



Identification of Amyotrophic Lateral Sclerosis Based on Diffusion Tensor Imaging and Support Vector Machine

Qiu-Feng Chen^{1†}, Xiao-Hong Zhang^{2†}, Nao-Xin Huang² and Hua-Jun Chen^{2*}

¹ College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou, China, ² Department of Radiology, Fujian Medical University Union Hospital, Fuzhou, China

Objectives: White matter (WM) impairments involving both motor and extra-motor areas have been well-documented in amyotrophic lateral sclerosis (ALS). This study tested the potential of diffusion measurements in WM for identifying ALS based on support vector machine (SVM).

Methods: Voxel-wise fractional anisotropy (FA) values of diffusion tensor images (DTI) were extracted from 22 ALS patients and 26 healthy controls and served as discrimination features. The revised ALS Functional Rating Scale (ALSFRS-R) was employed to assess ALS severity. Feature ranking and selection were based on Fisher scores. A linear kernel SVM algorithm was applied to build the classification model, from which the classification performance was evaluated. To promote classifier generalization ability, a leave-one-out cross-validation (LOOCV) method was adopted.

Results: By using the 2,400~3,400 ranked features as optimal features, the highest classification accuracy of 83.33% (sensitivity = 77.27% and specificity = 88.46%, $P = 0.0001$) was achieved, with an area under receiver operating characteristic curve of 0.862. The predicted function value was positively correlated with patient ALSFRS-R scores ($r = 0.493$, $P = 0.020$). In the optimized SVM model, FA values from several regions mostly contributed to classification, primarily involving the corticospinal tract pathway, postcentral gyrus, and frontal and parietal areas.

Conclusions: Our results suggest the feasibility of ALS diagnosis based on SVM analysis and diffusion measurements of WM. Additional investigations using a larger cohort is recommended in order to validate the results of this study.

Keywords: amyotrophic lateral sclerosis, diffusion tensor imaging, white matter, support vector machine, diagnosis

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a cryptogenetic and fatal neurodegenerative disorder that occurs in adults, involving the upper motor neurons as well as lower motor neurons. ALS is a heterogeneous disease and is generally difficult to diagnosis during the early stages. Most ALS patients die of respiratory failure. The median survival time of ALS is 3–5 years (1). Riluzole can only prolong survival time by 2–3 months (2). Therefore, early diagnosis of ALS is particularly important.

OPEN ACCESS

Edited by:

Hans-Peter Müller,
University of Ulm, Germany

Reviewed by:

Francesca Trojsi,
University of Campania Luigi
Vanvitelli, Italy
Alessia Sarica,
University of Magna Graecia, Italy

*Correspondence:

Hua-Jun Chen
chj0075@126.com

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Applied Neuroimaging,
a section of the journal
Frontiers in Neurology

Received: 15 January 2020

Accepted: 24 March 2020

Published: 28 April 2020

Citation:

Chen Q-F, Zhang X-H, Huang N-X and
Chen H-J (2020) Identification of
Amyotrophic Lateral Sclerosis Based
on Diffusion Tensor Imaging and
Support Vector Machine.
Front. Neurol. 11:275.
doi: 10.3389/fneur.2020.00275

Diffusion tensor imaging (DTI) is a non-invasive scanning procedure that provides exquisite details on white matter (WM) tissue microstructure (3) and thus plays a key role in investigating the pathology of neurological disorders. DTI parameters, such as fractional anisotropy (FA), axial diffusivity (AD), radial diffusivity (RD), and mean diffusivity (MD), provide information on the molecular diffusion in various directions. Several DTI studies have revealed that ALS patients have the decreased FA and increased diffusivity parameters in both motor and extra-motor areas (4–6). Of these changes, the alterations in DTI measurements in the corticospinal tract (CST) and corpus callosum are considered as the promising biomarker candidate for the diagnosis and evaluation of ALS (7). For instance, many studies have consistently demonstrated the reduction of CST FA value (6) and its high ability to differentiate ALS patients from healthy controls, at the group-level (8, 9). Also, several discrimination studies at the individual level have suggested that altered diffusion metrics in the CST and corpus callosum can be used as the sensitive marker to identify ALS (10, 11). Furthermore, DTI measurement may have an improved potential role in the correct discrimination of ALS when combined with other neuroimaging biomarkers (such as cortical thickness and functional measurement) (10, 12). In addition, adopting diffusion measuring and high-resolution volumetric/surface imaging simultaneously, researchers have demonstrated the feasibility of multi-model neuroimaging for predicting survival of ALS patients (13, 14).

Recently, the field of machine learning that holds promise to enable the computer-aided diagnosis of neuropsychological disorders has attracted widespread attentions. With the various model neuroimaging data (such as high-resolution T1-weighted image, magnetic resonance spectroscopy, and DTI), many machine learning approaches have been successfully implemented in the predictive modeling of ALS (15, 16). Of them, Random Forests method is a good case in point, which can incorporate multimodal imaging data and achieve quite promising result for individual identification of ALS (10, 12).

Support Vector Machine (SVM) is another algorithm that has been employed in assessing the discriminative brain map of patients diagnosed with ALS (15). In fact, there is increasing studies examining the application of SVM with neuroimaging data for clinical prediction of ALS. For example, a previous resting-state functional magnetic resonance imaging investigation has employed SVM to identify ALS based on the functional connectivity measurements in brain networks and achieved high accuracy for disease state classification (17). Recent studies have also demonstrated that ^{18}F -FDG PET (^{18}F -2-fluoro-2-deoxy-D-glucose Positron Emission Tomography) with SVM discriminant analysis can yield the promising result in differentiating ALS patients

Abbreviations: ALS, amyotrophic lateral sclerosis; ALSFRS-R, revised ALS Functional Rating Scale; DTI, diffusion tensor imaging; FA, fractional anisotropy; AD, axial diffusivity; RD, radial diffusivity; MD, mean diffusivity; SVM, support vector machine; ROC, receiver operating characteristic; LOOCV, leave-one-out cross-validation; CST, corticospinal tract; SLF, superior longitudinal fasciculus; ILF, inferior longitudinal fasciculus.

TABLE 1 | Demographic and clinical information of the study participants.

	Healthy controls (n = 26)	ALS patients (n = 22)	P-value
Age (years)	53.1 ± 6.4	55.4 ± 6.0	0.21*
Sex (male/female)	16/10	15/7	0.82#
Education (years)	8.2 ± 3.2	7.5 ± 3.3	0.63*
Site of onset (Bulbar/Cervical/ Thoracic/Lumbosacral)	—	1/14/1/6	—
Diagnostic category (Definite/Probable/Possible)	—	7/6/9	—
ALSFRS-R score	—	40.1 ± 7.2	—
Disease duration (months)	—	15.7 ± 13.2	—
Disease progression rate	—	0.66 ± 0.49	—

ALS, amyotrophic lateral sclerosis; ALSFRS-R, revised ALS Functional Rating Scale. Calculation of Disease Progression Rate was performed using the equation: $(48 - \text{ALSFRS-R}) / \text{Disease duration}$. “—” denotes no data available. P-values marked with “*” and “#” were calculated by the Student’s t-test and chi-square test, respectively.

from healthy controls. In this exploratory study, we were to make the attempt to test the potential of voxel-wise diffusion measurements in WM in identifying ALS, based on the SVM learning method.

