

Machine learning classificatory as a tool in the diagnosis of amyotrophic lateral sclerosis using diffusion tensor imaging parameters collected with 1.5T MRI scanner: A case study

Milosz Jamroz ^{1*}, Edyta Maj ¹, Maksymilian Bielecki ², Marta Bartoszek ³, Marek Golebiowski ¹,
Magdalena Kuzma-Kozakiewicz ⁴

¹ Department of Clinical Radiology, Medical University of Warsaw, Warsaw, POLAND

² Department of Psychology, SWPS University of Social Sciences and Humanities, Warsaw, POLAND

³ Department of Pediatric Radiology, University Clinical Center of the Medical University of Warsaw, Warsaw, POLAND

⁴ Department of Neurology, Medical University of Warsaw, Warsaw, POLAND

*Corresponding Author: miloszjamroz161@gmail.com

Citation: Jamroz M, Maj E, Bielecki M, Bartoszek M, Golebiowski M, Kuzma-Kozakiewicz M. Machine learning classificatory as a tool in the diagnosis of amyotrophic lateral sclerosis using diffusion tensor imaging parameters collected with 1.5T MRI scanner: A case study. *Electron J Gen Med.* 2023;20(6):em535. <https://doi.org/10.29333/ejgm/13536>

ARTICLE INFO

Received: 19 Jun. 2023

Accepted: 18 Jul. 2023

ABSTRACT

The relevance of the study lies in the need to improve the diagnosis of amyotrophic lateral sclerosis (ALS) by utilizing diffusion tensor imaging (DTI) obtained from conventional 1.5 Tesla MRI scanners. The study aimed to investigate the potential of using different machine learning (ML) classifiers to distinguish between individuals with ALS. In this study, five ML classifiers (“support vector machine (SVM)”, “k-nearest neighbors (K-NN)”, naïve Bayesian classifier, “decision tree”, and “decision forest”) were used, based on two DTI parameters: fractional anisotropy and apparent diffusion coefficient, obtained from two manually selected ROIs at the level of the brain pyramids in 47 ALS patients and 55 healthy subjects. The quality of each classifier was evaluated using the confusion matrix and ROC curves. The highest accuracy in differentiating ALS patients from healthy individuals based on DTI data was demonstrated by the radial kernel support vector method (77% accuracy [$p=0.01$]), while K-NN and “decision tree” classifiers had slightly lower performance, and “decision forest” classifier was overtrained to the training set (AUC=1). The authors have shown a sufficiently accuracy of ML classifier “SVM” in detecting radiological characteristics of ALS in pyramidal tracts.

Keywords: magnetic resonance imaging, 1.5 Tesla MRI scanner, fractional anisotropy, measured diffusion coefficient, machine learning classifiers, support vector method

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) or Lou Gehrig’s disease is a devastating neurodegenerative neuromuscular disease characterised by progressive degeneration of upper motor neurons in the motor cortex and lower motor neurons in the spinal cord and brain stem. This loss of motor neurons causes progressive muscle weakness and amyotrophy [1]. The prevalence of ALS varies across geographical clusters, ranging from 3.01 (per 100,000 people) in Asia (excluding Japan) to 6.22 (per 100,000 people) in Europe and 7.96 in Japan [2].

ALS is a heterogeneous disease whose mechanisms of pathogenesis and progression are not fully understood. There is currently no effective treatment able to halt the disease or reverse its symptoms, the available disease-modifying drugs can either slightly prolong survival or slow a functional decline. Currently, the average survival time of patients with ALS from the first symptoms to death is estimated at 36 months [3]. However, in some cases, symptomatic treatment including nutritional and respiratory support, and palliative care can alleviate its course, improve the patient’s quality of life and/or

prolong life [4]. It is in this context that timely and accurate diagnosis of ALS is crucial, which, firstly, excludes other disorders that may have similar manifestations but a more favourable prognosis/available treatment, and, secondly, allows for an earlier introduction of an optimal treatment strategy. The diagnosis of ALS can be difficult due to the heterogeneity of clinical manifestations of the disease, the presence of atypical phenotypes, and, most importantly, due to the similarity to other neurological disorders [5]. Up to 10% of patients initially diagnosed with ALS have other diseases with similar symptoms and, occasionally, a much better prognosis [5].

Until recently, the revised El Escorial criteria and the Awaji criteria remained the “gold standard” for the diagnosis of ALS [6]. In 2019, ALS experts established new Gold Coast criteria, which a large cross-sectional study confirmed to be more accurate and sensitive than their predecessors [7].

The examination tools used in patients with suspected ALS depend on the clinical picture and may include nerve conduction studies, needle electromyography, magnetic resonance imaging (MRI) and genetic testing, etc. Although neuroimaging methods are mainly used to exclude other

disorders of the nervous system, they can also provide additional information based on well-established radiological signatures of the disease, as presented in [8, 9]. Various MRI methods are used to track the functional and structural changes that characterise the onset and course of this disease, including voxel-based morphometry, functional MRI at rest (functional MRI) and magnetic resonance spectroscopy etc [10-12].

Diffusion tensor imaging (DTI) is considered to be the most promising candidate for a radiological biomarker in ALS. This method, in particular, allows measuring microstructural pathological abnormalities in the corticospinal tract (CST), which reflect the degree of brainstem and cerebral cortex atrophy characteristic of ALS [13]. The studies [14-17] demonstrate that diffusion-weighted MRI-based neuroimaging provides reliable measurements for assessing clinical severity and monitoring the progression of the disease and should be included in clinical evaluation protocols for patients with suspected ALS.

DTI provides unique information about the microstructure of white matter in the central nervous system. This method provides contrast images based on differences in the diffusion of water molecules in the brain. By measuring the orientation dependence of water molecule diffusion, DTI generates unique tissue contrasts that can be used to explore axonal organisation. Pathologic changes in white matter fibres *in vivo* are assessed by specific indicators, such as diffusion coefficients and fractional anisotropy (FA), whereby diffusion coefficients characterise changes in the degree and FA in the orientation of proton movement [18, 19]. However, measuring, calculating, and interpreting these indicators takes a lot of time, generates large amounts of information, and requires using complex data processing and analysis methods. Therefore, although potentially clinically useful, DTI is most commonly used in ALS research and has almost no place in routine clinical practice [8].

