A mobile application for detection of amyotrophic lateral sclerosis via voice analysis

Denis Likhachov[1],Maxim Vashkevich[1[0000-0002-1267-0795]], Elias Azarov[1], Katsiaryna Malhina[2], Yuliya Rushkevich[2[0000-0002-2529-8005]]

[1]Belarusian State University of Informatics and Radioelectronics, 6 P. Brovky str., 220013, Minsk,Belarus
[2]Republican Research and Clinical Center of Neurology and Neurosurgery. 24 F. Skoriny str., 220114,Minsk, Belarus
likhachov@bsuir.by

Abstract. Analysis of the state of the art in the field of treatment and diagnosis of a rapidly progressing neurological disease of the amyotrophic lateral sclerosis (ALS) revealed a lack of tools for early diagnosis and monitoring the course of the disease. This paper proposes a method for evaluating the state of the voice function of patients with ALS, intended for use in a mobile application. As a speech probe a long pronunciation of the vowel sound / a / is used from whichacoustic features (jitter, shimmer, degree of vibrato pathology, etc.) are extracted. Applying a classifier based on the method of linear discriminant analysis to the obtained vector of parameters, it is possible to obtain an estimate whether the presented voice is pathological and possible progress of the disease. For training and verification of the proposed method, a database of 64 voices (33 healthy, 31 patients with ALS), recorded at the Republican Research and Clinical Center of Neurology and Neurosurgery(Minsk, Belarus), was used. To test the proposed method, a prototype of the mobile application «ALS Expert»was developed with the ability to record, process voice and display the results of voice analysis. The results indicate that the acoustic analysis of the voice and the resulting objective parameters are a promising direction for solving this problem.

Keywords:speech processing, detection of pathology in the voice, amyotrophic lateral sclerosis.

1 The relevance of solving the problem of developing affordable non-invasive methods for early diagnosis of ALS

The most common type of motor neuron disease (MND) is amyotrophic lateral sclerosis (ALS). This disease is characterized by damage to the motor neurons of the brain and spinal cord, conducting corticospinal pathways with the development of muscle atrophy, bulbar and respiratory disorders. The rapid progression of the disease leads to a violation of the functions of breathing and swallowing, the ability to move independently. The life expectancy of patients with ALS is 3-5 years. According to the conducted studies, the one-year survival rate of patients with ALS is 80.3%[95% CI 76.6-84.1], the 2-year survival rate decreases to 51%[95% CI 46.3-55.8], the 3-year survival rate decreases to 30.9%[95%CI 26.2-35.5]. Currently, there is an increase in mortality rates from ALS and its prevalence, which is probably due to improved diagnostic capabilities and an increase in the number of patients seeking medical care [1]. Early diagnosis of ALS allows you to actively influence the course of the disease, preserving the quality of life of the patient.

Speech and swallowing disorders develop either at the onset of the disease or are associated with the progression of the disease, occurring in more than 80% of all cases of ALS. In the early stages of the disease, speech disorders are skipped by doctors due to the rarity of the pathology. The consequence of this is the wrong diagnosis and the appointment of treatment with the use of unnecessary, and sometimes dangerous for ALS types of therapy or even surgical treatment.

Thus, the main tasks of the doctor are a quick and objective diagnosis, monitoring the patient's status in dynamics for its timely correction, especially speech and swallowing [2]. The development and testing of tools for the diagnosis, monitoring and evaluation of patients' status is a necessary criterion for the implementation of effective actions on the part of the doctor.The process of speech production involves cortical and subcortical structures, and therefore is particularly sensitive to the death of motor neurons. This means that speech analysis for early diagnosis of ALS can be useful for identifying both people with the bulbar form of the disease and people with other forms of ALS who have concomitant bulbar disorders.

Currently, active research is underway aimed at developing a system for detecting speech disorders caused by ALS [3,4,5,6,7].

The aim of this paper is to develop a method for diagnosing and tracking speech disorders in people with ALS, adapted for use in a mobile application. The main advantages of creating a system for early diagnosis of the disease and monitoring its course in the form of a mobile application are its ease of use and high availability. It should also be noted that the mobile application does not impose strict restrictions on the place of testing and does not require it to be carried out in a specially equipped room. This allows testing both directly by the doctor and by the patient independently.

