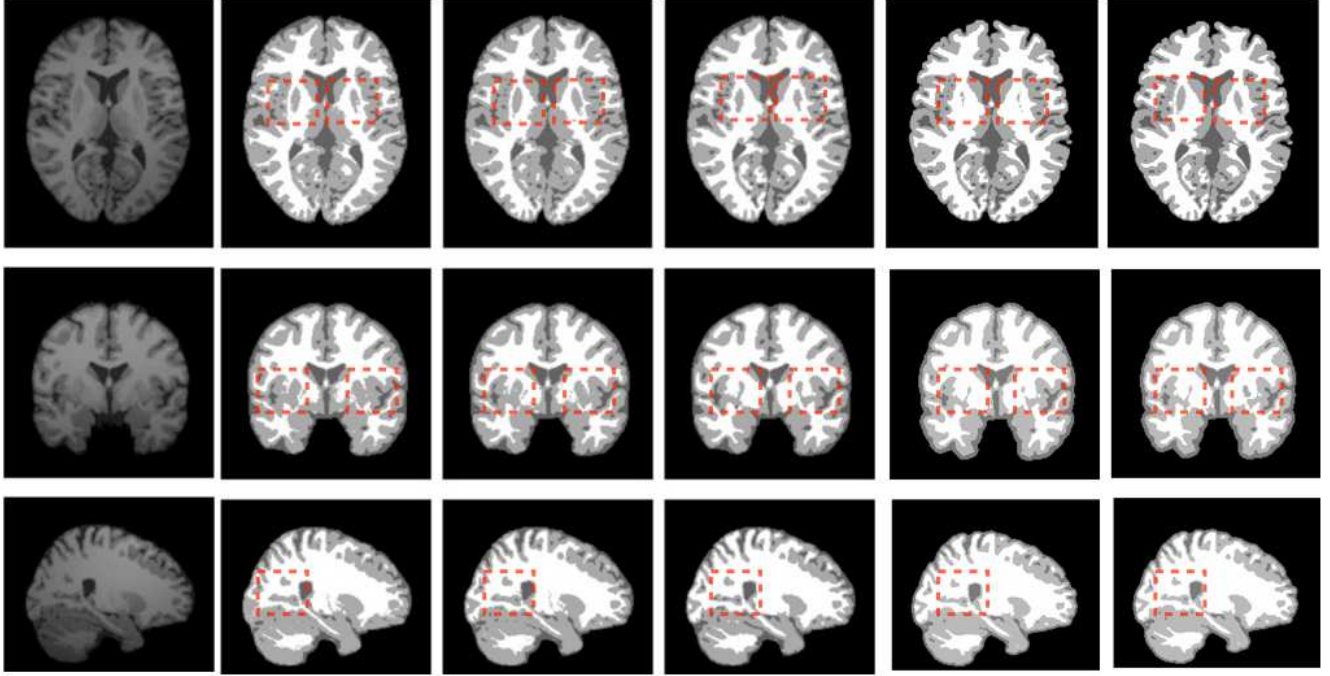
Figure 3. Qualitative Comparison with previous state of the art methods. Column 1: Original images. Column 2: Ground Truth. Column 3: Leemput et al.'s method [36]. Column 4: Wang et al.'s method [38]. Column 5: Madiraju et. al.'s method [20] and Column 6: Proposed TS-SHDL. The regions with the difference are highlighted with red boxes.

rameters λ. Further, F_λ being symmetric and positive definite, can be decomposed as $F_\lambda = L'_\lambda L_\lambda$, which inturn enables the Fisher Kernel to be re-written as the dot product of $H_\lambda = L_\lambda G^X_\lambda$. Here H_λ is the Fisher Vector corresponding to input X.

2.4. Supervised CRF Segmentation

The Fisher vector (FV) features are used by the Conditional Random Field (CRF) to achieve the desired brain matter segmentation. CRF is typically a probabilistic framework that establishes a relationship between the pixel labels and extracted features [8]. It is an undirected graphical model that uses 4 pairwise connected grids having finite vertices or nodes and corresponding edges connecting the vertices. Here, each node represents a random variable X and the edges define the neighbourhood relation between these unobserved random variables. To perform the image segmentation, a clique loss function is used with Tree-Reweighted [8] inference which uses LBFGS optimization algorithm. During training, after each iteration, the loss value is checked and if bad search direction is encountered, L-BFGS is reinitialized [8].

3. Experimental Results

This section provides details of the experiments performed on a 2D synthetic datasets (the McGill Brain Web [15]) and a 3D real dataset (MRBrainS website [22]). The

qualitative and quantitative comparison is presented with the state of the art brain segmentation methods on both datasets. The images used in the experimentation have a brain images of intensity non-uniformity(INU) = 100% and noise level of 3%.