The solution to this problem can be using machine learning (ML) methods, which are a form of artificial intelligence and provide tools for developing powerful classifiers: they transform existing knowledge into an algorithm that can be used by physicians. Currently, ML is widely used in diagnostics, risk prediction, severity assessment, drug dose calculation etc. They are increasingly becoming useful assistants for healthcare professionals, allowing them to go beyond their own experience, reducing the time required to diagnose patients, eliminating cognitive biases, and reducing the risk of human error.

This study aimed to demonstrate the effectiveness of ML methods in diagnosing ALS and distinguishing ALS patients from healthy individuals using routine MRI data. The research addresses key questions in ALS diagnosis, including the reliability of using FA and ADC as markers, compensating for low MRI resolution, the viability of an SVM classifier as an alternative to complex techniques, the complementarity of FA and ADC, classifier performance comparison, potential for routine use and screening, and identifying distribution trends for improved diagnostic tools.

MATERIALS AND METHODS

The study was conducted at the University Clinical Center of the Medical University of Warsaw.

The study involved 47 patients with clinically definite, probable, or possible ALS according to [6]. The average duration of the disease, i.e., the time from the onset of the first paresis to the time of the study, was 31 months. The patients included 30 men and 17 women aged 23 to 81, with an average age of 53.

The control group included 55 healthy volunteers without symptoms of ALS and without diagnosed neurodegenerative diseases. There were 29 men and 26 women aged 18 to 82. The average age of the control group was 50 years.

Thus, both groups were comparable in terms of gender and age. All patients and volunteers provided written informed consent prior to inclusion into the study.

For MRI examinations, a 1.5 Tesla scanner (Magnetom Avanto SQ Engine TIM 76×32) with a 12-channel main coil manufactured by Siemens (Erlangen, Germany) was used. The routine protocol consisted of an axial T2-weighted turbo spin echo (TSE) (TR/TE 4,650/85 ms; slice thickness 5 mm), axial T2-weighted FLAIR image (Fluid attenuated inversion recovery, FLAIR) (TR/TE/TI 9,000/89/2,500 ms; slice thickness 5 mm), coronary T2 TSE (TR/TE 4,790/77 ms; slice thickness 5 mm), axial T1-weighted spin echo (SE) (TR/TE 592/13 ms; slice thickness 5 mm), axial slice in the mode of magnetic susceptibility-weighted imaging (SWI) (TR/TE 49/40 ms, slice thickness 3 mm), sagittal T2-weighted TSE (TR/TE 3,000/111 ms, slice thickness 5 mm), axial diffusion-weighted imaging (DWI) (TR/TE 4,600/99 ms; slice thickness 5 mm, coefficient $b=0/1,000/2,000$ s/mm²), and a sagittal T1-weighted MPR (Magnetisation prepared rapid acquisition gradient, MPR) image (TR/TE 1720/2.92 ms; slice thickness 1 mm).

Diffusion-weighted measurements were performed based on a spin echo echo-planar pulse sequence (SE EPI) with diffusion gradients applied in 30 spatial areas with the following parameters: TR/TE 3100/86 ms; slice thickness 5 mm; coefficient $b=0/1,000$ s/mm²; field of view (FOV) 230×230 mm, matrix size 128×128, number of averages=4. The study lasted 6:45 minutes.

The free-hand regions of interest (ROI) method was used to measure three parameters: axial FA, apparent diffusion coefficient (ADC), and TRACE. FA indicates how much the diffusion anisotropy scales the feature in the range from 0 to 1, where a result of "0" means that the diffusion is isotropic; and a value of "1" is theoretical and represents the highest level of anisotropy.

ADC is equal to the average diffusion coefficient and represents the strength of diffusion. The association of FA with ADC allows identifying brain regions, where strong anisotropy accompanies high diffusion. TRACE parameter, or total diffusivity, is the sum of the three DTI eigenvalues, which when divided by three gives the mean diffusivity (MD).

The areas of interest were placed by a qualified researcher according to the scheme presented by [8]: two in the anterior and posterior thirds of the pedicle of the internal capsule; one in the middle part of the brain stem; one in the bridge and medullary pyramids of the medulla oblongata. All ROIs presented were set on both sides, i.e., a set of 10 ROIs was obtained for each measurement. After performing the measurement using two ROIs in each location, the results were averaged across the sides. To evaluate the statistical significance of the parameters, statistical analysis using ANOVA with a mixed design was applied.

Table 1. Summary results of two-factor mixed design ANOVA

Parameter	Group (ALS vs HCs)		Structure		Group*structure		Significance of simple effects of group				
	F(1, 100)=	p	F(4, 400)=	p	F(4, 400)=	p	ICPost	ICAnt	CP	Pons	Pyramids
FA	33.40	<.001	246.63	<.001	5.36	<.001	0.007	0.239	<.001	<.001	<.001
ADC	16.53	<.001	55.13	<.001	2.68	0.032	0.240	0.083	0.003	<.012	<.001
TRACE	2.85	0.094	38.01	<.001	3.56	0.007	0.014	0.357	0.137	0.084	0.065

Note. ALS group: Patients; HCs group: Control group; ICPost & ICAnt: Measurements in anterior & posterior third of internal capsule pole; CP: In middle part of brain pons; Pons: In bridge of medulla oblongata; & Pyramids: In medullary pyramids of medulla oblongata

The study involved evaluating and comparing five different classifiers to assess their ability to distinguish between individuals with ALS and healthy individuals. The classifiers examined were support vector machine (SVM) with various kernels (linear, polynomial, radial, and sigmoid), k-nearest neighbors (K-NN), naïve Bayesian, decision tree, and forest of solutions classifiers.

All 102 cases (47 patients and 55 control subjects) were randomly assigned with an appropriate proportion between ALS and control subjects for both the test and training sets. A training set of 76 samples was used to train the classifiers. The train function of the caret library provides a function for adjusting model hyperparameters. The training set was further split into 25 sets of 76 samples using the initial loading algorithm: each set was used to determine the optimal hyperparameter value from the hyperparameter tuning grid using Cohen's kappa coefficient as a metric.

For each classifier, a confusion matrix was calculated based on the predicted results of the test sets. For all analyses, it was used the R-statistical package, ver. 4.0.3 [20].

RESULTS

Statistical analysis using ANOVA with a mixed design was performed on the averages of the images obtained from both sides. In the "groups" category, only FA and ADC parameters demonstrated statistically significant differences ($p < 0.001$) between patients with ALS and healthy controls (**Table 1**).

