pattern recognition, supervised classification, k-NN rules,feature selection, respiration, ventilation, paralysis, diaphragm

Adam JÓŹWIK[*,**], Beata SOKOŁOWSKA[***]

SOME PROBLEMS WITH CONSTRUCTION OF THE K-NN CLASSIFIER FOR RECOGNITION OF AN EXPERIMENTAL RESPIRATION PATHOLOGY

An objective of the work is to demonstrate some difficulties with construction of a classifier based on the k-NN rule. The standard k-NN classifier and the parallel k-NN classifier have been chosen as the two most powerful approaches. This kind of classifiers has been applied to automatic recognition of diaphragm paralysis degree. The classifier construction consists in determination of the number of nearest neighbors, selection of features and estimation of the classification quality. Three classes of muscle pathology, including the control class, and five ventilatory parameters are taken into account. The data concern a model of the diaphragm pathology in a cat. The animals were forced to breathe in three different experimental situations: air, hypercapnic and hypoxic conditions. A separate classifier is constructed for each kind of the mentioned situations. The calculation of the misclassification rate is based on the *leave one out* and on the testing set method. Several computational experiments are suggested for the correct feature selection, the classifier type choice and the misclassification probability estimation.

1. INTRODUCTION

The diaphragm is the most important muscle involved in mammalian ventilation [3, 11]. Its weakness increases the demands on other respiratory muscles and may lead to decrease in ventilation [8, 9, 10]. The purpose of the present study was to examine if a diaphragm paralysis can be recognized on the basis of ventilatory parameters in a quiet and stimulated breathing. An influence of the hemidiaphragm paralysis on a respiratory activity was assessed in 12 anaesthetized, spontaneously breathing cats in supine position. The gradually diaphragmatic paralysis was induced by unilateral and bilateral section of phrenic nerves. Animals were observed both before and after an unilateral phrenicotomy (class 1 and 2, respectively) and after a bilateral phrenicotomy (class 3). The recognition possibility of the diaphragm paralysis was examined in three experimental conditions: before any stimulus, after an hypercapnic one (5% CO_2 in O_2) and after a hypoxic stimulus (11% O_2 in N_2). The ventilation measurements were performed 1 hour after the stimulus was given. The following five parameters were analyzed: a breathing frequency (feature 1), an inspiratory timing (feature 2), an expiratory timing (feature 3), a tidal volume (feature 4) and a minute ventilation (feature 5).

[*] Institute of Biocybernetics & Biomedical Engineering, PAS, Trojdena 4, 02-109 Warsaw, Poland
[**] Technical University of Łódź, Computer Engineering Department, Al. Politechniki 11, 90-924 Łódź, Poland
[***] Medical Research Centre, Department of Neurophysiology, PAS, Pawińskiego 5, 02-106 Warsaw, Poland

To analyze a dependence of respiratory parameters on the paralysis degree we applied the statistical methods offered by the pattern recognition theory which deals with object classification.

The objects are understood in a very general sense. It is assumed that objects are described by sets of parameters called features. Thus, each object is represented by a vector, or a point, in a feature space. Since now, the point or the vector in the feature space will be identified with the object. Methods of the statistical pattern recognition theory starts with a set of objects with known class membership. Such set, called the reference set or the training set, is then used for obtaining a decision rule that allows to classify new objects from outside the reference set. The larger is the reference set the more precisely the considered classes are defined. The numerical forces of the classes in the reference set ought to be approximately consistent with the frequencies how these classes appear. The experimental data that correspond to three different conditions of breathing: without, with hypercapnic and with hypoxic stimulus, were gathered in the sets X_1, X_2 and X_3, respectively. An analysis of the respiratory data, based on the statistical pattern recognition approach, was already used in the authors' earlier works [6, 13].

