# CLASSIFICATION OF DIFFUSION TENSOR METRICS FOR THE DIAGNOSIS OF A MYELOPATHIC CORD USING MACHINE LEARNING

**SHUQIANG WANG**

*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China*

**YONG HU**

*Department of Orthopaedic Surgery, The University of Hong Kong*

*12 Sandy Bay Road, Hong Kong , Duchess of Kent Children's Hospital, Pokfulam, Hong Kong, Tel: (852) 29740359,*

*Fax: (852) 29740335, E-mail: yhud@hku.hk (corresponding author)*

**YANYAN SHEN**

*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China*

**HANXIONG LI[a, b]**

*[a]Department of Systems Engineering and Engineering Management, City University of Hong Kong*

*[b]School of Mechanical and Electrical Engineering, Central South University, Changsha, China*

***Abstract*:** In this study, we propose an automated framework that combines diffusion tensor imaging (DTI) metrics with machine learning algorithms to accurately classify control groups and groups with cervical spondylotic myelopathy (CSM) in the spinal cord. The comparison between selected voxel-based classification and mean value based classification were performed. A support vector machine(SVM) classifier using a selected voxel-based dataset produced an accuracy of 95.73%, sensitivity of 93.41% and specificity of 98.64%. The efficacy of each index of diffusion for classification was also evaluated. Using the proposed approach, myelopathic areas in CSM are detected to provide an accurate reference to assist spine surgeons in surgical planning in complicated cases.

*Keywords:* Diffusion tensor imaging; machine learning; feature selection; cervical spondylotic myelopathy; diffusion indices

## 1. **Introduction**

Recently, machine learning techniques have been used in combination with multi-source medical data to provide efficient diagnostic tools for neurological diseases[1-6]. The diagnosis of Alzheimer's disease has been studied using probabilistic neural networks[7], the spatio-temporal wavelet-chaos approach[8], and support vector machines (SVMs)[9] using EEG data and fMRI data. Parkinson's disease has been modeled using the recently developed enhanced probabilistic neural network (EPNN)[10, 11]. Nonconvulsive seizures have been detected by a wavelet statistical feature approach using the EEG signal[12]. Scans of mild cognitive impairment have been classified using cross-sectional pattern analysis[13].

Cervical spondylotic myelopathy (CSM) is the most common type of spinal cord dysfunction in patients over 55 years old[14]. Although machine learning techniques are effective as a computer-aided diagnosis system[15-20], few researchers have studied the potential to classify myelopathic regions in the cervical spinal cord using diffusion tensor imaging (DTI) in conjunction with machine learning tools[21]. This combination of techniques is likely to be a useful diagnostic tool because DTI parameters are more sensitive in detecting the microstructural abnormalities of the cervical spinal cord compared with routine T1/T2 MRI techniques that only account for macroscopic information[22-25]. In our previous work[26], we proposed a method to identify myelopathic levels in the cervical spinal cord by combining DTI with a machine learning technique, and we demonstrated that machine learning-based classifiers using the eigenvalues of DTI can provide a direct measure of the level of myelopathy. However, there remain some issues in the previous work[26]: (1) One key step in the diagnosis of myelopathic levels by combining DTI with machine learning methods is to select those key features of DTI from the region of interest (ROI). However, for simplicity, the mean of the DTI eigenvalues is extracted from the ROI to identify the myelopathic regions. From the perspective of feature extraction[27], this treatment may lead to a loss of valid features that are essential for the classification of healthy controls and CSM subjects. This may be the reason that the true positive rate calculated by the SVM/Bayesian classifier was lower in a previous work[26]. (2) In clinical trials, the most common parameters used to characterize spinal cord tissue micro-architecture with DTI are fractional anisotropy (FA), axial diffusion (AD), mean diffusivity (MD), and radial diffusion (RD)[28]. To analyze the diffusion process of a water molecule in myelopathic regions, it is important to evaluate the efficacy of these indices (FA, AD, MD, and RD) to classify myelopathic regions in the spinal cord. In the previous work, we only used the eigenvalues of the diffusion tensor to identify the myelopathic regions and ignored the use of these diffusion indices (FA, AD, MD, and RD). (3) Anatomically, the white matter (WM) tracts of the ROI are subdivided into dorsal, lateral, and ventral sections[29]. The dorsal, lateral, and ventral gray matter sections contain the neuron cell bodies for the sensory, sympathetic, and motor systems, respectively[30]. It is necessary to clarify which section contributes most to spinal cord dysfunction and which section is more susceptible to the narrowing of the cervical spinal canal by degenerative and congenital issues, especially for prompt surgical intervention. However, in the previous work, we failed to determine the most important section in the ROI for CSM classification.

Our aim in the current study is to address the above issues. Our main contributions are the following: (1) The original features from the important voxel-based dataset are extracted to automate the identification of myelopathic regions in the cervical spinal cord. A voxel-based classification and mean value-based classification are compared to demonstrate the strength of key DTI features for detecting myelopathic regions. (2) The efficacy of each index of diffusion (FA, AD, MD, and RD) for CSM classification is assessed. (3) Based on the partition of the WM tracts of the ROI, the locations of the voxels (dorsal, lateral, and ventral) that are most useful for discriminating between healthy controls and CSM subjects are determined. The experimental results are consistent with the previous clinical trial.

