| | |
|---|---|
| Title | Prediction of myelopathic level in cervical spondylotic myelopathy using diffusion tensor imaging |
| Author(s) | Wang, SQ; Li, X; Cui, JL; Li, HX; Luk, KDK; Hu, Y |
| Citation | Journal of Magnetic Resonance Imaging, 2015, v. 41 n. 6, p. 1682-1688 |
| Issued Date | 2015 |
| URL | http://hdl.handle.net/10722/213825 |
| Rights | Journal of Magnetic Resonance Imaging. Copyright © John Wiley & Sons, Inc. |

# Prediction of Myelopathic Level in Cervical Spondylotic Myelopathy Using DTI

Shu-Qiang Wang[1,2,#], PhD, Xiang Li[1,#], MD, Jiao-Long Cui[1], MD, Han-Xiong Li[3], PhD, Keith DK Luk[1] ,MBBS, Yong Hu[1*], PhD

1. Department of Orthopaedics and Traumatology, Li Ka Shing Faculty of Medicine, the University of Hong Kong
2. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
3. Department of Systems Engineering and Engineering Management, City University of Hong Kong

#: These two authors contributed equally to this paper.

"*" Correspondence Author

Dr Yong Hu

Dept. of Orthopeadics and Traumatology,

The University of Hong Kong

Address: 12 Sandy Bay Road, Pokfulam, Hong Kong

Email address: yhud@hku.hk;

Tel: (852) 29740359; Fax: (852) 29740335

**Grant Support:**

The study is supported by a grant from the Research Grants Council of the Hong Kong (HKU 774211M).

Running Title at the bottom of the page: Prediction of Myelopathic

## Abstract

**Purpose:** To investigate the use of newly designed machine learning based classifier in automatic identification of myelopathic levels in cervical spondylotic myelopathy (CSM).

**Methods:** Fifty-eight normal volunteers and sixteen subjects with CSM were recruited for diffusion tensor imaging (DTI) acquisition. The eigenvalues were extracted as the selected features from DTI images. Three classifiers, naive Bayesian, support vector machine, and support tensor machine, and fractional anisotropy (FA) were employed to identify myelopathic levels The results were compared with clinical level diagnosis result and accuracy, sensitivity and specificity were calculated to evaluate the performance of the developed classifiers. .

**Results:** The accuracy by support tensor machine was the highest (93.62%) among the three classifiers. The support tensor machine also showed excellent capacity to identify true positives (sensitivity: 84.62%) and true negatives (specificity: 97.06%). The accuracy by FA value was the lowest (76%) in all the methods.

**Conclusion:** The classifiers based method using Eigenvalues had a better performance in identifying the levels with CSM than the diagnosis using FA values. The support tensor machine was the best among three classifiers.


***Keywords*:** Cervical spondylotic myelopathy, Spinal cord, Diffusion tensor imaging, Eigenvalue, Fractional anisotropy, Machine learning.

## INTRODUCTION

Cervical spondylotic myelopathy (CSM) is the most common type of spinal cord dysfunction in patients older than 55 years of age, and the most common cause of acquired spastic paraparesis in the middle and later years of life ((1) . CSM is the result of narrowing of the cervical spinal canal by degenerative and congenital causes. Surgical treatment is recommended for patients with moderate to severe function deficit and compatible imaging findings (2), and level diagnosis is pivotal for surgical planning (2). Neurologic level diagnosis in cervical myelopathy has been employed in performing surgery relatively early (within 1 year of symptom onset), but neurological level diagnosis is considered complicated and difficult in clinical practice (2). Magnetic resonance imaging (MRI) is now widely used for evaluating spinal cord parenchyma. However, conventional MRI, such as T1- and T2-weighted imaging, is limited to providing macroscopic information, including gross deformity and hemorrhage (3). Recently, diffusion tensor imaging (DTI) has allowed detection of microarchitecture of tissue based on a rank-two diffusion tensor model (4).