MATERIALS AND METHODS

Subjects

A total of 22 ALS patients (1 familial, 21 sporadic) as well as 26 healthy controls (HC) were enrolled in this study. We employed the El Escorial criteria (18) in diagnosing ALS, whereas the revised ALS Functional Rating Scale (ALSFRS-R) was utilized to assess their severity of disease. The clinical and demographic information of the study participants are presented in **Table 1**. No significant differences between the patient and control groups in terms of age, sex, or educational level were observed (the detailed information see **Table 1**). The exclusion criteria were as follows: (1) existence of other neuropsychiatric disorders, including Parkinson’s disease, Alzheimer’s disease, epilepsy, or depression; (2) receiving psychotropic drugs; (3) occurrence of respiratory failure or other severe conditions such as angiocardopathy or cancer; or (4) contraindication of MRI examination. Approval for this evaluation was obtained from the Research Ethics Committee of Fujian Medical University Union Hospital, China. All of the subjects provided their written informed consent.

MRI Data Acquisition

A 3T MRI scanner (Prisma, Siemens Medical Systems, Erlangen, Germany) was utilized in image acquisition. DTI data were gathered with a spin-echo single-shot echo-planar imaging sequence using the following parameters: b -value = 1,000 s/mm² and 64 encoding diffusion directions; repetition time = 2,500 ms; echo time = 81 ms; number of averages = 1; slice thickness = 2 mm without gaps; field of view = 260 × 260 mm; matrix = 130 × 130; flip angle = 90°; 72 axial slices; and multiband factor = 4.

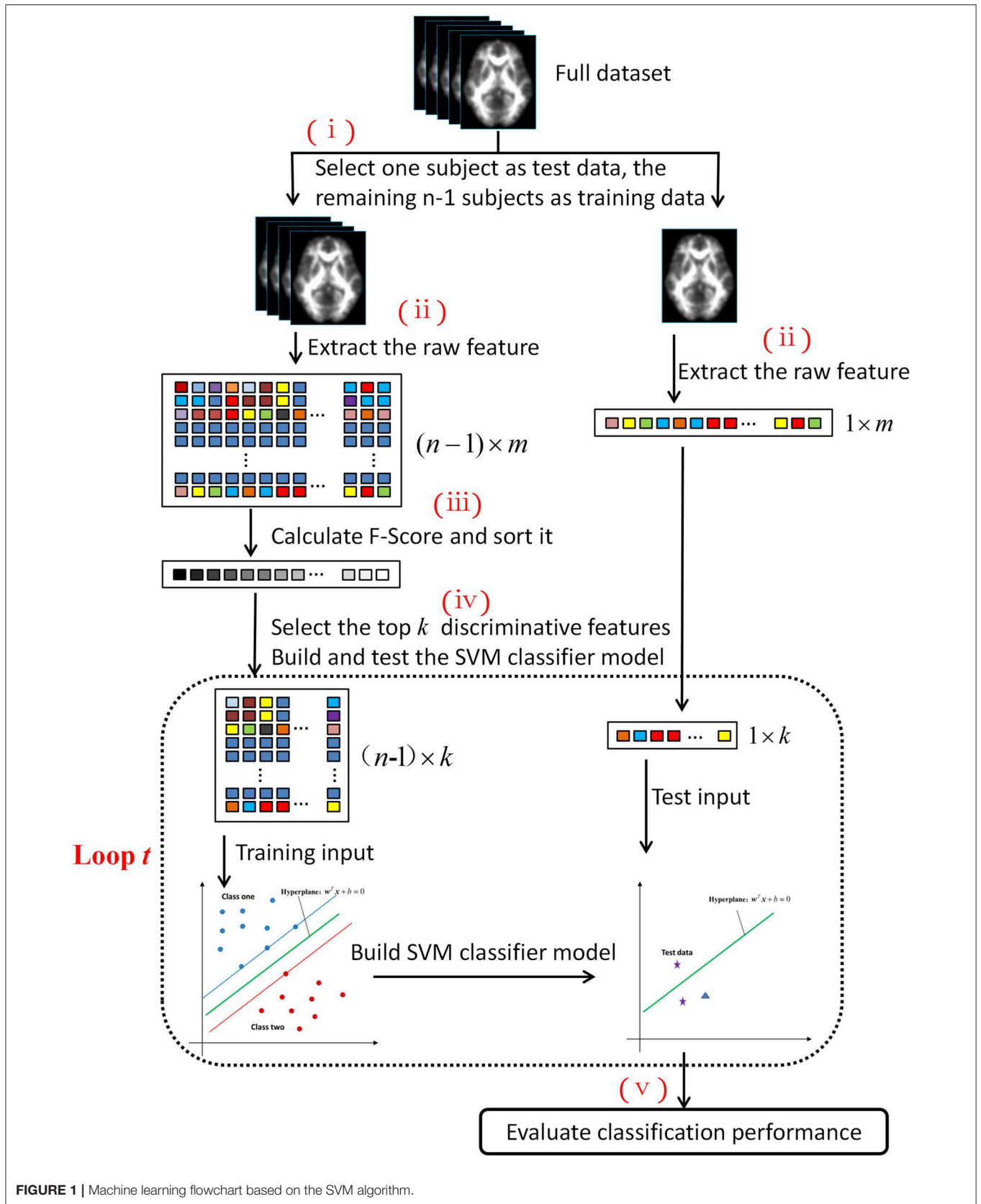


FIGURE 1 | Machine learning flowchart based on the SVM algorithm.

DTI Data Processing

DTI data were processed by an FSL-based pipeline (19). We corrected the raw DTI data for head movement and eddy-current distortion, and we then fitted a diffusion tensor model independently for each voxel. Thereafter, FA images of each subject were obtained. All of the FA images of the study participants were aligned to an FMRIB-58 FA template in Montreal Neurological Institute (MNI) space by a non-linear registration algorithm. Then, all of the images were smoothed with a 6 mm full width at half maximum (FWHM) Gaussian kernel.

Machine Learning Classification

Figure 1 shows the five steps of the machine learning process: (i) splitting the entire dataset into two parts: one subject served as the test set, whereas the remaining comprised the training set; (ii) extracting the features from the DTI images; (iii) calculating the Fisher scores of each feature and ranking the features, (iv) selecting the top k discriminative features to build the SVM classifier model and test it; and (v) evaluating the performance of the machine learning classification. The code was implemented in MATLAB (release 2016a, MathWorks, Natick, MA, USA), based on the LibSVM toolbox (20).

To assess the generalization performance of the developed SVM classifier, we adopted the leave-one-out cross-validation (LOOCV) method. If the number of subjects was n , then this cross-validation process is conducted n times (i.e., the flowchart in **Figure 1** is repeated n times). In each fold, one subject was selected as the test dataset, and the remaining $n-1$ subjects constituted the training dataset.

In step (ii), the feature vector was extracted from the FA image for each subject to form the raw feature matrixes. The dimensions of the feature matrixes were $(n-1) \times m$ and $1 \times m$ for the training and test datasets, respectively.

Due to the presence of the unrelated or redundant features, the learning model will tend to overfit, which in turn will degrade the classification performance (21). The Fisher score algorithm is a supervised feature selection method that can effectively measure data discrimination from two classes and assign the higher score to the feature, which has more discriminative ability (22).