3.1. Synthetic Image Dataset

A total of 300 brain MRI images are chosen from the McGill Brain Web [1] [15]. The dataset, split into 140 | 60 | 100 images for $train$ | $validation$ | $test$ sets, contains realistic MRI images of the human brain produced by an MRI simulator based on two anatomical models: normal and multiple sclerosis (MS). We provide both qualitative (Fig. 1) and quantitative (Table. 1) comparison to evaluate the performance of our proposed method with the state of the art.

3.2. Real Image Dataset

Multisequence 3T MRI scans of twenty people available on the MRBrainS website [2] [22] is used as the real image dataset. In the dataset, the subjects were selected to have varying degrees of atrophy and white matter lesions. All the scans have manual segmentations into Grey (GM) and White Matter (WM). 5 subject's scans are used as the train

---

[1]http://mcgill.ca/bic
[2]http://mrbrains13.isi.uu.nl/

Table 1. The table shows comparison of average Jaccard ($\mathcal{J}$) values of grey matter and white matter of our algorithm with current state of the art algorithms on the 2D synthetic MRI brain dataset [15]. Grey: Best Results and Bold: TS-SHDL Results.

Method	FSL [41]	SPM [2]	MAP with histograms [39]	Decision Forest Classifier [41]	HDF [20]	TS-SHDL
$\mathcal{J}$ of Grey Matter	0.756^2	0.790^2	0.814	0.838	0.89	**0.907**
$\mathcal{J}$ of White Matter	NA	NA	0.710	0.731	0.91	**0.921**

set and the remaining used as test set. We provide extensive quantitative results on this dataset as presented in Table. 2.

3.3. TS-SHDL Parameters

This section presents the parameters for the (i) hand-crafted, (ii) unsupervised ,and (iii) supervised block used for the experiments.

- *Hand-crafted Module*: The ScatterNet used 2 DTCWT scales, 10 scales were used for the vesselness filter while 3 scales and 6 orientations were used for the texture filter.

- *Unsupervised Learning Module*: The number of filter learned in the first, second, third and fourth, PCA-Net layer are 40, 30, 20 and 10, respectively.

- *Supervised Learning Module*: The number of GMM cluster for the Fisher encoding are chosen as 5 and the dimension of the final feature vector to be encoded is chosen to be 500.

3.4. Qualitative Analysis

Figure 3 provides the qualitative comparison of our proposed method with other state of the art methods. We perform these analyses on the synthetic image dataset. Being 2-dimensional images, it aids visualization and qualitative inspection. Upon close observation, it can be inferred that our method provides superior segmentation and better correlation with ground truth. In the axial image section (first row), it can be seen that our method correctly segments the central grey lobes which was missing in Madiraju et. al.'s method [20]. Also in the coronal (second row) image, the grey matter segmentation is best represented in our method as compared to the other methods. This shows a qualitative improvement over both grey and white matter segmentation.

3.5. Quantitative Analysis

To provide emphasis over the effectiveness of our method, we also provide quantitative evidence using the Jaccard similarity [37] metric on the 2D synthetic brain MRI dataset and more robust Dice coefficient [7] (DC), 95th percentile of Hausdorff distance [7] (HD), and Absolute Volume Difference [7] (AVD) metrics on the 3D real brain MRI [22] dataset for both grey and white matter.

3.5.1 2D MRI Synthetic Dataset

Jaccard similarity is used to evaluate the performance of the proposed TS-SHDL network on the 2D MRI Synthetic Dataset [15]. Jaccard similarity measures the similarity of the segmented regions with the ground truth:

$$\mathcal{J}(G,S) = \frac{|G \cap S|}{|G \cup S|} \qquad (9)$$

value of $\mathcal{J}$ lies between 0 and 1. Higher the value of $\mathcal{J}$ accurate the segmentation.

Table. 1 presents an extensive comparison with other brain matter segmentation algorithms applied on the synthetic dataset. Here our method outperforms the state of the art methods in both grey and white matter segmentation. The model also surpasses the recent state of the art method of Madiraju et. al. [20] thus showing its superior performance.

The proposed TS-SHDL network is also compared with the supervised deep Fully Convolutional Network (FCN) [17] with 8-pixel stride for Brain matter segmentation. The training data fine-tunes the FCN trained on ImageNet and then test on the brain MR image test to achieved a Jaccard index of 0.87 and 0.88 for the grey and white matter respectively. The FCN underperformed because the training data is not sufficient to fine-tune the FCN learned on the ImageNet.