DTI and fiber tractography have advanced the scientific understanding of numerous neurological and psychiatric disorders (5). The most common parameters employed in delineating spinal cord tissue microarchitecture include fractional anisotropy (FA), mean diffusivity, and apparent diffusion coefficient. All of these parameters are derived from eigenvalues to evaluate the scalar properties of water molecule diffusion (6). Eigenvectors and eigenvalues derived from the diffusion tensor matrix reflect the direction and strength of the movement of water molecules (6). Routine T1/T2 MRI techniques only provide macroscopic level information, while DTI parameters are more sensitive in showing microstructural abnormalities in cervical myelopathy (7). There is a growing interest in the application of DTI to evaluate the spinal cord microarchitecture. For example, Cui et al. (8) employed entropy-based principal eigenvector (8)for the evaluation of microstructural changes after cervical

myelopathy. Facon et al. (9) also reported that DTI can detect myelopathic cord with higher sensitivity and specificity compared with the conventional anatomical MR images. Moreover, Uda et al. (10) demonstrated a decrease in FA in most patients with cervical spondylosis.

To date, machine learning techniques have been applied to a range of MRI modalities in an effort to automate the diagnosis of mild cognitive impairment and Alzheimer's disease (11,12). However, few studies have examined the potential for DTI in conjunction with machine learning algorithms to automate the identification of a myelopathic spinal cord. Achieving the automatic classification of healthy levels and myelopathic levels should facilitate the level diagnosis of CSM and clinical determination of surgical strategy.

## MATERIALS & METHODS

### *Subjects*

This study was approved by institutional ethics committee. Written informed consent forms were signed by all subjects prior to participat this study. A total of 74 volunteers, including 58 healthy people and 16 CSM patients aged from 21 to 84 years, were recruited in this study with informed consent. The mean age of the healthy group is 40.1 with 31 males and 27 females, and the mean age of the CSM group is 67.7 with 9 males and 7 females. 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 score system (13), and negative Hoffman's sign under physical examination. Exclusion criteria included the presence of neurological signs and symptoms, or a past history of neurological injury, diseases, and operations. All the recruited patients were confirmed with the diagnosis of CSM by senior spine surgeons (two surgeons with thirty years of experience and ten years of experience respectively). The clinical diagnosis of CSM in this study are based on the neurological examination (two experienced surgeons with thirty years of

experience and ten years of experience respectively) and imaging findings (image analysis by XL with three years of experience in DTI image analysis and YH with ten years of experience in DTI image analysis). The inclusion criteria included 1）numbness or paresthesias in the upper extremities; 2） sensory changes in the lower extremities; 3） motor weakness in the upper or lower extremities; 4） gait difficulties; 5） myelopathic or "upper motor neuron" findings (ie, spasticity, hyperreflexia, clonus, Babinski and Hoffman signs, and bowel and bladder dysfunction); 6） cervical spondylosis and cord compression on conventional MRI. The exclusion criteria are patients with ossification of posterior longitudinal ligament, ossification of ligamentum flavum, congenital stenosis and other acquired compressive pathology (e.g., tumor and calcification) as well as other neurological disorders (e.g., multiple sclerosis, amyotrophic lateral sclerosis and peripheral neuropathy)(14).

***Clinical level diagnosis***

We employed the index developed by Seichi et al.(2) to define the topography of sensory disturbance, levels of segmental motor innervations and localization of the reflex center, and made level diagnosis from sensory disturbance, tendon reflexes and MMT respectively. Sensory disturbance was defined as at least one of the following three: patient-perceived numbness or sensory disturbance detected by light touch or by pinprick. Due to the possible symptom overlap of higher affected level with the lower ones, the neurological signs may be only able to detect the highest impaired level, and sometimes one or two severely impaired levels underneath. So we combined the result from sensory, motor and reflex and made the clinical level diagnosis.

***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 motion artifacts. The subject was then scanned with anatomical T1-weighted (T1W), T2-weighted (T2W) imaging, DTI.

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: 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 $mm^3$ (T1W) and 0.78×1.01×3.0 $mm^3$ (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. The parameters used in axial imaging include as follows: FOV = 80×80 mm, slice thickness = 7 mm, slice gap = 2.2 mm, fold-over direction = anterior/posterior (AP), NEX = 3, resolution =0.63×0.68×7.0 $mm^3$ (T1W) and 0.63×0.67×7.0 $mm^3$ (T2W), recon resolution = 0.56×0.56×7.0 $mm^3$ (T1W) and 0.63×0.63×7.0 $mm^3$ (T2W) and TE/TR = 8/1000 ms (T1W) and 120/4000 ms (T2W). Cardiac vector cardiogram (VCG) triggering was used to minimize the impact of the pulsation artifact from cerebrospinal fluid. Image acquisition began as soon as the rise of the wave of QRS complex. A total of 12 transverse images covering the cervical spinal cord from C1 to C7 were acquired, each of which was placed at the center of either a vertebra or an intervertebral disk. Diffusion encoding was performed in 15 non-collinear and non-coplanar diffusion directions with b-value = 600 $s/mm^2$. The parameters employed in imagine acquisition were as follows: FOV = 80×80 mm, image matrix, 128×128, slice thickness = 7 mm, slice gap = 2.2 mm, fold-over direction = AP, NEX = 3, resolution =1×1.26×7.0 $mm^3$, recon resolution = 0.63×0.63×7.0 $mm^3$ and TE/TR = 60 ms/5 heartbeats. The image slice planning was the same as in the anatomical axial T1W and T2W images, with 12 slices covering the cervical spinal

