Applied Computer Science, vol. 18, no. 2, pp. 71–85
doi: 10.35784/acs-2022-14

Submitted: 2022-05-31 | Revised: 2022-06-23 | Accepted: 2022-06-27

Keywords: acoustic emission, machine learning, osteoarthritis, knee joint, kinetic chain

Robert KARPIŃSKI [0000-0003-4063-8503]*, **

KNEE JOINT OSTEOARTHRITIS DIAGNOSIS BASED ON SELECTED ACOUSTIC SIGNAL DISCRIMINANTS USING MACHINE LEARNING

Abstract

This paper presents the results of a preliminary study on simplified diagnosis of osteoarthritis of the knee joint based on generated vibroacoustic processes. The analysis was based on acoustic signals recorded in a group of 50 people, half of whom were healthy, and the other half – people with previously confirmed degenerative changes. Selected discriminants of the signals were determined and statistical analysis was performed to allow selection of optimal discriminants used at a later stage as input to the classifier. The best results of classification using artificial neural networks (ANN) of RBF (Radial Basis Function) and MLP (Multilevel Perceptron) types are presented. For the problem involving the classification of cases into one of two groups HC (Healthy Control) and OA (Osteoarthritis) an accuracy of 0.9 was obtained, with a sensitivity of 0.885 and a specificity of 0.917. It is shown that vibroacoustic diagnostics has great potential in the non-invasive assessment of damage to joint structures of the knee.

1. INTRODUCTION

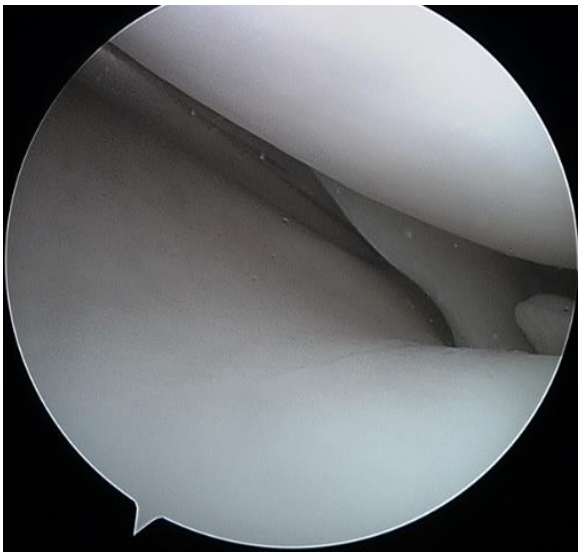
The knee joint is the largest and one of the most complex joints in the human body. During everyday activities, it is exposed to significant loads, often several times the weight of the human body, which makes it extremely vulnerable to injury and the possibility of osteoarthritis (Arendt, Miller & Block, 2014; Będziński, 1997; Krakowski, Karpiński, Maciejewski & Jonak, 2021). It is one of the most susceptible to the loss of articular cartilage and, as a result, to the development of degenerative changes in the joints of the human body (Cross et al., 2014; Krakowski, Gerkowicz, et al., 2019; Kyu et al., 2018). Osteoarthritis is a clinical syndrome characterized by progressive degeneration of articular cartilage that may involve the entire joint, including synovial membranes and fragments of subchondral bone (Befrui et al., 2018; Krakowski, Karpiński, Jonak & Maciejewski, 2021; Krishnan et al., 2000). It is now one of the most common causes of pain and disability in middle-aged and elderly people and a major cause of impairment of activities of daily living (Hunter & Bierma-

* Department of Machine Design and Mechatronics, Faculty of Mechanical Engineering, Lublin University of Technology, Lublin, Poland, r.karpinski@pollub.pl
** Department of Psychiatry, Psychotherapy, and Early Intervention, Medical University of Lublin, Lublin, Poland, robertkarpinski@umlub.pl

Zeinstra, 2019; McDonough & Jette, 2010). The incidence of osteoarthritis is influenced by a number of factors such as work, sports, musculoskeletal injuries, obesity and gender (Felson, 2004; Johnson & Hunter, 2014; Krakowski, Karpiński, Maciejewski, Jonak & Jurkiewicz, 2020; Krakowski, Karpiński, Jojczuk, Nogalska & Jonak, 2021; Reyes et al., 2015).

Non-invasive diagnosis of knee osteoarthritis is most often carried out based on physical examination and typical imaging techniques such as radiography (X-ray) (Hayashi, Roemer & Guermazi, 2019; Richette & Latourte, 2019), computed tomography (CT) (Ahn & El-Khoury, 2006), magnetic resonance imaging (MRI) (Krakowski, Nogalski, Jurkiewicz, Karpiński, Maciejewski & Jonak, 2019) and also ultrasound (US) (Mathiessen et al., 2016; Möller et al., 2008). Invasive methods include arthroscopy of the joint during which intraoperative assessment is performed, but this procedure carries a risk of complications as with any surgery. The main limitations of diagnostic imaging for the assessment of osteo-arthritis undoubtedly include the low sensitivity of detecting lesions at an early stage of their (Krakowski, Nogalski, et al., 2019) and the limited accessibility especially for patients living in smaller towns. These methods are also associated with high costs related to the use of specialized equipment and the need to employ highly qualified staff. Considering the above facts, it seems extremely important to develop an inexpensive, non-invasive method of assessing damage to joint structures, allowing diagnostics to be performed in the orthopedic office during routine examinations.

