World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:7, No:7, 2013

A Serial Hierarchical Support Vector Machine and 2D Feature Sets Act for Brain DTI Segmentation

Mohammad Javadi

***Abstract*—**Serial hierarchical support vector machine (SHSVM) is proposed to discriminate three brain tissues which are white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). SHSVM has novel classification approach by repeating the hierarchical classification on data set iteratively. It used Radial Basis Function (rbf) Kernel with different tuning to obtain accurate results. Also as the second approach, segmentation performed with DAGSVM method. In this article eight univariate features from the raw DTI data are extracted and all the possible 2D feature sets are examined within the segmentation process. SHSVM succeed to obtain DSI values higher than 0.95 accuracy for all the three tissues, which are higher than DAGSVM results.

***Keywords*—**Brain segmentation, DTI, hierarchical, SVM.

I. Introduction

SEGMENTATION is known as one of the main techniques used for investigating the brain images. In segmentation, the objective is to partition the brain volume into a number of predefined tissues. Diffusion tensor imaging (DTI) is a popular modality which is used in brain imaging. Due to the complexity form of the brain tissues, different methods have been used to obtain accurate segmentation [1]-[4]. In this research for brain DTI data segmentation, support vector machine (SVM) with a novel serial hierarchical classifier (SHSVM) is proposed.

SVM gives ability to map the data into higher dimensional space. Thus in the brain DTI data which has nonconvex distribution, SVM works properly. Different Kernel functions are used in SVM to perform space transformation such as linear, polynomial, Radial Basis Function (rbf) [5], etc.

The SVM was used for one-class discrimination with quadratic classifiers in [6]. Also it can be used for multiclass segmentation. The most common used methods for the multiclass SVM segmentation are: one-against-one [7], one-against-all [8], and DAGSVM [9].

In this work, we focus on brain tissue segmentation using DTI data, where all three major brain tissues including white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) are taken into account. As discriminatory parameters we consider eight DTI-based features extracted from diffusion tensor data (eigenvalues, scalar invariants, relative and fractional anisotropy) and from these univariate features all the possible 2D feature sets are extracted and used.

As the segmentation methods we investigate SHSVM and DAGSVM. The experiences are executed with different σ values. Dice similarity index (DSI) is used as the evaluation technique [10]. Also 3D scatterplots for manual and SVM segmentations are used for investigating the results visually.

M. Javadi was with the Department of Signals and Systems, Chalmers University of technology. He is now doing research in medical image analysis and image processing (e-mail: javadim@student.chalmers.se).

II. Methods

A. Preprocessing and Feature Extraction

From the DTI raw data by the help of FDT-FSL three eigen values ($\lambda1,\lambda2,\lambda3$) were extracted. These eigenvalues used to extract fractional anisotropy (FA), relative anisotropy (RA), and scalar invariants (I1,I2,I3) according to (1)-(5) [11]-[15].

$$FA = \sqrt{\frac{3}{2}} \sqrt{\frac{\sum_{i=1,2,3} (\lambda_i - \bar{\lambda})^2}{\sum_{i=1,2,3} \lambda_i^2}} \tag{1}$$

$$RA = \frac{1}{\sqrt{6}} \frac{\sqrt{\sum_{i=1,2,3} (\lambda_i - \bar{\lambda})^2}}{\bar{\lambda}} \tag{2}$$

where in FA and RA,

$$\bar{\lambda} = \frac{1}{3} \sum_{i=1}^{3} \lambda_i$$

$$I_1 = \lambda_1 + \lambda_2 + \lambda_3 \tag{3}$$

$$I_2 = \lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1 \tag{4}$$

$$I_3 = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \tag{5}$$

Our previous research examined almost all the univariate and different multifeature sets [16]. The obtained results showed the multifeature sets, had the highest accuracy. Performing all the combination of the features is out of the scope of this article. Here the experiences are performed on all the possible 2D feature sets (e.g. FA-RA, I1-I2, …) and the best results are presented.