## 2. Materials and Methods

### *2.1 Subjects*

A protocol was approved specifically for this study by the ethics committee of The University of Hong Kong. All subjects signed informed consent forms prior to their participation in this study. Seventy volunteers were recruited, including 40 healthy subjects and 30 CSM patients, from 22 to 84 years old. The control group consisted of 21 males and 19 females with a mean age of 40.2, and the CSM group consisted of 17 males and 13 females with a mean age of 57.4. Volunteers with complete motor and sensory function confirmed by the Japanese Orthopaedic Association score system[31] were considered to be eligible for the control group. Volunteers with abnormal neurological symptoms and/or a history of neurological diseases, injury, and operations were excluded from the control group. CSM subjects diagnosed by senior spinal surgeons were considered to be eligible for the CSM group. The neurological examinations by senior spinal surgeons included an investigation of deep tendon reflexes, manual muscle

strength testing, and sensory disturbance areas. Patients who had sensory disturbances experienced sensory disturbance detected by a pinprick or light touch, patient-perceived numbness, or a combination of these symptoms.

### *2.2 Imaging methods and DTI processing*

Imaging was performed using a Philips Achieva 3.0 Tesla MR system. Each subject was placed supine and instructed not to swallow to minimize motion artifacts during the acquisition process. An SNV head and neck coil was positioned to enclose the cervical region, and then each subject was scanned with anatomical T1-weighted (T1W) imaging, T2-weighted (T2W) imaging, and DTI sequentially. This imaging protocol was approved by the Institutional Review Board of The University of Hong Kong

Sagittal and axial T1W and T2W images were acquired for each subject using a fast spin-echo sequence. The parameters employed for sagittal imaging were 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 $mm^3$ (T1W) and 0.78 × 1.01 × 3.0 $mm^3$ (T2W), recon resolution = 0.49 × 0.49 × 3.0 $mm^3$, and echo time / repetition time (TE/TR) = 7.2/530 ms (T1W) and 120/3314 ms (T2W). The parameters used for axial imaging were as follows: FOV = 80 × 80 mm, 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 vectorcardiogram triggering was used to minimize cerebrospinal fluid pulsation artifacts. Image acquisition was triggered by the rise of the wave of the QRS complex. Twelve transverse images were acquired with spinal cord from C1 to C7 acquired. For more details on the experimental parameters, please see a previous article[26].

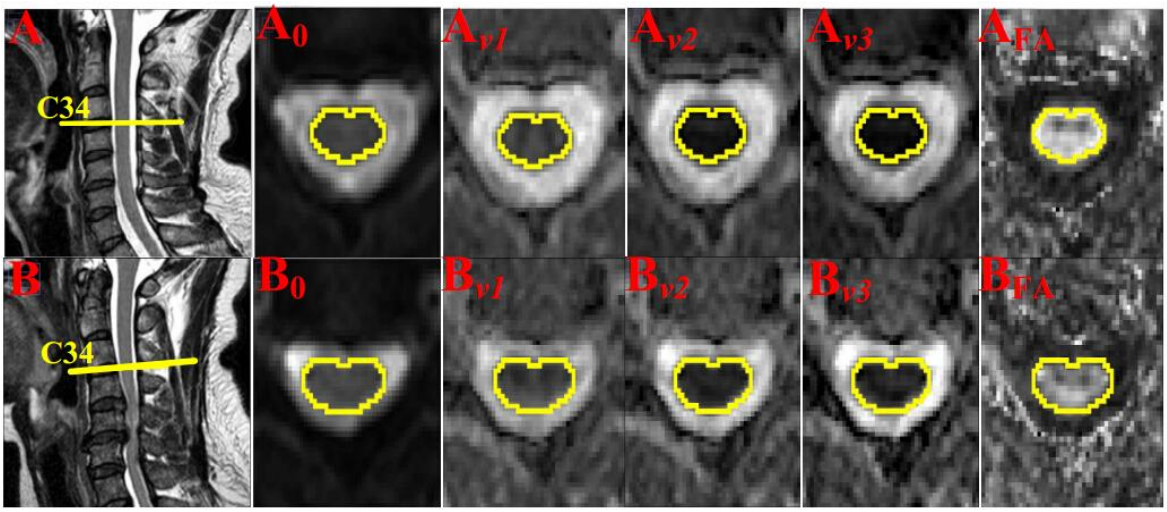


Figure 1. The representative images showing sagittal T2W, B0, three principal eigenvector images ($\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$) and FA in the healthy cord (A, $A_0$, $A_{v1}$, $A_{v2}$, $A_{v3}$, $A_{FA}$) and myelopathic cord (B, $B_0$, $B_{v1}$, $B_{v2}$, $B_{v3}$, $B_{FA}$). The region of interest (ROI) was defined by B0 image to cover the spinal cord.

To minimize the influence of echo planar imaging distortion, a distortion correction tool[32, 33] based on parallel imaging and reversed gradient polarity was used. To correct motion artifacts in addition to the gradient table for slice prescription, the Automated Image Registration (AIR) (Laboratory of Neuroimaging, UCLA, Los Angeles, California, USA) tool[34] was used. To ensure image quality, the registered and realigned diffusion-weighted datasets were double-checked. The processed datasets were then used to calculate the diffusion tensors, including three eigenvalues ($\lambda_1$, $\lambda_2$, and $\lambda_3$) and their corresponding eigenvectors ($\boldsymbol{v}_1$, $\boldsymbol{v}_2$, and $\boldsymbol{v}_3$). Then, the diffusion tensor measurement was performed using DTI Studio software (Version 2.4.01 2003, Johns Hopkins Medical Institute, Johns Hopkins University, Baltimore, Maryland, USA). The ROIs were defined manually draw on the B0 images that covered the spinal cord. The ROI definition and morphometry measurements were performed using ImageJ (National Institute of Health, USA). The mean value of FA, AD, and RD within the ROI were derived from the diffusion matrix accordingly.

The ROI was specified using B0 images that covered the spinal cord, as shown in Fig. 1. For each voxel in the ROI, the FA, MD: $(\lambda_1 + \lambda_2 + \lambda_3)/3$, AD: $\lambda_1$, and RD: $(\lambda_2 +$

$\lambda_3)/3$ were calculated along the entire length of the cervical spinal cord (C2-3, C3-4, C4-5, C5-6, C6-7, and C7-8). The FA was calculated according to the following equation[35]:

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - \lambda_0)^2 + (\lambda_2 - \lambda_0)^2 + (\lambda_3 - \lambda_0)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \quad (1)$$

where $\lambda_0 = (\lambda_1 + \lambda_2 + \lambda_3)/3$.