An alternative to the currently used methods may be vibroacoustic diagnostics. This method is widely used in non-invasive diagnostics of machines where it gives very good results (Figlus, Kozioł & Kuczyński, 2019; Jedliński et al., 2015; Jedliński & Jonak, 2015). For the diagnosis of the knee joint, this method was first proposed by Blodgett in 1902 (Blodgett, 1902). Depending on the efficiency of lubrication, the joint systems generate vibrations and noises. The same is true for the knee joint and other joints in the human body. Changes in the mechanical properties and surface structure of the articular cartilage, such as the appearance of bumps, cracks or cartilage defects in the successive stages of degenerative changes, influence the vibroacoustic signals recorded during movement of the knee joint (Karpiński, Machrowska & Maciejewski, 2019; Prior et al., 2010; van den Borne et al., 2007). Fig. 1 shows images taken during arthroscopy of the knee showing healthy (a) and degenerated (b) articular cartilage. Analysis of the differences in the characteristic parameters of the recorded signals may allow for a preliminary assessment of the condition of the joint and allow for accelerated diagnosis and selection of an appropriate therapy. In addition, this method provides information about the mutual cooperation of intra- and extra-articular structures of the moving joint, in most other cases we obtain information about the joint in a static position (Kernohan et al., 1990; Rangayyan & Wu, 2008; Walters, 1929; Wu, 2015). The application of machine learning methods can be extremely useful in the analysis of issues related to non-invasive diagnosis of articular cartilage damage based on vibroacoustic signals.



(a)

(b)

Fig. 1. Arthroscopic images of healthy (a) and degenerated (b) articular cartilage

The purpose of this study is to evaluate the usefulness of selected discriminants determined for acoustic signals generated in the knee joint, recorded in the control group (healthy subjects) and the study group (confirmation of degenerative changes), in the diagnosis of cartilage damage of the knee joint using machine learning.

2. MATERIALS AND METHODS

2.1. Participants and study protocol

The study was conducted on a group of 50 volunteers, half of whom were healthy people (HC group) and the other half, people previously qualified by an orthopedic surgeon, for surgery including arthroscopy or total knee replacement due to the presence of degenerative changes in the knee joint. All surgeries were performed by a surgeon specializing in TKR or arthroscopy according to a standard protocol depending on the type of surgery. During surgery, the extent of cartilage damage and the location of cartilage damage were accurately assessed. The International Cartilage Repair Society (ICRS) scale was used to assess cartilage (Brittberg & Winalski, 2003; Cameron, Briggs & Steadman, 2003). The control group, on the other hand, consisted of people without any complaints occurring in the knee joints. Signal recording in this group was carried out in laboratory conditions in the premises of the Lublin University of Technology. In the case of the study group, the recording of signals took place in the conditions of the orthopaedic department on the day of admission to hospital preceding the surgery. The examination in both groups was preceded by physical tests and a questionnaire examination, allowing to determine the condition of the joint. The protocol was identical in both groups and was conducted by a specialist in orthopedics and traumatology of the musculoskeletal system. In the control group, the signal recorded for one randomly selected limb was analyzed, while in the case of subjects in the study group, the signal acquired from the limb qualified for surgery was analyzed. The whole procedure was conducted in accordance with the principles of good clinical practice and was approved by the bioethics committee of the Medical University of Lublin (consent no. KE-0254/261/2019). Details of the groups are summarized in Table 1.

Tab. 1. Characteristics of study participants

Study group	N	Males/ Females	Age (years ± SD)	Heigh (m ± SD)	Weight (kg ± SD)	BMI	Tegner-Lyshom score
Healthy control (HC)	25	9/16	25.04 ± 5.76	1.72 ± 0.09	68.00 ± 15.33	22.69 ± 2.98	100 ± 0.0
Osteoarthritis (OA)	25	9/16	61.40 ± 9.11	1.68 ± 0.08	90.84 ± 14.36	32.38 ± 4.82	34.00 ± 9.54

To record acoustic signals the author's measurement system was used, consisting of a microcontroller with peripheral devices for signal acquisition in the form of a CM01b piezoelectric microphone, a rotary encoder installed on a typical knee orthosis and a computer with software for signal recording. The signals were recorded with the microphone, while the encoder was used to measure the flexion angle and limb position. The sampling frequency was 1400 Hz and the resolution was 10 bits. Details of the measurement system are described in previous work (Karpiński, Machrowska & Maciejewski, 2019; Karpiński et al., 2022b). Signals were recorded for a closed kinetic chain (CKC) where standing up from a sitting position to full upright was realized. A movement in the range of 90° – 0° – 90° was considered as one cycle. The cycle duration was approximately 2s and depended on the individual predisposition of the patient. The measurement was conducted for ten complete cycles. The signals obtained in this way contained some artifacts including mainly the

registration of time series before and after cycles of movements, random noise, and electrical network disturbances. In order to remove these disturbances, the signals were subjected to a preprocessing procedure involving the removal of insignificant fragments of the time series before and after the analyzed movements. For this purpose, a procedure was used to semi-automatically remove irrelevant signal samples from the encoder data. A trend line was also removed from the signals using the EEMD procedure, which is described in more detail in the papers (Huang et al., 1998; Jonak et al., 2019; Karpiński et al., 2022b; Machrowska & Jonak, 2018; Wu, 2015). For the signals prepared in this way, selected discriminants were determined.

2.2. Feature extraction

In diagnostic practice it is very important to use simple discriminants of vibration acoustic signals, which are easy to determine and allow assessment of the type of damage present (Madej, Czech & Konieczny, 2003). An observable quantity associated with a technical condition may be classified in the group of diagnostic symptoms, provided that the relationships describing this phenomenon are determined (Stanik et al., 2013). The changes occurring in the observed signal in a unit of time take many instantaneous values, which often, due to the dynamics of changes, makes it impossible to observe each change of the instantaneous value in the analysis (Cempel, 2005; Stanik et al., 2013). However, it is important to define the values of signals describing the characteristic points of greatest importance for the overall signal analyzed. In this study, the following signal discriminants were determined for the analyzed acoustic signals:

1. Straightened average value (SA):

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{1}$$

where: x_i – the value of the discrete signal at the nth point, $n = 1, \dots, N$,
N – number of samples in the signal.