We calculated the score for each feature independently based on the Fisher criterion. Then, we selected the top k features with the highest ranking scores. The Fisher score for the q th feature is defined as follows:

$$F(q) = \frac{(\bar{x}_q^{(1)} - \bar{x}_q)^2 + (\bar{x}_q^{(2)} - \bar{x}_q)^2}{\frac{1}{n_1-1} \sum_{p=1}^{n_1} (x_{p,q}^{(1)} - \bar{x}_q^{(1)})^2 + \frac{1}{n_2-1} \sum_{p=1}^{n_2} (x_{p,q}^{(2)} - \bar{x}_q^{(2)})^2}, \quad (1)$$

where \bar{x}_q , $\bar{x}_q^{(1)}$, and $\bar{x}_q^{(2)}$ are the mean values of the q th feature of the entire cohort and the HC and patient group, respectively; n_1 and n_2 are the number of the HC and patient subjects, respectively; $x_{p,q}^{(1)}$ denotes the q th feature of the p th subject in HC group; and $x_{p,q}^{(2)}$ represents the q th feature of the p th subject in the patient group. The numerator represents interclass variance, whereas the denominator signifies intraclass variance.

Apparently, the greater the Fisher score, the stronger the discriminative power of the feature.

The features were sorted in descending order according to the pre-calculated Fisher scores. Then, some of these features were selected as the input to build and test the SVM classifier model. **Figure 1** shows that the dotted box shows a loop within the machine learning process. For each loop, the ranked features were selected with an increase in length of 100 (i.e., for loop t , the number of the selected feature (denoted as k) was equal to $t \times 100$). The total number of cycles was 200; therefore, the length of the selected feature ranged from 100 to 20,000. Feature sorting and selection were built in a nested leave-one-out procedure to promote model generalization, by which we obtained the training input and test input (whose dimensions were $(n-1) \times k$ and $1 \times k$, respectively).

Based on these inputs, the SVM classifier model was trained and tested. The SVM is a classification algorithm that separates two classes by means of the maximal hyper-plane margin. Its main task is to establish a discriminant decision function f from the training input, so that for the test input \mathbf{x} , this decision function can predict the class label through $y = f(\mathbf{x})$. The decision function is defined in this form:

$$y = f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b, \quad (2)$$

where \mathbf{w} is the weight vector perpendicular to the decision hyper-plane; T represents the matrix transpose manipulation; b is the offset of the hyper-plane (bias parameter); and ϕ is the transformation function that converts the input vector \mathbf{x} into some other feature space where the SVM algorithm can provide a linear separation for the training input. By applying the kernel method and duality theorem, the predicted function value for the test input \mathbf{x} can be written as follows:

$$y = \sum_{j=1}^{n-1} \alpha_j y_j \phi(\mathbf{x}_j)^T \phi(\mathbf{x}) + b = \sum_{j=1}^{n-1} \alpha_j y_j K(\mathbf{x}_j, \mathbf{x}) + b, \quad (3)$$

where α_j is the Lagrange multiplier; and $K(\mathbf{x}_j, \mathbf{x})$ is the kernel function that computes the dot product $\phi(\mathbf{x}_j)^T \phi(\mathbf{x})$ and can be generalized to the non-linear case. When the feature dimensionality is high, the data tend to be linearly separable; therefore, we selected the linear kernel other than the commonly used RBF kernel (23). There was only one parameter C that controlled the misclassification penalty for the linear SVM model. This parameter was fixed to 1, as previously described (23, 24). When using the linear kernel, the weight vector would be $\mathbf{w} = \sum_{j=1}^{n-1} \alpha_j y_j \mathbf{x}_j$, and the decision function was $y = f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$, according to Equation (3). According to the SVM optimization theory, the weight vector \mathbf{w} represents the direction along which the feature differed most between two groups. Hence, it was used to produce the discrimination brain map. A positive value in the discrimination map indicated that the patient's FA value was relatively lower than the HC's, whereas a negative value indicated that the HCs had lower FA values. The magnitude of the absolute value of the weight vector \mathbf{w} represented the intensity

of between-group differences in FA (25). Given the length of the top ranking features as $k = t \times 100$, we could obtain the weight vector w (whose length = k) and the bias parameter b for each LOOCV fold. Once w and b were calculated, we could predict the function value y for test input x . If the function value y was >0 , then the test subject was assigned to the HC group (label 1); otherwise, the test subject was assigned to the patient group (label 2). The accuracy, sensitivity, and specificity of the classifier model were computed based on these class labels. It should be noted that the mean centering and normalization operations were pre-conducted on the selected features before establishing the SVM model.

As the weight vector w slightly varied from fold to fold, we calculated the average weight vector to generate the discriminative map. The quantitative analysis of the classifier performance was made by the receiver operating characteristic (ROC), and the permutation test was conducted to evaluate the statistical significance (denoted by P -value) of classification accuracy (26, 27). We permuted the class labels of the input subjects randomly and repeated the classification analysis procedure (=10,000 times). We counted the number when the classification accuracy in the permutation test was no less than the real classification accuracy of 83.33%; then, we divided this number by 10,000 to obtain the P -value.

In addition, the other DTI-derived parameters (i.e., RD, AD, and MD) were also calculated and analyzed by the above processes. Overall, the machine learning classification performance based on the FA feature was systematically better than those based on the other DTI parameters (Supplementary Table 1 and Supplementary Figures 1–3).

RESULTS

Figure 2 shows the detailed results on the accuracy, sensitivity, and specificity based on the distinct number of FA features. When the size of the feature set was small, the input data could not provide enough information to train a reliable model; and when the size was too large, there would be redundant or irrelevant information in the input data, which degraded the machine learning performance. The classification algorithm could attain the optimal performance only when we selected the appropriate dataset. The best classification accuracy was acquired when a specific amount (from 1,400 to 3,400) of features was selected. Thus, we selected the top 2,400 (middle value between 1,400 and 3,400) ranked features as the optimal features. Using these discriminative features, we determined that SVM classifier accuracy can be as high as 83.33% (sensitivity = 77.27% and specificity = 88.46%, $P = 0.0001$). Taking the generalization accuracy as the statistical variable, the estimated permutation distribution is shown in Figure 3, which shows that the probability (when the classification accuracy $> 83.33\%$) is very low ($P = 0.0001$), indicating that our results are highly reliable.

In addition, the predicted function value of each test subject was acquired using the SVM classifier (Figure 4). The circle and triangle points represented the HCs and patients, respectively.

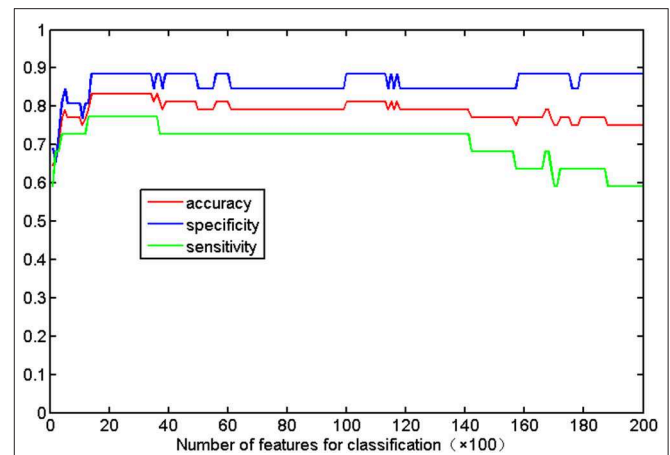


FIGURE 2 | Classification accuracy with respect to the distinct number of FA features. The range of features (from 1,400 to 3,400) can result in the highest accuracy (83.33%), with a sensitivity of 77.27% and specificity of 88.46%.