### *2.3. Classification analysis based on machine learning algorithms*

The ROI contained 400 voxels in the WM skeleton (Fig. 1). The identification of myelopathic areas in the spinal cord can be considered as a bi-classification problem, which may be solved by introducing machine learning-based classifiers. The rationale of this method is illustrated in Fig. 2. We used machine learning algorithms, including an SVM[36, 37], support tensor machine (STM)[38], and naive Bayes (NB)[39]. The SVM and NB programs were from the freely available WEKA[40] software package (http://www.cs.waikato.ac.nz/ml/weka). The STM program was written by the present author and is available on request.

In the following, we consider the SVM as an example to analyze the classification using DTI data. First, the skeletonized FA, DA, RD, and MD data was analyzed in MATLAB (a program written by SQ Wang[26] and available on request), which extracted the diffusion values from the WM skeleton and transformed them into a WEKA compatible format. The classification between groups was undertaken using each index of diffusion separately to determine the most efficient index for classification. The most efficient index from each location (the dorsal, lateral, and ventral aspects of the ROI) was then used. Finally, multiple classifiers were used to distinguish between healthy areas and areas with CSM.

To mitigate the effects of noise, the number of voxels was first reduced to include only those that were most relevant for distinguishing between control groups and CSM groups. This step can delete non-discriminative voxels that may lead to loss of accuracy. In this study, the ReliefF algorithm[41, 42] was used to select the most relevant voxels from the full datasets for each subject. The ReliefF algorithm is a general and successful attribute estimator. It can detect conditional dependencies between attributes and provide a unified view of attribute estimation regarding regression and classification. The aim of the ReliefF algorithm is to estimate the quality of voxels according to how well the value of a voxel distinguishes between instances that are near to each other. The algorithm works on the assumption that the voxels of nearby individuals with different diagnoses are the most useful for assessing the predictive ability of the voxel. According to the degree for which the value of a voxel can distinguish between samples that are near to each other, the ReliefF algorithm can estimate the quality of the voxels. In the current work, for each index of diffusion, 10 reduced datasets were created with 20, 30, 40, 50, 60, 80, 120, 170, 230, and 280 voxels. The sizes of the above reduced datasets were chosen based on a previous article[26]. To date, approximately 40–80 voxels provide optimal classification results.

By training the classifier on labeled data, the SVM using the linear model can assign labels to new samples with nonlinear class boundaries using kernel transformation. The optimal kernel function is usually determined by trial and error. In the current work, mapping the original samples to a higher-dimensional space was implemented using a radial basis function (RBF) kernel. RBF kernels use two parameters: C and GAMMA. GAMMA represents the width of the RBF, and C represents the error/trade-off parameter that adjusts the importance of the separation error in the creation of the separation surface. In the current work, C was set to one and GAMMA was set to 0.05. Once the SVM was trained, a new test subject could be labeled based on the distance between the subject and separating hyperplane.

The classification accuracy was evaluated using a 10 times 10-fold cross validation method[43] to ensure performance generalization. For each run of the 10-fold cross validation, the data was randomly divided into 10 parts in which each class was represented in approximately the same proportions as in the full dataset. Each fold was held out in turn and the learning scheme trained on the remaining nine-tenths. The error rate was then calculated on the 10th fold. Thus the learning procedure was executed a total of 10 times on different training sets. The 10 error estimates were averaged to yield an overall error estimate. This procedure was repeated 10 times, which resulted in the learning algorithm being implemented 100 times on datasets that were all nine-tenths the size of the original. This is a standard procedure in machine learning that reduces the variation related to data selection and allows results to be averaged to yield robust calculations of the performance of classifiers. To analyze the results, measures of sensitivity, specificity, accuracy, and the area under the curve for the receiver operating characteristic (ROC) curve are given. Accuracy is defined as (TP + TN)/(TP + TN + FN + FP), where TN = true negative, TP = true positive, FN = false negative, and FP = false positive. Sensitivity is defined as TP/(TP + FN) and specificity is defined as TN/(FP + TN).