2 Specificity of ALS diagnostics by voice analysis

2.1 Speech tasks used in ASL detection systems

Currently available research systems for detecting speech disorders associated with ALS differ in the speech tasks used in them. Thus, in [8, 9], a count from one to ten was used as a speech task. In [10-12], it is proposed to use several speech tests consisting in the utterance of monosyllabic words, notes, and a long sound [and]. In [13], the test signal contained the short phrase "one, two, three". Phrases of general content (such as "I need help", etc) were used as a speech task in the works [14, 15].

In [5], the participants were asked to read a specially designed short text passage ("Bamboo passage")[16]. This passage is designed to make it easier to automatically search for pauses between words. To do this, the voiced consonants in it are placed on the borders of words, since they better emphasize the border between the word and the pause than the deaf consonants.

Some researchers use shorter, but specially selected sentences. Thus, in [6] only the first sentence from the "bamboo passage" was used, and in [17] the speech task consisted of repeating the sentence "Buy Bobby a puppy" ten times. It can be noted that the sentence uses explosive consonants [p] and [b], the pronunciation of which requires a coordinated work of articulatory organs. Other examples of specially developed proposals can be found in [18].

In some studiesdedicated to the detection of speech disorders in ALS, a special type of speech task is used – the diadochokinetic (DDC) test [15, 19]. The DDC test consists of quickly pronouncing syllables (for example, "pa - / ta/ - ka" or "bu - / ter/ - kap") with maximum speed and accuracy in one breath. This test is widely used in differential diagnosis and for detecting disorders in the muscular apparatus of speech [20].

Sometimes a long-drawn-out vowel sound is used as a speech task. It can be included in the test of several speech tasks [19, 20] or can act as a single source of information [21, 22].

In [14], the repetition of listened phrases and spontaneous speech were used as speech tasks.

Summarizing the above data, we can say that the speech tasks used in the systems for detecting speech disorders caused by ALS are divided into the following groups: :

1) simple short phrases of general meaning [13-15];
2) short specially designed sentences [6, 17];
3) specially designed passages of the text [5];
4) diadochokinetic (DDC) test [4, 15, 19];
5) prolonged pronouncing of a vowel sound [23, 24];
6) a complex test consisting of several types of speech tasks [10-12, 21, 22].

2.2 Acoustic voice and speech features used in ALS detection systems

The choice of acoustic features for detecting speech disorders in ALS is an open question, and researchers solve it in different ways.

In [6], for the purpose of early diagnosis of ALS, a set of a very large number of acoustic features (6861 features) was used, which were extracted from the analyzed signal using a special OpenSMILE toolkit. For classification, all features were grouped into seven categories: cepstral coefficients, formants, energy parameters, pitch frequency, spectral parameters, time parameters, and parameters that evaluate the quality of the speech signal. The OpenSMILE library was also used in [25, 15]

In [5], the researchers used the speed of speech (number of words per minute), the speed of articulation (number of syllables per second), and the ratio of the duration of pauses to the duration of the speech task to differentiate people with and without symptoms of the disease. It has been experimentally shown that the use of indicators of speech speed and articulation is a fairly promising approach for the diagnosis of ALS.

To identify changes in speech motor skills that can be associated with ALS, it was proposed in [4] to use four speech features: 1) coordination; 2) consistency; 3) speed; 4) precision. The coordination and consistency of speech with multiple repetition of certain syllables were evaluated using the voice onset time – the length of time that passes between the release of a stop consonant and the onset of voicing. The speed and accuracy were evaluated by analyzing the frequency response of the second formant (F2). The experimental verification of the approach proposed in [4] was performed by analyzing speech recordings from two groups: a control group (18 people) and those with ALS symptoms (14 people). The data obtained were analyzed using a linear mixed model (LMM). With a sufficient degree of confidence, it was shown that coordination, speed and accuracy of speech can be associated with the deterioration of speech function in patients with ALS.

2.3 Classification methods used in systems for detecting speech disorders in ALS

The efficiency of the currently existing classifier models depends on the features, structure, and patterns [26] that are present in the analyzed data. For the task of detecting speech disorders in ALS, classification based on the support vector machine (SVM – support vector machine) is most often used [6,7,15,17,25]. Sometimes researchers use deep neural networks (DNN-deep neural network) and convolutional neural networks (CNN-convolutional neural network) [7,14,15]. In [19], the XGBoost (Extreme Gradient Boosting) method was used for classification. In [9,13,23,24], the classification was based on the linear discriminant analysis (LDA) method.