2. METHODS

2.1. THE BAYE'S AND THE K-NN CLASSIFIERS

The best possible decision rule, i.e. the one that offers minimum misclassification rate (error rate) is based on the Bayes formula: $p(j/\mathbf{x})=p(j)\cdot f(\mathbf{x}/j)/f(\mathbf{x})$, where $p(j/\mathbf{x})$ is a probability of the class j under assumption of the feature vector $\mathbf{x}$, $f(\mathbf{x}/j)$ denotes the density of the probability distribution for the class j and $f(\mathbf{x})$ is the density of the probability distribution of the feature vector $\mathbf{x}$. The classified point $\mathbf{x}$ ought to be assigned to the class j that corresponds to the greatest value of $p(j/\mathbf{x})$. All functions, which appear on the right hand side in the mentioned above Bayes formula, are unknown. To approximate them on the basis of the training set, one can use the neighborhood containing k points nearest to the classified point $\mathbf{x}$. Thus, $p(j)\approx m_j/m$, $f(\mathbf{x}/j)\approx k_j/(m_j\cdot V(\mathbf{x},k))$, $f(\mathbf{x})\approx k/(m\cdot V(\mathbf{x},k))$, where m_j is a number of points $\mathbf{x}$ from the class j in the training set, m is a numerical force of the training set, k_j means a number of points from the class j among k nearest points (neighbors) of the classified point $\mathbf{x}$, and $V(\mathbf{x},k)$ is a volume occupied by a smallest sphere containing these k nearest neighbors. Hence, the probability function that occurs on the left hand side of the Bayes formula can be approximated by the ratio k_j/k, i.e. $p(j/\mathbf{x})\approx k_j/k$. In this way the k nearest neighbor (k-NN) decision rule has been defined. The classified point $\mathbf{x}$ is assigned to the class j that corresponds to the highest value of k_j/k. It is proved by others authors [2] that if the size m of the training set gets larger ($m\to\infty$), $k\to\infty$ and $k/m\to 0$ then the performance of the k-NN rule converges to the performance of the Bayes classifier. This is why we decided to deal with the classifiers based on this rule and its modifications. As a distance function the city distance measure is assumed. The k-NN was introduced half century ago in the work [4].

The training set containing points with known class membership is the set that is used for the construction of the classifier. In case of the k-NN rule it may serve for experimental determination of the parameter k. The value of k should be determined in a way that offers the smallest probability of misclassification. Such a k is called an optimum one. The choice of k is based on misclassification rate estimation for k=1,2,...,m-1, where m is the numerical force of the reference set.

2.2. METHODS OF ESTIMATION OF THE MISCLASSIFICATION RATE

The misclassification rate can be estimated by the separate testing data set T, or by the *leave one out* method [1] on the basis of the reference set R. The method using the reference set consists in that all points from the testing set T are classified by the k-NN rule based on the reference set R. A probability of misclassification is estimated by an error rate $er=r/m$, where r is a number of misclassified objects and m is a number of classified points, *i.e.* the number of points in the set T. The *leave-one-out* method consists in classification of each object **x** of the reference set by the k-NN rule with the reference set decreased by the currently classified object, *i.e*. the object **x** is classified by the k-NN classifier with the set R-{**x**}as the reference set. Also in this case, the error rate $er=r/m$ estimates the probability of misclassification, where r means the number of misclassified objects and m is this time the number of points in the set R. We do not need to be in possession of the testing set in such an approach. For this reason the whole data set X can be used as the reference set, *i.e.* R=X. Random division of the whole data set X is the most common way of creating the sets T and R.

2.3. THE FEATURE SELECTION PROBLEM

It is difficult to extract the features that ought to be used for the object description. No theory exists in this matter. The feature extraction is based on our experience and feeling. So, very often some features, not relevant to the considered classification problem, are present in the reference set. The redundant features make the classification more complex and spoil its performance. For this reason the feature selection is strongly recommended. We should select the feature set, out of all available features, which offers the minimum value of the error rate for the optimum k-NN rule.

2.4. THE PARALLEL K-NN CLASSIFIER

A multi-class problem can be decomposed to some two-decision tasks. One of the possible solutions is the construction of a parallel net of two-decision classifiers, a separate classifier for each pair of classes, and then forming the final decision by voting of these two-decision classifiers. Such kind of approach was first used in [7] and later theoretically analyzed in [12]. This parallel network of two-decision classifiers should offer better performance than the standard k-NN classifier. It results from the geometrical interpretation of both discussed types of classifiers. In case of the standard classifier, the boundary separating any pair of classes i and j depends also on the samples from the remaining classes. They have influence on the value k and on the selected features. These samples may act as noise. The parallel net may reduce this noise effect. By using the error rate estimated by the *leave one out* method as a criterion, we can find an optimum numbers of k for the k-NN rules and perform the feature selection separately for each of the component classifiers. In spite of the above considerations, the parallel net of two-decision k-NN classifiers does not ensure better results as compared to the standard k-NN classifier.