cord from C1 to C7. Due to the nature curvature of the cervical spine, it is not possible to set one stack of DTI scan with every scan slice in vertical with the course of the cord. In this study, we used three stacks to fit the curvature of cervical spinal cord while DTI scan in each stack lasted about 8 min. The average duration of the whole DTI scan was 24 min per subject, with an average heart rate of 60 beats per minute. To reduce the impact of the fold-over, spatial saturation with spectral presaturation with inversion recovery was employed. The distortion correction method based on reversed gradient polarity and parallel imaging was used to reduce the EPI distortion impact caused by increased magnetic susceptibility at 3.0-T (15,16).

***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 DTI processing is to correct the gradient table for slice prescription and correct images for any residual eddy current distortions and motion artifacts using a nonlinear two-dimensional registration and a three-dimensional 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, USA) 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 (for further information about DTI processing see Soares et al. (17)). The region of interest (ROI) was defined by B0 images to cover the spinal cord (Fig. 1). The FA values were calculated and averaged over all selected voxels in the cord for all subjects using Image J (National Institute of Health, Bethesda, MD, USA).

***Machine Learning Methods***

Given the DTI data, the identification of myelopathic levels can be defined as a bi-classification problem. This problem can be solved by introducing some machine learning based classifiers. The logic behind the method of identifying myelopathic levels is illustrated in Figure 2. The following classification methods are considered in this task: Naive Bayesian (18), SVM (19,20), and STM (21). Considering that both Naive Bayesian and SVM are relatively mature methods, we focused on the STM, as follows.

In STM, the discriminant function can be given by (21):

$$y = sign\left[ X\prod_{k=1}^{M} w_k + b \right], \qquad [1]$$

where $w_k$ indicates the weight tensor and b the offset.

In this study, X is defined as a second-order tensor with the mode-1 fiber indicating the eigenvalues of the diffusion tensor and mode-2 fiber as the region of interest (the dorsal, lateral, and ventral region of both left and right sides). The discriminant function can be described by:

$$y = sign\left[ w^T{}_1 X w_2 + b \right], \qquad [2]$$

where $w$ and $b$ can be calculated by solving the following constrained optimization problem:

$$\min_{w_k|_{k=1}^{2}\ b\ \xi} \mathrm{J}(w_k|_{\mathrm{k}=1}^{2}\ b\ \xi) = \frac{1}{2}\left\| \bigotimes_{k=1}^{2} w_k \right\|_{Fro}^{2} + c\sum_{i=1}^{N}\xi_i \qquad [3]$$

$$s.t. \quad y_i\left(w^T{}_1 X w_2 + b\right) \geq 1-\xi_i \quad \xi_i \geq 0,\ 1 \leq i \leq N,$$

where $\xi$ is introduced as the slack variable to deal with noise in DTI data.

By introducing the positive Lagrange multipliers α and λ, the above optimization problem can be rewritten as follows:

$$\max_{\lambda,\ \alpha} \min_{w_k|_{\mathrm{k}=1}^{\mathrm{M}}\ b\ \xi} L(w_k|_{\mathrm{k}=1}^{\mathrm{M}}, b,\ \xi,\ \lambda,\ \alpha) \qquad [4]$$

with

$$L(w_k \mid_{k=1}^{M}, b, \xi, \lambda, \alpha) = \frac{1}{2}\left\| \bigotimes_{k=1}^{2} w_k \right\|_{Fro}^{2} + c\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\lambda_i(y_i\left(w^T{}_1 X w_2 + b\right) - 1 + \xi_i\ ) - \sum_{i=1}^{N}\alpha_i\xi_i$$

where $\alpha_i \geq 0,\ \lambda_i \geq 0\ (1 \leq i \leq N)$.