Therewith, all DTI parameters divided into "structures" demonstrated high statistical significance (all $p < 0.001$), which was further confirmed by comparing the mean values between "groups" and "structures" (all $p < 0.001$). It suggests that the difference between the control subjects and the ALS group was markedly different due to the "structures". The interaction effects were even larger when interpreted using simple group effects, which were estimated for each of the respective structures. The values of these comparisons are presented in the last five columns of **Table 1**.

The most pronounced differences were observed in FA values in the middle part of the brain stem, in the bridge of the medulla oblongata and the medullary pyramids of the medulla oblongata, and ADC values at the level of the pyramids. It confirms that they can be useful biomarkers for daily practice.

There are several ML algorithms or classifiers that allow searching for boundaries between groups and determining whether an object belongs to one of them.

SVM method is considered one of the most reliable forecasting methods based on a statistical learning system. Considering a set of training examples, each of which is labelled as belonging to one of two categories, this algorithm establishes a model that assigns new examples to one or the other category.

In its original form, SVM is a linear classification algorithm, but using the so-called "kernel trick" it can be transformed into a non-linear classifier. In this approach, instead of learning a fixed set of parameters that correspond to the features of the input data, the algorithm "remembers" specific training examples and the weight of each of them. A prediction for new input data is established by applying a similarity function (called a kernel) to the training data. The SVM kernel is usually chosen depending on the nature of the phenomenon being studied.

K-NN assumes that a new case should be classified into the nearest neighbor group. Thus, each new object is classified using the set of its neighbors (nearest training examples) and belongs to the class most common among them (k training samples).

A naïve Bayes classifier checks the probability of a new object belonging to each group under the assumption of (naïve) independence between the features.

"Decision tree" and "decision forest" algorithms divide data sets into subsets (groups) based on their most important attribute in several decision-making steps.

All of the above classifiers can be used to achieve the stated purpose of this study, i.e., to analyse the data obtained by neuroimaging to separate ALS patients from individuals without neurodegenerative diseases. However, considering such features of ALS as heterogeneity of clinical manifestations, atypical phenotypes, etc., when applying ML algorithms to analyse DTI results, there is a high risk of overfitting – a phenomenon when the built model explains well the examples from the training sample, but performs poorly on new examples due to some random patterns. In other words, due to excessive complexity and too many parameters relative to the number of observations, the statistical model describes random error or noise instead of the underlying relationship. A perfect match to the training data is manifested when the Area under the ROC curve (AUC) is close to one, but this value demonstrates the classifier's low efficiency, which is associated with a loss of generalisation.

The results of the confusion matrix for all classifiers using the test set are presented in **Table 2**.

Three classifiers produced statistically significant results: SVM with a radial kernel ($p < 0.01$), K-NN classifier ($p < 0.04$) and "decision tree" classifier ($p < 0.04$). SVM classifier with a radial kernel demonstrated the highest accuracy (> 0.77) and the lowest value of p. It was characterised by a sensitivity of 0.5, the highest specificity of 1.0, and the narrowest confidence interval.

Using SVM classifier with a radial kernel to distinguish between ALS patients and healthy controls based on DTI results are demonstrated in **Figure 1** and **Figure 2**, classifier performance with the training set and classifier performance with the test set, respectively.

Table 2. Confusion matrix results for all classifiers

Classifier	TP	TN	FP	FN	Acc	95% CI	p [Acc>NIR]	Kappa	McNemar's test: p	TPR	TNR	PPV	NPV	DR	DP	BA
SVM linear	6	12	2	6	0.69	0.48 0.86	0.08	0.37	0.29	0.50	0.86	0.75	0.67	0.23	0.31	0.68
SVM radial	6	14	0	6	0.77	0.56 0.91	0.01	0.52	0.04	0.50	1.00	1.00	0.70	0.23	0.23	0.75
SVM polynomial	6	12	2	6	0.69	0.48 0.86	0.08	0.37	0.29	0.50	0.86	0.75	0.67	0.23	0.31	0.68
SVM sigmoid	6	12	2	6	0.69	0.48 0.86	0.08	0.37	0.29	0.50	0.86	0.75	0.67	0.23	0.31	0.68
K-NN	6	13	1	6	0.73	0.52 0.88	0.04	0.44	0.13	0.50	0.93	0.86	0.67	0.23	0.27	0.71
Naïve Bayes	6	12	2	6	0.69	0.48 0.86	0.08	0.37	0.29	0.50	0.86	0.75	0.67	0.23	0.31	0.68
Decision tree	8	11	3	4	0.73	0.52 0.88	0.04	0.46	1.0	0.67	0.79	0.73	0.73	0.31	0.42	0.73
Decision forest	7	12	2	5	0.73	0.52 0.88	0.04	0.45	0.45	0.58	0.86	0.78	0.71	0.27	0.35	0.72

Note. SVM linear: Support vector machine with a linear kernel; SVM radial: Support vector machine with a radial kernel; SVM polynomial: Support vector machine with a polynomial kernel; SVM sigmoid: Support vector machine with a sigmoid kernel; K-NN: k-nearest neighbors classifier; Naïve Bayes: Naïve Bayes classifier; Decision tree: “Decision tree”; Decision forest: “Decision forest”; TP: True positive; TN: True negative; FP: False positive; FN: False negative; Acc: Accuracy; CI: confidence interval; NIR: No information rate–level of lack of information; TPR (true positive rate): Sensitivity; TNR (true negative rate): Specificity; PPV: Positive predictive values; NPV: Negative predictive values; DR: Detection rate; DP: Detection prevalence; & BA: Balanced accuracy

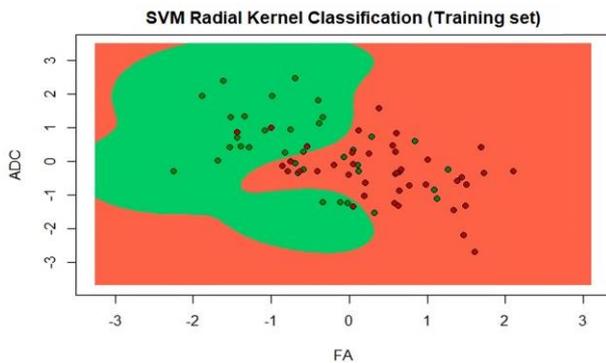


Figure 1. Radial kernel SVM classifier with training set (Source: Authors’ own elaboration)