Despite the growing popularity of convolutional neural networks and deep learning methods, the application of these approaches to the task of detecting speech disorders associated with ALS has not yet shown a significant advantage over other classification methods. For example, the researchers in [14] obtained a detection accuracy of 75-77 % when using convolutional neural networks. The popularity of classification based on the support vector machine can be explained by its relatively low computational complexity (compared to a neural network) and its freely distributed software implementation in the form of the libSVM library. The advantages of the method are its resistance to retraining and high efficiency. For example, in [6], a classifier based on the support vector machine with a linear kernel function was used. The classifier was trained using speech recordings from 123 people divided into three groups: control group, with pronounced symptoms and with initial symptoms of the disease. To identify patients with severe ALS symptoms, a high classification accuracy index (AUC = 0.91 – 0.99) was obtained. However, in the context of the problem considered in this paper, the support vector method has a serious drawback. The result of his work is very difficult to interpret [26]. For a doctor, it is not enough just to know what result the detection system gave after analyzing the voice, but it is important to understand on the basis of what the decision was made. In this respect, the LDA-based classification approach is in a better position. This method is close to the linear regression model and its result can be easily interpreted.

3 The proposed method for evaluating the state of the voice function in patients with ALS

3.1 The used acoustic features

This paper is based on the results obtained earlier in [23, 24]. These papers describe systems for detecting speech disorders in patients with ALS based on linear discriminant analysis (LDA). The simplest test for long-term phonation of sound (a) was used as a speech task. The following acoustic features were extracted from the phonation record.

Jitter. Jitter – a measure of the change in the pitch frequency.It is an indicator of the pronator system stability of the speech formation subsystem (when the neuromotor function is impaired, this indicator increases).In a simplest case jitter is defined as

$$J_{loc} = \frac{1}{N-1}\sum_{i=2}^{N}|T_i - T_{i-1}| \,/\, \frac{1}{N}\sum_{i=1}^{N}|T_i| \,, \tag{1}$$

where T_i – is length of the i-th pitch period, N is total number of pitch periods in the voice signal.

According to (1) jitter estimates short-term variations in pitch period duration. This small variation cannot be accounted to voluntary changes of pitch and therefore attributed to violation of neuromotor control of the phonatory system.

In general case, jitter estimated as a period perturbation quotient (PPQ) on L consecutive pitch periods:

$$J_{ppqL} = \frac{\frac{1}{N-L+1}\sum_{i=1+(L-1)/2}^{N-(L-1)/2}\left|T_i - \frac{1}{L}\sum_{n=i-(L-1)/2}^{i+(L-1)/2} T_n\right|}{\frac{1}{N}\sum_{i=1}^{N}|T_i|}\,, \tag{2}$$

As a rule, parameter L is takes the values $3, 5$ and 55 [29, 30].As a measure of pitch frequency instability directional perturbation factor (DPF) is used.

Shimmer. Shimmer – a measure of the variation in the acoustic wave amplitude during phonation. An increase in this parameter, considering the natural decline in the intensity of the voice, indicates the presence of speech disorders. In a simplest case shimmer is defined as:

$$S_{loc} = \frac{1}{N-1}\sum_{i=2}^{N}|A_i - A_{i-1}| \,/\, \frac{1}{N}\sum_{i=1}^{N}|A_i| \,, \tag{3}$$

where A_i – is amplitude of the *i*-th pitch period.

In general case, shimmer estimated as amplitude perturbation quotient (APQ) on L consecutive pitch periods:

$$S_{apqL} = \frac{\frac{1}{N-L+1}\sum_{i=1+(L-1)/2}^{N-(L-1)/2}\left|A_i - \frac{1}{L}\sum_{n=i-(L-1)/2}^{i+(L-1)/2} A_n\right|}{\frac{1}{N}\sum_{i=1}^{N}|A_i|}, \quad (4)$$

As a rule, parameter L is takes the values $3, 5, 11$ and 55 [29, 30].