We compared the result of machine learning methods to clinical level diagnosis and calculated accuracy, sensitivity and specificity to evaluate the performance of the employed classifiers. Accuracy is calculated by (TP+TN)/(TP+TN+FN+FP), where TP = True Positive, TN =True Negative, FP = False Positive and FN = False Negative. Sensitivity is defined as TP/(TP+FN) and Specificity is defined as TN/(FP+TN).

***Model Evaluation: Cross-Validation***

The problem with evaluating a proposed model is that it may demonstrate adequate prediction capability on the training data, but might fail to predict future unseen data. Cross-validation is a procedure for estimating the generalization performance in this context (22). In this study, there were two goals for cross-validation:

(i) To estimate performance of the learned model from available data using one algorithm; i.e., to gauge the generalizability of an algorithm.

(ii) To compare the performance of naive Bayesian, SVM, and STM, and determine the best algorithm for the available data.

In typical cross-validation, the training and validation sets must crossover in successive rounds such that each data point has a chance of being validated. The basic form of cross-validation is k-fold cross-validation. Other forms of cross-validation are special cases of k-fold cross-validation or involve repeated rounds of k-fold cross-validation (23). In this study, we employed two methods for evaluating the classifiers: 10-fold cross-validation and holdout validation (the traditional validation method). In K-fold cross-validation, the subset size, n, can be optimized by the following steps:

i. Divide the data into K roughly equal parts

ii. For each $k = 1, 2, \ldots. K$, fit the model with parameter $n$ to the other $K$-$1$ parts, and calculate its error in predicting the $k_{th}$ part. This gives the cross-validation error:

$$s(n)=\frac{1}{K}\sum_{k=1}^{K} e_k(n) \qquad [5]$$

iii. Do this for many values of $n$ and choose the value of $n$ that makes $s(n)$ smallest.

## RESULTS

### *Identification of Myelopathic Level Using Naive Bayesian, SVM, and STM*

To perform holdout validation, the DTI dataset from 20 normal people and 16 CSM patients was divided into two parts. The data from 12 normal people and 8 CSM patients were used for training classifiers, while the data from 8 normal subjects and 8 CSM patients were used for validation (Table 1). The class labels were from neurological diagnosis by senior spine surgeons. In this study, the level with CSM was defined as positive and the healthy level as negative. The neurology result in table 1 indicates the confirmed diagnosis by senior spine surgeons, which was the benchmark in our study. The subjects from case 1 to case 8 are with cervical spinal stenosis and the subjects from case 9 to case 16 are normal people. The underline in the table 1 indicates that there is no any level with CSM. The subject of case 13 has no any CSM level in the view of senior spine surgeons. But the classifier based method identified C3-4 as a CSM level. This was a false positive given by the classifier.

The statistical results of identification of myelopathic level from the three classifiers are shown in Table 2. The accuracy by STM was the highest (93.62%) of the three classifiers. STM also showed excellent capacity to identify true positives (sensitivity: 84.62%) and true negatives (specificity: 97.06%). Next, to compare the performance of naive Bayesian, SVM, and STM, we used the 10-fold cross-validation method. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and

validation, and each observation is used for validation exactly once. The statistical results of identifying myelopathic level with 10-fold cross-validation are shown in Table 3, where STM performed the best among the three classifiers (accuracy: 94.54%; sensitivity: 89.75%; specificity: 98.85%).

***Comparison with Fractional Anisotropy (FA) Identification of Myelopathic Level***

Fractional anisotropy (FA) is one of the most common parameters in DTI. FA is calculated from the eigenvalues of the diffusion tensor, with values between 0 (perfectly isotropic diffusion) and 1 (the hypothetical case of an infinite cylinder). We used FA values from 50 healthy subjects to create a threshold to detect the level with CSM. The means and standard variations of the FA for each level were: C23: 0.7286±0.0602, C34: 0.6754±0.0657, C45: 0.6877±0.0716, C56: 0.6420±0.0755, C67: 0.6518±0.0525 and C78: 0.6471±0.0767. We defined threshold of low FA value (LFV) as $\mathrm{LFV} = (\mathrm{Mean\ FA\ value} - 2.5 * \mathrm{SD})$. The level with FA value below the threshold was defined as a myelopthatic level.

The FA values from the first eight cases in Table 1 were used for myelopathic level diagnosis. The identified levels with CSM are listed in column 8 in Table 1, with the statistical results shown in row 5 (accuracy: 76.0%; sensitivity: 30.77%; specificity: 91.89%). The experimental results demonstrate that the classifiers using eigenvalues had a better ability to identify the levels with CSM than the level diagnosis by FA values.