2. Root mean square (RMS):

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} {x_i}^2} \tag{2}$$

3. Peak value (PV):

$$\hat{x} = max|x_i| \tag{3}$$

4. Peak to peak value (PPV):

$$x_{PPV} = |x_{max} - x_{min}| \tag{4}$$

5. Crest Factor (CF):

$$x_{CF} = \frac{\hat{x}}{x_{RMS}} \tag{5}$$

6. Impact Factor (IF):

$$x_I = \frac{\hat{x}}{\bar{x}} \tag{6}$$

7. Shape factor (SF):

$$x_{SF} = \frac{x_{RMS}}{\bar{x}} \tag{7}$$

8. Variance (VAR):

$$x_{VAR}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{8}$$

9. Kurtosis (KUR):

$$x_{KUR} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^4}{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2} \tag{9}$$

A detailed description of the designated discriminants was presented in the author's previous work (Karpiński et al., 2021a, 2021b, 2022a, 2022b) and other studies (Jedliński & Jonak, 2020; Stanik i in., 2013).

2.3. Statistical Analysis

Statistical analysis to determine whether there were statistically significant differences between the determined signal discriminants for signals recorded in the HC and OA groups was performed using Statistica 13.3. The significance level was taken as $\alpha = 0.05$.

The first step in the analysis was to test whether the values of the individual discriminants of the acoustic signals in the both groups have a normal distribution. For this purpose, three tests were performed: Kolmogorov-Smirnov test, Lilliefors test and Shapiro-Wilk test (Karpiński, Szabelski & Maksymiuk, 2019a; Rabiej, 2018; Szabelski, 2018).

Analysis of equality of variance was performed using three tests: F (Fisher's), Levene's, and Brown and Forsyth. For results characterized by normality of distribution and equality of variance – Student's t-test was used to analyze the equality of the resulting means of individual signal discriminants, at the assumed level of significance. For the results characterized by normality of distribution but lack of equality of variance – to analyze the equality of mean values of signal discriminants we used Student's t-test with separate variance estimation (Cochran-Cox test) (Karpiński, Szabelski & Maksymiuk, 2019b; Rabiej, 2018). For discriminants that did not have a normal distribution, the Mann-Whitney U test with a continuity correction was used. This correction is used to ensure that the test statistic can accommodate all real number values as assumed by the normal distribution. Discriminants showing statistically significant differences were determined from the analyses and were used as inputs to the neural classifiers.

2.4. Machine learning

Machine learning (ML) is a subset of artificial intelligence (AI). It is an area devoted to algorithms whose performance improves automatically based on collected experience, or exposure to data. A subset of machine learning is deep learning (DL), within which are artificial neural networks (ANNs) (Machrowska, Karpiński, et al., 2020; Machrowska,

Szabelski, et al., 2020). A neural network is a collection of appropriately connected neurons arranged in layers. Artificial neural networks are one type of highly parameterized statistical models. They have been developed based on information about the functioning of the nervous system of living organisms and are capable of mapping extremely complex functions (Bauer, Stütz & Kley, 2021; Dudek-Dyduch, Tadeusiewicz & Horzyk, 2009; Tadeusiewicz, 1993). Technically, an artificial neuron is an element whose properties correspond to selected properties of its biological counterpart. Each neuron has at least one input and one output. The signals at the inputs are multiplied by factors referred to as synaptic weights. Each neuron calculates a weighted sum of its inputs, which are then summed. The activation level determined in this way becomes the argument of the transition function (activation function), which calculates the output of the neuron. The values of weights can be changed, which allows the network to learn and adapt to the problem being solved. ANNs find application in solving problems related to data processing and analysis, prediction and classification especially when the analyzed issues involve poorly known phenomena and processes (Badurowicz, 2022; Kosicka, Krzyzak, Dorobek & Borowiec, 2022; Rogala, 2020; Szabelski, Karpiński & Machrowska, 2022).

In the analyzed problem classification was carried out using the Statistica 13.3 package (Tulsa, OK, USA), which includes modules covering machine learning and artificial neural networks. A built-in algorithm was used to automatically search for multilevel perceptron (MLP) and radial basis function (RBF) neural networks with the best possible parameters. Acoustic signal discriminants that had statistically significant differences between the study groups in the statistical analysis were used as input data. The data were randomly divided 70% for training, 15% for testing and 15% for validation. The outputs of the neural networks, on the other hand, proposed a simplified partitioning that allowed assigning the cases to one of the two groups OA and HC. Results for sets of the three best MLP and RBF networks are presented. The parameters for automatic design of MLP and RBF neural networks in Statistica software are shown in Figure 2.

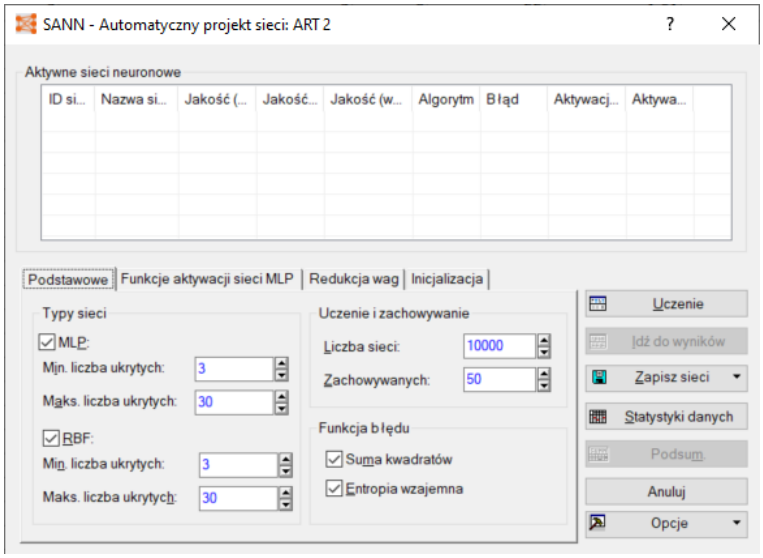


Fig. 2. The parameters for automatic design of MLP and RBF neural networks

3. RESULTS

3.1. Statistical analysis

Checking whether the values of individual discriminants of acoustic signals have normal distribution was carried out using Kolmogorov-Smirnov test, Lilliefors test and Shapiro-Wilk test. The results of the performed analyses are summarized in Table 2. The following hypotheses were assumed in the analysis:

H_0 – Variables within the analyzed groups have normal distribution.

H_1 – Variables within the analyzed groups do not have normal distribution.