Pitch periods entropy (PPE). PPE [24] is used to evaluate the violation of the ability to control the stability of the pitch frequency during the prolonged utterance of sounds. An increase in this parameter indicates that the natural level of pitch variation is exceeded.

The pathology vibrato index (PVI). Vibrato is the rapid and regular oscillation of the pitch frequency during prolonged phonation. In patients with ALS, this parameter is characterized by high-frequency components (about 9-14 Hz). The algorithm for calculating PVI is given in [23-24].

Noise parameters. We also used parameters HNR (*harmonic noise ratio*) and GNE (*glottal noise excitation ratio*) that characterized noise component of the voice signal [Michaelis-97]. HNR and GNE calculated on a frame-by-frame basis, therefore, as a result of such an analysis, a set of values of each parameter are obtained, which can be considered as a random variable. For the HNR and GNE values, the median (HNR_{me} and GNE_{me} and the interquartile range (HNR_{IRQ} and GNE_{IRQ} were calculated. **Table 1** summarized features used for train LDA classifier model.

Table 1. Features extracted from vowel /a/.

Parameter group	Number of features	List of features
Frequency perturbation	6	$J_{loc}, J_{ppq\,3}, J_{ppq\,5}, J_{ppq\,55} J_{ppq\,63}$, DPF
Amplitude perturbation	5	$S_{loc}, S_{apq\,3}, S_{ppq\,5}, S_{ppq\,11}, S_{ppq\,55}$
Based on F0-contour	2	PPE, PVI
Noise	4	$\mathrm{GNE_{me}}, \mathrm{GNE_{IQR}}, \mathrm{HNR_{me}}, \mathrm{HNR_{IQR}}$
Total	17	

Jitter and Shimmer parameters optimization. In the literature described the cases of use J_{ppqL} and S_{apqL} parameters with fixed values of L (as mentioned above). In this work we carried out optimization of this parameters in order to find the optimal value of L. The following criterion was used:

The J_{ppqL} and S_{apqL} parameters can be applied at fixed values of L (as mentioned above). In this paper, these parameters were optimized in order to find the optimal value of L. The following criterion was used:

$$C(L) = \frac{\left(\mathrm{mean}\left(J_{ppqL}{}^{(H)}\right) - \mathrm{mean}\left(J_{ppqL}{}^{(P)}\right)\right)^2}{\mathrm{var}\left(J_{ppqL}{}^{(H)}\right) + \mathrm{var}\left(J_{ppqL}{}^{(P)}\right)}, \tag{5}$$

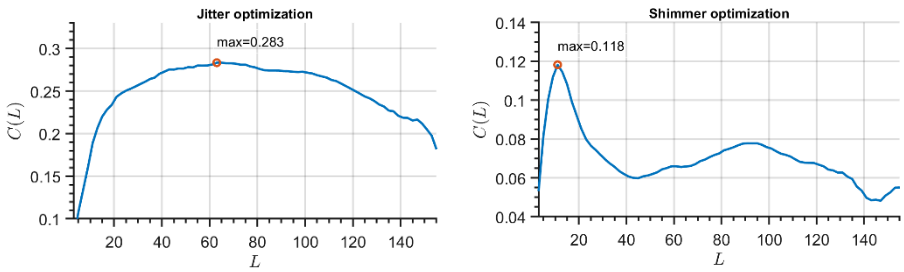
where $J_{ppqL}{}^{(H)}$ is the values of jitter calculated for the control group with parameter L, $J_{ppqL}{}^{(P)}$ is the values of jitter calculated for the ALS group with parameter L, $\mathrm{mean}(\cdot)$ is the operator of calculating the average value, $var(\cdot)$ is the operator variance.

The problem of finding the optimal value of L is formulated as

$$L_{opt} = \underset{L\in[3,\ 155]}{\mathrm{argmax}}\ C(L), \tag{6}$$

According to criterion (16), in (17) the value of L is searched for which the difference of the average value of J_{ppqL} between the control group and the group of patients with ALS is maximal, and the variation of J_{ppqL} within each group is minimal. For the S_{apqL} parameter, a similar procedure was used to find the optimal value of L.