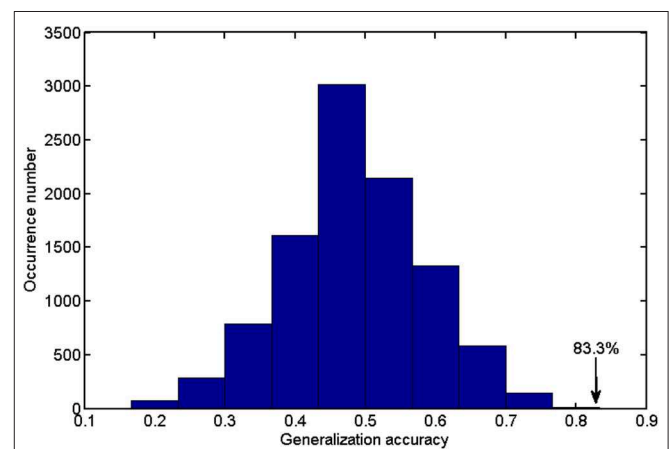


FIGURE 3 | The estimated permutation distribution using the linear SVM classifier (number of repetitions = 10,000), when the 2,400 most representative features were selected. The x- and y-axes denote the generalization accuracy and occurrence number, respectively. This figure demonstrated that the proposed method was unlikely to exceed the optimal accuracy of 83.33% that was obtained from the real class labels.

The circles at the right were the correctly labeled HC subjects, whereas the triangles at the left were correctly labeled patients. According to the distribution of these points, we found that most of the subjects (40/48) were assigned to the correct labels by the classifier.

By taking each subject's predicted function value as an indicator, we generated a receiver operating characteristic (ROC) curve for the classifier (Figure 5). The area under the ROC curve (AUC) of the developed method was 0.862, illustrating the relatively strong power for classification. The Spearman correlation coefficient was calculated to assess the correlation between the predicted function value and the ALSFRS-R score (Figure 6). We observed a positive correlation between these parameters ($r = 0.0397$, $P = 0.034$).

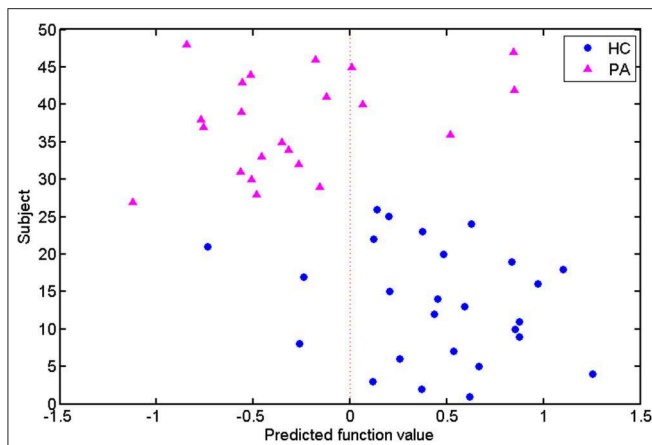


FIGURE 4 | Predicted function values of the test subjects. The healthy controls and patients are indicated by circle and triangle points, respectively. The points on the two sides of the dotted line are labeled as different classes.

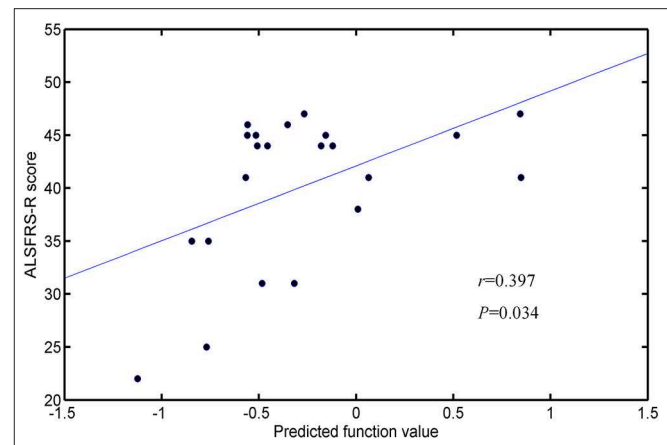


FIGURE 6 | Correlation between the predicted function values and ALSFRS-R scores.

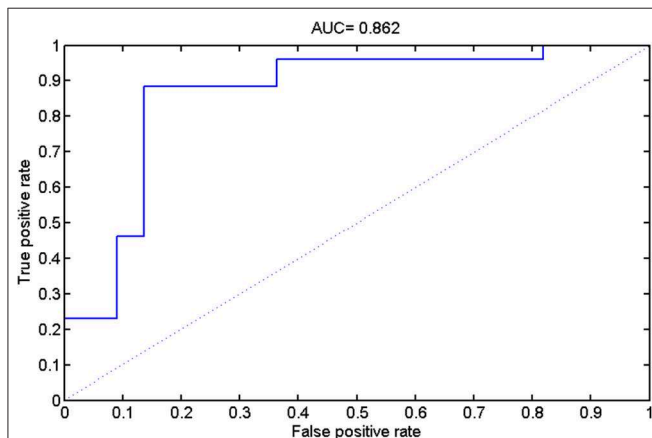


FIGURE 5 | The results of receiver operating characteristic (ROC) curve analysis.

The weight vector w was used to indicate a subset of features that were most discriminative between the two groups. Thus, the WM regions whose FA value contributed mostly to the discrimination were identified, using a threshold of $\geq 10\%$ of the maximum weight vector score. The details on these regions are presented in **Table 2** and we sorted them according to the contribution for between-group discrimination (i.e., the weight vector w). The relevant distribution of attribute weights generated by SVM analysis is presented in **Figure 7**. The WM regions with relatively increased FA in the HC group were located in several areas, such as bilateral corona radiata and precentral gyrus, right postcentral gyrus, right posterior limb of internal capsule, left superior frontal gyrus, left angular gyrus, left middle temporal gyrus, left middle occipital gyrus, bilateral midbrain, and bilateral pons and medulla, whereas the WM regions with relatively increased FA in the patient group were located in the left frontal lobe, left inferior parietal lobule, and right superior parietal lobule.

DISCUSSION

In the present study, we combined DTI with SVM to classify ALS patients and HCs. The high classification accuracy of 83.33% can be obtained in the optimized SVM model. The permutation statistics further validated the reliability of our SVM classifier. The FA values from both motor and extra-motor areas contributed to the classification, which may indicate that ALS is a multi-system neurodegenerative disease. Further ROC analysis also indicated the high potential of FA measurement in the accurate discrimination of ALS. Moreover, the predicted function value of classifier was correlated with ALS disease severity. These results suggested the promising perspective related to the application of SVM approach to ALS identification, based on the DTI measurement in WM.

The impairments of WM integrity (as reflected by reduced FA) have been well-documented in ALS, which may be due to the destruction of the axons and myelin (6, 28, 29). An early neuropathological feature of ALS is defective axonal transport, which may contribute to distal axon energy deficiency and dying-back axonopathy (30, 31). Oligodendrocytes myelinate the central nervous system (CNS) axons and support the function and survival of axons (32, 33). The pathological abnormalities in oligodendrocytes (e.g., oligodendrocyte death and impaired maturation of new oligodendrocytes) have been reported in ALS (34, 35), which could contribute to axonal demyelination (36).