B. Support Vector Machine

Support vector machine (SVM) is a supervised learning method which has been illustrated perfectly in [17]. The SVM attempts to perform linear classification on nonlinear distributed data. It has two main stages: training and test. Assume a training set with N samples, each one represent as (x_i , y_i), where x_i is feature vector and y_i is class labels which are $+1$ and -1. In SVM a linear decision boundary or hyperplane is considered as (6),

$$w^T x + b = 0 \tag{6}$$

where x is an input data (vector), w is the hyperplane orientation (vector) and b is a bias parameter. Training process attempts to obtain perfect w and b to maximize distance between two classes that so called margin. After training process the test stage or the classification is carried out. Each input vector fulfilled one of the following equations and leads to be classified.

$$w^T x_k + b > 0 \quad then \quad x_k \in class\ (+1) \tag{7}$$

$$w^T x_k + b < 0 \quad then \quad x_k \in class\ (-1) \tag{8}$$

The major part in SVM which makes it a strong method against the nonconvex distribution and nonseparable data is space transformation (mapping). Different Kernel functions are used to transfer data into higher dimensional space to separate the data. Therefore, linear classifier works for the mapped data where the original data need non-linear classifier. In this article Radial Basis Function (rbf) Kernel (9) is examined [5].

$$K(x_1, x_2) = exp\left(\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right) \tag{9}$$

where x_1 and x_2 are input vector space and σ is rbf Kernel function parameter. Selecting the σ value is important and some investigation in [18]. Here we examined different σ values and the best results presented.

C. Segmentation Approach

Two SVM segmentation methods are proposed in this research: SHSVM and DAGSVM. SHSVM is a series of hierarchical classification using SVM. In this method in each iteration $k(k-1)/2$ classifiers are used, where k is number of classes. SHSVM has three steps. First step, data is classified with $k(k-1)/2$ classifiers. Each classifier segmented data into class/non-class. Second step, the segmented results which are k classified data are merged together to obtain one sort of data set. This data set contains k classes and some unclassified or misclassified regions. Third step, the unclassified data is considered as the input data set and the procedure from the first step is repeated. This process is continued until all data are classified.

In this method when data are merged together. Some of the class regions overlap each other. In addition, some of the data are not classified into any of the classes. The process separates these data and the SHSVM is performed on the separated data sets. SHSVM is executed for six different σ values. These values are: $\sigma = \{0.001, 0.005, 0.01, 0.05, 0.1, 0.3\}$.

III. Experimental Results

In this research 28 numbers of 2D feature sets from the eight univariate features were extracted. Where each 2D feature set in each slice contains about 4000 pixels, in total about 120000 pixels in each slice were examined. And the selective results are presented as follows.

Fig. 1 shows 3D scatterplot of FA-I3 feature set, obtained with the SHSVM segmentation vs. manual segmentation.

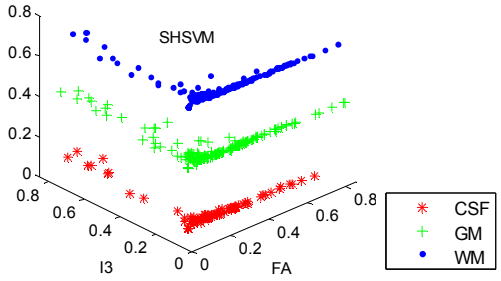

Fig. 1 Scatterplot of SHSVM with $\sigma = 0.01$

Fig. 2 shows 3D scatterplot of FA-I3 feature set, obtained with the DAGSVM vs. manual segmentation.

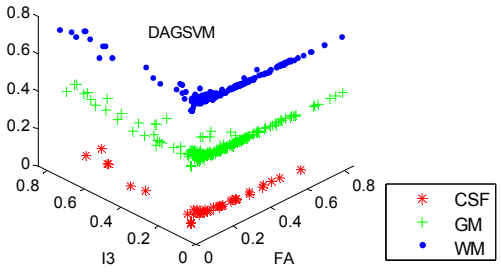

Fig. 2 Scatterplot of DAGSVM with $\sigma = 0.01$

Fig. 3 shows 3D scatterplot of FA-I3 feature set, obtained manually. The 3D scatterplots in Figs. 1 and 2 are compared with Fig. 3 and discussed in Section IV.

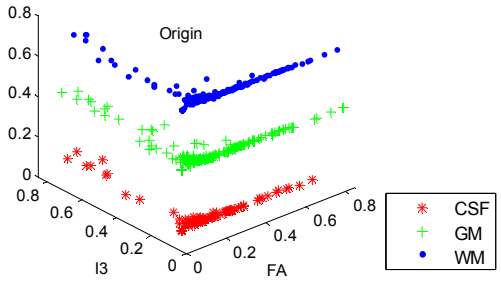

Fig. 3 Scatterplot of manual segmentation with $\sigma = 0.01$

Fig. 4 shows DSI values of WM discrimination using SHSVM. The results are obtained from seven 2D feature sets

(FA-I3, RA-I3, λ1-I3, λ2-I3, λ3-I3, I1-I3, I2-I3) vs. variations of σ in this set: σ ={0.001, 0.005, 0.01, 0.05, 0.1, 0.3}.

Fig. 5 shows the same experiences in Fig. 4, unless it used DAGSVM for segmentation.