3. RESULTS

3.1. THE SCHEME OF EXPERIMENTAL COMPUTATIONS

In the section 2.2 we have described the two ways of misclassification rate estimation, the testing set approach and the *leave one out* method. If the size of the reference set is not sufficiently large in relation to the number of features then the obtained error rate, calculated with the use of the *leave one out* method, may be too optimistic [5]. There is no rule to evaluate a priori the sufficiency of the reference size. The more features are used, the larger reference set is required. If we decide to apply the testing set approach, then our data set X ought to be divided into the reference set R and the testing set T. Then the constructed classifier can be determined by the set R and next tested with the use of the set T. To verify how confident is this approach in case of the data X we have performed such an experiment few times. For this aim we have created randomly 10 pairs of sets (A_i, B_i), i=1,2,..10, where $A_i \cup B_i = X$. Then 20 new pairs (R_j,T_j), j=1,2,…,20, such that: $R_1=A_1$, $T_1=B_1$, $R_2=B_1$, $T_2=A_1$,…,$R_{19}=A_{10}$, $T_{19}=B_{10}$, $R_{20}=B_{10}$, $T_{20}=A_{10}$, were considered. Below we present the results of these experiments done separately for each data set X_1, X_2, X_3.

3.2. BREATHING WITHOUT STIMULUS

The data set X_1, as it was described in the paragraph 1, concerns a situation when the cats were breathing in the atmospheric air. A classification task consists in recognition the paralysis degree on the basis of 5 previously particularized features, not necessary using all of them. We knew a priori that only the features 1, 2 and 4 are independent. The values of the remaining features 3 and 5 can be derived from these 3 features by virtue of a mathematical formula. Results gathered in the Tab. 1 suggest that the smallest error rate, equal 0.0358 was offered by the parallel classifier. Another interesting phenomenon is that the 3 independent features 1, 2 and 4 were ones being the most frequently selected as by the standard *k*-NN as well as by the parallel net of two-decision *k*-NN classifiers.

Standard *k*-NN, no feat. selection	**Standard *k*-NN, with feat. select.**	**Parallel *k*-NN, no feat. selection**	**Parallel *k*-NN, with feat. selection**	**Standard *k*-NN, features 1,2,4**	**Parallel *k*-NN, features 1,2,4**
1	2	3	4	5	6
0.0542	0.0450	0.0542	0.0358*	0.0596	0.0581
In the standard classifier the features **1,2**,3,**4**,5 appeared **16, 18,** 13, **15** and 7 times respectively					
In the parallel classifier the features **1,2**,3,**4**,5 appeared **20, 20,** 17, **20** and 17 times respectively					

Tab.1. Mean error rates and results of feature selection obtained by *leave one out* method for the sets R_1,R_2,…,R_{20} randomly chosen from the data set X_1

The error rate estimation by the *leave one out method* is not quite honest since the values of *k* were fitted to receive the best result. Also the features (see columns 2 and 4 in Tab.1) were selected to minimize the error rate. In case of the standard *k*-NN classifier the value of *k* and the features were chosen to minimize the final misclassification rate. When the parallel classifier was analyzed, the values of *k* and the features were selected to minimize the error rates separately for the component classifiers. More desired would be to minimize the global error rate.

Problem 1: No methods have been developed till now to select the numbers of k and to perform feature selection to minimize the error rate for the whole parallel net.

To evaluate all classifiers presented in the Tab. 1 in a more honest way, the testing set approach was applied in each of 20 experiments. The optimum numbers of k and the feature sets were established with the use of the sets $R_1,R_2,\ldots,R_{20}$, but the error rates, see Tab. 2, were calculated by classification the objects from the corresponding testing sets $T_1,T_2,\ldots,T_{20}$. Comparing the columns 2 and 1, and next 4 and 3, we can see that feature selection caused an increase of the error rates in the both types of classifiers.

Basic charact.	**Stand. *k*-NN, no feat. sel.**	**Stand. *k*-NN, with feat. sel.**	**Parallel *k*-NN, no feat. sel.**	**Parallel *k*-NN, with feat. sel.**	**Stand. *k*-NN, features 1,2,4**	**Parallel *k*-NN, features 1,2,4**
0	1	2	3	4	5	6
minimum	0.0231	0.0154	0.0231	0.0308	0.0308	0.0308
mean	**0.0631***	**0.0708**	**0.0823**	**0.0869**	**0.0677**	**0.0761**
maximum	0.1308	0.1385	0.1769	0.1846	0.1385	0.1615
stand. dev.	0.0285	0.0329	0.0445	0.0421	0.0275	0.0333

Tab.2. Error rates computed by the use of the testing set method for the data X_1