Fig. 1shows the results of optimization jitter and shimmer parameters. It turned out that optimal window size for jitter is equal to 63 pitch periods. In turn, for the shimmer, the optimal window size is equal to 11 pitch periods.



a) b)

Fig. 1. The result of voice perturbation parameters optimization:
a) Jitter, $L_{opt} = 63$; b) Shimmer, $L_{opt} = 11$.

3.2 Classifier models

In this work we used two popular machine learning techniques to build classifier model: linear discriminant analysis and k nearest-neighbor. In this section we briefly describe those approaches.

Using LDA analysis, the vector of input parameters is transformed into a single value z, on the basis of which a decision on the presence of a disease is made.

Let there be a training data set $\left\{\left(\mathbf{x}^{(i)}, y_i\right)\right\}$, $i = 1,2, \dots m.$. In this case, $\mathbf{x}^{(i)}$ is a d-dimensional vector of acoustic features derived from analysis of the recording of the i-th voice from the database, y_i is the corresponding class label (for a healthy voice use the label "-1", for the voice of the patient with als – "1").

In the method of LDA data are interpreted as points in d-dimensional space, that which is sought in the one-dimensional projection:

$$z = \mathbf{w}^{\mathrm{T}}\mathbf{x}, \tag{7}$$

where $\mathbf{w}$ is the vector defining the normal to the hyperplane to be projected onto. The idea of the LDA method is to find a hyperplane for which, as a result of projecting a training set on it, the intra-class variance would be minimal, and the inter-class difference would be maximum (so-called Fisher criterion [27]).

The classification function in the LDA method has the following form:

$$f(\mathbf{x}) = \mathrm{sign}(\mathbf{w}^T\mathbf{x} - b), \tag{8}$$

where $\mathbf{x}$is the vector of acoustic features, $b$is the offset, and $\mathrm{sign}(a)$ is the function that returns the sign of the number a. If $f(\mathbf{x}) = 1$, then vector $\mathbf{x}$ belongs to the class of patients with ALS, andif $f(\mathbf{x}) = -1$, it belongs to the class of healthy patients. The offset$b$ is a threshold value, so for the input$\mathbf{x}$ vector, if $\mathbf{w}^T\mathbf{x} > b$, then it is classified as belonging to the class of ALS patients, otherwise to the class of healthy ones.

The selection of the offset b can be done as follows. The first stage is the projection of all vectors from the training data set into a one-dimensional space using the expression (7). At the second stage, a value of b is produced that would allow optimal separation of the marked data in a one-dimensional space according to a pre-specified criterion (the criterion can be obtaining maximum sensitivity, specificity, etc.).

The idea of k nearest-neighbor (kNN) classification is to assigninglabel to data sample based on the k closest points $\mathbf{x}^{(i)}$ in the training set. Classification function of the kNN method is defined as follows:

$$f(\mathbf{x}) = \frac{1}{k} \sum_{\mathbf{x}^{(i)} \in N_k(\mathbf{x})} y_i, \tag{9}$$

where$N_k(\mathbf{x})$ – k nearest neighbor from training set closest to $\mathbf{x}$.We used Mahalanobis distance to measure the distance between data samples.

3.3 The database of voices

To train the LDA model proposed in this paper, a database of voices was collected at the Republican Research and Clinical Center of Neurology and Neurosurgery (Minsk, Belarus). The voices of 64 people were recorded, of which 33 were healthy (13 men, 20 women) and 31 patients with ALS with signs of bulbar disorders (17 men, 14 women). The average age in the healthy group was 50.2 years (SD 13.8) for men and 56.1 (SD 9.7) for women. The average age in the ALS group was 61.1 years (SD7.7) for men and 57.3 (SD 7.8) for women. Each participant was asked to pronounce the long vowel sound /a/ in one breath for as long as possible with a comfortable pitch and volume. The voice was recorded using a smartphone with a headset (sampling rate 44.1 kHz) and stored as uncompressed 16-bit wav files. The average duration of the recordings was 4.1 s. The Matlab functions used for voice analysis are hosted in a public repository[1]. The voice database is available in public GitHub repository[2].

3.4 Train and validation of classifier models

Our goal was to obtain a classifier with a small number of features. It was assumed that no more than one feature from each group can be selected into the feature vector. An exception was made only for features based on F0-contour: the feature vector must include at least one of these features or both. According to this formulation of the problem, there are 7 × 6 × 3 × 5 = 630 variants of feature vector that satisfy the requirements. This number of options is small, so all of them can be verified by direct search.