3.5.2 3D MRI Real Dataset

The performance of the proposed TS-SHDL is evaluated on the 3D MRI Real Dataset [22] using three robust evaluations metrics: (i) Dice coefficient [7], (ii) Hausdorff distance [7], and (iii) Absolute Volume Difference [7].

Dice coefficient [7] measures the spatial overlap between the ground truth and the predicted segmentation. It is defined as,

$$D(G,S) = \frac{2|G \cap S|}{|G| + |S|}.100\% \qquad (10)$$

Here, G is the ground truth while S is the segmentation. A higher value of the Dice coefficient signifies higher segmentation accuracy.

Hausdorff distance [7] measures the distance between segmentation results and ground truth. In-order to be in-

Table 2. The table shows comparison of different methods (DC: %, HD: mm, AVD: %) of grey matter and white matter of our algorithm with current state of the art algorithms on the 3D real MRI brain dataset [22]. Grey: Best Results and Bold: TS-SHDL Results.

Method		3D U-net [7]	PyraMID [35]	VoxResNet [6]	Mahbod et. al. [21]	Pereira et. al. [26]	TS-SHDL
Grey Matter	DC	85.44	84.82	86.15	84.77	84.50	**86.01**
	HD	1.58	1.70	1.45	1.71	1.70	**1.40**
	AVD	6.60	6.77	6.60	6.02	7.10	**6.46**
White Matter	DC	88.86	88.33	89.46	88.45	88.04	**89.49**
	HD	1.95	2.08	1.90	2.34	2.12	**1.88**
	AVD	6.47	7.05	6.05	7.67	7.74	**5.99**

sensitive to outliers, we use the Kth ranked distance,

$$h_{95}(S,G) =^{95} K_{s \in S}^{th} \quad \min_{g \in G} ||g - s|| \qquad (11)$$

$$HD(G,S) = max\{h_{95}(S,G), h_{95}(G,S)\} \qquad (12)$$

Unlike Dice coefficient, a smaller HD denotes higher accuracy.

Finally, we calculate Absolute Volume Difference [7] which is defined as,

$$AVD(G,S) = \frac{|V_s - V_g|}{V_g}.100\% \qquad (13)$$

where, V_s, V_g are the volume of segmentation results and ground truth. Similar to HD, lesser value of AVD signifies better accuracy. All the three metrics are applied for both grey and white matter in the real image dataset.

We have compared our method with both state of the art deeplearning based models that include 3D U-net [7], PyraMID-LSTM [35], VoxResNet [6] and hand-crafted feature based models such as [21, 26] etc. 3D U-net extracts volumetric feature representation from the 3D images, PyraMID paralellises multi-dimensional RNNs in a pyramid fashion and VoxResNet provides a voxelwise residual network for performing the segmentation. On the other hand, methods based on hand-crafted features use different techniques that involve histogram based features [21] and gradients [26] etc. As seen in Table 2, our model is able to outperform all these methods in the real image dataset.

4. Conclusion

The paper presents the ScatterNet Hybrid Deep Learning Network (SHDL) with Texture and Structure Features that capture the intensity and structural properties of the brain matter. Four layers of the PCANet are used to learn hierarchical invariant mid-level features from the concatenated hand-crafted features. Fisher Encoding is used on the mid-level feature to obtain a compact and dense representation that is finally used by the CRF is used to achieve the desired brain matter segmentation. The proposed network outperformed the state-of-the-art on both qualitative and quantitative measures. In addition, the proposed network can be the potential model choice for applications with small datasets as deep supervised learning network may not be trainable because of insufficient the training dataset.

References

[1] V. Angoth, C. Dwith, and A. Singh. A novel wavelet based image fusion for brain tumor detection. *International Journal of computer vision and signal processing*, 2(1):1–7, 2013.

[2] J. Ashburner and K. Friston. Multimodal image coregistration and partitioninga unified framework. *Neuroimage*, 6(3):209–217, 1997.

[3] S. R. Aylward and E. Bullitt. Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. *IEEE transactions on medical imaging*, 21(2):61–75, 2002.

[4] J. C. Bezdek, L. Hall, and L. Clarke. Review of mr image segmentation techniques using pattern recognition. *Medical physics*, 20(4), 1993.

[5] J. Bruna and S. Mallat. Invariant scattering convolution networks. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 35:1872 –1886, 2013.

[6] H. Chen, Q. Dou, L. Yu, and P.-A. Heng. Voxresnet: Deep voxelwise residual networks for volumetric brain segmentation. *arXiv preprint arXiv:1608.05895*, 2016.

[7] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 424–432. Springer, 2016.

[8] J. Domke. Learning graphical model parameters with approximate marginal inference. *IEEE transactions on pattern analysis and machine intelligence*, 35(10):2454–2467, 2013.

