A Disease Diagnostic Assistant System Using DTI and Extreme Learning Machine

Shuqiang Wang, Dewei Zeng, Yanyan Shen, and Jinxing Hu
Institute of Advanced Computing and Digital Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Science
Email: sq.wang@siat.ac.cn

***Abstract*—There is a growing interest in applying diffusion tensor imaging (DTI) to the evaluation of brain and spinal cord related disease. In the present study, the DTI data and extreme learning machine were employed to identify the levels with cervical spondylotic myelopathy (CSM) in spinal cord. In this work, there are 40 volunteers including 20 healthy people and 20 patients with CSM ranging from 24 to 81 years old. Experiment results show that the extreme learning machine based classifier performs well in detecting the patients with CSM (accuracy 93.6%, sensitivity: 91.2%, specificity 94.7%). The current work reveals the potential of using diffusion tensor imaging in conjunction with extreme learning machine to automate the classification of healthy subjects and subjects with brain and spinal cord related disease.**

***Index Terms*—diagnostic assistant system, extreme learning machine, diffusion tensor imaging, cervical spondylotic myelopathy**

I. Introduction

Disease diagnosis is essential in medicine, and the accuracy of diagnosis achieved in certain historical periods determines mainly the state-of-the-art in medical science. Computer-based methods are increasingly used to improve the quality of medical services. Artificial intelligence is the area of computer science focusing on creating expert machines that can engage on behaviors that humans consider intelligent [1]. There are some medical diagnostic expert systems in the literature: like MYCIN [2], PERFEX [3] and EasyDiagnosis [4]. As the first well known medical expert system, MYCIN was developed by Stanford University to help people, not expert in antimicrobial drugs, prescribe such drugs for blood infections. One short-side of MYCIN is that its knowledge base is uncompleted, since it can not cover anything like the full spectrum of infectious diseases.

Recently, diffusion tensor imaging (DTI) is developed to detect the micro-architecture of tissue based on a rank-two diffusion tensor model [5] and it is a promising method for characterizing micro-structural changes or differences with neuropathology and treatment. DTI and fiber tractography have already advanced the scientific understanding of many neurologic and psychiatric disorders [6]. The most common parameters employed in delineating the spinal cord tissue microarchitecture include fractional anisotropy (FA), mean diffusivity (MD) and apparent diffusion coefficient (ADC). All of these parameters are derived from eigenvalues to evaluate the scalar properties of water molecule diffusion [7]. Eigenvectors and eigenvalues derived from the diffusion tensor matrix respectively reflect the direction and strength of the movement of water molecules [7]. There is a growing interest in applying DTI to evaluate the spinal cord microarchitecture. The entropy-based principal eigenvector analysis has been introduced into the cervical spinal cord by Cui et al. for the evaluation of microstructural changes after cervical myelopathy [8].

In disease diagnostic assistant system, machine learning methods have been applied to a range of MRI modalities in an effort to automate the diagnosis of Mild cognitive impairment and Alzheimer's disease [9], [10]. However, few studies have looked at the potential of using DTI in conjunction with machine learning algorithms to automate the level diagnosis of cervical myelopathy. Nowdays, Neurologic examination by senior spine surgeons is dominant methods to detect the levels with CSM in spinal cord [11-14]. However, there are some disadvantages with this method. Comparing with the proposed classifier based method, neurologic examination costs a lot of time and the efficiency is not high. This will cause troubles when the patients swarm. Besides, the result of neurologic examinations is strongly dependent on spine surgeons. However, Spine surgeons are human beings. Diagnostic errors may happen. In this work, we wish to develop a method of combining DTI together with extreme learning machine to automate the classification of healthy subjects and subjects with CSM.

II. Materials and Methods

A. Subjects

Position a total of 40 volunteers, including 20 healthy subjects and 20 cervical spondylotic myelopathy patients, ranging from 24 to 81 years old, were recruited. All volunteers were screened to confirm their eligibility. The inclusion criteria of healthy subjects were intact sensory and motor function evaluated by the Japanese Orthopaedic Association (JOA) score system, and negative Hoffman's sign under physical examination.

Manuscript received January 5, 2015; revised May 1, 2015.

doi: 10.12720/jait.7.2.129-133

Exclusion criteria included the presence neurological signs and symptoms, or a past history of neurological injury, diseases and operations. The CSM patients were recruited with confirmed diagnosis by senior spine surgeons. We performed neurologic examinations, including manual muscle-strength testing, investigation of deep tendon reflexes, and sensory disturbance areas.