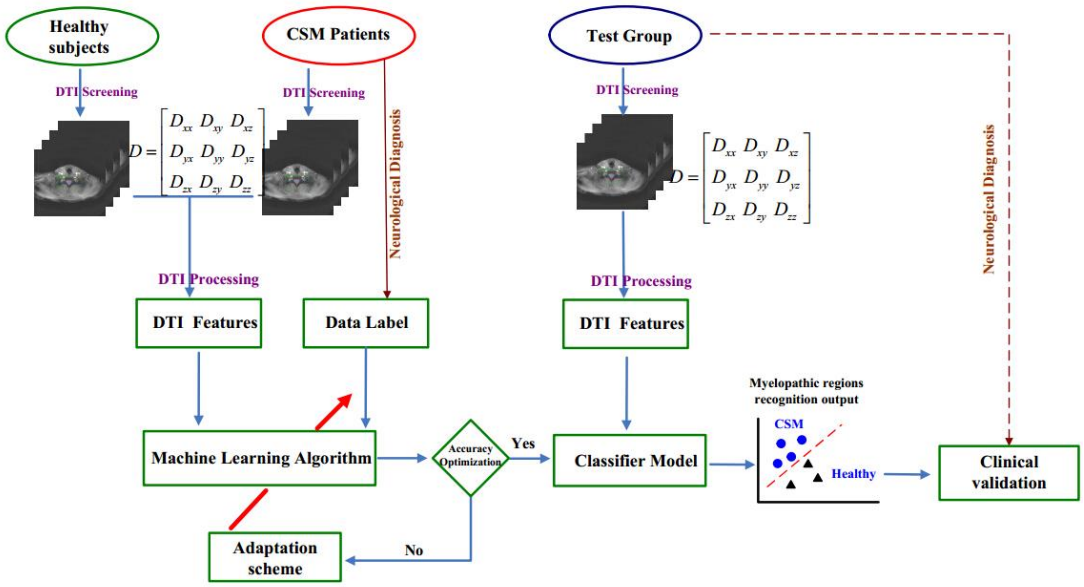


Figure 2. Framework of a machine learning-based classifier of myelopathic levels. The labeled DTI data is employed as train data. The machine learning algorithms employed in this work are support vector machine, support tensor machine, naive Bayes. The red arrow indicates the fine-tuning of model parameters using the labeled data

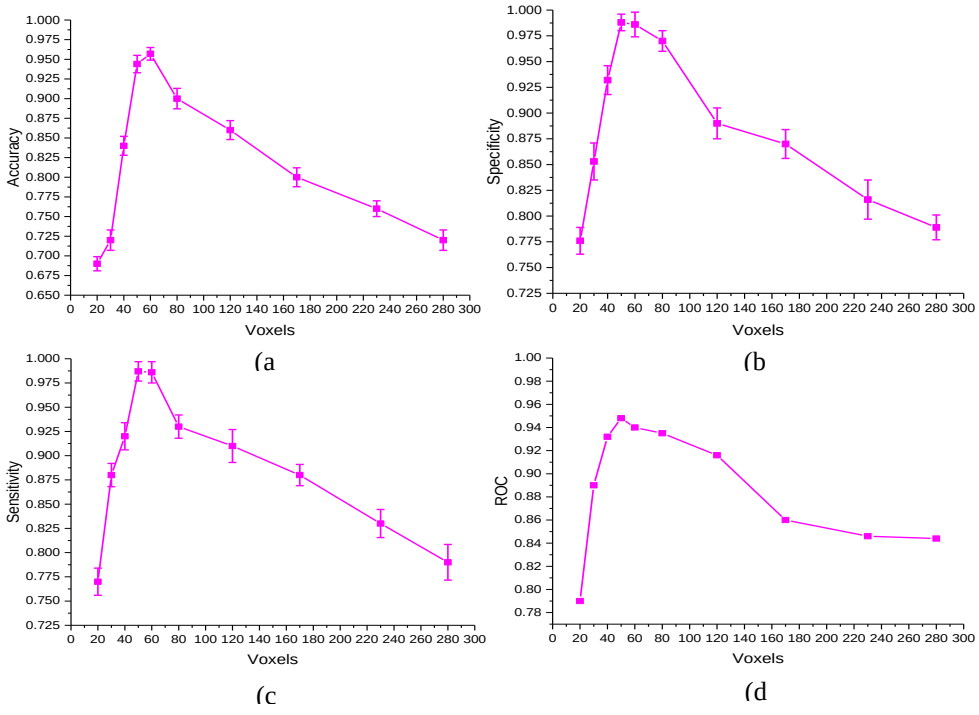


Figure 3. Accuracy, sensitivity, specificity and the area under the curve for a receiver operating characteristic curve (ROC AUC) for control and CSM classification using SVM. The values indicated are weighted averages for the two classes under consideration. Results are shown for ten datasets – 20 voxel dataset, 30 voxel dataset, 40 voxel dataset, 50 voxel dataset, 60 voxel dataset, 80 voxel dataset, 120 voxel dataset, 170 voxel dataset, 230 voxel dataset and 280 voxel dataset. The voxels comprising these reduced datasets were selected by the ReliefF algorithm.

## 3. Results

### *3.1. Identifying myelopathic levels using eigenvalues of a selected voxel dataset*

The myelopathic levels were first identified using the eigenvalues of a selected voxel dataset using the SVM classifier. The number of selected voxels influenced the performance of the classifier regarding accuracy (Fig. 3(a)), sensitivity (Fig. 3(b)), specificity (Fig. 3(c), and the area under the ROC curve (ROC AUC) (Fig. 3(d)). The error bars using standard deviations are also shown in Fig. 3 to indicate the variability of the classifications. Results are shown for 10 datasets for 20, 30, 40, 50, 60, 80, 120, 170, 230, and 280 voxels. The voxels that comprised these reduced datasets were selected using the ReliefF algorithm. Fig. 3 shows that the best performance of the SVM was obtained for datasets from 50 to 60 voxels. Table 1 shows that the classification using the selected voxel dataset performed better than that with the mean value, especially for the classification using the SVM. The SVM classifier using the 60 voxel dataset produced an accuracy of 95.73%, sensitivity of 93.41%, and specificity of 98.64%, whereas the classification using the mean value only achieved an accuracy of 81.52%, sensitivity of 54.33%, and specificity of 92.43%. The classification results for orientation entropy[44] (OE) are also presented in Table 1. It can be observed that the proposed classifiers using the selected voxel-based dataset outperformed the OE method. The Wilcoxon test was conducted to compare the performance values of the SVM, STM, and NB when used with the selected 60 voxel dataset. The results are presented in Table 2. For all the comparisons in this study, the significance level α was set to 0.05. The *p*-values were calculated to highlight the significance of the differences. The results demonstrate that the SVM and STM outperformed NB significantly, whereas the SVM and STM did not have significantly different performance.