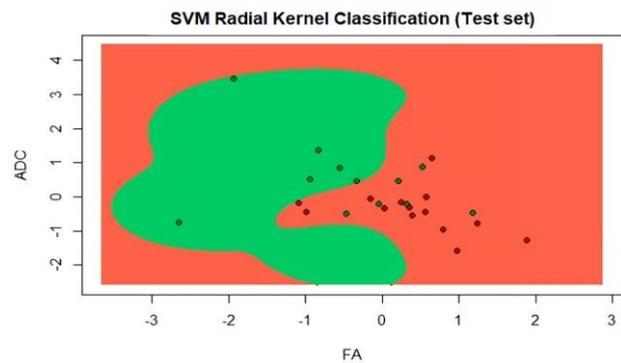


Figure 2. Radial kernel SVM classifier with a test set (Source: Authors’ own elaboration)

Table 4. AUC values for ROC curves

Classifier	SVM linear	SVM radial	SVM polynomial	SVM sigmoid	K-NN	Naïve Bayes	Decision trees	Decision forest
	AUC	0.73	0.77	0.70	0.70	0.70	0.70	0.76
	0.68	0.75	0.68	0.68	0.71	0.68	0.73	0.72

Note. SVM linear: Support vector machine with a linear kernel; SVM radial: Support vector machine with a radial kernel; SVM polynomial: Support vector machine with a polynomial kernel; SVM sigmoid: Support vector machine with a sigmoid kernel; K-NN: k-nearest neighbors classifier; Naïve Bayes: Naïve Bayes classifier; Decision tree: “Decision tree”; & Decision forest: “Decision forest”

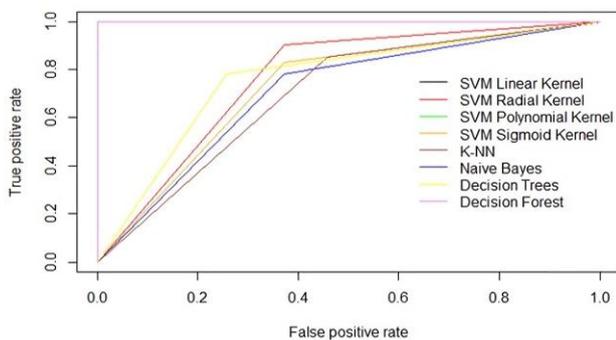


Figure 3. ROC curves with training set (Source: Authors’ own elaboration)

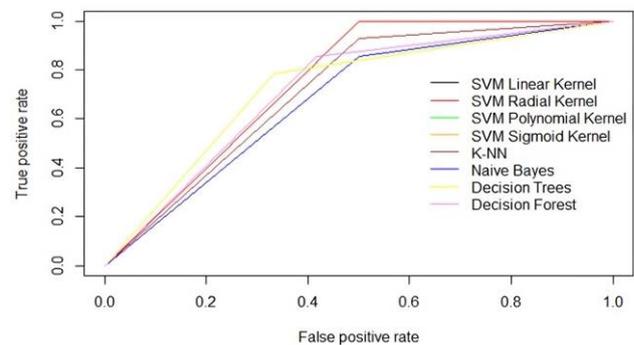


Figure 4. ROC curves with test set (Source: Authors’ own elaboration)

The results of calculating area under the curve (AUC) in the analysis of ROC curve, which visually displays the effectiveness of binary classification models, are presented in **Table 4**.

Generally, when plotting ROC curve for different models, the model with the largest AUC is considered the best. Based on this, the best classifiers are decision forest and radial kernel SVM, as they have the highest AUC values for both the test and training datasets. However, with the training set, the AUC value

for the “decision forest” classifier is one, which indicates that this classifier has been overtrained and disqualified.

The ROC curves with the training set are presented in **Figure 3**, and the ROC curves calculated with the test set are presented in **Figure 4**.

The graphical display of the data obtained during the analysis of the confusion matrix indicates that the “decision forest” classifier is overtrained and unusable.