Consistent with previous studies (37, 38), we found that the WM regions with decreased FA in ALS involved the bilateral precentral gyrus and the CST pathway, such as bilateral corona radiata, right posterior of internal capsule, bilateral midbrain, and bilateral pons, and medulla. The precentral gyrus is part of the primary motor cortex (PMC). The degenerative alterations of the PMC such as significantly decreased Betz cells and cortical thinning (39, 40) have been reported in ALS. The CST, which connects the neurons in the motor cortex and spinal cord, is the fibers that are associated with cortical control of spinal cord activity (41). Degeneration of the CST is also a hallmark of ALS

TABLE 2 | White matter regions in which the FA feature largely contributed to the classification.

Cluster size (number of voxels)	White matter region	MNI coordinates			Peak <i>W_i</i>
		x	y	z	
HC group > ALS group					
31	Left middle occipital gyrus	-24	-84	10	0.158
23	Left angular gyrus	-34	-62	38	0.146
39	Left precentral gyrus	-30	-14	56	0.132
297	Right corona radiate/precentral and postcentral gyrus	22	-20	40	0.120
27	Left superior frontal gyrus	-18	0	54	0.106
125	Right midbrain and posterior limb of internal capsule	12	-12	-12	0.105
48	Left midbrain	-8	-26	-18	0.099
58	Left corona radiate	-22	-22	42	0.095
46	Bilateral pons and medulla	-6	-34	-40	0.089
20	Left middle temporal gyrus	-54	-34	-12	0.018
HC group < ALS group					
27	Right superior parietal lobule	26	-42	52	-0.335
21	Left inferior parietal lobule	-28	-64	28	-0.241
51	Left frontal lobe	-30	-2	32	-0.137

The above brain regions were identified by using a threshold of $\geq 10\%$ of the maximum weight vector score. The first column only lists clusters larger than 20 voxels. *W_i* (reported in the last column) is the weight of each cluster centroid, i.e., the value that indicates the relative contribution of the FA feature to the SVM-based classification.

(42). In sum, damage to these motor-related regions could lead to motor neuron dysfunction and is responsible for the relevant symptoms (e.g., muscle weakness and loss of voluntary control) observed in ALS patients (43).

The regions with decreased FA also included several extra-motor areas such as the right postcentral gyrus, left superior longitudinal fasciculus (SLF) that involves the left superior frontal gyrus, left angular gyrus, and left middle temporal gyrus, and left inferior longitudinal fasciculus (ILF) that involves the left middle temporal gyrus and left middle occipital gyrus, which agrees with the findings of previous studies (44–46). For example, it has been demonstrated that the significant cortical thinning of the postcentral gyrus, namely, primary somatosensory cortex (47), occurring in ALS is correlated with disease severity (48, 49). In addition, the SLF, which connects the frontal, parietal, and temporal lobes and plays a key role in language function (50, 51), is disrupted in ALS (52, 53). Meanwhile, damage to the left ILF, the fiber that is primary associated with visual processing, language/semantic function, and regulation of emotion (54, 55), has been reported in ALS patients (45, 56). Therefore, damage to these extra-motor regions, as reflected by decreased FA, may be associated with the non-motion dysfunctions that have been reported in ALS, such as sensory deficits, language dysfunction, and behavioral and psychiatric abnormalities (57–59).

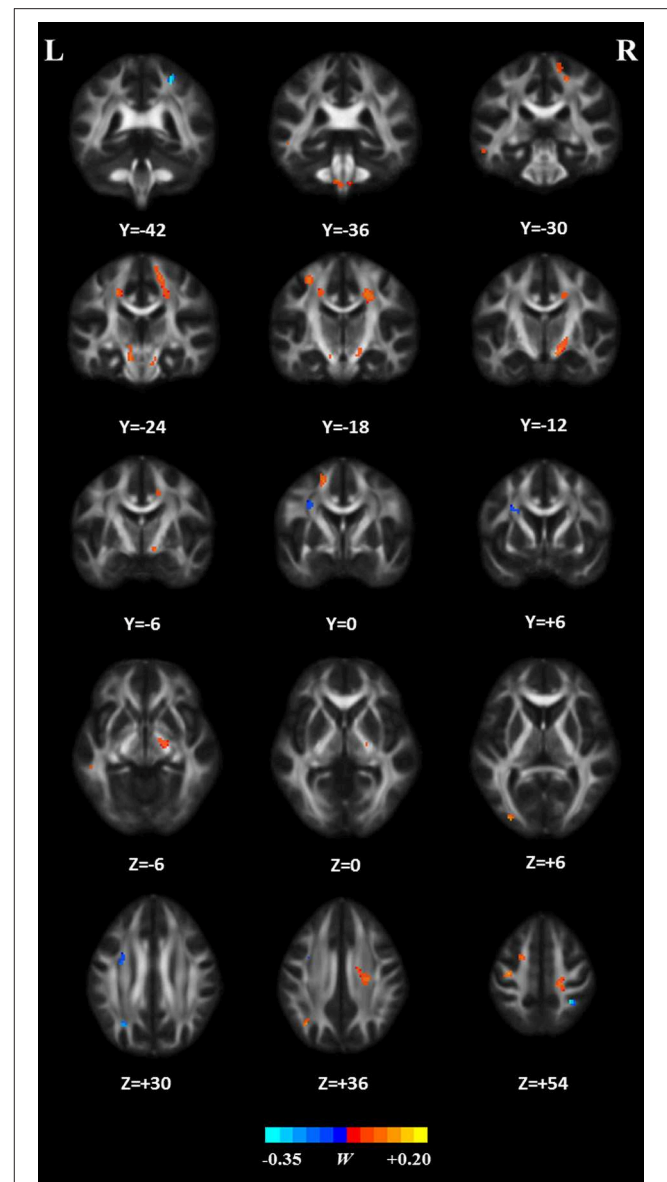


FIGURE 7 | White matter regions whose FA features mostly contribute to classification. These regions were identified by using a threshold of $\geq 10\%$ of the maximum weight vector score and their cluster size were larger than 20 voxels.

In contrast, our results showed several brain areas, including the left frontal lobe, left inferior parietal lobule, and right superior parietal lobule, with relatively increased FA in ALS patients. It is speculated that these changes may be related to the functional compensation or the reorganization of the cerebral structure in ALS (60, 61). In agreement with this speculation, several compensatory phenomena, as reflected by the increased gray matter volume or the overactivation of specific brain regions (62, 63), have been demonstrated in ALS.

The accurate recognition and characterization of ALS remain challenging due to its low incidence [$\approx 5/100,000$ (64)] and

heterogeneous nature (1, 15). Our results suggest that DTI measurement in the white matter could be utilized as an alternative biomarker of ALS based on machine learning method. Consistently, previous studies have also demonstrated the potential of diffusion measurement in identifying ALS, at the group and individual levels (6, 15, 16, 65, 66). The ROC analysis and permutation statistics further verified the reliability of our classification results. Moreover, our results suggested that FA was the most promising biomarker for ALS identification, relative to other DTI metrics, which keeps in line with previous evidences highlighting FA change is the consistent hallmark of ALS (6). In addition, the correlation between the SVM predicted function value and the ALSFRS-R score was observed. As the predicted function value was computed by projecting the optimal features onto the weight vector of the hyper-plane, a larger absolute value meant that the subject was situated farther away from the hyper-plane and more significantly contributed to the classification. Thus, we could deduce that when the ALSFRS-R score was higher (i.e., higher disease severity), the subject is less likely to be misclassified.