For each possible variant of the feature vector, the quality of the classifier model was assessed using the *k-fold cross-validation method* [28] (for k = 9). Classifiers were found with the highest accuracy, sensitivity, specificity and AUC. The results for the LDA-based classifiers are shown in **Table 2**.

Table 2. Results of LDA-based classification.

Feature vector	Accuracy,%	Sensitivity, %	Specificity, %	AUC, \%
J_{loc}, PPE, PVI, $\mathrm{GNE_{me}}$	**77.7 ± 2.5**	84.7 ± 2.8	70.4 ± 4.3	0.81 ± 0.02
$J_{ppq\,63}$, PPE	66.7 ± 2.0	**95.0 ± 2.2**	37.6 ± 3.5	0.77 ± 0.02
PVI	60.7 ± 1.4	32.9 ± 1.9	**89.4 ± 1.8**	**0.83 ± 0.01**

The classifier that uses the parameters jitter, PPE, PVI and the noise parameter GNE has the highest accuracy. The greatest specificity is obtained by using the optimized jitter and PPE parameter. The PVI-based classifier showed the highest specificity and AUC.

For kNN-based classifier were tested different values of parameter $k = \{1,3,5,7,9\}$. It was found that the best performance characteristics obtained for value $k = 1$. For 3-component feature vector $J_{ppq\,63}$, PPE, PVI the kNN-classifier give the following results: accuracy 83,9±2,5, sensitivity 82,7±4,9, specificity 85,2±3,4, AUC 0.84±0.03.

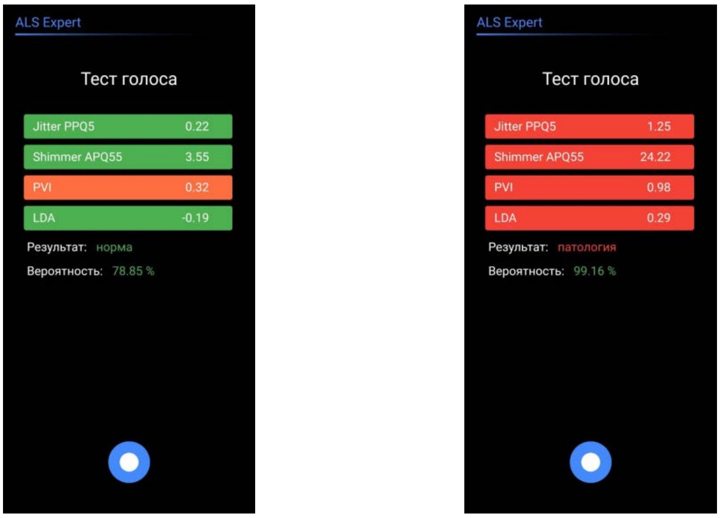
For the kNN-based classifier, the accuracy results were slightly better than for the LDA-based classifier. However, the preference was given to the LDA-based classifier, since it has a better specificity parameter,thus there will be fewer cases of healthy people being classified as ALS patients.

3.5 ALS Expert mobile application

To test the proposed method for assessing the state of the voice function, a prototype of the ALS Expert mobile application was created (see**Fig. 2**), which is focused on performing two main functions: early diagnosis of ALS and monitoring the state of speech apparatus for ALS patients in dynamics.

[1]https://github.com/Mak-Sim/Troparion

[2]https://github.com/Mak-Sim/Minsk2020_ALS_database



Norm Pathology

Fig. 2. Examples of output results in the ALS Expert application.

ALS Expert[3] is an application for the Android operating system. The computational core of the application analyzes the digitized speech signal in accordance with the mathematical model described above. It was implemented in the C++ programming language. The UI and application logic are implemented with Flutter, a custom UI toolkit for building mobile apps from Google.

The procedure for evaluating the status of the patient's voice function is conducted by a doctor and consists of the following. The patient performs a speech task (drawls the sound /a/). The result of uttering the sound /a/ is recorded using the built-in microphone or headset microphone and processed by the application's computing core. The result of the application is the values of the acoustic parameters, which are further analyzed by the doctor.