Table 1. Classifier performance of SVM, Bayesian, STM and OE using mean values and selected voxel-based dataset.

| **Method** | | **Accuracy** | **Sensitivity** | **Specificity** |
|---|---|---|---|---|
| The selected 60 voxel dataset | Bayesian | 90.83 % | 84.52% | 95.93% |
| | SVM | 95.73% | 93.41% | 98.64% |
| | STM | 95.31% | 92.27% | 98.96% |
| Mean of ROI | Bayesian | 80.16 % | 61.29% | 87.81% |
| | SVM | 81.52% | 54.33% | 92.43% |
| | STM | 92.84% | 84.16% | 96.56% |
| Reference [44] | OE | 80.42% | 83.51% | 91.47% |

Table 2. Wilcoxon test comparison for the performance of SVM, STM and NB using the selected 60 voxel dataset. $R^+$ corresponds to the sum of the ranks for the method on the left and $R^-$ for the right.

| Method | Accuracy | | | Sensitivity | | | Specificity | | |
|---|---|---|---|---|---|---|---|---|---|
| | $R^+$ | $R^-$ | p-value | $R^+$ | $R^-$ | p-value | $R^+$ | $R^-$ | p-value |
| SVM vs. STM | 41 | 14 | 0.169 | 42 | 13 | 0.313 | 38 | 17 | 0.527 |
| SVM vs. NB | 55 | 0 | **0.005** | 55 | 0 | **0.005** | 55 | 0 | **0.005** |
| STM vs. NB | 52 | 3 | **0.012** | 55 | 0 | **0.005** | 52 | 3 | **0.012** |

### *3.2. Identifying myelopathic levels with multiple indices*

For the classification of myelopathic levels, various DTI indices demonstrated a difference in sensitivity from 69.3%–88.7% (Fig. 4(c)), specificity from 83.6%–95.7% (Fig. 4(b)), accuracy from 71.5%–86.9% (Fig. 4(a)), and area under the ROC curve (Fig. 4(d)), while FA demonstrated the highest value in all performances. The error bars using standard deviations are also shown in Fig. 4 to indicate the variability of the classifications.

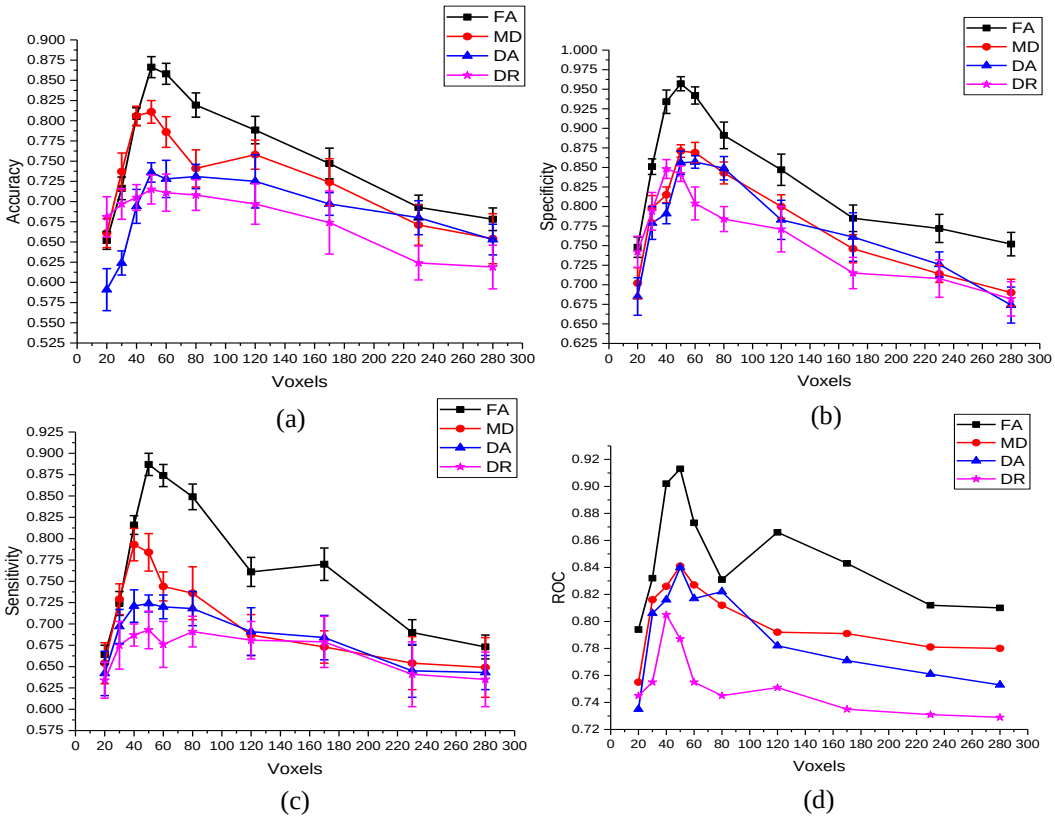


Figure 4. Accuracy, specificity, sensitivity and the area under the curve for a receiver operating characteristic curve (ROC AUC) for control and CSM classification. The values indicated are weighted averages for the two classes under consideration. Results are shown for ten datasets – 20 voxel dataset, 30 voxel dataset, 40 voxel dataset, 50 voxel dataset, 60 voxel dataset, 80 voxel dataset, 120 voxel dataset, 170 voxel dataset, 230 voxel dataset and 280 voxel dataset. The voxels comprising these reduced datasets were selected by the ReliefF algorithm

### *3.3. Regions most influential for detecting areas with CSM*

In the spinal cord, WM and gray matter are macroscopically organized in a "butterfly" or "H-shaped" pattern, with the interior gray matter surrounded by WM, as seen in an axial section. The WM consists of axonal bundles that ascend or descend the spinal cord. Anatomically, the WM tracts may be subdivided into dorsal, lateral, and ventral sections[29] (Fig. 5(a)). For each section, six reduced datasets were created with 20, 40, 60, 80, and, 100 voxels. FA was used to identify the regions that were most relevant for distinguishing between control groups and CSM groups. The classification was performed using the SVM. The highest sensitivity (73.6%), specificity (78.1%), and accuracy (72.9%) were achieved using the 60 voxel dataset from the dorsal section (Figs. 5(b), 5(c), and 5(d)). Fig. 5 shows that CSM may be most easily detected in the dorsal section and least easily detected in the lateral section, in agreement with Vedantam et al.[45].