From a data-driven methodological perspective, this study employed SVM method to establish the predicting model of ALS, given that SVM can accommodate all of voxel-wise DTI measures simultaneously and can model their interactions in high dimensional feature space to optimize between-group classification. As a supervised learning model, the advantage of SVM relies on its regularization parameter that is helpful in preventing model overfitting and SVM also have good performance and generalization capability when processing small-sample data (67, 68), such that SVM formalism was preferred for this exploratory work with the limited sample size. Different from our study, recent machine learning researches using DTI metrics of the pre-defined region of interest (e.g., CST) have applied Random Forests approach to build discrimination model between ALS and healthy control (10–12). Random Forests holds the advantages over other methods in the aspects of ability to handle highly non-linear biological data, robustness to noise, and tuning simplicity (12, 69). Taken together, it is noted that SVM and other methods (such as Random Forests) can be used in the implementation of prediction modeling of ALS and show the distinct methodological advantages, thereby, a systemic comparison of the diagnostic performances of various machine learning methods is recommended in the future.

This study has a number of limitations. First, unlike other previous studies on ALS, we did not consider disease heterogeneity (70, 71) and did not perform any investigation to examine the potential of DTI measurements in distinguishing ALS subtypes. Second, despite the prevalence of the SVM algorithm for medical data analysis, other machine learning algorithms could also be explored to promote the performance of the classification model and seek more reliable biomarkers for ALS patients. Third, other neuroimaging features (e.g., functional connectivity) can also contribute to ALS identification (17), so the combination of DTI and other modalities of MRI data should be considered in future classification studies. Fourth, further

investigations using a larger cohort are recommended to further validate the findings of our study.

Our results suggest the feasibility of ALS diagnosis based on SVM analysis and diffusion measurements of WM. The WM regions whose FA values mostly contributed to SVM classification involved motor as well as extra-motor areas, thereby supporting the notion that ALS is a disease involving multi-system neurodegeneration. In addition to the existing studies, our findings further provided the confirmatory evidences that the application of machine learning method such as SVM with neuroimaging data holds the promising perspective for the prediction of ALS. However, in consideration of the small sample size and other limitations, this study represented an exploratory work in nature. The further investigation using a larger cohort is warranted to validate the generality of our results and the future studies are recommended to verify the added diagnostic value of the inclusion of other modality neuroimaging data.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Files**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Approval for this evaluation was obtained from the Research Ethics Committee of Fujian Medical University Union Hospital, China. All of the subjects provided their written informed consent. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

H-JC, Q-FC, and X-HZ conceived and designed the study, acquired and analyzed the data, and wrote the manuscript. H-JC, Q-FC, and N-XH contributed to data analysis. All authors have read and approved the manuscript.

FUNDING

This study was funded by the grants from the National Natural Science Foundation of China (No. 81501450), Fujian Provincial Science Fund for Distinguished Young Scholars (No. 2018J06023), Fujian Provincial Program for Distinguished Young Scholars (No. 2017B023), and Fujian Provincial Health Commission Project for Scientific Research Talents (2018-ZQN-28).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fneur.2020.00275/full#supplementary-material>