## DISCUSSION

In this study, we proposed a data driven-based method to identify the spinal cord levels with CSM.. The eigenvalues of the DTI data were used to train the proposed classifiers, and we compared the eigenvalue-based machine learning method with the FA values. We found that the machine learning-based classifiers were excellent for identifying the levels with CSM

in spinal cord. FA is one of the most commonly used indices in DTI analysis (6,24). However, although Uda et al. (10) demonstrated a decrease in FA in most patients with CSM, the use of FA only is insufficient to detect the levels with CSM. The current data suggest that the eigenvalues from DTI data can provide more useful information in identifying the levels with CSM in spinal cord than for FA.

CSM is a degenerative disease of cervical spine, which is usually of extensive range of lesion involving multiple segments. Multilevel affected CSM is complex with clinical manifestation and difficult to precisely localize all the involved levels by neurological examination. However, not all the myelopathic levels appear with the high signal intensity on MRI (25). The myelopathic levels identified by MRI findings are usually mismatching with those from neurological examination (2,25). The surgical outcome of CSM still varies a lot and can't be predicted precisely by either neurologic deficit or any existing imaging method. Therefore, from spinal cord compression to functional deficit, there is a gap in between and that is the pathological change of spinal cord tissue, which is undetectable by conventional methods but DTI images.

In clinical practice, level diagnosis is critical in determining decompression levels and surgical approach (anterior or posterior). In case of multilevel CSM, conventional MRI may detect some mild compression levels beside a severest compressed level. If the neurological deficit does not clearly point to the mildly compressed levels, it will be controversial for the surgical planning. Surgical decision made purely on conventional MRI will put the patients at the risk of 'over-killing' or inadequate decompression. DTI could disclose the micro-structure impairment and reflect the pathological changes of spinal cord based on its unique principle. With the method introduced in present study, the myelopathic levels could be identified in CSM patients based on the detection of demyelination of white matter. It could provide the pathological condition of impaired cervical spinal cord in an efficient way.

In the present study, STM was especially suitable for identification of spinal cord levels with CSM. As a relatively new learning approach, STM has some potential advantages in dealing with DTI data. While traditional linear classification algorithms like SVM find a classifier in $R^m$, STM finds a classifier in tensor space $R^{m1} \otimes R^{m2}$, which provides a structured classification. Therefore, STM can use the DTI data structure, while SVM often results in data structure loss when translating a tensor into a vector. The number of independent unknown parameters in STM is also less than that of SVM. For example, a vector $X \in R^n$ can be transformed to a second order tensor $X \in R^{n_1} \otimes R^{n_2}$, where $n \approx n_1 \times n_2$. In SVM, a linear classifier can be represented as $k^T x+b$ in which there are $n+1$ ($\approx n_1 \times n_2+1$) independent parameters ($b$, $k_i$, $i$=1, 2…n). In STM, a linear classifier can be represented as $\boldsymbol{\omega}_1^T \mathbf{X} \boldsymbol{\omega}_2+b$ where $\boldsymbol{\omega}_1 \in \mathbf{R}^{n1}$ and $\boldsymbol{\omega}_1 \in \mathbf{R}^{n2}$. Thus, there are only $n_1+n_2+1$ parameters. This property makes STM especially suitable for small sample cases and robust against over-fitting.

Prior to the use of advanced imaging techniques, neurological examination was the main approach to estimate the level of myelopathy, and remains an essential method for evaluating the severity and location of the lesion. The cervical cord segments approximately correspond to one or two higher intervertebral levels in CSM (2), owing to the different anatomical relationship between cord segments and spinal roots with regard to intervertebral levels. Although it is difficult to distinguish all the myelopathic versus normal levels in cases of multilevel involvement, neurologic examination can provide the most direct evidence that certain levels of the cervical spinal cord exhibit myelopathy. Thus, neurological examination remains a benchmark for identifying cervical spinal cord myelopathy for comparisons to DTI methods.

In this study, we demonstrated that the machine learning-based classifiers using eigenvalues of DTI can provide a direct measure of the level of myelopathy, with the STM-

based classifier providing the optimal detection method. There are several advantages of the STM-based classifier in identifying the levels with CSM in the spinal cord. First, compared with the vector space model, STM can exploit the DTI data structure, as well as correlations in the original data. Thus, the STM-based classifier allows the detection of CSM levels with higher accuracy and sensitivity. Second, the number of independent parameters in STM is less than that of the vector-based classifier, which makes the STM-based classifier more robust against over-fitting compared with the vector space model, such as the SVM-based classifier. This also allows the STM-based classifier to deal with small sizes, which are very common in medical science. Finally, the STM-based classifier is more cost effective than neurological examination by senior spine surgeons, and this classifier can work efficiently as long as it is well trained. Therefore, the proposed STM-based classifier would provide additional diagnosis of myelopathic levels for surgeons to make the most appropriate surgical plan.