3.6 Conclusion

In this paper, a method was proposed for evaluating the state of the voice function in patients with ALS. The method is based on LDA classifier that used to determine if a voice belongs to a group of healthy people or a group of ALS patients. Our goal was to obtain the classifier with a small number of features. To study the efficiency of the classification, four groups of acoustic parameters were used: frequency perturbation, amplitude perturbation, features based on F0-contour and noise parameters. It was found that the classifier that uses the parameters jitter, PPE, PVI and GNE has the highest accuracy. The greatest specificity is obtained by using the optimized jitter and PPE parameter. The PVI-based classifier showed the highest specificity and AUC.

---

[3]https://drive.google.com/file/d/14U-t2ajWX5ILfNWx9hBZG1UfpHlhW197/view?usp=sharing

The obtained results indicate that the developed method can be used for evaluating of the voice function state in patients with ALS. The proposed method was developed considering the specifics of use in a mobile application (limited computing power, the impossibility of performing complex combined voice tests, etc.). The strengths of this approach: 1) ease of organization and execution of the test – a simple test for the long pronunciation of the vowel / a / is used; 2) the result is available immediately after the test; 3) possibility of automatic data accumulation; 4) relative low cost of the technical solution. However, the proposed method and its software implementation are experimental and require further clinical testing and verification.

References

1. Rushkevich Yu.N., Likhachev S.A. Modern concepts of motor neuron disease.Meditsinskie novosti. – 2019. – №1. – P. 23–31.
2. Rushkevich Yu.N., Likhachev S.A. Epidemiology of motor neurone disease in the Republic of Belarus.Neurology and neurosurgery. The Eastern Europe. –2018. –№4. – P. 551–561.
3. Eshghi M. et al. Reduced task adaptation in alternating motion rate tasks as an early marker of bulbar involvement in amyotrophic lateral sclerosis // Proceedings of Interspeech. – 2019. – P. 4524-4528.
4. Rowe H. P., Green J. R. Profiling speech motor impairments in persons with amyotrophic lateral sclerosis: an acoustic-based approach //Proc. Interspeech. – 2019. – С. 4509-4513.
5. Connaghan K. P. et al. Use of Beiwe Smartphone App to Identify and Track Speech Decline in Amyotrophic Lateral Sclerosis //Proc. Interspeech 2019. – 2019. – С. 4504-4508.
6. Gutz S. E. et al. Early Identification of Speech Changes Due to Amyotrophic Lateral Sclerosis Using Machine Classification //Proc. Interspeech. – 2019. – С. 604-608.
7. Suhas B. N. et al. Comparison of Speech Tasks and Recording Devices for Voice Based Automatic Classification of Healthy Subjects and Patients with Amyotrophic Lateral Sclerosis //Proc. Interspeech. – 2019. – С. 4564-4568.
8. RushkevichYu.N., etal.Clinical experience of speech signal time-frequency analysis in the diagnostics and monitoring of bulbar dysfunction. TheEasternEurope. – 2017. – №3. – P. 429–439.
9. Gvozdovich A.D., Rushkevich Y.N., Vashkevich M.I. Detection of bulbar amyotrophic lateral sclerosis based on speech analysis. Doklady BGUIR. 2018;(6):52-58. (In Russ.)
10. Mobile app to collect diagnostic information through the registration of speech signals / A. N. Osipov, et al. // BIGDATAAdvancedAnalytics: collectionofmaterialsofthefourthinternationalscientificandpracticalconference, Minsk, Belarus, May 3 – 4, 2018 / editorialboard: M. Batura [etc.]. – Minsk, BSUIR, 2018. – P. 343 – 347.
11. Kul T.P., Rushkevich Y.N., Lihachev S.A. Adaptation of digital signal processing methods to the analysis of speech in neurological pathologies. Doklady BGUIR. 2018;(7):128-132.
12. Kul T.P., Mezhennaya M.M., Rushkevich Yu.N., Osipov A.N., Likhachev S.A., Rushkevich I.V. Methodical and hardware-software for recording and processing speech signals for diagnosis of neurological diseases. Informatics. 2019;16(2):27-39. (In Russ.)
13. Vashkevich M., Gvozdovich A., Rushkevich Y. Detection of Bulbar Dysfunction in ALS Patients Based on Running Speech Test //International Conference on Pattern Recognition and Information Processing. – Springer, Cham, 2019. – P. 192-204.