REFERENCES

- Hardiman O, van den Berg LH, Kiernan MC. Clinical diagnosis and management of amyotrophic lateral sclerosis. *Nat Rev Neurol*. (2011) 7:639–49. doi: 10.1038/nrneuro.2011.153
- Miller RG, Mitchell JD, Moore DH. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst Rev*. (2012) CD001447. doi: 10.1002/14651858.CD001447.pub3
- Le Bihan D, Mangin JF, Poupon C, Clark CA, Pappata S, Molko N, et al. Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging*. (2001) 13:534–46. doi: 10.1002/jmri.1076
- Sage CA, Peeters RR, Görner A, Robberecht W, Sunaert S. Quantitative diffusion tensor imaging in amyotrophic lateral sclerosis. *NeuroImage*. (2007) 34:486–99. doi: 10.1016/j.neuroimage.2006.09.025
- Li J, Pan P, Song W, Huang R, Chen K, Shang H. A meta-analysis of diffusion tensor imaging studies in amyotrophic lateral sclerosis. *Neurobiol Aging*. (2012) 33:1833–8. doi: 10.1016/j.neurobiolaging.2011.04.007
- Zhang F, Chen G, He M, Dai J, Shang H, Gong Q, et al. Altered white matter microarchitecture in amyotrophic lateral sclerosis: a voxel-based meta-analysis of diffusion tensor imaging. *Neuroimage Clin*. (2018) 19:122–9. doi: 10.1016/j.nicl.2018.04.005
- Mazon M, Vazquez Costa JF, Ten-Esteve A, Marti-Bonmati L. Imaging biomarkers for the diagnosis and prognosis of neurodegenerative Diseases. The example of amyotrophic lateral sclerosis. *Front Neurosci*. (2018) 12:784. doi: 10.3389/fnins.2018.00784
- Graham JM, Papadakis N, Evans J, Widjaja E, Romanowski CA, Paley MN, et al. Diffusion tensor imaging for the assessment of upper motor neuron integrity in ALS. *Neurology*. (2004) 63:2111–9. doi: 10.1212/01.WNL.0000145766.03057.E7
- Agosta F, Galantucci S, Riva N, Chio A, Messina S, Iannaccone S, et al. Intrahemispheric and interhemispheric structural network abnormalities in PLS and ALS. *Hum Brain Mapp*. (2014) 35:1710–22. doi: 10.1002/hbm.22286
- Ferraro PM, Agosta F, Riva N, Copetti M, Spinelli EG, Falzone Y, et al. Multimodal structural MRI in the diagnosis of motor neuron diseases. *Neuroimage Clin*. (2017) 16:240–7. doi: 10.1016/j.nicl.2017.08.002
- Sarica A, Cerasa A, Valentino P, Yeatman J, Trotta M, Barone S, et al. The corticospinal tract profile in amyotrophic lateral sclerosis. *Hum Brain Mapp*. (2017) 38:727–39. doi: 10.1002/hbm.23412
- Fratello M, Caiazzo G, Trojsi F, Russo A, Tedeschi G, Tagliaferri R, et al. Multi-view ensemble classification of brain connectivity images for neurodegeneration type discrimination. *Neuroinformatics*. (2017) 15:199–213. doi: 10.1007/s12021-017-9324-2
- van der Burgh HK, Schmidt R, Westeneng HJ, de Reus MA, van den Berg LH, van den Heuvel MP. Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis. *Neuroimage Clin*. (2017) 13:361–9. doi: 10.1016/j.nicl.2016.10.008
- Agosta F, Spinelli EG, Riva N, Fontana A, Basaia S, Canu E, et al. Survival prediction models in motor neuron disease. *Eur J Neurol*. (2019) 26:1143–52. doi: 10.1111/ene.13957
- Grollemund V, Pradat PF, Querin G, Delbot F, Le Chat G, Pradat-Peyre JF, et al. Machine learning in amyotrophic lateral sclerosis: achievements, pitfalls, and future directions. *Front Neurosci*. (2019) 13:135. doi: 10.3389/fnins.2019.00135
- Agosta F, Spinelli EG, Filippi M. Neuroimaging in amyotrophic lateral sclerosis: current and emerging uses. *Expert Rev Neurother*. (2018) 18:395–406. doi: 10.1080/14737175.2018.1463160
- Welsh RC, Jelsone-Swain LM, Foerster BR. The utility of independent component analysis and machine learning in the identification of the amyotrophic lateral sclerosis diseased brain. *Front Hum Neurosci*. (2013) 7:251. doi: 10.3389/fnhum.2013.00251
- Brooks BR, Miller RG, Swash M, Munsat TL, World Federation of Neurology Research Group on Motor Neuron Diseases. El escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord*. (2000) 1:293–9. doi: 10.1080/146608200300079536
- Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. *Fsl. Neuroimage*. (2012) 62:782–90. doi: 10.1016/j.neuroimage.2011.09.015
- Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol*. (2011) 2:27. doi: 10.1145/1961189.1961199
- Cui Y, Wen W, Lipnicki DM, Beg MF, Jin JS, Luo S, et al. Automated detection of amnesic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. *Neuroimage*. (2012) 59:1209–17. doi: 10.1016/j.neuroimage.2011.08.013
- Liu F, Guo W, Fouche J-P, Wang Y, Wang W, Ding J, et al. Multivariate classification of social anxiety disorder using whole brain functional connectivity. *Brain Struct Funct*. (2015) 220:101–15. doi: 10.1007/s00429-013-0641-4
- Li F, Huang X, Tang W, Yang Y, Li B, Kemp GJ, et al. Multivariate pattern analysis of DTI reveals differential white matter in individuals with obsessive-compulsive disorder. *Hum Brain Mapp*. (2014) 35:2643–51. doi: 10.1002/hbm.22357
- Mourao-Miranda J, Friston KJ, Brammer M. Dynamic discrimination analysis: a spatial-temporal SVM. *Neuroimage*. (2007) 36:88–99. doi: 10.1016/j.neuroimage.2007.02.020
- Ecker C, Rocha-Rego V, Johnston P, Mourao-Miranda J, Marquand A, Daly EM, et al. Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. *Neuroimage*. (2010) 49:44–56. doi: 10.1016/j.neuroimage.2009.08.024
- Golland P, Fischl B. Permutation tests for classification: towards statistical significance in image-based studies. In: *Biennial International Conference on Information Processing in Medical Imaging*. Ambleside: Springer (2003). p. 330–41. doi: 10.1007/978-3-540-45087-0_28
- Etzel JA, Braver TS. MVPA permutation schemes: permutation testing in the land of cross-validation. In: *2013 International Workshop on Pattern Recognition in Neuroimaging*. IEEE (2013). p. 140–3. doi: 10.1109/PRNI.2013.44
- Cirillo M, Esposito F, Tedeschi G, Caiazzo G, Sagnelli A, Piccirillo G, et al. Widespread microstructural white matter involvement in amyotrophic lateral sclerosis: a whole-brain DTI study. *AJNR Am J Neuroradiol*. (2012) 33:1102–8. doi: 10.3174/ajnr.A2918
- Prudlo J, Bissbort C, Glass A, Grossmann A, Hauenstein K, Benecke R, et al. White matter pathology in ALS and lower motor neuron ALS variants: a diffusion tensor imaging study using tract-based spatial statistics. *J Neurol*. (2012) 259:1848–59. doi: 10.1007/s00415-012-6420-y
- Ferraiuolo L, Kirby J, Grierson AJ, Sendtner M, Shaw PJ. Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis. *Nat Rev Neurol*. (2011) 7:616–30. doi: 10.1038/nrneuro.2011.152
- Moller A, Bauer CS, Cohen RN, Webster CP, De Vos KJ. Amyotrophic lateral sclerosis-associated mutant SOD1 inhibits anterograde axonal transport of mitochondria by reducing Miro1 levels. *Hum Mol Genet*. (2017) 26:4668–79. doi: 10.1093/hmg/ddx348
- Philips T, Bento-Abreu A, Nonneman A, Haack W, Staats K, Geelen V, et al. Oligodendrocyte dysfunction in the pathogenesis of amyotrophic lateral sclerosis. *Brain*. (2013) 136:471–82. doi: 10.1093/brain/aws339
- Nave KA. Myelination and support of axonal integrity by glia. *Nature*. (2010) 468:244–52. doi: 10.1038/nature09614
- Zhou T, Ahmad TK, Gozda K, Truong J, Kong J, Namaka M. Implications of white matter damage in amyotrophic lateral sclerosis (Review). *Mol Med Rep*. (2017) 16:4379–92. doi: 10.3892/mmr.2017.7186
- Nonneman A, Robberecht W, Van Den Bosch L. The role of oligodendroglial dysfunction in amyotrophic lateral sclerosis. *Neurodegener Dis Manag*. (2014) 4:223–39. doi: 10.2217/nmt.14.21
- Benarroch EE. Oligodendrocytes: Susceptibility to injury and involvement in neurologic disease. *Neurology*. (2009) 72:1779–85. doi: 10.1212/WNL.0b013e3181a6b123
- Prell T, Peschel T, Hartung V, Kaufmann J, Klauschies R, Bodammer N, et al. Diffusion tensor imaging patterns differ in bulbar and limb onset amyotrophic lateral sclerosis. *Clin Neurol Neurosurg*. (2013) 115:1281–7. doi: 10.1016/j.clineuro.2012.11.031
- Zhang J, Yin X, Zhao L, Evans AC, Song L, Xie B, et al. Regional alterations in cortical thickness and white matter integrity in amyotrophic lateral sclerosis. *J Neurol*. (2014) 261:412–21. doi: 10.1007/s00415-013-7215-5
- Mochizuki Y, Mizutani T, Shimizu T, Kawata A. Proportional neuronal loss between the primary motor and sensory cortex in amyotrophic lateral sclerosis. *Neurosci Lett*. (2011) 503:73–5. doi: 10.1016/j.neulet.2011.08.014