Several issues in the STM-based classifier should be considered. The first is how to sort the features in the tensor. In SVM, we implicitly assume that the features are independent (19). A classifier in vector space can be written as $k^TX+b$. Obviously, the change of the order of the features has no impact on training the classifier. In the tensor space model, a linear classifier is represented as $V^TXU+b$. Therefore, the independence assumption for features no longer holds in training the tensor-based classifier, and different feature sorting will lead to different training results in the tensor space model. In this study, we sorted the features into descending order of eigenvalues. The second issue is the loss of features caused by the ROI. Note that the eigenvalue employed in this study is the average of those from the voxels drawn within the ROI. Thus, the definition of the ROI is important. Unsuitable drawing of the ROI may lead to feature loss in training the classifier. In future studies, we will consider training the classifier using DTI data from all voxels of the whole spinal cord. Finally, identifying the

levels with CSM is only the first step. Estimating the pathological severity at each level is necessary in clinical practice.

There are some limitations in the present study. Since CSM has a vast array of signs and symptoms and there are no pathognomonic findings, selection bias may happen in the inclusion of patients by experienced surgeons. Besides, different disease severity and compression pattern in multilevel cases may also bias the result. Another limitation of the present study is the lack of neurological examination results from CSM patients after treatment, which may provide sufficient information of justifying the classifiers. Although neurology could not reveal the precise information of myelopathy along cervical cord segments, it is the only available and acceptable benchmark after a careful clinical diagnosis by experienced spine surgeons. A more appropriate reference standard should be based on the surgical outcome following level diagnosis, which need a large scale clinical trial in the future study.

In conclusion, the proposed machine learning-based method might provide a valuable method for predicting the changes of clinical symptoms and the estimated pathological severity at each level over time. With the proposed classifiers, we could detect the myelopathic levels in CSM and give useful reference to spine surgeons in decision making of surgical plan in some complicated cases.

## REFERENCE


1. Montgomery D, Brower R. Cervical spondylotic myelopathy. Clinical syndrome and natural history. The Orthopedic clinics of North America 1992;23(3):487-493.
2. Seichi A, Takeshita K, Kawaguchi H, et al. Neurologic level diagnosis of cervical stenotic myelopathy. Spine 2006;31(12):1338-1343.
3. Baron EM, Young WF. Cervical spondylotic myelopathy: a brief review of its pathophysiology, clinical course, and diagnosis. Neurosurgery 2007;60(1 Supp1 1):S35-41.
4. Thurnher M, Law M. Diffusion-weighted imaging, diffusion-tensor imaging, and fiber tractography of the spinal cord. Magnetic resonance imaging clinics of North America 2009;17(2):225-244.
5. Mukherjee P, Berman J, Chung S, Hess C, Henry R. Diffusion tensor MR imaging and fiber tractography: theoretic underpinnings. AJNR American journal of neuroradiology 2008;29(4):632-641.
6. Hagmann P, Jonasson L, Maeder P, Thiran JP, Wedeen VJ, Meuli R. Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. Radiographics : a review publication of the Radiological Society of North America, Inc 2006;26 Suppl 1:S205-223.
7. Demir A, Ries M, Moonen C, et al. Diffusion-weighted MR imaging with apparent diffusion coefficient and apparent diffusion tensor maps in cervical spondylotic myelopathy. Radiology 2003;229(1):37-43.
8. Cui J-L, Wen C-Y, Hu Y, Li T-H, Luk K. Entropy-based analysis for diffusion anisotropy mapping of healthy and myelopathic spinal cord. Neuroimage 2011;54(3):2125-2131.
9. Facon D, Ozanne A, Fillard P, Lepeintre J-F, Tournoux-Facon C, Ducreux D. MR diffusion tensor imaging and fiber tracking in spinal cord compression. AJNR American journal of neuroradiology 2005;26(6):1587-1594.
10. Uda T, Takami T, Tsuyuguchi N, et al. Assessment of cervical spondylotic myelopathy using diffusion tensor magnetic resonance imaging parameter at 3.0 tesla. Spine 2013;38(5):407-414.
11. Desikan R, Cabral H, Settecase F, et al. Automated MRI measures predict progression to Alzheimer's disease. Neurobiology of aging 2010;31(8):1364-1374.
12. Magnin B, Mesrob L, Kinkingnéhun S, et al. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology 2009;51(2):73-83.
13. Yonenobu K, Abumi K, Nagata K, Taketomi E, Ueyama K. Interobserver and intraobserver reliability of the japanese orthopaedic association scoring system for evaluation of cervical compression myelopathy. Spine 2001;26(17):1890.