## 4. Discussion

In this study, we used a machine learning-based method to identify spinal cord regions with CSM. Our results demonstrate that an automated procedure that combines DTI with machine learning algorithms accurately classifies control groups and groups with CSM.

Compared with our previous work[26], in the current study, we make three main contributions. First, we analyzed the contribution of each voxel of the ROI for classification and selected the valid voxel dataset that contained the most important features for detecting the CSM group, whereas in the previous work we only performed classification using the mean values, which led to key information loss. Experimental results demonstrate that classification using a selected voxel dataset performed better than that using the mean value, especially for classification using the SVM. The SVM classifier using the 60 voxel dataset produced an accuracy of 95.73%, sensitivity of 93.41%, and specificity of 98.64, whereas

classification using the mean value only achieved an accuracy of 81.52%, sensitivity of 54.33%, and specificity of 92.43%. The results indicate that a selected voxel dataset enables more valid features than the mean values of the ROI. Second, the classification efficiencies of the four diffusion indices (FA, MD, AD, and RD) of DTI were discussed. Experimental results demonstrate that FA performed best in detecting the CSM regions in the spinal cord with a sensitivity of 88.7%, accuracy of 86.9%, and specificity of 95.7%. The RD index produced the worst performance for detecting the CSM, with a sensitivity of 69.3%, accuracy of 71.5%, and specificity of 83.6% using the 60 voxel dataset. Third, in this study, we identified the areas of the ROI that were most useful for accurate classification. Performance was best when using the dorsal area, and a sensitivity of 73.6.2%, accuracy of 72.9%, and specificity of 78.1% were achieved. The ventral area occupied second place, whereas the lateral area was the worst performing. This finding is consistent with the study of Vedantam[45] et al.

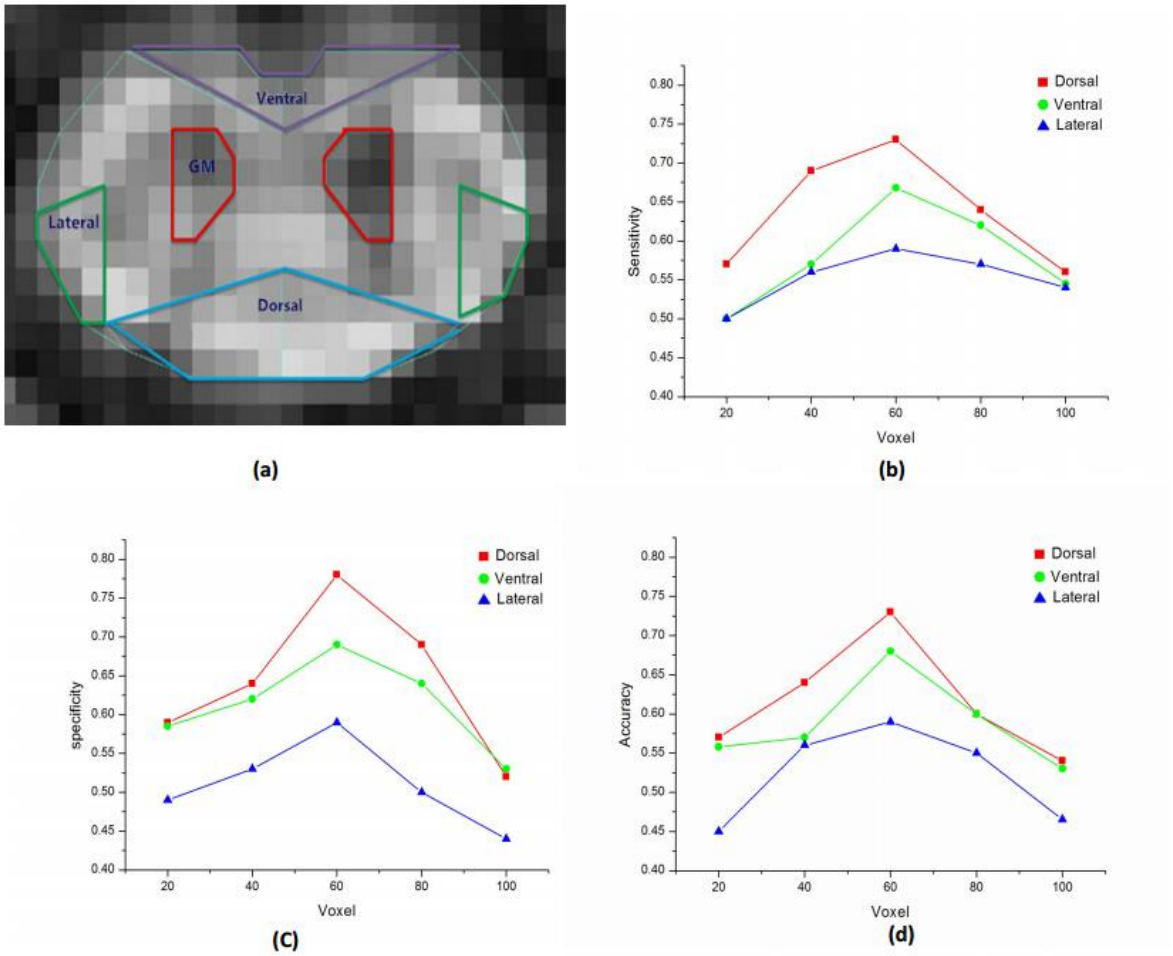


Figure 5. (a) The three sections in ROI: dorsal, lateral, and ventral section. (b) (c) (d) Sensitivity, specificity and accuracy derived from FA with SVM. Results are shown for 6 datasets – 20 voxel dataset, 40 voxel dataset, 60 voxel dataset, 80 voxel dataset and 100 voxel dataset. The voxels comprising these reduced datasets were selected by the ReliefF algorithm.