40. Verstraete E, Veldink JH, Hendrikse J, Schelhaas HJ, van den Heuvel MP, van den Berg LH. Structural MRI reveals cortical thinning in amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry*. (2012) 83:383–8. doi: 10.1136/jnnp-2011-300909
41. Welniarz Q, Dusart I, Roze E. The corticospinal tract: Evolution, development, and human disorders. *Dev Neurobiol*. (2017) 77:810–29. doi: 10.1002/dneu.22455
42. Schuster C, Elamin M, Hardiman O, Bede P. The segmental diffusivity profile of amyotrophic lateral sclerosis associated white matter degeneration. *Eur J Neurol*. (2016) 23:1361–71. doi: 10.1111/ene.13038
43. Esquenazi A, Mayer NH, Elia AE, Albanese A. Botulinum toxin for the management of adult patients with upper motor neuron syndrome. *Toxicol*. (2009) 54:634–8. doi: 10.1016/j.toxicol.2009.01.022
44. Trojsi F, Corbo B, Caiazzo G, Piccirillo G, Monsurro MR, Cirillo S, et al. Motor and extramotor neurodegeneration in amyotrophic lateral sclerosis: a 3T high angular resolution diffusion imaging (HARDI) study. *Amyotroph Lateral Scler Frontotemporal Degener*. (2013) 14:553–61. doi: 10.3109/21678421.2013.785569
45. Crespi C, Cerami C, Dodich A, Canessa N, Arpone M, Iannaccone S, et al. Microstructural white matter correlates of emotion recognition impairment in amyotrophic lateral sclerosis. *Cortex*. (2014) 53:1–8. doi: 10.1016/j.cortex.2014.01.002
46. Trojsi F, Caiazzo G, Corbo D, Piccirillo G, Cristillo V, Femiano C, et al. Microstructural changes across different clinical milestones of disease in amyotrophic lateral sclerosis. *PLoS ONE*. (2015) 10:e0119045. doi: 10.1371/journal.pone.0119045
47. Kropf E, Syan SK, Minuzzi L, Frey BN. From anatomy to function: the role of the somatosensory cortex in emotional regulation. *Braz J Psychiatry*. (2019) 41:261–9. doi: 10.1590/1516-4446-2018-0183
48. Cosottini M, Pesaresi I, Piazza S, Diciotti S, Cecchi P, Fabbri S, et al. Structural and functional evaluation of cortical motor areas in Amyotrophic Lateral Sclerosis. *Exp Neurol*. (2012) 234:169–80. doi: 10.1016/j.expneurol.2011.12.024
49. Thorns J, Jansma H, Peschel T, Grosskreutz J, Mohammadi B, Dengler R, et al. Extent of cortical involvement in amyotrophic lateral sclerosis—an analysis based on cortical thickness. *BMC Neurol*. (2013) 13:148. doi: 10.1186/1471-2377-13-148
50. Jang SH, Hong JH. The anatomical characteristics of superior longitudinal fasciculus I in human brain: diffusion tensor tractography study. *Neurosci Lett*. (2012) 506:146–8. doi: 10.1016/j.neulet.2011.10.069
51. Conner AK, Briggs RG, Rahimi M, Sali G, Baker CM, Burks JD, et al. A connectomic atlas of the human cerebrum—chapter 10: tractographic description of the superior longitudinal fasciculus. *Oper Neurosurg*. (2018) 15:S407–22. doi: 10.1093/ons/opy264
52. Du XQ, Zou TX, Huang NX, Zou ZY, Xue YJ, Chen HJ. Brain white matter abnormalities and correlation with severity in amyotrophic lateral sclerosis: An atlas-based diffusion tensor imaging study. *J Neurol Sci*. (2019) 405:116438. doi: 10.1016/j.jns.2019.116438
53. Ash S, Olm C, McMillan CT, Boller A, Irwin DJ, McCluskey L, et al. Deficits in sentence expression in amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener*. (2015) 16:31–9. doi: 10.3109/21678421.2014.974617
54. Herbet G, Zemmoura I, Duffau H. Functional anatomy of the inferior longitudinal fasciculus: from historical reports to current hypotheses. *Front Neuroanat*. (2018) 12:77. doi: 10.3389/fnana.2018.00077
55. Shin J, Rowley J, Chowdhury R, Jolicœur P, Klein D, Grova C, et al. Inferior longitudinal fasciculus' role in visual processing and language comprehension: a combined MEG-DTI study. *Front Neurosci*. (2019) 13:875. doi: 10.3389/fnins.2019.00875
56. Sedda A. Disorders of emotional processing in amyotrophic lateral sclerosis. *Curr Opin Neurol*. (2014) 27:659–65. doi: 10.1097/WCO.0000000000000147
57. Iglesias C, Sangari S, El Mendili MM, Benali H, Marchand-Pauvert V, Pradat PF. Electrophysiological and spinal imaging evidences for sensory dysfunction in amyotrophic lateral sclerosis. *BMJ Open*. (2015) 5:e007659. doi: 10.1136/bmjopen-2015-007659
58. Vucic S. Sensory and autonomic nervous system dysfunction in amyotrophic lateral sclerosis. *Neuropathol Appl Neurobiol*. (2017) 43:99–101. doi: 10.1111/nan.12336
59. Strong MJ, Abrahams S, Goldstein LH, Woolley S, McLaughlin P, Snowden J, et al. Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. *Amyotroph Lateral Scler Frontotemporal Degener*. (2017) 18:153–74. doi: 10.1080/21678421.2016.1267768
60. Schoenfeld MA, Tempelmann C, Gaul C, Kühnel GR, Düzel E, Hopf JM, et al. Functional motor compensation in amyotrophic lateral sclerosis. *J Neurol*. (2005) 252:944–52. doi: 10.1007/s00415-005-0787-y
61. Konrad C, Jansen A, Henningsen H, Sommer J, TurSKI PA, Brooks BR, et al. Subcortical reorganization in amyotrophic lateral sclerosis. *Exp Brain Res*. (2006) 172:361–9. doi: 10.1007/s00221-006-0352-7
62. Poujois A, Schneider FC, Faillenot I, Camdessanche JP, Vandenberghe N, Thomas-Anterion C, et al. Brain plasticity in the motor network is correlated with disease progression in amyotrophic lateral sclerosis. *Hum Brain Mapp*. (2013) 34:2391–401. doi: 10.1002/hbm.22070
63. Qiu T, Zhang Y, Tang X, Liu X, Wang Y, Zhou C, et al. Precentral degeneration and cerebellar compensation in amyotrophic lateral sclerosis: a multimodal MRI analysis. *Hum Brain Mapp*. (2019) 40:3464–74. doi: 10.1002/hbm.24609
64. Marin B, Boumédiène F, Logrosino G, Couratier P, Babron MC, Leutenegger AL, et al. Variation in worldwide incidence of amyotrophic lateral sclerosis: a meta-analysis. *Int J Epidemiol*. (2016) 46:dyw061. doi: 10.1093/ije/dyw061
65. Liu C, Jiang R, Yi X, Zhu W, Bu B. Role of diffusion tensor imaging or magnetic resonance spectroscopy in the diagnosis and disability assessment of amyotrophic lateral sclerosis. *J Neurol Sci*. (2015) 348:206–10. doi: 10.1016/j.jns.2014.12.004
66. Chio A, Pagani M, Agosta F, Calvo A, Cistaro A, Filippi M. Neuroimaging in amyotrophic lateral sclerosis: insights into structural and functional changes. *Lancet Neurol*. (2014) 13:1228–40. doi: 10.1016/S1474-4422(14)70167-X
67. Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*. (2009) 45:S199–209. doi: 10.1016/j.neuroimage.2008.11.007
68. Vapnik VN. *The Nature of Statistical Learning Theory*. New York, NY: Springer-Verlag (1995). doi: 10.1007/978-1-4757-2440-0
69. Lebedev AV, Westman E, Van Westen GJ, Kramberger MG, Lundervold A, Aarsland D, et al. Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *Neuroimage Clin*. (2014) 6:115–25. doi: 10.1016/j.nicl.2014.08.023
70. Kim HJ, de Leon M, Wang X, Kim HY, Lee YJ, Kim YH, et al. Relationship between clinical parameters and brain structure in sporadic amyotrophic lateral sclerosis patients according to onset type: a voxel-based morphometric study. *PLoS ONE*. (2017) 12:e0168424. doi: 10.1371/journal.pone.0168424
71. Chang JL, Lomen-Hoerth C, Murphy J, Henry RG, Kramer JH, Miller BL, et al. A voxel-based morphometry study of patterns of brain atrophy in ALS and ALS/FTLD. *Neurology*. (2005) 65:75–80. doi: 10.1212/01.wnl.0000167602.38643.29

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Chen, Zhang, Huang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.