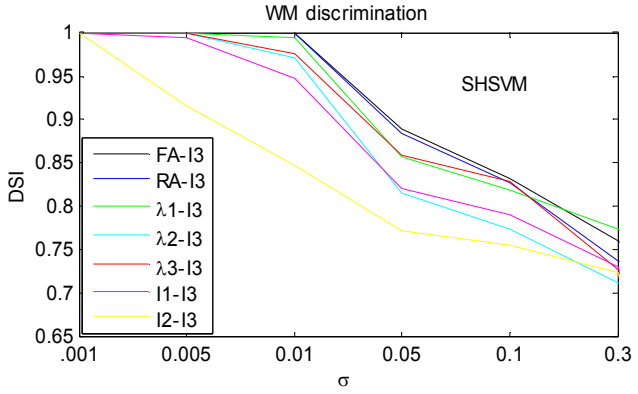

Fig. 4 WM discriminations of seven 2D features by SHSVM

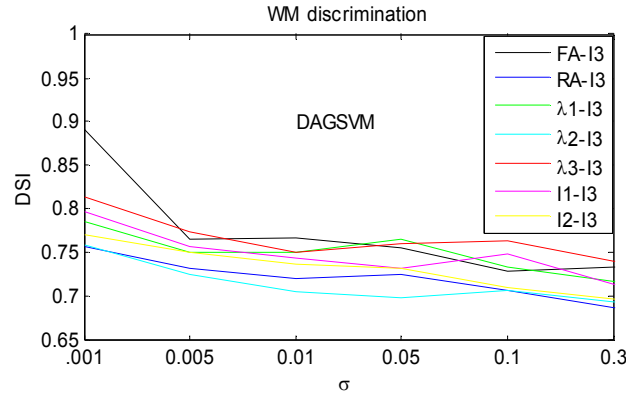

Fig. 5 WM discriminations of seven 2D features by DAGSVM

The results of SHSVM and DAGSVM for σ = 0.1 are collected in Table I. These results were sketched in Figs. 4 and 5.

TABLE I
DSI VALUES OF CSF, GM, AND WM DISCRIMINATIONS USING SHSVM AND DAGSVM (σ =0.1)

	SHSVM			DAGSVM		
	CSF	GM	WM	CSF	GM	WM
FA-I3	0.95	0.80	0.83	0.68	0.66	0.73
RA-I3	0.94	0.78	0.83	0.55	0.43	0.71
λ1-I3	0.89	0.80	0.82	0.70	0.67	0.73
λ2-I3	0.90	0.68	0.77	0.57	0.38	0.71
λ3-I3	0.93	0.80	0.83	0.67	0.68	0.76
I1-I3	0.88	0.67	0.79	0.57	0.61	0.75
I2-I3	0.86	0.65	0.76	0.58	0.43	0.71

IV. DISCUSSION

The results show each 2D feature set is appropriate for especial tissue discrimination (see Figs. 4, 5).

The accurate result using the DAGSVM with σ = 0.1 for CSF discrimination (0.70) was obtained by λ1-I3. That for GM discrimination (0.68) and WM discrimination (0.76) was obtained by λ3-I3. Where the accurate results using the SHSVM with σ = 0.1 for CSF discrimination (0.95), GM discrimination (0.80), and for WM discrimination (0.83) were obtained by FA-I3. Therefore, there is an interaction between feature sets and segmentation methods. To obtain accurate discrimination for each tissue, it needs a tradeoff between different feature sets and SVM methods. Kernel function tuning effects have to be considered as well.

As it was seen, the accuracy of the results was influenced by decreasing the σ value. In SHSVM for the CSF, decreasing the σ amount values caused to increase the segmentation accuracy from 0.62 up to 1. For GM that is from 0.40 up to 1 and for WM that is from 0.68 up to 0.89. These σ variations in DAGSVM for the CSF caused to increase the segmentation accuracy from 0.54 up to 0.97. For GM that is from 0.17 up to 0.85 and for WM that is from 0.69 up to 0.89.

The variations of the σ values have effect on segmentation accuracy. However, for some of the features decreasing the σ value from especial value has no effect on result's accuracy. For instance, in FA-I3, decreasing the σ lower than 0.01 has no effect on accuracy. Also selecting the low σ values caused to increase the execution time. Therefore, the σ tuning for segmentation accuracy and execution time is quite important.

In addition, 3D scatterplots show that the input data have nonconvex distribution (see Fig. 3). Visual investigation of the 3D scatterplots of SHSVM in Fig. 1 and DAGSVM in Fig. 2 and comparing these results with that of the manual segmentation in Fig. 3 show SHSVM has more accurate segmentation results than DAGSVM.