B. Imaging Methods

Imaging was conducted with a Philips Achieva 3.0 Tesla MR system (Best, The Netherlands). During the acquisition process, the subject was placed supine with the SNV head and neck coil enclosing the cervical region, and was instructed not to swallow to minimize the motion artifacts. The subject was then scanned with the anatomical T1-weighted (T1W), T2-weighted (T2W) images and diffusion tensor images (DTI). The standardized procedures in this study were approved by the authors' Institutional Review Board. Sagittal and axial T1W and T2W images were acquired for each subject using a fast spin-echo sequence. The parameters employed in sagittal imaging include as follows: field of view (FOV) = 250×250 mm, slice gap = 0.3 mm, slice thickness = 3 mm, fold-over direction = feet/head, number of excitation (NEX) = 2, resolution =0.92×1.16×3.0 mm3 (T1W) and 0.78×1.01×3.0 mm3 (T2W), recon resolution = 0 .49×0. 49×3.0 mm3, and echo time /repetition time = 7.2/530 ms (T1W) and 120/3314 ms (T2W). A total of 11 sagittal images covering the whole cervical spinal cord were acquired.

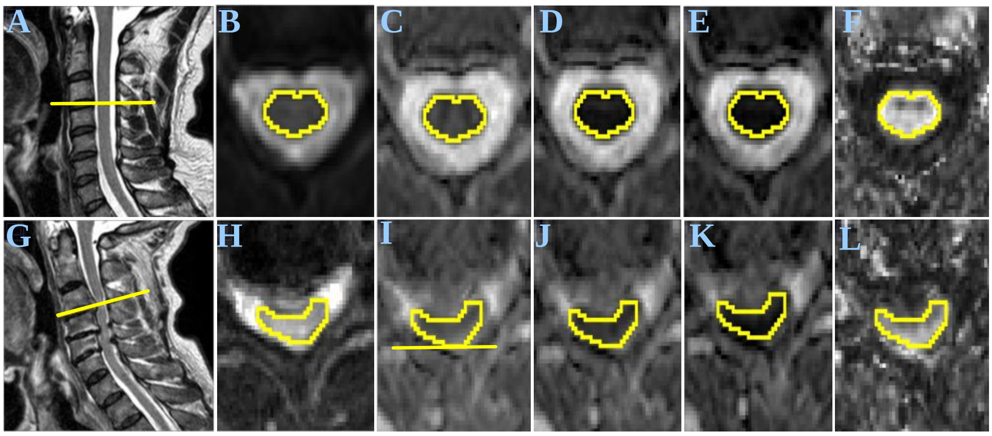

Figure 1. The typical images showing sagittal T2W, three principal eigenvector images (e0, e1, e2), B0 and FA in the healthy cord (A, B, C, D, E, F) and myelopathic cord (G, H, I, J, K, L). The ROI was defined by B0 image to cover the spinal cord.

C. DTI Processing

In raw DTI images, diffusion-weighting gradients can lead to eddy currents, which results in artifacts. Such artifacts may include shear, false fiber tracking, enhanced background, image intensity loss, and image blurring. These distortions are different for different gradient directions. The goal of the DTI processing is as follows: correct the gradient table for slice prescription and correct images for any residual eddy current distortions and motion artifacts using a nonlinear 2D registration and a 3D rigid body registration. In this study, the Automated Image Registration (AIR) program (a source code embedded in DTI Studio software, Version 2.4.01 2003, Johns Hopkins Medical Institute, Johns Hopkins University, Baltimore, MD) was employed to reduce the effect of artifact. The realigned and co-registered diffusion weighted data sets were double checked for image quality, and then used for estimation of diffusion tensors, including three eigenvalues and the corresponding eigenvectors. The region of interest (ROI) was defined to cover the spinal cord (Fig. 1).

D. Extreme Learning Methods

The diagnosis of CSM can be regarded as a classification problem. The logic behind the constructing classifier model is illustrated in Fig. 2. In this work, the extreme learning (ELM) algorithm is employed to construct the classifier model.