The standard k-NN classifier, which operates with all 5 features, promises the best result. Slightly worse is the standard k-NN classifier based on the features 1, 2 and 4. The final classifier should be rather based on the whole data set X_1. Let us consider two situations: the first one with k=1 and all 5 features for k-NN classifier and the second one with k=1 for all component classifiers and 5 features in the parallel net of two-decision classifiers. The *leave one out* method applied to the whole data set X_1 estimates the same error rate equal 0.0308 for both types of classifiers. These two classifiers offer also the same error rate, equal 0.0403, in case of 3 features: 1, 2 and 4. The *leave one out* method, used for feature selection with X_1 as the reference set, has chosen the 4 features 1, 2, 3 and 4 for the standard classifier and all features for the parallel one. Different feature sets were selected for each of the component classifiers. All features selected for the component classifiers, taken together, are treated as a feature set selected the parallel net of the classifiers.

All experiments presented above suggest that the most reasonable decision could be the use of the standard 1-NN classifier based on all 5 features and the whole data set X_1 as the reference set. The error rate associated with this classifier, equal 0.0308, may be a little too optimistic although the value of k as well as the feature set were fixed before the *leave one out* application. This phenomenon is a property of the applied method for misclassification rate estimation. A serious problem arises with evaluating the expected error rate for the k-NN classifier. Which of the two misclassification rate values is closer to the true one, the mean one equal 0.0631 or the one equal 0.0308?

Problem 2: It is difficult to find the value of the expected error rate, i.e., which estimate is more precise, the one obtained as a mean from several experiments with the testing set method or the one calculated by the leave one out method with the use of the whole data sets?

3.3. BREATHING WITH THE HYPERCAPNIC STIMULUS

The similar experiments have been done for the data set X_2, which concern the hypercapnic stimulus. This time, the parallel classifier *k*-NN outperformed the standard *k*-NN one in a sense of the mean error rate and for all analyzed feature sets as it was shown in the Tab. 3.

The features 1, 2 and 4 were also analysed separately. Other 3 features 2, 3 and 4, that were most frequently chosen while performing feature selection, were additionally studied too. For the same reason as it took place in case of the set X_1, all error rates shown in the Tab. 3 are also too optimistic, mainly the ones derived after feature selection. So, it is reasonable to treat more seriously the misclassification rates obtained by the testing set method.

Std. *k*-NN, no feat. sel.	**Std. *k*-NN, with feat. sel.**	**Par. *k*-NN, no feat. sel.**	**Par. *k*-NN, with feat. sel.**	**Std. *k*-NN, feat. 1,2,4**	**Par. *k*-NN, feat. 1,2,4**	**Std. *k*-NN, feat. 2,3,4**	**Par. *k*-NN, feat. 2,3,4**
1	2	3	4	5	6	7	8
0.0632	0.0457	0.0564	0.0301*	0.0723	0.0654	0.0789	0.0444
In the standard classifier the features 1,**2,3,4**,5 appeared 6, **18**, **19**, **15** and 6 times respectively							
In the parallel classifier the features 1,2,3,4,5 appeared 20, 20, 17, 20 and 17 times respectively							

Tab.3. Mean error rates and results of feature selection obtained by *leave one out* method for the sets $R_1,R_2,\ldots,R_{20}$ randomly chosen from the data set X_2

In this task the feature selection significantly improved the results for both classifier types, see columns 2 and 1, and next the columns 4 and 3, in the Tab. 4. Also in this case the features selected for each of the component classifiers were different.

Basic charact.	**Std. *k*-NN, no f. sel.**	**Std. *k*-NN, with f. sel.**	**Par. *k*-NN, no f. sel.**	**Par. *k*-NN, with f. sel.**	**Std. *k*-NN, feat. 1,2,4**	**Par. *k*-NN, feat. 1,2,4**	**Std. *k*-NN, feat. 2,3,4**	**Par. *k*-NN, feat. 2,3,4**
0	1	2	3	4	5	6	7	8
minimum	0.0385	0.0321	0.0385	0.0385	0.0513	0.0577	0.0321	0.0321
Mean	**0.0770**	**0.0673**	**0.0773**	**0.0722**	**0.0935**	**0.0867**	**0.0541***	**0.0557**
maximum	0.1538	0.1667	0.1410	0.1569	0.1667	0.1410	0.0897	0.0897
std. dev.	0.0280	0.0296	0.0225	0.0326	0.0296	0.0187	0.0168	0.0170

Tab.4. Error rates computed by the use of the testing set method for the data X_2

The results for three basic features 1, 2, and 4 were remarkable worse than those received for all 5 features. However, the triplet of the features 2, 3 and 4 offers the best performance, see the columns 7 and 8, the smallest mean values and the smallest standard deviations.