The pathophysiology of CSM was considered with static factors, dynamic factors, and ischemia[46]. DTI has been recommended as a precise measurement to detect the microstructure impairment of nerve bundles and diagnose a myelopathic cord with higher sensitivity and specificity than conventional anatomical MR images[24]. However, it is a technical challenge for clinicians and surgeons to read and interpret DTI in a clinical diagnosis of CSM, whereas the results of the current study would be a satisfactory solution using machine learning technology to provide a powerful tool for the classification of CSM.

## 5. Conclusion

In this work, we demonstrated that the use of DTI in conjunction with machine learning algorithms is a powerful strategy for CSM classification and is potentially useful in a clinic. The protocol outlined in the

current study is a blueprint for creating a machine learning-based classifier that successfully learns the structural differences between CSM and healthy regions. Our results suggest that this framework may provide novel and efficient assistance for the clinical diagnosis of myelopathic regions in the cervical spinal cord.

**Acknowledgment**

This work was supported in part by Natural Science Foundation of Guangdong Province (Grant No.2014A030310154 and No. 2016A030313176) and National Natural Science Foundations of China (Grants No. 61503368 and No. 61502473) and Special Funding for Science and Technology Development of Guangdong Province (Grant No. 2016A050503035) and a grant from the Research Grants Council of the Hong Kong (GRF 771608M). Authors thank for professional English editing from Liwen Editing service.

**REFERENCES**

1. H. Adeli and S. Ghosh-Dastidar, *Automated EEG-based Diagnosis of Neurological Disorders Inventing the Future of Neurology* (CRC Press, Taylor & Francis, Boca Raton, Florida, 2010).
2. L. O'Dwyer, F. Lamberto, A. L. W. Bokde and M. Ewers et al., Using Support Vector Machines with Multiple Indices of Diffusion for Automated Classification of Mild Cognitive Impairment, *PLoS ONE*. 7(2) (2012) e32441.
3. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer's Disease and Models of Computation: Imaging, Classification, and Neural Models, *J Alzheimers Dis*. 7(3) (2005) 187-199.
4. S. Ghosh-Dastidar and H. Adeli, Improved Spiking Neural Networks for EEG Classification and Epilepsy and Seizure Detection, *Integrated Computer-Aided Engineering.* 14(3) (2007) 187-212.
5. A. Ortiz, J. Munilla, J. M. Górriz, et al., Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer's Disease, *Int J Neural Syst.* 26(07) (2016)1650025.
6. H. Adeli, S. Ghosh-Dastidar, A New Supervised Learning Algorithm for Multiple Spiking Neural Networks with Application in Epilepsy and Seizure Detection, *Neural Networks*. 22(10) (2009) 1419-1431.
7. Z. Sankari and H. Adeli, Probabilistic Neural Networks for EEG-based Diagnosis of Alzheimer's Disease Using Conventional and Wavelet Coherence, *J Neurosci Methods.* 197(1) (2011) 165-170.
8. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A Spatio-temporal Wavelet-Chaos Methodology for EEG-based Diagnosis of Alzheimer's Disease, *Neurosci Lett.* 444(2) (2008) 190-194.
9. B. Magnin, L. Mesrob, S. Kinkingne´hun and M. Pe´le ´grini-Issac, Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI, *Neuroradiology.* 51(2) (2009) 73–83.
10. T. Hirschauer, H. Adeli and T. Buford, Computer-Aided Diagnosis of Parkinson's Disease using an Enhanced Probabilistic Neural Network, *J Med Syst*. 39(11) (2015) 179.
11. M. Ahmadlou and H. Adeli, Enhanced Probabilistic Neural Network with Local Decision Circles: A Robust Classifier, *Integrated Computer-Aided Engineering.* 17(3) (2010) 197.
12. P. Sharma, Y. U. Khan, O. Farooq, M. Tripathi and H. Adeli, A Wavelet-Statistical features Approach for Non-convulsive Seizure Detection, *Clin EEG Neurosci.* 45(4) (2014) 274-284.
13. C. Davatzikos, Y. Fan, X. Wu and D. Shen, Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging, *Neurobiol Aging.* 29(4) (2008) 514–523.
14. W. R. Brain, D. Northfield and M. Wilkinson, The Neurological manifestations of cervical spondylosis, *Brain*. 75(2) (1952) 187–225.
15. M. Grana, M. Termenon, A. Savio, A. Gonzalez-Pinto, Computer aided diagnosis system for alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation, *Neurosci Lett*. 502(3) (2011) 225–229.
16. K. Herholz, E. Salmon, D. Perani and J. C. Baron, et al., Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET, *Neuroimage*. 17(1) (2002) 302–316.
17. Y. Zhang, W. Zhou and S. Yuan, Multifractal Analysis and Relevance Vector Machine-based Automatic Seizure Detection in Intracranial, *Int J Neural Syst.* 25(6) (2015) 1550020.
18. V. Jumutc, P. Zayakin and A. Borisov, Ranking-based kernels in applied biomedical diagnostics using a support vector machine, *Int J Neural Syst.* 21(6) (2011) 459-73.
19. P. Piaggi, D. Menicucci, C. Gentili, G. Handjaras, A. Gemignani and A. Landi, Singular spectrum analysis and adaptive filtering enhance the functional connectivity analysis of resting state fMRI data, *Int J Neural Syst.* 24(3) (2014) 1450010.
20. S. Ferdowsi, S. Sanei and V. Abolghasemi, A predictive modeling approach to analyze data in EEG-fMRI experiments, *Int J Neural Syst.* 25(1) (2015) 1440008.
21. M. Thurnher and M. Law, Diffusion-weighted imaging, diffusion-tensor imaging, and fiber tractography of the spinal cord, *Magn Reson Imaging C.* 17(2) (2009) 225-244.
22. A. Demir, M. Ries, C. T. Moonen and J. M. Vital, Diffusion-weighted MR imaging with apparent diffusion coefficient and apparent diffusion tensor maps in cervical spondylotic myelopathy, *Radiology.* 229(1) (2003) 37-43.
23. J. L. Cui, C. Y. Wen, Y. Hu, T. H. Li and K. Luk, Entropy-based analysis for diffusion anisotropy mapping of healthy