Tab. 2. Results of statistical analysis to test the normality of distribution

Discriminant	Study group	N	maks D	K-S	Lillief.	W	p
Straightened average value (SA)	OA	25	0.168	p > .20	p < .10	0.932	0.098
Root mean square (RMS)	OA	25	0.159	p > .20	p < .15	0.942	0.165
Peak value (PV)	OA	25	0.266	p < .10	**p < .01**	**0.773**	**0.000**
Peak to peak value (PPV)	OA	25	0.192	p > .20	**p < .05**	**0.838**	**0.001**
Crest Factor (CF)	OA	25	0.275	**p < .05**	**p < .01**	**0.804**	**0.000**
Impact Factor (IF)	OA	25	0.226	p < .15	**p < .01**	**0.753**	**0.000**
Shape factor (SF)	OA	25	0.314	**p < .05**	**p < .01**	**0.569**	**0.000**
Variance (VAR)	OA	25	0.165	p > .20	p < .15	0.934	0.105
Kurtosis (KUR)	OA	25	0.214	p < .20	**p < .01**	**0.889**	**0.010**
Straightened average value (SA)	HC	25	0.104	p > .20	p > .20	0.986	0.973
Root mean square (RMS)	HC	25	0.134	p > .20	p > .20	0.961	0.428
Peak value (PV)	HC	25	0.235	p < .15	**p < .01**	**0.789**	**0.000**
Peak to peak value (PPV)	HC	25	0.266	p < .10	**p < .01**	**0.524**	**0.000**
Crest Factor (CF)	HC	25	0.190	p > .20	**p < .05**	**0.782**	**0.000**
Impact Factor (IF)	HC	25	0.309	**p < .05**	**p < .01**	**0.522**	**0.000**
Shape factor (SF)	HC	25	0.326	**p < .01**	**p < .01**	**0.454**	**0.000**
Variance (VAR)	HC	25	0.110	p > .20	p > .20	0.964	0.500
Kurtosis (KUR)	HC	25	0.166	p > .20	p < .10	**0.813**	**0.000**

On the basis of the analyses performed, it can be concluded that in the case of discriminants such as Straightened average value (SA), Root mean square (RMS) and Variance (VAR), there are no grounds to reject the hypothesis of normality of distribution, because for these variables for the W Shapiro-Wilk test obtained levels of $p > \alpha = 0.05$. In the case of other discriminants, the results obtained give grounds to reject the hypothesis of normality of distribution and assume that these discriminants do not have normal distribution. The next step in the analyses conducted for discriminants having a normal distribution was to conduct an analysis of equality of variance using three tests: F (Fisher's), Levene, and Brown and Forsyth. Depending on the results of equality of variance analysis, the analysis of equality of means was conducted using Student's t-test (in case of equality of variance) and Student's t-test with separate variance estimation (in case of non-equality of variance). The results of the analyses of equality of variance and equality of means are presented in Table 3.

Tab. 3. Results of Student's t-test and Student's t-test with the separate estimation of variance

Discriminant	Mean OA	Mean HC	t	p	t sep.	p	Standard deviation	Standard deviation	F quotient	p	Levene's	p	Brn-Fors	p
Straightened average value (SA)	119.96	224.50	-5.66	0.00	-5.66	0.000	65.72	64.99	1.02	0.957	0.133	0.717	0.118	0.733
Root mean square (RMS)	157.24	263.53	-5.04	0.00	-5.04	0.000	74.75	74.27	1.01	0.975	0.120	0.731	0.062	0.804
Variance (VAR)	25919.25	59888.86	-6.19	0.00	-6.19	0.000	17297.74	21301.67	1.52	0.314	0.743	0.393	0.646	0.426

The results obtained show that the mean values for individual discriminants have statistically significant differences between HC and OA groups. In the analyzed case, all discriminants were characterized by equality of variance, so the results of Student's t-test were taken into account in the analysis of equality of means.

For discriminants that did not have a normal distribution, the analyses used the Mann-Whitney U test with continuity correction. The results obtained are presented in Table 4. On the basis of Mann-Whitney U-test it can be stated that in case of discriminants such as Peak value (PV) ($p = 0.004$), Peak to peak value (PPV) ($p = 0.008$), Crest Factor (CF) ($p = 0.000$) and Kurtosis (KUR) ($p = 0.000$) between the tested groups there are statistically significant differences on the assumed significance level. These differences were not found for discriminants such as Impact Factor (IF) ($p = 0.535$) and Shape factor (SF) ($p = 0.561$), due to the lack of statistically significant differences these discriminants were not included in further considerations.

Tab. 4. Mann-Whitney U test results

Discriminant	Sum of ranks	Sum of ranks	U	Z	p	2*1str.
Peak value (PV)	488.00	787.00	163.00	**-2.89**	**0.004**	**0.003**
Peak to peak value (PPV)	501.00	774.00	176.00	**-2.64**	**0.008**	**0.008**
Crest Factor (CF)	873.00	402.00	77.00	**4.56**	**0.000**	**0.000**
Impact Factor (IF)	605.00	670.00	280.00	-0.62	0.535	0.538
Shape factor (SF)	607.00	668.00	282.00	-0.58	0.561	0.564
Kurtosis (KUR)	899.00	376.00	51.00	**5.06**	**0.000**	**0.000**

3.2. Classification

The results for the most accurate classifiers proposed by the automatic neural network selection algorithm of the Statistica package are presented below, three cases each for multilayer perceptron (MLP) and radial basis function (RBF) networks. Seven discriminants having statistically significant differences between the analyzed groups were used as input data, viz: Straightened average value (SA), Root mean square (RMS), Variance (VAR), Peak

value (PV), Peak to peak value (PPV), Crest Factor (CF), Kurtosis (KUR). The detailed results of learning, testing and validation accuracy for each network in all analyzed variants are shown in Table 5.