Finally, it is interesting to know the values of the error rates for both types of classifiers, without feature selection and k=1. In this case no classifier parameters are fitted to receive the minimum misclassification rate. The use of 5 features offered the error rates equal 0.0647 and 0.0518 for the standard and the parallel classifier respectively. If 3 features 1, 2 and 4 were used, the standard *k*-NN classifier offered the error rate was equal 0.0777 while for the parallel classifier 0.0637. The features 2, 3 and 4 formed again the most attractive feature set and the error rates for the standard and the parallel classifier equaled <u>0.0550</u> and <u>0.0485</u>, respectively. Now, we have the situation when each of the two error rate evaluation methods indicate the different type of the classifier as the best one. The feature selection based on the whole data set and the *leave one out*

method points to 3 features: 2, 3 and 5 for the standard classifier and all 5 features for the parallel one. Results of feature selection were different for each of the component classifiers.
Problem 3: How one can choose the better classifier type if the different misclassification probability estimation methods favor different classifiers?

3.4. BREATHING WITH THE HYPOXIC STIMULUS

As in the previous two data sets also in case of the set X_3 the error rates found by the *leave one out* method gave more optimistic results than the testing set approach, what can be seen by comparing the results gathered in the Tab. 5 and 6. In this case 3 features, 1, 3 and 5, were chosen as the ones which were most frequently selected during all 20 experiments. These features promise the lowest error rate according to the results shown in the Tab. 5.

Analyzing the results received for the data set X_2 and X_3, we can notice that the feature selection based on the series of experiments can be very successful.

Std. *k*-NN, no feat. sel.	**Std. *k*-NN, with feat. sel.**	**Par. *k*-NN, no feat. sel.**	**Par. *k*-NN, with feat. sel.**	**Std. *k*-NN, feat. 1,2,4**	**Par. *k*-NN, feat. 1,2,4**	**Std. *k*-NN, feat. 1,3,5**	**Par. *k*-NN, feat. 1,3,5**
1	2	3	4	5	6	7	8
0.1103	0.0951	0.1066	0.0803	0.1370	0.1318	0.1064*	0.1007
In the standard classifier the features **1**,2,**3**,4,**5** appeared **13**, 11, **19,** 9 and **15** times respectively							
In the parallel classifier the features **1**,2,**3**,4,**5** appeared **20**, 18, **20,** 17 and **19** times respectively							

Tab.5. Mean error rates and results of feature selection obtained by *leave one out* method for the sets $R_1,R_2,\ldots,R_{20}$ randomly chosen from the data set X_3

The standard 1-NN classifier and the parallel one with k=1 for all component classifiers, with X_3 as the reference set and all 5 features, offered the error rate equal 0.0936 and 0.0517, respectively. Of course, these both error rates were estimated by the *leave one out* method. For the features 1, 2 and 4 the errors were equal 0.1108 and 0.0665, and for the features 1, 3 and 5 they equaled 0.0813 and 0.0419 for the standard and the parallel classifiers respectively. The values of k were equal 1 as for the features 1, 2 and 4 as well as for the features 1, 3 and 5.

Basic charact.	**Std. *k*-NN, no f. sel.**	**Std. *k*-NN, with f. sel.**	**Par. *k*-NN, no f. sel.**	**Par. *k*-NN, with f. sel.**	**Std. *k*-NN, f. 1,2,4**	**Par. *k*-NN, feat. 1,2,4**	**Std. *k*-NN, feat. 1,3,5**	**Par. *k*-NN, feat. 1,3,5**
0	1	2	3	4	5	6	7	8
min.	0.0931	0.0980	0.0882	0.0842	0.1029	0.1275	0.0784	0.0792
Mean	**0.1309**	**0.1217**	**0.1390**	**0.1256**	**0.1579**	**0.1660**	**0.1165**	**0.1298**
max.	0.1589	0.1634	0.1912	0.1618	0.2304	0.2304	0.1569	0.1832
std. dev.	0.0189	0.0188	0.0246	0.0190	0.0306	0.0250	0.0181	0.0252

Tab.6. Error rates computed by the use of the testing set method for the data X_3

It is very unusual that for the whole reference set X_3, feature selection resulted also in the 3 features 1, 3 and 5. Furthermore, for each of the component classifiers all these 3 features were chosen. The next problem arises. This time the situation with the result of feature selection was very comfortable.