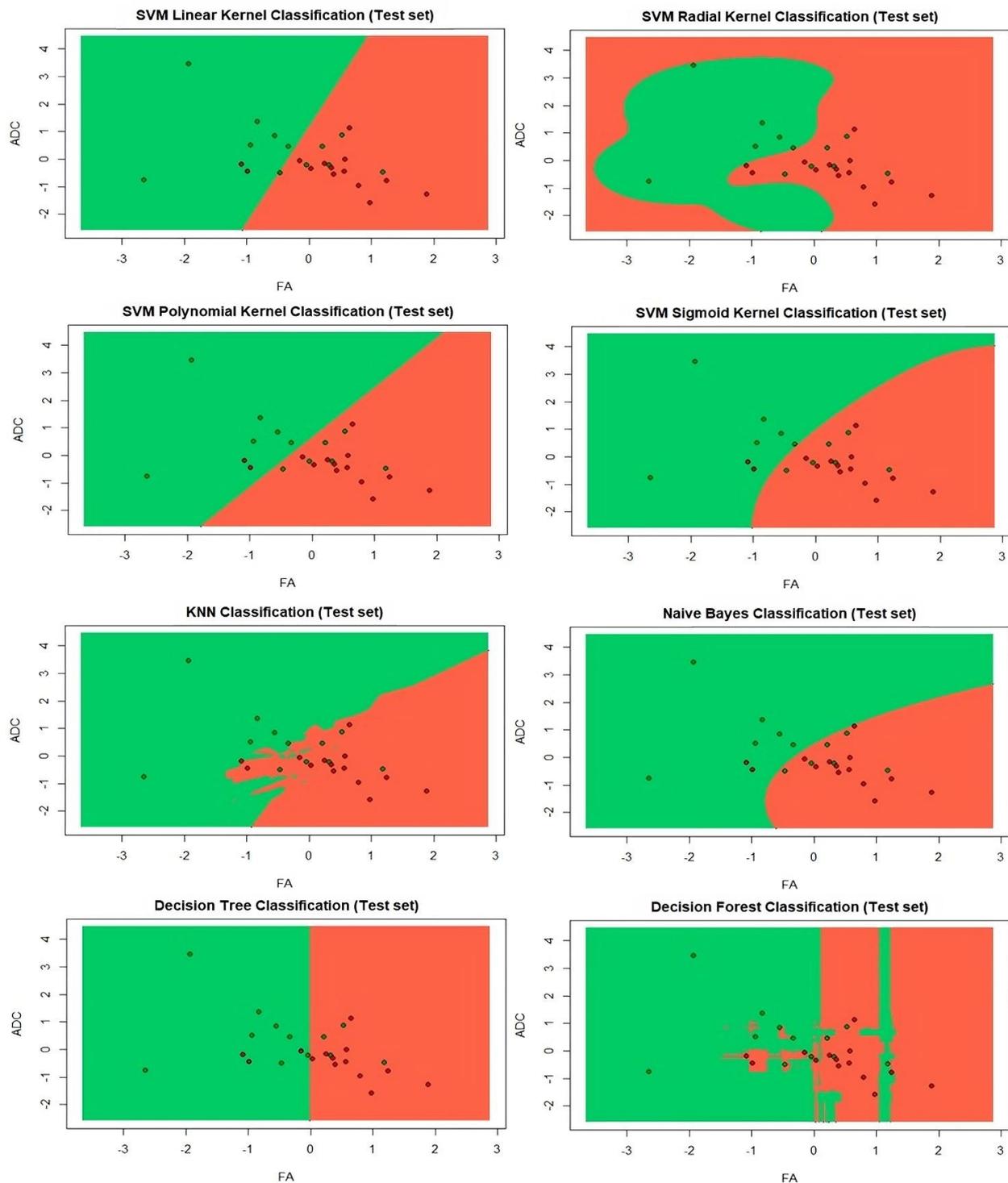


Figure 5. All classifiers with a test set (Source: Authors' own elaboration)

For comparison, **Figure 5** presents all the classifiers with the test set.

In all figures, ADC and FA values are scaled, green dots represent patients with ALS, red dots represent controls, the green box represents classification into ALS group, the red box represents assignment to the control group. Thus, the more points that match the colour of the field on which they are placed, the more accurate the classifier is.

DISCUSSION

Recently, ML has been increasingly used in evidence-based medicine (EBM). EBM is an experimental pool based on hypotheses and protocol experiments with well-defined populations and pre-selected variables, which is largely based on conventional statistical models [21-23]. In contrast, a form of artificial intelligence such as ML does not require a specific hypothesis and protocols, but transforms existing knowledge and data into algorithms, identifying patterns among several variables that are not defined in advance. However, as I. Scott

[24] notes, despite the fundamental difference between these two approaches, they can work in parallel, complementing each other, as some patterns that advanced computing programs (machines) can see cannot be detected by conventional biostatistics.

For a while, it was believed that ML could only complicate the interpretation of results by the National Institutes of Health, as its methods, such as deep learning, usually require a very large amount of data. However, it was later established that some less demanding ML methods based on simple rules that work very efficiently even on small datasets have great potential for solving EBM problems [25, 26]. This approach both expanded the practical use of ML in medical research and allowed its adaptation to explore rare and complex diseases, including ALS.

Currently, ML is widely used in ALS research, for example, to find and analyse biomedical signals inherent in this disease, identify and predict clinical subgroups, evaluate the results of muscle ultrasound in the early stages of ALS [27-30] etc. In addition, there are several publications devoted to using ML to analyse images of the brain affected by the disease obtained by neuroimaging and, in particular, diffusion tensor images [24, 31-35]. This particular MRI method can detect changes in the white matter that occur in ALS by measuring differences in restrictions on water diffusion in the brain. Summarizing the results of a systematic review of using ML algorithms to analyse data obtained by DTI-based neuroimaging, A. Behler et al. [31] note their enormous academic and clinical potential in developing biomarkers for this disease, which can be useful both at the group and individual level, for example, to improve individual differential diagnosis or serve as endpoints in clinical trials.

Considering that ALS is a rare disease, the vast majority of researchers have to work with limited samples. Thus, in the works devoted to using ML for this disease, they use only those algorithms that work effectively on small data sets. An example of such algorithms is the Random Forest classifier, which proved to be a powerful tool for complex classification in a small group of 24 ALS patients in the study by I. Scott [24] and 48 patients in A. Sarica et al. [35], and a linear SVM classifier that was able to distinguish between ALS patients and healthy volunteers in a relatively large group of 502 people in the study of T. Kocar et al. [32] and in a small group of 63 patients in the study of R. Welsh et al. [33].

Despite successful examples of ML applications for the interpretation of DTI images in ALS, it is still considered mainly as a research tool, which is quite far from daily clinical practice. It is largely explained by some technical difficulties of DTI, such as high requirements for scanning equipment, and qualification of radiologists or clinicians performing post-processing. Some studies use ultra-high-field MRI scanners up to 7 Tesla to measure microstructural pathological abnormalities in CST, and complex data processing is required to assess axonal damage, which is not available in clinical settings [5, 8, 34, 35].

Some difficulties occur when applying ML algorithms. In particular, they need training sets with high representativeness, as the accuracy and quality of classifier generalisation depend on it. In the case of such a complex disease as ALS, with a variety of clinical manifestations and phenotypes, obtaining such sets can be a significant challenge [36, 37]. Insufficient or excessive training of classifiers, which makes them unsuitable for analysis, is another threat, and

therefore, ML methods should be used carefully and cautiously to explore ALS.

To bridge the gap between science and daily clinical practice, i.e., to ensure that advanced methods and technologies are used as quickly and as widely as possible, it is necessary to develop ways to combine their effectiveness with ease of use. The purpose of the current study was precisely to solve this problem for which high-resolution scanners were replaced with conventional 1.5 Tesla MRI machines, and instead of advanced post-processing methods, two manually selected ROI and ML classifiers were used. The method proposed in this study significantly reduced the requirements for the qualifications of specialists, since to apply it, they were only required to mark the two areas being explored in the DTI image and then transcribe the two digits into a simple classifier.

Based on previous studies [14-17], the parameters of axial FA and measured ADC were chosen as markers that allow distinguishing patients with ALS from healthy individuals based on DTI images, as demonstrated by the studies of A. Behler et al. [31], S. H. Baek et al. [15], J. Li et al. [16], it is the association of FA with ADC that allows the best identification of brain regions, where strong anisotropy accompanies high diffusion, i.e., to identify those functional and structural changes that are considered pathological signs of the disease.