[9] C. Dwith, V. Angoth, and A. Singh. Wavelet based image fusion for detection of brain tumor. *International Journal of Image, Graphics and Signal Processing*, 5(1):25, 2013.

[10] T.-H. C. et al. Pcanet: A simple deep learning baseline for image classification? *ArXiv:1404.3606*, 2014.

[11] M. e. a. Havaei. Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35:18–31, 2017.

[12] S. Jain, S. Gupta, and A. Singh. A novel method to improve model fitting for stock market prediction. *International Journal of Research in Business and Technology*, 3(1), 2013.

[13] T. Kapur E et al. Segmentation of brain tissue from magnetic resonance images. *Medical image analysis*, 1(2):109–127, 1996.

[14] N. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Applied and computational harmonic analysis*, 10:234–253, 2001.

[15] R.-S. e. a. Kwan. Mri simulation-based evaluation of image-processing and classification methods. *IEEE transactions on medical imaging*, 18, 1999.

[16] T. Lindeberg E et al. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2), 1998.

[17] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional network for semantic segmentation. *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2015.

[18] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[19] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[20] N. Madiraju, A. Singh, and S. Omkar. Level set segmentation of brain matter using a trans-roto-scale invariant high dimensional feature. In *Asian Conference on Computer Vision (ACCV) workshops*, pages 595–609. Springer, 2016.

[21] A. Mahbod. Structural brain mri segmentation using machine learning technique, 2016.

[22] A. M. Mendrik et al. Mrbrains challenge: online evaluation framework for brain image segmentation in 3t mri scans. *Computational intelligence and neuroscience*, 2015:1, 2015.

[23] S. Nadella, A. Singh, and S. Omkar. Aerial scene understanding using deep wavelet scattering network and conditional random field. *European Conference on Computer Vision (ECCV) workshops*, 9913:205–214, 2016.

[24] E. Oyallon and S. Mallat. Deep roto-translation scattering for object classification. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2865–2873, 2015.

[25] A. Pasini. Artificial neural networks for small dataset analysis. *Journal of Thoracic Disease*, 7(11):2278–2324, 2015.

[26] S. e. a. Pereira. Automatic brain tissue segmentation in mr images using random forests and conditional random fields. *Journal of neuroscience methods*, 270:111–123, 2016.

[27] F. Perronnin and C. Dance. Fisher kenrels on visual vocabularies for image categorizaton. *In Proc. CVPR*, 2006.

[28] A. e. a. Pitiot. Expert knowledge-guided segmentation system for brain mri. *NeuroImage*, 23:S85–S96, 2004.

[29] J. Plaza, A. Plaza, R. Perez, and P. Martinez. On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images. *Pattern Recognition*, 42:3032–3045, 2009.

[30] A. Singh, S. Karanam, S. Bajpai, A. Choubey, and T. Raviteja. Malignant brain tumor detection. In *International Journal of Computer Theory and Engineering*, volume 4, 2012.

[31] A. Singh and N. Kingsbury. Multi-resolution dual-tree wavelet scattering network for signal classification. *11th International Conference on Mathematics in Signal Processing*, 2016.

[32] A. Singh and N. Kingsbury. Dual-tree wavelet scattering network with parametric log transformation for object classification. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2622–2626, 2017.

[33] A. Singh and N. Kingsbury. Scatternet hybrid deep learning (shdl) network for object classification. *Machine Learning for Signal Processing (MLSP) workshop*, 2017.

[34] R. Smith E et al. Sift: spherical-deconvolution informed filtering of tractograms. *Neuroimage*, 67:298–312, 2013.

[35] M. F. Stollenga, W. Byeon, M. Liwicki, and J. Schmidhuber. Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation. In *Advances in Neural Information Processing Systems*, pages 2998–3006, 2015.

[36] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of mr images of the brain. *Medical Imaging, IEEE Transactions on*, 18(10):885–896, 1999.

[37] U. Vovk, F. Pernus, and B. Likar. A review of methods for correction of intensity inhomogeneity in mri. *Medical Imaging, IEEE Transactions on*, 26(3):405–421, 2007.

[38] L. Wang, L. He, A. Mishra, and C. Li. Active contours driven by local gaussian distribution fitting energy. *Signal Processing*, 89(12), 2009.

[39] Z. Yi E et al. Discriminative, semantic segmentation of brain tissue in mr images. In *Medical Image Computing and Computer-Assisted Intervention*. 2009.

[40] W. e. a. Zhang. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage*, 108:214–224, 2015.

[41] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *Medical Imaging, IEEE Transactions on*, 2001.