14. McCormick WE, Steinmetz MP, Benzel EC. Cervical spondylotic myelopathy: make the difficult diagnosis, then refer for surgery. Cleveland Clinic journal of medicine 2003;70(10):899-904.
15. Chuang T-C, Huang T-Y, Lin F-H, et al. PROPELLER-EPI with parallel imaging using a circularly symmetric phased-array RF coil at 3.0 T: application to high-resolution diffusion tensor imaging. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 2006;56(6):1352-1358.
16. Morgan P, Bowtell R, McIntyre D, Worthington B. Correction of spatial distortion in EPI due to inhomogeneous static magnetic fields using the reversed gradient method. J Magn Reson Imaging 2004;19(4):499-507.
17. Soares J, Marques P, Alves V, Sousa N. A hitchhiker's guide to diffusion tensor imaging. Frontiers in neuroscience 2013;7:31.
18. Domingos P, Pazzani M. Beyond independence: Conditions for the optimality of the simple bayesian classi er. Proc 13th Intl Conf Machine Learning 1996.
19. Cortes C, Vapnik V. Support-vector networks: 1995.
20. Steinwart I, Christmann A. Support vector machines. 2008.
21. Kotsia I, Guo W, Patras I. Higher rank support tensor machines for visual recognition. Pattern Recognition 2012.
22. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI 1995.
23. Stone M. Asymptotics for and against cross-validation: 1977.
24. Budzik JF, Balbi V, Le Thuc V, Duhamel A, Assaker R, Cotten A. Diffusion tensor imaging and fibre tracking in cervical spondylotic myelopathy. European radiology 2011;21(2):426-433.
25. Zhang L, Zeitoun D, Rangel A, Lazennec JY, Catonné Y, Pascal-Moussellard H. Preoperative evaluation of the cervical spondylotic myelopathy with flexion-extension magnetic resonance imaging: about a prospective study of fifty patients. Spine 2011;36(17):9.

Table 1. Results of identification of myelopathic level from neurology, SVM, Bayesian and STM and FA value.

| Case no. | Gender | Age | Identification of Myelopathic Level | | | | |
|---|---|---|---|---|---|---|---|
| | | | Neurology | Bayesian | SVM | STM | FA |
| 1 | F | 79 | C34, C45 | C45, C56 | C34, C56 | C34, C45, C56 | - |
| 2 | M | 83 | C34, C56 | C34, C45, C56 | C45, C56 | C34, C56 | - |
| 3 | F | 62 | C34, C45 | C34, C56 | C45, C67 | C34, C45, C56 | C45,C78 |
| 4 | M | 65 | C34 | C34, C56 | C34,C45 | C34, C56 | C34,C45, C67 |
| 5 | F | 43 | C34 | C45, C56 | C56 | C34, C45, C67 | C34,C45 |
| 6 | M | 80 | C34, C45 | C45, C56 | C34, C56 | C34, C45, C56, | C23, C45,C56,C67 |
| 7 | F | 63 | C34, C45 | C45 | C45,C56 | C56, C67 | C67 |
| 8 | M | 60 | C56 | C34, C56 | C45,C56 | C56, C67 | - |
| 9 | M | 29 | - | - | - | - | - |
| 10 | M | 30 | - | - | - | - | - |
| 11 | F | 49 | - | - | - | - | - |
| 12 | M | 26 | - | - | - | - | - |
| 13 | F | 59 | - | C34,C45 | C34 | C34 | - |
| 14 | F | 52 | - | - | - | - | - |
| 15 | M | 50 | - | C56 | - | - | C34, C56 |
| 16 | M | 54 | - | - | - | - | - |

"-" indicates there is no level with CSM for the subject.