14. An, K. Automatic early detection of amyotrophic lateral sclerosis from intelligible speech using convolutional neural networks/ K. An, M. Kim, K. Teplansky, J. Green, T. Campbell, Y. Yunusova, D. Heitzman, J. Wang // Proceedings of Interspeech 2018 – P. 1913-1917.
15. Wang, J., Towards Automatic Detection of Amyotrophic Lateral Sclerosis from Speech Acoustic and Articulatory Samples/ J. Wang, P.V. Kothalkar, B. Cao, D. Heitzman// Proc. Interspeech 2016 – P. 1195-1199.
16. Yunusova, Y. et al. Profiling speech and pausing in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) //PloS one. – 2016. – Т. 11. – №. 1. – P. 1-18.
17. Bandini, A., Classification of Bulbar ALS from Kinematic Features of the Jaw and Lips: Towards Computer-Mediated Assessment / A. Bandini, J.R. Green, L. Zinman , Y. Yunusova// Proc. Interspeech 2017 – P.1819-1823.
18. Liss, J. M. Discriminating dysarthria type from envelope modulation spectra / J.M. Liss, S. LeGendre, A. J. Lotto // Journal of Speech, Language, and Hearing Research. – 2011. – № 5 (53). – P. 1246-1255.
19. Spangler, T. Fractal features for automatic detection of dysarthria / T. Spangler, N. V. Vinodchandran, A. Samal, J. R. Green // IEEE EMBS International Conference on Biomedical Health Informatics (BHI), Orlando, USA, Feb. 16-19, 2017. – Orlando, 2017 – pp. 437–440.
20. Wang, Y.-T. Analysis of diadochokinesis in ataxic dysarthria using the motor speech profile programTM/ Y.-T. Wang, R. D. Kent, J. R. Duffy, and J. E. Thomas// Folia phoniatrica et logopaedica: official organ of the International Association of Logopedics and Phoniatrics (IALP). – 2009.– vol. 61, no. 1. – P. 1–11.
21. Green, J.R. Bulbar and speech motor assessment in ALS: challenges and future directions / J.R. Green, Y. Yunusova, M.S. Kuruvilla, J. Wang, G.L. Pattee, L. Synhorst, L. Zinman, J.D. Berry // Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration. – 2013. – № 14. –P. 494-500.
22. Yunusova, Y. Detection of bulbar ALS using a comprehensive speech assessment battery/ Y. Yunusova, J. S. Rosenthal, J. R. Green, S. Shellikeri, P. Rong, J. Wang, L. H. Zinman // in Proc. of the International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications, 2013, P. 217-220.
23. Vashkevich M., Petrovsky A., Rushkevich Y. Bulbar ALS Detection Based on Analysis of Voice Perturbation and Vibrato // 2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA). – P. 267-272.
24. VashkevichM., RushkevichY. Detection of impairment of speech function in patients with als based on analysis of voice signal // Digitalsignalprocessing. – 2019. – №4. – P. 33 – 41.
25. Norel, R., Detection of amyotrophic lateral sclerosis (ALS) via acoustic analysis / R. Norel, M. Pietrowicz, C. Agurto , S. Rishoni, G. Cecchi // Proc. Of Interspeech 2018.– P. 377–381.
26. Kelleher J. D., Mac Namee B., D'arcy A. Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. – MIT press, 2020.
27. Bishop, C. M. Patternrecognitionandmachinelearning. – Springer-Verlag. – NewYork, 2006.
28. Flach P. Machine learning: the art and science of algorithms that make sense of data. – Cambridge University Press, 2012.
29. Baken, R.J., R.F. Orlikoff Clinical measurement of speech and voice, 2nd edition / R.J. Baken, Orlikoff R.F. // Thomson Learning, 2000. – 864 p.

30. Telephony-based voice pathology assessment using automated speech analysis / R. J. Moran [et al.] // IEEE Transactions on Biomed. Engineering – 2006. – vol. 53, no. 3. – P. 468-477.
31. Michaelis D., Gramss T., Strube H. W. Glottal-to-noise excitation ratio–a new measure for describing pathological voices. Acta Acustica united with Acustica 1997; 83; 4; 700-706.