A 1.5 Tesla MRI scanner with a relatively low resolution was used for the scan, and to obtain better data, a manual selection of the region of interest (ROI) was used. This approach has already been tested in the studies of A. T. Toosy et al. [38], M. Cosottini et al. [39] and the authors earlier studies [17] and confirmed its effectiveness.

To analyse the data obtained with DTI, several ML algorithms were tested: SVM, K-NN method, naïve Bayesian classifier, "Decision Tree" and "Decision Forest". The effectiveness of each classifier was tested using a confusion matrix.

Three classifiers produced statistically significant results: SVM with the radial kernel ($p < 0.01$), K-NN ($p < 0.04$), and Decision Tree classifier ($p < 0.04$), and the best results among them were demonstrated by SVM with the radial kernel, reaching 77% accuracy ($p = 0.01$). In addition, this algorithm could separate ALS patients from the control group using the ROC curve, demonstrating an AUC of 0.75. It suggests that a supervised ML model, such as an SVM with a radial kernel, can be useful for separating ALS patients from healthy individuals based on DTI data obtained in a clinical setting.

The results of the current study are in line with the conclusions of other authors who note that SVM is a robust classification algorithm that provides high accuracy by transforming data into a high-dimensional feature space, where there is a maximum margin for class separation [31, 40-42]. While the accuracy of ML classifiers is lower than conventional threshold-based ROC, their usefulness is better due to the generalisation function: although the expansion of the tested group can give a significantly lower result than when using threshold-based ROC, ML classifiers can maintain their accuracy and generalisation quality, which depend only on the representativeness of the training set.

Another significant advantage of this classifier is its ease of use. For example, the methodology developed by [43] to identify patients with schizophrenia using MRI results and ML achieves accuracy in a similar range but has a much higher

complexity compared to using an SVM classifier and the simple procedure for measuring and preparing data for analysis proposed in this study.

Slightly better results when using the SVM classifier (AUC 0.87-0.88) to separate patients with ALS from the control group were obtained by [32], which can probably be explained by a much larger sample size (404 patients with ALS and 98 healthy controls). Therewith, in a smaller group of 32 patients diagnosed with ALS and 31 healthy controls, the study in [33] obtained results that are somewhat inferior to the current study. Using SVM, it was achieved 71% accuracy in classifying disease states.

In the current study, the specificity of the radial kernel SVM criterion reached 100%, but the sensitivity was only 50%. According to the latter indicator, it was significantly ahead of “decision forest” classifier, but it was “disqualified” due to overtraining, demonstrating AUC=1 on the training dataset. It excluded the possibility of generalising and correctly assessing the usefulness of the classifier.

Other classifiers tested in the study that demonstrated statistically significant results were less effective. Thus, the accuracy of K-NN, “decision tree”, and “decision forest” algorithms reached 73%, and the AUC with the training data set was 0.70, 0.76, and 1.00, respectively. The exclusion of the “decision forest” classifier from the study due to retraining once again confirmed the necessity to be cautious when applying ML methods in practice.

CONCLUSIONS

The study confirmed that FA and ADC are reliable markers to distinguish ALS patients from healthy individuals using DTI. Manual selection of the region of interest compensates for low MRI resolution, and a radial kernel SVM classifier showed the highest accuracy, providing an alternative to complex post-processing techniques. This easy-to-use classifier combined with a simple measurement protocol offers an affordable tool for routine ALS diagnosis with high accuracy, expediting the diagnostic process. The classifiers demonstrated similar boundary patterns, indicating the potential of FA and ADC for complementary diagnosis. Further research with larger datasets can explore trends and enhance the development of an effective ALS diagnostic tool using ML.

Limitations

A limitation of the current study is that it did not test random forest classifier, which has demonstrated good results in previous studies [24, 31, 35].

It was a single center study that was based on a simplified procedure for measuring only three DTI parameters. In addition, this should be considered a limitation, as the unifying theme of ML applications is the integration of many multidimensional data sources that provide different perspectives on diseases. As noted by [31], the perfect identification of all patients with ALS in a group mixed with healthy people remains unattainable solely by DTI, but the combination of DTI parameters with those of other MRI methods can significantly increase the diagnostic sensitivity. Corrections for the lifestyle of the control group, the duration of the disease, and the addition of other parameters to the CST diffusion metrics, such as motor cortical diffusion or structural parameters such as cortical thickness or texture properties,

can increase the diagnostic accuracy of SVMs [31, 32, 43], but significantly complicate using ML methods, particularly in daily clinical practice.

Future Directions and Clinical Implications

Future research can focus on integrating multimodal data to improve the accuracy and reliability of ALS diagnosis. Combining DTI parameters with other imaging modalities such as functional MRI (fMRI), structural MRI, genetic markers, or clinical biomarkers could provide a more comprehensive diagnostic assessment. This multimodal approach has the potential to enhance the understanding of ALS pathology and improve diagnostic accuracy. The development of advanced ML algorithms, particularly deep learning models, could contribute to improving the diagnostic capabilities in ALS. These algorithms can leverage the complex patterns present in DTI data to extract valuable features and provide more accurate and precise diagnostic predictions.

Longitudinal monitoring and disease progression in ALS patients can be another focus of future research. ML approaches can be applied to analyze longitudinal DTI data, enabling the tracking of disease progression and the monitoring of treatment response. By identifying changes in DTI parameters over time, ML models may provide valuable prognostic information and aid in personalized treatment planning.

Ultimately, the translation of ML-based diagnostic tools into clinical decision support systems can revolutionize ALS diagnosis. Integrating these algorithms into clinical workflows would assist clinicians in interpreting DTI data and making accurate ALS diagnoses, leading to improved diagnostic efficiency and patient outcomes.