Tab. 5. Accuracy of learning, testing and validation for selected MLP and RBF neural networks

Network name	Accuracy (learning) %	Accuracy (testing) %	Accuracy (validation) %	Learning algorithm	Error function	Activation (hidden)	Activation (output)
MLP 7-22-2	88.89	100.00	85.71	BFGS 10	Entropy	Tanh	Softmax
MLP 7-10-2	88.89	100.00	85.71	BFGS 14	Entropy	Tanh	Softmax
MLP 7-7-2	83.33	100.00	85.71	BFGS 14	Entropy	Exponential	Softmax
RBF 7-9-2	88.89	85.71	85.71	RBFT	Entropy	Gauss	Softmax
RBF 7-10-2	86.11	85.71	85.71	RBFT	Entropy	Gauss	Softmax
RBF 7-5-2	83.33	85.71	85.71	RBFT	Entropy	Gauss	Softmax

The highest learning accuracy of 88.89% was observed for MLP 7-22-2, MLP 7-10-2 and RBF 7-9-2, while the lowest accuracy of 83.33% was observed for MLP 7-7-2 and RBF 7-5-2, respectively. In the case of testing accuracy, all MLP networks scored 85.71% while RBF networks scored 85.71%. The validation accuracy was identical for all the networks presented and was 85.71% respectively.

Information on the classification results for each group and all cases is shown in Table 6. The highest classification accuracy of 90% was obtained for MLP-type networks having 10 and 22 neurons in the hidden layer, respectively. For the best RBF-type network, the obtained classification accuracy was slightly lower than that of MLP-type networks at 88%. The obtained results show that both networks of multilayer perceptron type and networks with radial basis functions perform well in the analyzed problem and give comparable results of classification accuracy.

Tab. 6. Summary of classification accuracy of MLP and RBF

Network name		HC	OA	Total
MLP 7-22-2	Total	25.00	25.00	50.00
	Correct	23.00	22.00	45.00
	Correct (%)	92.00	88.00	90.00
MLP 7-10-2	Total	25.00	25.00	50.00
	Correct	22.00	23.00	45.00
	Correct (%)	88.00	92.00	90.00
MLP 7-7-2	Total	25.00	25.00	50.00
	Correct	21.00	22.00	43.00
	Correct (%)	84.00	88.00	86.00
RBF 7-9-2	Total	25.00	25.00	50.00
	Correct	23.00	21.00	44.00
	Correct (%)	92.00	84.00	88.00
RBF 7-10-2	Total	25.00	25.00	50.00
	Correct	21.00	22.00	43.00
	Correct (%)	84.00	88.00	86.00
RBF 7-5-2	Total	25.00	25.00	50.00
	Correct	21.00	21.00	42.00
	Correct (%)	84.00	84.00	84.00

There are many tools to assess the validity of classifiers among others are ROC curves, and metrics such as F1-score and Matthews correlation coefficient (MCC). ROC curves are a visualization of the relationship between the effectiveness of positive classifiers (sensitivity) and the ineffectiveness of negative case classification (1-specificity) at each probability level. A detailed summary of the performance of each classifier is shown in Table 7. The highest AUC value of 0.941 was obtained for the MLP 7-10-2 network with a sensitivity of 0.917 and a specificity of 0.885. For the RBF-type network, the highest value of 0.931 was obtained for the network with 9 neurons in the hidden layer with a sensitivity of 0.852 and a specificity of 0.913. A summary of the ROC curves for all the proposed classifiers is shown in Figure 3.

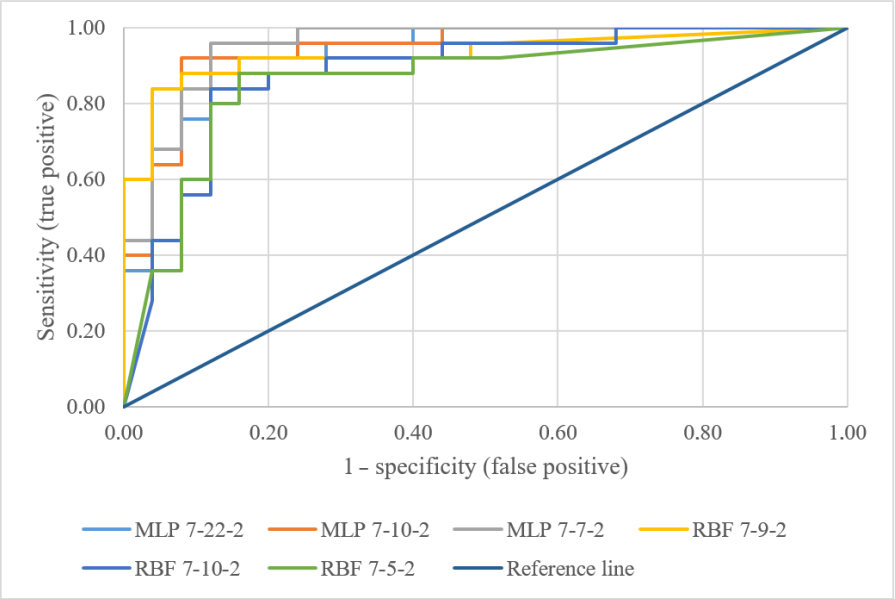


Fig. 3. Comparison of ROC curves for selected classifiers.

Tab. 7. Performance comparison of selected MLP and RBF classifiers

Network name	Sensitivity	Specificity	AUC	ROC Treshold	Accuracy	Precision	Recall	F1 score	MCC
MLP 7-22-2	0.885	0.917	0.933	0.564	0.900	0.920	0.885	0.902	0.801
MLP 7-10-2	0.917	0.885	0.941	0.485	0.900	0.880	0.917	0.898	0.801
MLP 7-7-2	0.875	0.846	0.954	0.490	0.860	0.840	0.875	0.857	0.721
RBF 7-9-2	0.852	0.913	0.931	0.563	0.880	0.920	0.852	0.885	0.762
RBF 7-10-2	0.875	0.846	0.881	0.285	0.860	0.840	0.875	0.857	0.721
RBF 7-5-2	0.840	0.840	0.860	0.556	0.840	0.840	0.840	0.840	0.680

In order to further compare the proposed classifiers, measures such as F1 score and Matthews correlation coefficient (MCC) were determined. F1 score is a measure of model precision on a dataset and it is most commonly used in evaluating binary classification systems. It is a way to combine model precision and recall. It is defined as the harmonic mean of model precision and recall (Chicco & Jurman, 2020). In contrast, the Matthews

correlation coefficient is one of the most informative single parameters for determining the prediction quality of a binary classifier in the context of a confusion matrix. It can be used even in cases where two classes have very different counts (Matthews, 1975; Powers, 2020).