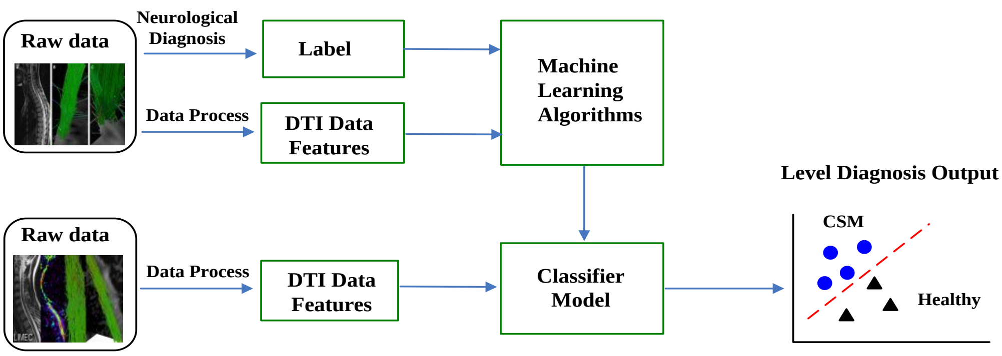

Figure 2. CSM diagnosis with a machine learning based model

ELM works for the generalized single-hidden layer feed-forward networks but the hidden layer in ELM need not to be tuned [15]. Such single-hidden layer feed-forward networks include but are not limited to polynomial network, support vector machine and RBF networks. As a new proposed machine learning algorithm, extreme learning machine has been applied to many filed [16-24]. The biggest characteristic of ELM is that all the hidden node parameters are independent from the training data-sets and target function. Therefore, these parameters can be generated at random. Given a training set D={(x_i, t_i)|i=1, 2 ...n}, activation function g(x) and the number of hidden nodes, the hidden layers can be mathematically modeled as

$$\sum_{i=1}^{n}\beta_i g(w_i \cdot x_j + b_i) = o_j,\ \text{j=1,....n}, \tag{1}$$

where w_i is the weight vector connecting the hidden layer and the input layer, β_i is the weight vector connecting the hidden layer and the output layer and bi is the threshhold of hidden layer. The SLFN with n hidden neurons with activation function g(x) can approximate these n samples with zero satifing that

$$\sum_{j=1}^{n}\left\| o_j - t_j \right\| = 0 \tag{2}$$

Given w_i, β_i and b_i, the following can be obtain

$$\sum_{i=1}^{n}\beta_i g(w_i \cdot x_j + b_i) = t_j,\ \text{j=1,...n} \tag{3}$$

The above n equations can be written as

$$H\beta = T \tag{4}$$

where

$$H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) \ .... \ g(w_n \cdot x_1 + b_n) \\ . \qquad\qquad .... \qquad . \\ . \qquad\qquad .... \qquad . \\ g(w_1 \cdot x_n + b_1) \ .... \ g(w_n \cdot x_n + b_n) \end{bmatrix},\quad \beta = \begin{bmatrix} \beta^T{}_1 \\ . \\ . \\ \beta^T{}_n \end{bmatrix},$$

$$T = \begin{bmatrix} t^T{}_1 \\ . \\ . \\ t^T{}_n \end{bmatrix}.$$

The algorithm of ELM can be summarized as the following three steps:

Step 1: Assign input weight w_i and bias bi randomly, i=1, 2 ...m.

Step 2: Calculate the hidden layer output matrix H.

Step 3: Calculate the output weight β.

III. EXPERIMENTS & RESULTS

TABLE I. RESULTS OF IDENTIFICATION OF MYELOPATHIC LEVEL FROM NEUROLOGY AND ELM

Case no.	Gender	Age	Identification of Myelopathic Level: Neurology	Identification of Myelopathic Level: ELM
1	M	76	C34, C56	C34, C45, C56
2	F	80	C34, C45	C34, C56
3	F	65	C34	C34, C45, C56
4	M	65	C34	C34, C56
5	F	43	C34	C34, C45, C67
6	M	80	C34, C45	C34, C45, C56,
7	F	63	C34, C45	C56, C67
8	M	66	C45	C45, C56
9	F	59	C34, C56	C34, C56
10	M	71	C34, C45	C34, C45
11	M	76	C34, C45	C34, C56
12	F	61	C45	C45
13	M	74	C45	C34
14	M	81	C45	C34, C45
15	F	58	C45, C56	C45
16	F	79	C34, C56	C34, C56
17	M	81	C56	C56
18	M	46	C34, C45	C34, C45
19	F	64	C34, C45	C34, C45
20	F	72	C34	C34, C45

TABLE II. TESTING RESULTS OF LEVEL DIAGNOSIS

Method	Accuracy	Sensitivity	Specificity
ELM	93.6 %	91.2%	94.7%
Control chart	76.5%	72.7%	74.2%