Problem 4: What should we do if the scheme with the multiple experiments and the testing set method offers different features than the feature selection based on the whole data set?

4. CONCLUSIONS

All 4 problems mentioned in the previous sections appear when data sets are not sufficiently large. Otherwise, both methods, the *leave one out* method and the testing set method, would give similar estimates of the misclassification probability for both kinds of the classifier. This results from the fact that both considered types of classifier converge to the Bayes classifier.

Three other problems were omitted in our considerations. First one concerns the way of error rate calculation. We have presented the practical use of the *leave one out* method and the testing set method. It is easy to notice that both these methods can be combined. The data set *X* can be divided into *n* parts treating the *n*-1 of them as the reference set and the remaining one as the testing set. Thus, *n* such experiments are recommended.

Problem 5: How can we determine the number n to save the best precision of error rate computation?

It is desired to divide the data set into such *n* parts that in each of these parts the density of the probability distribution would be similar. Thus, the sequential problem has been formulated.

Problem 6: How to divide the data set into *n* parts to receive the best precision of the misclassification probability estimation?

Maybe it would be better not to divide the data set into *n* parts, but to make several experiments dividing the data sets only into two parts?

Problem 7: What is the best proportion of the reference set size to the testing set size?

The authors intend to deal with all the mentioned problems in their future research work.

BIBLIOGRAPHY

[1] DEVIJVER, P.A., KITTLER, J., Pattern recognition: a statistical approach, Prentice Hall International, London, pp.356-357, 1982.

[2] DUDA R., HART P., Pattern classification and scene analysis, Wiley Interscience, New York, 1973.

[3] EPSTEIN S.K., An overview of respiratory muscle function, Clinics Chest Medicine, Vol.15, No.4, pp.619-639, 1994.

[4] FIX E., HODGES J. L., Discriminatory analysis: nonparametric discrimination small sample performance, project 21-49-004, report number 11, USAF school of aviation medicine, randolph field, Texas, pp. 280-322, 1952, reprinted in the book: Dasarathy B. V., NN pattern classification techniques, IEEE Computer Society Press, pp. 40-56, 1991.

[5] JÓŹWIK A., Eksperymentalne badanie zależności pomiędzy prognozowaną jakością klasyfikacji a wielkością zbioru uczącego i liczbą cech, Materiały IX Krajowej Konferencji Biocybernetyki i Inżynierii Biomedycznej, Warszawa, pp.28-30, 2001.

[6] JÓŹWIK A., SOKOŁOWSKA B., BUDZIŃSKA K., An example of computer aided decision-making system for recognition of respiration pathology, Medical Informatics and Technologies, Ustroń , pp. MI42-MI48, 2001.

[7] JÓŹWIK A., VERNAZZA G., Recognition of leucocytes by a parallel k-NN classifiers, Lecture Notes of ICB Seminar, Warsaw, pp.138-153, 1988.

[8] KATAGIRI M., YOUNG R.N., PLATT R.S.KIESER T.M., EASTON P.A., Respiratory muscle compensation for unilateral or bilateral hemidiaphragm paralysis in awake canines, Journal of Applied Physiology, Vol.77, No.4, pp.1972-1981, 1994.

[9] POLE D.C., SEXTON W.L., FARKAS G.A., POWERS S.K.,REID M.B., Diaphragm structure and function in health and disease, Medicine & Science in Sports & Exercise, Vol. 29, No.6, pp.738-754, 1997.

[10] ROCCO, P.R.M., FAFFE, D.S., FEIJOO, M., MENEZES, S.L., VASCONCELLOS, F.P., ZIN, W.A., Effects of uni- and bilateral phrenicotomy on active and passive respiratory mechanics in rats, Respiration Physiological. Vol. 110, pp. 9-18, 1997.

[11] SIECK, G.C., Diaphragm muscle: structural and functional organization, Clinics in Chest Medicine, Vol.9, No.2, pp.195-210, 1988.

[12] SIEDLECKI W., A formula for multi-class distributed classifiers, Pattern Recognition Letters, Vol. 15, pp. 739-742, 1994.

[13] STAWSKA Z., JÓŹWIK A., SOKOŁOWSKA B., BUDZIŃSKA K., A multistage classifier based on distance measure and its use for detection of respiration pathology, Komputerowe Systemy Rozpoznawania (KOSYR2001), Wrocław, pp. 67-71, 2001.