The highest F1 score of 0.902 was obtained for the MLP-type network with twenty-two neurons in the hidden layer at a precision of 0.920 and recall of 0.885, while the lowest value of 0.840 was obtained for the RBF-type network with five neurons in the hidden layer at a precision and recall of 0.840. In the case of MCC, the highest value of 0.801 was obtained for the MLP-type network with twenty-two and ten neurons in the hidden layer, respectively, while the lowest value of 0. 680 as in the case of the F1 index was obtained for the RBF7-5-2 network. Detailed results for the parameters of the individual classifiers are presented in Table 7. The results obtained show that both multilayer perceptron (MLP) type networks and networks with radial basis functions give good results for the analyzed problem with a slight advantage of MLP type networks in most of the analyzed parameters of the classifiers studied.

A limitation of this study is undoubtedly the use of a simplified classification model, involving the assignment of results to two groups. To the limitations should also be added the fact of large differences in personal characteristics between the analyzed groups.

Further studies are planned, including the development of a method to precisely determine the degree of damage according to the ICRS scale and the approximate location of degenerative changes.

4. CONCLUSIONS

The obtained results confirm the effectiveness of the proposed method for diagnosing articular cartilage damage, which involves the selection of optimal discriminants of acoustic signals recorded for knee joints based on the results of statistical analyses and subsequent classification using artificial neural networks such as MLP and RBF. Vibroacoustic assessment may be a cheap and non-invasive alternative to typical imaging diagnostic methods, which will complement the physical examination performed during a standard visit at an orthopedic clinic.

REFERENCES

Ahn, J. M., & El-Khoury, G. Y. (2006). Computed Tomography of Knee Injuries. *Imaging Decisions MRI*, *10*(1), 14–23. https://doi.org/10.1111/j.1617-0830.2006.00063.x

Arendt, E. A., Miller, L. E., & Block, J. E. (2014). Early knee osteoarthritis management should first address mechanical joint overload. *Orthopedic Reviews*, *6*(1). https://doi.org/10.4081/or.2014.5188

Badurowicz, M. (2022). Detection of source code in internet texts using automatically generated machine learning models. Applied Computer Science. *Applied Computer Science*, *18*(1), 89–98. https://doi.org/10.23743/acs-2022-07

Bauer, L., Stütz, L., & Kley, M. (2021). Black box efficiency modelling of an electric drive unit utilizing methods of machine learning. *Applied Computer Science*, *17*(4), 5–19. https://doi.org/10.23743/acs-2021-25

Będziński, R. (1997). *Biomechanika inżynierska: Zagadnienia wybrane*. Oficyna Wydawnicza Politechniki Wrocławskiej.

Befrui, N., Elsner, J., Flesser, A., Huvanandana, J., Jarrousse, O., Le, T. N., Müller, M., Schulze, W. H. W., Taing, S., & Weidert, S. (2018). Vibroarthrography for early detection of knee osteoarthritis using normalized frequency features. *Medical & Biological Engineering & Computing*, *56*(8), 1499–1514. https://doi.org/10.1007/s11517-018-1785-4

Blodgett, W. E. (1902). Auscultation of the Knee Joint. *The Boston Medical and Surgical Journal*, *146*(3), 63–66. https://doi.org/10.1056/NEJM190201161460304

Brittberg, M., & Winalski, C. S. (2003). Evaluation of cartilage injuries and repair. *The Journal of Bone and Joint Surgery. American Volume*, *85-A Suppl 2*, 58–69.

Cameron, M. L., Briggs, K. K., & Steadman, J. R. (2003). Reproducibility and Reliability of the Outerbridge Classification for Grading Chondral Lesions of the Knee Arthroscopically. *The American Journal of Sports Medicine*, *31*(1), 83–86. https://doi.org/10.1177/03635465030310012601

Cempel, C. (2005). Diagnostyka wibroakustyczna maszyn-historia, stan obecny, perspektywy rozwoju. *Problemy Eksploatacji*, *3*, 7–25.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Cross, M., Smith, E., Hoy, D., Nolte, S., Ackerman, I., Fransen, M., Bridgett, L., Williams, S., Guillemin, F., Hill, C. L., Laslett, L. L., Jones, G., Cicuttini, F., Osborne, R., Vos, T., Buchbinder, R., Woolf, A., & March, L. (2014). The global burden of hip and knee osteoarthritis: Estimates from the Global Burden of Disease 2010 study. *Annals of the Rheumatic Diseases*, *73*(7), 1323–1330. https://doi.org/10.1136/annrheumdis-2013-204763

Dudek-Dyduch, E., Tadeusiewicz, R., & Horzyk, A. (2009). Neural network adaptation process effectiveness dependent of constant training data availability. *Neurocomputing*, *72*(13-15), 3138–3149. https://doi.org/10.1016/j.neucom.2009.03.017

Felson, D. T. (2004). Obesity and vocational and avocational overload of the joint as risk factors for osteoarthritis. *The Journal of Rheumatology Supplement*, *70*, 2–5.

Figlus, T., Kozioł, M., & Kuczyński, Ł. (2019). The Effect of Selected Operational Factors on the Vibroactivity of Upper Gearbox Housings Made of Composite Materials. *Sensors*, *19*(19), 4240. https://doi.org/10.3390/s19194240

Hayashi, D., Roemer, F. W., & Guermazi, A. (2019). Imaging of Osteoarthritis by Conventional Radiography, MR Imaging, PET–Computed Tomography, and PET-MR Imaging. *PET Clinics*, *14*(1), 17–29. https://doi.org/10.1016/j.cpet.2018.08.004

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *454*(1971), 903–995.