The DTI dataset from 20 normal people and 20 CSM patients is divided into two parts. The data from 15 control subject and 15 CSM subjects is employed for training classifiers and the remaining data is for validation. In this work, the class labels are from neurological diagnosis by senior spine surgeons. The level with CSM is defined as positive and the healthy level negative. The testing results are shown in Table I. To highlight the performance of ELM, we employ control chart method to detect the CSM levels. FA (fractional anisotropy) is the most important parameter in DT and It is derived from the eigenvalues of the diffusion tensor. In this work, the FA values from healthy subjects are employed to create a rough Six Sigma control chart then the chart is used to determine the CSM levels. The mean and standard variation of the FA for each level was calculated. In this work, standard variation multiplying 2.5 is employed to get the upper control limit and lower control limit. We employ sensitivity, accuracy, and specificity to evaluate the ELM classifier and control chart method. The experimental results are shown in Table II. From Table II, we can find that that the ELM classifiers using eigen-values have better capability in identifying the levels with CSM than control chart with FA

IV. CONCLUSIONS

In this work, a disease diagnostic assistant system using DTI and extreme learning machine is proposed to identify the levels with CSM. The Eigen values of the DTI are employed to train the ELM classifier. We compare the eigen values based classifiers with the FA values based control chart. The experimental results shows that the machine learning based classifiers have excellent capability in identifying the levels with CSM in spinal cord. The pathophysiology of cervical spondylotic myelopathy has been uncertain owing to the complexity of cervical spinal cord architecture. Neurologic examination by senior spine surgeons is one of the current methods to detect the levels with CSM in spinal cord. However, there is a disadvantage with this method. Comparing with the proposed classifier based method, neurologic examination costs a lot of time and the efficiency is not high. This will cause troubles when the patients swarm. It is necessary to develop disease diagnostic assistant system in medical science.

Comparing with neurologic examination by senior spine surgeons, ELM-based classifier has a dominant advantage in predicting the levels with CSM in spinal cord. ELM-based classifier means lower cost. The classifier can work efficiently as long as it is well trained. The proposed machine learning based method might provide a valuable view in predicting the changes of clinical symptoms and the estimated pathological severity at each level over time. In future work, we will consider training the classifier using the DTI data from all voxels of the whole spinal cord.

ACKNOWLEDGMENT

This work was supported in part by China Postdoctoral Science Foundation (Grant No. 2014M562223 and No. 2015T80925), in part by Shenzhen Basic Research Project (Grant No. JCYJ20140610151856729), in part by Natural Science Foundation of Guangdong Province (Grant No. 2014A030310154), and in part by National Natural Science Foundations of China (Grants No. 61503368 and No. 61502473).

REFERENCES

[1] R. Stuart and N. Peter, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2nd ed. 2002.

[2] S. Abidi, "Knowledge management in healthcare: Towards knowledge-driven decision-support services," *International Journal of Medical Informatics*, vol. 63, no. 1-2, pp. 5-18, 2001.

[3] N. Ezquerra, R. Mullick, E. Garcia, C. Cooke, and E. Kachouska, "PERFEX: An expert system for interpreting 3D myocardial perfusion," Georgia Institute of Technology, 1992.

[4] F. Martin, "Medical diagnosis: Test first, talk later?" *Mathemedics Inc*, vol. 1, no. 1, 2004.

[5] M. M. Thurnher and M. Law, "Diffusion-weighted imaging, diffusion-tensor imaging, and fiber tractography of the *spinal* cord," *Magn Reson Imaging Clin N Am*, vol. 17, no. 2, pp. 225–244, 2009.

[6] P. Mukherjee, J. I. Berman, S. W. Chung, C. P. Hess, and R. G. Henry, "Diffusion tensor MR imaging and fiber tractography: Theoretic underpinnings," *AJNR Am J Neuroradiol*, vol. 29, no. 4, 2008.

[7] P. Hagmann, L. Jonasson, P. Maeder, J. P. Thiran, V. J. Wedeen, and R. Meuli, "Understanding diffusion MR imaging techniques: From scalar diffusion-weighted imaging to diffusion tensor *imaging* and beyond," *Radiographics*, vol. 26, no. 1, pp. S205-23, 2006.

[8] J. L. Cui, C. Y. Wen, Y. Hu, T. H. Li, and K. D. Luk, "*Entropy*-based analysis for diffusion anisotropy mapping of healthy and myelopathic spinal cord," *Neuroimage*, vol. 54, no. 3, pp. 2125-2131, 2011.

[9] R. S. Desikan, H. J. Cabral, C. P. Hess, W. P. Dillon, C. M. *Glastonbury*, et al., "Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease," *Brain*, vol. 132, no. 8, pp. 2048–2057, 2009.