Summarize, future research can focus on exploring alternative ML algorithms, such as random forests or deep learning models, to improve the classification of ALS patients. Integration of multimodal data, validation on independent cohorts, refinement of feature selection techniques, and practical implementation of the classifiers in clinical settings are also important areas for further investigation.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the work complies with ethical standards. The study was conducted in accordance with the Declaration of Helsinki approved by the Human Experiments Ethics Committee of Medical University of Warsaw, Warsaw, Poland.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

1. Meyer T. Amyotrophic lateral sclerosis (ALS)–diagnosis, course of disease and treatment options. *Dtsch Med Wochenschr.* 2021;146(24-25):1613-8. <https://doi.org/10.1055/a-1562-7882> PMID:34879411
2. Brown CA, Lally C, Kupelian V, Flanders WD. Estimated prevalence and incidence of amyotrophic lateral sclerosis and SOD1 and C9orf72 genetic variants. *Neuroepidemiol.* 2021;55(5):342-53. <https://doi.org/10.1159/000516752> PMID:34247168

3. Corcia P, Beltran S, Bakkouche SE, Couratier P. Therapeutic news in ALS. *Rev Neurol*. 2021;177(5):544-9. <https://doi.org/10.1016/j.neurol.2020.12.003> PMID:33781562
4. Oskarsson B, Gendron TF, Staff NP. Amyotrophic lateral sclerosis: An update for 2018. *Mayo Clin Proc*. 2018;93(11):1617-28. <https://doi.org/10.1016/j.mayocp.2018.04.007> PMID:30401437
5. Kwan J, Vullaganti M. Amyotrophic lateral sclerosis mimics. *Muscle Nerve*. 2022;66(3):240-52. <https://doi.org/10.1002/mus.27567> PMID:35607838
6. Ido B, J, Kacem I, Ouedraogo M, et al. Sensitivity of Awaji criteria and revised El Escorial criteria in the diagnosis of amyotrophic lateral sclerosis (ALS) at first visit in a Tunisian cohort. *Neurol Res Int*. 2021;2021:8841281. <https://doi.org/10.1155/2021/8841281> PMID:33552600 PMCid:PMC7847325
7. Turner MR, UK MND Clinical Studies Group. Diagnosing ALS: The Gold Coast criteria and the role of EMG. *Pract Neurol*. 2022;22(3):176-8. <https://doi.org/10.1136/practneurol-2021-003256> PMID:34992096 PMCid:PMC9120398
8. Cosottini M, Donatelli G, Costagli M, et al. High-resolution 7T MR imaging of the motor cortex in amyotrophic lateral sclerosis. *Am J Neuroradiol*. 2016;37(3):455-61. <https://doi.org/10.3174/ajnr.a4562> PMID:26680464 PMCid:PMC7960124
9. Roccatagliata L, Bonzano L, Man-Cardi G, Canepa C, Capon-Netto C. Detection of motor cortex thinning and corticospinal tract involvement by quantitative MRI in amyotrophic lateral sclerosis. *Amyotroph Lateral Scler*. 2009;10(1):47-52. <https://doi.org/10.1080/17482960802267530> PMID:18622772
10. Mohammadi B, Kollewe K, Samii A, Krampfl K, Dengler R, Münte TF. Changes of resting state brain networks in amyotrophic lateral sclerosis. *Exp Neurol*. 2009;217(1):147-53. <https://doi.org/10.1016/j.expneurol.2009.01.025> PMID:19416664
11. Foerster BR, Callaghan BC, Petrou M, Edden RAE, Chenevert TL, Feldman EL. Decreased motor cortex-Aminobutyric acid in amyotrophic lateral sclerosis. *Neurology*. 2012;78(20):1596-600. <https://doi.org/10.1212/wnl.0b013e3182563b57> PMID:22517106 PMCid:PMC3348851
12. Trojsi F, Di Nardo F, Siciliano M, et al. Frontotemporal degeneration in amyotrophic lateral sclerosis (ALS): A longitudinal MRI one-year study. *CNS Spectr*. 2021;26(3):258-67. <https://doi.org/10.1017/s109285292000005x> PMID:32089134
13. Li H, Zhang Q, Duan Q, Jin J, Hu F, Dang J, Zhang M. Brainstem involvement in amyotrophic lateral sclerosis: A combined structural and diffusion tensor MRI analysis. *Front Neurosci*. 2021;15:675444. <https://doi.org/10.3389/fnins.2021.675444> PMID:34149349 PMCid:PMC8206526
14. Kassubek J, Pagani M. Imaging in amyotrophic lateral sclerosis: MRI and PET. *Curr Opin Neurol*. 2019;32(5):740-6. <https://doi.org/10.1097/wco.0000000000000728> PMID:31335337
15. Baek SH, Park J, Kim YH, et al. Usefulness of diffusion tensor imaging findings as biomarkers for amyotrophic lateral sclerosis. *Sci Rep*. 2020;10:5199. <https://doi.org/10.1038/s41598-020-62049-0> PMID:32251314 PMCid:PMC7090054
16. Li J, Pan P, Song W, Huang R, Chen K, Shan H. A meta-analysis of diffusion tensor imaging studies in amyotrophic lateral sclerosis. *Neurobiol Aging*. 2012;33(8):1833-8. <https://doi.org/10.1016/j.neurobiolaging.2011.04.007> PMID:21621298
17. Maj E, Jamrozý M, Bielecki M, Bartoszek M, Gołębowski M, Wojtaszek M, Kuźma-Kozakiewicz M. Role of DTI-MRI parameters in diagnosis of ALS: Useful biomarkers for daily practice? Tertiary centre experience and literature review. *Neurol Neurochir Pol*. 2022;56(6):490-8. <https://doi.org/10.5603/pjnns.a2022.0070> PMID:36426927
18. Keil C, Prell T, Peschel T, Hartung V, Dengler R, Grosskreutz J. Longitudinal diffusion tensor imaging in amyotrophic lateral sclerosis. *BMC Neurosci*. 2012;13:141. <https://doi.org/10.1186/1471-2202-13-141> PMID:23134591 PMCid:PMC3531302
19. Alexander AL, Lee JE, Lazar M, Field AS. Diffusion tensor imaging of the brain. *Neurotherapeutics*. 2007;4(3):316-29. <https://doi.org/10.1016%2Fj.nurt.2007.05.011> PMID:17599699 PMCid:PMC2041910
20. R Core Team. R-project. Available at: <https://www.R-project.org/> (Accessed: 12 May 2023).
21. Koshbayeva L, Medelova A, Hailey D, Yermukhanova L, Uraz R, Aitmanbetova A. Influence of a health technology assessment on the use of pediatric cochlear implantation in Kazakhstan. *Health Policy Technol*. 2018;7(3):239-42. <https://doi.org/10.1016/j.hlpt.2018.06.002>
22. Serniak YP, Sagalevych AI, Frolov OS, Serniak PY, Kryvopustov MS. Extraperitoneoscopic radical prostatectomy after pelvic surgery procedures. *Wiad Lek*. 2020;73(6):1093-6. <https://doi.org/10.36740/WLek202006102> PMID:32723932
23. Okasova AK, Ilderbayev OZ, Nursafina AZ, et al. Evaluation of lipid peroxidation under immobilization stress in irradiated animals in experiment. *Open Access Macedon J Med Sci*. 2021;9:119-22. <https://doi.org/10.3889/oamjms.2021.5781>
24. Scott IA. Machine learning and evidence-based medicine. *Ann Intern Med*. 2018;169(1):44-6. <https://doi.org/10.7326/M18-0115> PMID:29710098
25. Akhmetova KM, Vochshenkova TA, Dalenov ED, Abduldajeva AA, Benberin VV. The interconnection of metabolic disorders and carotid atherosclerosis in the Kazakh population. *Syst Rev Pharm*. 2020;11(12):2152-9.
26. Molcan J, Dobrovanov A, Koren R, Kralinsky K, Balaz V. Unilateral scrotal hernia with dual ureter herniation: The first experience of successful surgical correction. *Pediatrics Zh. im G.N. Speranskogo*. 2021;100(4):171-5. <https://doi.org/10.24110/0031-403X-2021-100-4-171-175>
27. Fernandes F, Barbalho I, Barros D, et al. Biomedical signals and machine learning in amyotrophic lateral sclerosis: A systematic review. *Biomed Eng OnLine*. 2021;20(1):61. <https://doi.org/10.1186/s12938-021-00896-2> PMID:34130692 PMCid:PMC8207575
28. Faghri F, Brunn F, Dadu A, et al. Identifying and predicting amyotrophic lateral sclerosis clinical subgroups: A population-based machine-learning study. *Lancet Digit Health*. 2022;4(5):359-69. [https://doi.org/10.1016/s2589-7500\(21\)00274-0](https://doi.org/10.1016/s2589-7500(21)00274-0) PMID:35341712
29. Fukushima K, Takamatsu N, Yamamoto Y, et al. Early diagnosis of amyotrophic lateral sclerosis based on fasciculations in muscle ultrasonography: A machine learning approach. *Clin Neurophysiol*. 2022;140:136-44. <https://doi.org/10.1016/j.clinph.2022.06.005> PMID:35772191