and myelopathic spinal cord, *Neuroimage*. 54(3) (2011) 2125-2131.
24. D. Facon, A. Ozanne, P. Fillard and J. F. Lepeintre, MR diffusion tensor imaging and fiber tracking in spinal cord compression, *Am J Neuroradiol*. 26(6) (2005) 1587-1594.
25. T. Uda, T. Takami, N. Tsuyuguchi and S. Sakamoto, Assessment of cervical spondylotic myelopathy using diffusion tensor magnetic resonance imaging parameter at 3.0 tesla, *Spine*. 38(5) (2013) 407-414.
26. S. Q. Wang, X. Li, J. L. Cui and H. X. Li, et al., Prediction of Myelopathic Level in the Cervical Spinal Cord Using DTI, *J Magn Reson Imaging.* 41(6) (2015) 1682-1688.
27. F.C. Garcia, M. Garcia-Torres, B. Melian and J. A. Moreno-Perez, Solving feature subset selection problem by a Parallel Scatter Search, *EJOR.* 169(2) (2006) 477–489.
28. P. Hagmann, L. Jonasson, P. Maeder, J. P. Thiran, Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond, *Radiographics.* 26(1) (2006) 205-223.
29. D. Schwartz, J. Duda, J. S. Shumsky and E. T. Cooper, et al., Spinal cord diffusion tensor imaging and fibber tracking can identify white matter tract disruption and glial scar orientation following lateral funiculotomy, *J Neurotrauma.* 22(12) (2005) 1388-1398.
30. B. M. Ellingson, S. N. Kurpad and B. D. Schmit, Functional correlates of diffusion tensor imaging in spinal cord injury, *Biomed Sci Instrum.* 44 (2008) 28-33.
31. K. Yonenobu, K. Abumi, K. Nagata and E. Taketomi, Interobserver and intraobserver reliability of the japanese orthopaedic association scoring system for evaluation of cervical compression myelopathy, *Spine.* 26(17) (2001) 1890-1894.
32. T. C. Chuang, T. Y. Huang, F. H. Lin and F. N. Wang, PROPELLER-EPI with parallel imaging using a circularly symmetric phased-array RF coil at 3.0 T: application to high-resolution diffusion tensor imaging, *Magn reson med.* 56(6) (2006) 1352-1358.
33. P. Morgan, R. Bowtell, D. McIntyre and B. Worthington, Correction of spatial distortion in EPI due to inhomogeneous static magnetic fields using the reversed gradient method, *J Magn Reson Imaging*. 19(4) (2004) 499-507.
34. J. Soares, P. Marques, V. Alves and N. Sousa, A hitchhiker's guide to diffusion tensor imaging, *Front Neurosci.* 7 (2013) 31.
35. C. F. Westin, S. E. Maier, H. Mamata and A. Nabavi, Processing and visualization of diffusion tensor MRI, *Med Image Anal.* 6(2) (2002) 93-108.
36. C. Cortes and V. Vapnik, Support-vector networks, *Mach Learn.* 20(3) (1995) 273-297.
37. E. Castillo, D. Peteiro-Barral, B. G. Berdiñas and O. Fontenla-Romero, Distributed One-Class Support Vector Machine, *Int J Neural Syst.* 25(7) (2015) 1550029.
38. I. Kotsia, W. Guo and I. Patras, Higher rank support tensor machines for visual recognition, *Pattern Recog.* 45(12) (2012) 4192–4203.
39. D. J. Hand and K. Yu, Idiot's Bayes — not so stupid after all, *Int Stat Rev.* 69(3) (2001) 385–399.
40. G. Holmes, A. Donkin and I.H. Witten, Weka: A machine learning workbench, *in Proc. 2nd ANZIIS,* (Brisbane, Australia, 1994), pp. 357–361.
41. M. Robnik-Sikonja and I. Kononenko, Theoretical and Empirical Analysis of ReliefF and RReliefF, *Mach Learn.* 53(suppl_1) (2003) 23–69.
42. Z. Zhao and H. Liu, Spectral feature selection for supervised and unsupervised learning, *in Proc. 24th Int. Conf. Mach Learn,* (Corvalis, Oregon, 2007), pp. 1151–1157.
43. M. Stone, Asymptotics for and against cross-validation, *Biometrika.* 64(1) (1977) 29-35.
44. J.L. Cui, C.Y. Wen, Y. Hu, K.C. Mark and K.D Luk, Orientation entropy analysis of diffusion tensor in healthy and myelopathic spinal cord, Neuroimage. 258 (2011) 1028–1033.
45. A. Vedantam, M. B. Jirjis, B. D. Schmit and M. C. Wang et al., Diffusion tensor imaging of the spinal cord: insights from animal and human studies, *Neurosurgery.* 74(1) (2014) 1-8.
46. C. Wen, J. Cui, H. Liu and K. Mak, et al., Is Diffusion Anisotropy a Biomarker for Disease Severity and Surgical Prognosis of Cervical Spondylotic Myelopathy, *Radiology.* 270(1) (2013) 197-204.