[10] B. Magnin, L. Mesrob, S. Kinkingne hun, M. Pe le ́grini-Issac, O. Colliot, et al., "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI," *Neuroradiology*, vol. 51, no. 2, pp. 73–83, 2009.

[11] A. Seichi, K. Takeshita, H. Kawaguchi, et al., "*Neurologic* level diagnosis of cervical stenotic myelopathy," *Spine*, vol. 31, no. 12, pp. 1338-1343, 2006.

[12] D. M. Montgomery, and R. S. Brower, "Cervical spondylotic myelopathy: Clinical syndrome and natural history," *Orthop Clin North Am*, vol. 23, no. 3, pp. 487–493, 1992.

[13] W. I. McDonald and T. A., Sears, "The effects of experimental demyelination on conduction in the central nervous system," *Brain*, vol. 93, no. 3, pp. 583-98, 1970.

[14] W. E. McCormick, M. P. Steinmetz, and E. C. Benzel, "Cervical spondylotic myelopathy: make the difficult diagnosis, then refer for surgery," *Cleve Clin J Med*, vol. 70, no. 10, pp. 899–904, 2003.
[15] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: A new learning scheme of feed forward neural *networks*," *IJCNN*, vol. 2, pp. 25-29, 2004.
[16] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, 2006.
[17] H. J. Rong, G. B. Huang, P. Saratchandran, and N. *Sundararajan*, "Online sequential fuzzy extreme learning machine for function approximation and classification problems," *IEEE Trans Syst. Man Cybern B Cybern*, vol. 39, no. 4, pp. 1067-1072, 2009.
[18] G. Feng, G. B. Huang, Q. Lin, and R. Gay, "Error minimized extreme learning machine with growth of hidden nodes and incremental learning," *IEEE Trans Neural Netw*, vol. 20, no. 8, pp. 1352-1357, 2009.
[19] Y. Lan, Y. C. Soh, and G. B. Huang, "Ensemble of online sequential extreme learning machine," *Neurocomputing*, *vol*. 72, no. 13-15, pp. 3391-3395, 2009.
[20] G. B. Huang, X. Ding, and H. Zhou, "Optimization method based extreme learning machine for classification," *Neurocomputing*, vol. 74, no. 1-3, pp. 155-163, 2010.
[21] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "*Extreme* learning machine for regression and multiclass classification," *IEEE Trans Syst Man Cybern B Cybern*, vol. 42, no. 2, pp. 513-529, 2012.
[22] L. L. C. Kasun, H. Zhou, G. B. Huang, and C. M. Vong, "Representational learning with extreme learning machine for big data," *IEEE Intell Syst*, vol. 28, no. 6, pp. 31-34, 2013.
[23] X. J. Lu and M. H. Huang, "A simple online *modeling* approach for a time-varying forging process," *Int J Adv Manuf Tech*, vol. 75, no. 5-8, pp. 1197-1205, 2014.

Shuqiang Wang received his Ph.D degree in System Engineering and Engineering Management from City University of Hong Kong,in 2012. He is currently an Associate Professor with Institute of Advanced Computing and Digital Engineering, Shenzhen Institute of Advanced Technology (SIAT), Chinese Academy of Science. Before join SIAT, He was a postdoctoral fellow with Department of Orthopaedic Surgery, The University of Hong Kong and a research engineer with The Noah's Ark Lab, Huawei Technologies. His current research interests include machine learning, data mining and systems biology.

Dewei Zeng received the B. S. degree in Statistics form Guangzhou Medical University School, Guangzhou China, in 2015. He is currently a Guest student in the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His current research interests including two aspects. First including tensor learning, machine learning and pattern recognition and their application in the fields of computer vision. The other part relate to robust regression method in clinical research.

Yanyan Shen received the B.S. and M. Eng. degrees in Electrical Engineering from Yanshan University, Qinhuangdao, China, in 2006 and 2009, respectively, and the Ph.D degree in the Department of Mechanical and Biomedical Engineering from City University of Hong Kong, Hong Kong SAR, China in 2012. From 2013 to 2014, she was a postdocal research fellow in Inha University in South Korea. She is currently an assistant professor in the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. Her current research interest includes energy harvesting for wireless networks, resource optimization, and cognitive radio networks.

Jinxing Hu graduated with PhD from the School of Earth and Space Sciences, Peking University in 2003. His current research areas include: 3S application integration, digital city framework and core applications, disaster monitoring and analysis, traffic data mining, etc. He hosted more than 20 research projects; published more than 30 academic papers and applied for more than 10 patents for inventions. His current research interests include 3S application integration, Digital city framework and core applications, Disaster monitoring and analysis, Traffic data mining.