Hunter, D. J., & Bierma-Zeinstra, S. (2019). Osteoarthritis. *The Lancet*, *393*(10182), 1745–1759. https://doi.org/10.1016/S0140-6736(19)30417-9

Jedliński, Ł., Caban, J., Krzywonos, L., Wierzbicki, S., & Brumerčík, F. (2015). Application of vibration signal in the diagnosis of IC engine valve clearance. *Journal of Vibroengineering*, *17*(1), 175–187.

Jedliński, Ł., & Jonak, J. (2015). Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. *Applied Soft Computing*, *30*, 636–641. https://doi.org/10.1016/j.asoc.2015.02.015

Jedliński, Ł., & Jonak, J. (2020). *Kontrola montazu zebatych przekladni stozkowych metoda bezdemontazowa*. Wydawnictwo Politechniki Lubelskiej.

Johnson, V. L., & Hunter, D. J. (2014). The epidemiology of osteoarthritis. *Best Practice & Research Clinical Rheumatology*, *28*(1), 5–15. https://doi.org/10.1016/j.berh.2014.01.004

Jonak, J., Karpinski, R., Machrowska, A., Krakowski, P., & Maciejewski, M. (2019). A preliminary study on the use of EEMD-RQA algorithms in the detection of degenerative changes in knee joints. *IOP Conference Series: Materials Science and Engineering*, *710*, 012037. https://doi.org/10.1088/1757-899X/710/1/012037

Karpiński, R., Krakowski, P., Jonak, J., Machrowska, A., Maciejewski, M., & Nogalski, A. (2021a). Analysis of differences in vibroacoustic signals between healthy and osteoarthritic knees using EMD algorithm and statistical analysis. *Journal of Physics: Conference Series*, *2130*(1), 012010. https://doi.org/10.1088/1742-6596/2130/1/012010

Karpiński, R., Krakowski, P., Jonak, J., Machrowska, A., Maciejewski, M., & Nogalski, A. (2021b). Estimation of differences in selected indices of vibroacoustic signals between healthy and osteoarthritic patellofemoral joints as a potential non-invasive diagnostic tool. *Journal of Physics: Conference Series*, *2130*(1), 012009. https://doi.org/10.1088/1742-6596/2130/1/012009

Karpiński, R., Krakowski, P., Jonak, J., Machrowska, A., Maciejewski, M., & Nogalski, A. (2022a). Diagnostics of Articular Cartilage Damage Based on Generated Acoustic Signals Using ANN-Part II: Patellofemoral Joint. *Sensors*, *22*(10). https://doi.org/10.3390/s22103765

Karpiński, R., Krakowski, P., Jonak, J., Machrowska, A., Maciejewski, M., & Nogalski, A. (2022b). Diagnostics of Articular Cartilage Damage Based on Generated Acoustic Signals Using ANN-Part I: Femoral-Tibial Joint. *Sensors*, *22*(6), 2176. https://doi.org/10.3390/s22062176

Karpiński, R., Machrowska, A., & Maciejewski, M. (2019). Application of acoustic signal processing methods in detecting differences between open and closed kinematic chain movement for the knee joint. *Applied Computer Science*, *15*(1), 36–48. https://doi.org/10.23743/acs-2019-03

Karpiński, R., Szabelski, J., & Maksymiuk, J. (2019a). Effect of Physiological Fluids Contamination on Selected Mechanical Properties of Acrylate Bone Cement. *Materials*, *12*(23), 3963. https://doi.org/10.3390/ma12233963

Karpiński, R., Szabelski, J., & Maksymiuk, J. (2019b). Seasoning Polymethyl Methacrylate (PMMA) Bone Cements with Incorrect Mix Ratio. *Materials*, *12*(19), 3073. https://doi.org/10.3390/ma12193073

Kernohan, W. G., Beverland, D. E., McCoy, G. F., Hamilton, A., Watson, P., & Mollan, R. A. B. (1990). Vibration arthrometry. *Acta Orthopaedica Scandinavica*, *61*(1), 70–79.

Kosicka, E., Krzyzak, A., Dorobek, M., & Borowiec, M. (2022). Prediction of Selected Mechanical Properties of Polymer Composites with Alumina Modifiers. *Materials*, *15*(3), 882. https://doi.org/10.3390/ma15030882

Krakowski, P., Gerkowicz, A., Pietrzak, A., Krasowska, D., Jurkiewicz, A., Gorzelak, M., & Schwartz, R. A. (2019). Psoriatic arthritis – new perspectives. *Archives of Medical Science*, *15*(3), 580–589. https://doi.org/10.5114/aoms.2018.77725

Krakowski, P., Karpiński, R., Jojczuk, M., Nogalska, A., & Jonak, J. (2021). Knee MRI Underestimates the Grade of Cartilage Lesions. *Applied Sciences*, *11*(4), 1552. https://doi.org/10.3390/app11041552

Krakowski, P., Karpiński, R., Jonak, J., & Maciejewski, R. (2021). Evaluation of diagnostic accuracy of physical examination and MRI for ligament and meniscus injuries. *Journal of Physics: Conference Series*, *1736*, 012027. https://doi.org/10.1088/1742-6596/1736/1/012027

Krakowski, P., Karpiński, R., Maciejewski, R., & Jonak, J. (2021). Evaluation of the diagnostic accuracy of MRI in detection of knee cartilage lesions using Receiver Operating Characteristic curves. *Journal of Physics: Conference Series*, *1736*, 012028. https://doi.org/10.1088/1742-6596/1736/1/012028

Krakowski, P., Karpiński, R., Maciejewski, R., Jonak, J., & Jurkiewicz, A. (2020). Short-Term Effects of Arthroscopic Microfracturation of Knee Chondral Defects in Osteoarthritis. *Applied Sciences*, *10*(23), 8312. https://doi.org/10.3390/app10238312

Krakowski, P., Nogalski, A., Jurkiewicz, A., Karpiński, R., Maciejewski, R., & Jonak, J. (2019). Comparison of Diagnostic Accuracy of Physical Examination and MRI in the Most Common Knee Injuries. *Applied Sciences*, *9*(19), 4102. https://doi.org/10.3390/app9194102