Table 2 Statistical results of identification of myelopathic level

| Method | Accuracy | Sensitivity | Specificity |
|---|---|---|---|
| Bayesian | 80.85 % | 61.54% | 88.24% |
| SVM | 82.48% | 53.85% | 94.12% |
| STM | 93.62% | 84.62% | 97.06% |
| FA | 76.0% | 30.77% | 91.89% |

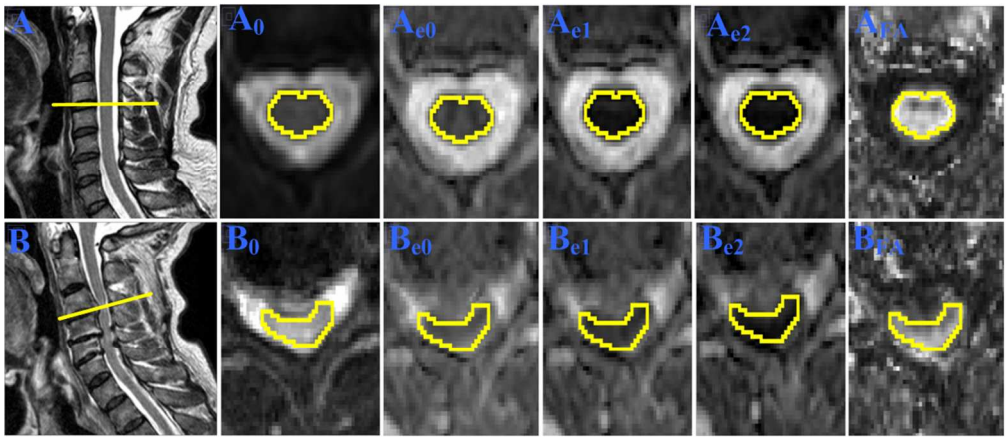
Table 3 Statistical results of identification of myelopathic level with 10-fold cross-validation

| Method | Accuracy | Sensitivity | Specificity |
|---|---|---|---|
| Bayesian | 83.51 % | 66.47% | 90.31% |
| SVM | 83.27% | 59.39% | 96.48% |
| STM | 94.54% | 89.75% | 98.85% |

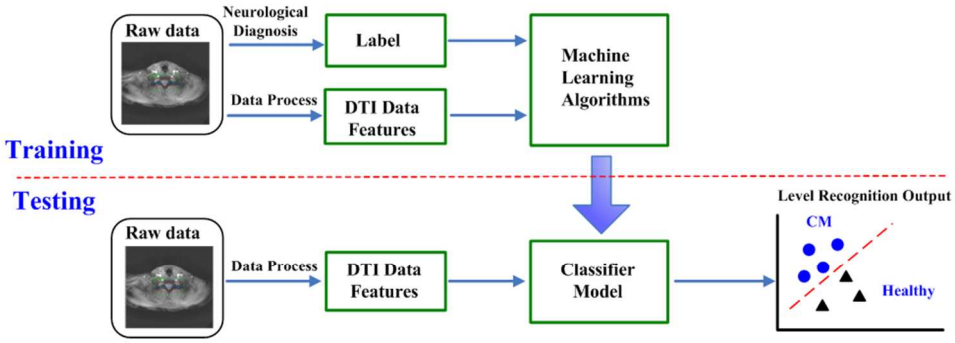
Figure legends

Figure 1. The representative images showing sagittal T2W, B0, three principal eigenvector images (e0, e1, e2) and FA in the healthy cord (A, $A_0$, $A_{e0}$, $A_{e1}$, $A_{e2}$, $A_{FA}$) and myelopathic cord (B, $B_0$, $B_{e0}$, $B_{e1}$, $B_{e2}$, $B_{FA}$). The ROI was de fined by B0 image to cover the spinal cord.

Figure 2. The framework of identifying myelopathic levels using machine learning based classifier model. In training step, the labeled DTI data is employed to classifiers. The machine learning algorithms employed in this work are naive Bayesian, support vector machine and support tensor machine.




The representative images showing sagittal T2W, B0, three principal eigenvector images (e0, e1, e2) and FA in the healthy cord (A, A0, Ae0, Ae1, Ae2, AFA) and myelopathic cord (B, B0, Be0, Be1, Be2, BFA). The ROI was de fined by B0 image to cover the spinal cord.
327x144mm (96 x 96 DPI)




The framework of identifying myelopathic levels using machine learning based classifier model. In training step, the labeled DTI data is employed to classifiers. The machine learning algorithms employed in this work are naive Bayesian, support vector machine and support tensor machine.
269x104mm (96 x 96 DPI)