30. Tursynova A, Omarov B, Sakhipov A, Tukenova N. Brain stroke lesion segmentation using computed tomography images based on modified U-net model with ResNet blocks. *Int J Online Biomed Engin.* 2022;18(13):97-112. <https://doi.org/10.3991/ijoe.v18i13.32881>
31. Behler A, Müller HP, Ludolph AC, Kassubek J. Diffusion tensor imaging in amyotrophic lateral sclerosis: Machine learning for biomarker development. *Int J Mol Sci.* 2023;24(3):1911. <https://doi.org/10.3390/ijms24031911> PMID:36768231 PMCID:PMC9915541
32. Kocar TD, Behler A, Ludolph AC, Müller HP, Kassubek J. 2021. Multiparametric microstructural MRI and machine learning classification yields high diagnostic accuracy in amyotrophic lateral sclerosis: Proof of concept. *Front Neurol.* 2021;12:745475. <https://doi.org/10.3389/fneur.2021.745475> PMID:34867726 PMCID:PMC8637840
33. 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.* 2023;7:251. <https://doi.org/10.3389/fnhum.2013.00251> PMID:23772210 PMCID:PMC3677153
34. Li W, Wei Q, Hou Y, et al. Disruption of the white matter structural network and its correlation with baseline progression rate in patients with sporadic amyotrophic lateral sclerosis. *Transl Neurodegener.* 2021;10(1):35. <https://doi.org/10.1186/s40035-021-00255-0> PMID:34511130 PMCID:PMC8436442
35. Sarica A, Cerasa A, Valentino P, et al. The corticospinal tract profile in amyotrophic lateral sclerosis. *Hum Brain Map.* 2017;38(2):727-39. <https://doi.org/10.1002/hbm.23412> PMID:27659483 PMCID:PMC6867092
36. Stadnik SN. Effect of statinotherapy on the cognitive functions of patients with disturbances of cardiac rhythm and conduction. *Azerb Pharm Pharmacother J.* 2021;21(2):61-9.
37. Atamanyuk IP, Kondratenko YP. Calculation method for a computer's diagnostics of cardiovascular diseases based on canonical decompositions of random sequences. *CEUR Workshop Proceed.* 2015;1356:108-20.
38. Toosy AT, Werring DJ, Orrell RW, et al. Diffusion tensor imaging detects corticospinal tract involvement at multiple levels in amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry.* 2003;74(9):1250-7. <https://doi.org/10.1136/jnnp.74.9.1250> PMID:12933929 PMCID:PMC1738665
39. Cosottini M, Giannelli M, Siciliano G, et al. Diffusion-tensor MR imaging of corticospinal tract in amyotrophic lateral sclerosis and progressive muscular atrophy. *Radiology.* 2005;237:258-64. <https://doi.org/10.1148/radiol.2371041506> PMID:16183935
40. Dobrovanov O, Kralinsky K, Molcan J, Kovalchuk VP. Relevance of ultrasound neonatal screening of the urinary system. *Ross Vest Perinat Pediatr.* 2019;64(2):68-72. <https://doi.org/10.21508/1027-4065-2019-64-3-68-72>
41. Shkorkorbatov Y, Pasiuga V, Kolchigin N, Batrakov D, Kazansky O, Kalashnikov V. Changes in the human nuclear chromatin induced by ultra wideband pulse irradiation. *Cent Eur J Biol.* 2009;4(1):97-106. <https://doi.org/10.2478/s11535-008-0051-4>
42. Schapovalova O, Gorlova A, de Munter J, et al. Immunomodulatory effects of new phytotherapy on human macrophages and TLR4- and TLR7/8-mediated viral-like inflammation in mice. *Front Med.* 2022;9:952977. <https://doi.org/10.3389/fmed.2022.952977> PMID:36091684 PMCID:PMC9450044
43. Yendiki A, Panneck P, Srinivasan P, et al. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Front Neuroinform.* 2011;5:23. <https://doi.org/10.3389/fninf.2011.00023> PMID:22016733 PMCID:PMC3193073