Krishnan, S., Rangayyan, R. M., Bell, G. D., & Frank, C. B. (2000). Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology. *IEEE Transactions on Biomedical Engineering*, *47*(6), 773–783. https://doi.org/10.1109/10.844228

Kyu, H. H., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, M., Abebe, Z., Abil, O. Z., Aboyans, V., Abrham, A. R., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., … Murray, C. J. L. (2018). Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1859–1922. https://doi.org/10.1016/S0140-6736(18)32335-3

Machrowska, A., & Jonak, J. (2018). xEMD procedures as a data—Assisted filtering method. *AIP Conference Proceedings*, 1922, 120007. https://doi.org/10.1063/1.5019122

Machrowska, A., Karpiński, R., Jonak, J., Szabelski, J., & Krakowski, P. (2020). Numerical prediction of the component-ratio-dependent compressive strength of bone cement. *Applied Computer Science*, *16*(3), 88–101. https://doi.org/10.23743/acs-2020-24

Machrowska, A., Szabelski, J., Karpiński, R., Krakowski, P., Jonak, J., & Jonak, K. (2020). Use of Deep Learning Networks and Statistical Modeling to Predict Changes in Mechanical Parameters of Contaminated Bone Cements. *Materials*, *13*(23), 5419. https://doi.org/10.3390/ma13235419

Madej, H., Czech, P., & Konieczny, Ł. (2003). Wykorzystanie dyskryminant bezwymiarowych w diagnostyce przekładni zębatych. *Diagnostyka*, *28*, 17–22.

Mathiessen, A., Cimmino, M. A., Hammer, H. B., Haugen, I. K., Iagnocco, A., & Conaghan, P. G. (2016). Imaging of osteoarthritis (OA): What is new? *Best Practice & Research Clinical Rheumatology*, *30*(4), 653–669. https://doi.org/10.1016/j.berh.2016.09.007

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) – Protein Structure*, *405*(2), 442–451. https://doi.org/10.1016/0005-2795(75)90109-9

McDonough, C. M., & Jette, A. M. (2010). The Contribution of Osteoarthritis to Functional Limitations and Disability. *Clinics in Geriatric Medicine*, *26*(3), 387–399. https://doi.org/10.1016/j.cger.2010.04.001

Möller, I., Bong, D., Naredo, E., Filippucci, E., Carrasco, I., Moragues, C., & Iagnocco, A. (2008). Ultrasound in the study and monitoring of osteoarthritis. *Osteoarthritis and Cartilage*, *16*, S4–S7. https://doi.org/10.1016/j.joca.2008.06.005

Powers, D. M. W. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv:2010.16061 [Cs, Stat]*. http://arxiv.org/abs/2010.16061

Prior, J., Mascaro, B., Shark, L.-K., Stockdale, J., Selfe, J., Bury, R., Cole, P., & Goodacre, J. A. (2010). Analysis of high frequency acoustic emission signals as a new approach for assessing knee osteoarthritis. *Annals of the Rheumatic Diseases*, *69*(5), 929–930. https://doi.org/10.1136/ard.2009.112599

Rabiej, M. (2018). *Analizy statystyczne z programami Statistica i Excel*. Wydawnictwo Helion.

Rangayyan, R. M., & Wu, Y. F. (2008). Screening of knee-joint vibroarthrographic signals using statistical parameters and radial basis functions. *Medical & Biological Engineering & Computing*, *46*(3), 223–232. https://doi.org/10.1007/s11517-007-0278-7

Reyes, C., Garcia-Gil, M., Elorza, J. M., Mendez-Boo, L., Hermosilla, E., Javaid, M. K., Cooper, C., Diez-Perez, A., Arden, N. K., Bolibar, B., Ramos, R., & Prieto-Alhambra, D. (2015). Socio-economic status and the risk of developing hand, hip or knee osteoarthritis: A region-wide ecological study. *Osteoarthritis and Cartilage*, *23*(8), 1323–1329. https://doi.org/10.1016/j.joca.2015.03.020

Richette, P., & Latourte, A. (2019). Osteoarthritis: Value of imaging and biomarkers. *La Revue Du Praticien*, *69*(5), 507–509.

Rogala, M. (2020). Neural Networks in Crashworthiness Analysis of Thin-Walled Profile with Foam Filling. *Advances in Science and Technology Research Journal*, *14*(3), 93–99. https://doi.org/10.12913/22998624/120989

Stanik, Z. (2013). *Diagnozowanie lozysk tocznych pojazdów samochodowych metodami wibroakustycznymi*. Wydawnictwo Naukowe Instytutu Technologii Eksploatacji – Państwowego Instytutu Badawczego.

Szabelski, J. (2018). Effect of incorrect mix ratio on strength of two component adhesive Butt-Joints tested at elevated temperature. *MATEC Web of Conferences*, *244*, 01019. https://doi.org/10.1051/matecconf/201824401019

Szabelski, J., Karpiński, R., & Machrowska, A. (2022). Application of an Artificial Neural Network in the Modelling of Heat Curing Effects on the Strength of Adhesive Joints at Elevated Temperature with Imprecise Adhesive Mix Ratios. *Materials*, *15*(3), 721. https://doi.org/10.3390/ma15030721

Tadeusiewicz, R. (1993). *Sieci neuronowe* (Vol. 110). Akademicka Oficyna Wydawnicza.

Van den Borne, M. P. J., Raijmakers, N. J. H., Vanlauwe, J., Victor, J., de Jong, S. N., Bellemans, J., & Saris, D. B. F. (2007). International Cartilage Repair Society (ICRS) and Oswestry macroscopic cartilage evaluation scores validated for use in Autologous Chondrocyte Implantation (ACI) and microfracture. *Osteoarthritis and Cartilage*, *15*(12), 1397–1402. https://doi.org/10.1016/j.joca.2007.05.005

Walters, C. F. (1929). The value of joint auscultation. *The Lancet*, *213*(5514), 920–921.

Wu, Y. (2015). *Knee joint vibroarthrographic signal processing and analysis*. Springer.