V. CONCLUSION

In this research SHSVM with rbf Kernel was used. It was seen that segmentation accuracy influenced by decreasing the σ values. Also there is a limitation for σ value which from that value we could not obtain higher accuracy. Therefore, perfect Kernel function tuning in addition of obtaining the accurate results cause to save the time.

The SHSVM in comparison with DAGSVM has the better results. Thus it is introduce for the brain segmentation. Moreover, for the future work we purpose to use different DTI data and using SHSVM with different Kernel functions. Applying different tuning on Kernel functions is other approach which might be considered to improve the accuracy and speed of the process.

REFERENCES

[1] E.D. Angelini, T. Song, B.D. Mensh, and A.F. Laine, "Brain MRI segmentation with multiphase minimal partitioning: A comparative study," *Int. Journal of Biomedical Imaging,* vol. 2007, pp. 1–15.

[2] G. Heckenberg, Y. Xi, Y. Duan, and J. Hua, "Brain structure segmentation from MRI by geometric surface flow," *International Journal of Biomedical Imaging,* vol. 2006, pp. 1–6, 2006.

[3] O. Commowick, S.K. Warfield, "A continuous STAPLE for scalar, vector, and tensor images: An application to DTI analysis," *IEEE Transactions on Medical Imaging,* vol. 28 , pp. 838–846, 2009.

[4] C. Lenglet, M. Rousson, and R. Deriche, "DTI segmentation by statistical surface evolution", *IEEE Trans. on Medical Imaging,* vol. 25, pp. 685–700, 2006

[5] I. Steinwart, A. Christmann, "Support Vector Machines (Book style)," *Springer*, LA, USA, 2008, pp. 112-116.

[6] Y. Washizawa, "Feature extraction using constrained approximation and suppression," *IEEE Trans. on Neural Networks*, vol. 21, pp. 201–210, 2010.
[7] S. Knerr, L. Personnaz, and G. Dreyfus, "Single-layer learning revisited: A stepwise procedure for building and training a neural network," in *Neurocomputing: Algorithms, Architectures and Applications*, J. Fogelman, Ed. New York: Springer-Verlag, 1990.
[8] J. Friedman. *Another approach to polychotomous classification*. Dept. Statist., Stanford Univ., Stanford, CA, 1996. [Online]. Available: http://www-stat.stanford.edu/reports/friedman/poly.ps.Z
[9] J. C. Platt, N. Cristianini, and J. Shawe-Taylor, "Large margin DAG's for multiclass classification," in *Advances in Neural Information Processing Systems. Cambridge*, MA: MIT Press, 2000, vol. 12, pp. 547–553.
[10] S. Miri, N. Passat, and J.P. Armspach, "Topology-preserving discrete deformable model: Application to multi-segmentation of brain MRI," In *Image and Signal Processing,* pp. 67-75, Springer, 2008.
[11] C. Pierpaoli and P. J. Basser, "Toward a quantitative assessment of diffusion anisotropy," *Magnetic Resonance in Medicine,* vol. 36, pp. 893–906, 1996.
[12] P.J. Basser, J. Mattiello, and D. Lebihan, "MR diffusion tensor spectroscopy and imaging," *Biophysical Journal,* vol. 66, pp. 259-267, 1994.
[13] A.L. Alexander, K. Hasan, G. Kindlmann, D.L. Parker, and J.S. Tsuruda, "A Geometric Analysis of Diffusion Tensor Measurements of The Human Brain," *Magnetic Resonance in Medicine,* vol. 44, pp. 283–291, 2000.
[14] L. Rittner and R. Lotufo, "Diffusion tensor imaging segmentation by watershed transform on tensorial morphological gradient," in *Computer Graphics and Image Processing, 2008. SIBGRAPI'08. XXI Brazilian Symposium on,* pp. 196–203, 2008.
[15] E.M. Akkerman, "Efficient measurement and calculation of MR diffusion anisotropy images using the platonic variance method," *Magnetic Resonance in Medicine*, vol. 49, pp. 599–604, 2003.
[16] M. Javadi, *Brain tissue segmentation using DTI data*, Dept. Chalmers University of Technology, EX007, 2013. [Online]. Available: http://publications.lib.chalmers.se/records/fulltext/176182/176182.pdf
[17] V. Vapnik, "Statistical learning theory," *John Wiley and Sons, Inc.*, NY, USA, 1998.
[18] Long Han, Mark J. Embrechts, Boleslaw Szymanski, "Sigma Tuning of Gaussian Kernels: Detection of Ischemia from Magnetocardiograms," *IGI Global 2011*, pp. 206–223.