Abstract

Osteoarthritis is one of the leading causes of disability around the globe. Up to this date there is no definite cure for cartilage lesions. Only fast and accurate diagnosis enables prolonging joint survivor time. Available diagnostic methods have disadvantages such as high price, radiation, need for experienced radiologists or low availability in some regions. The present study evaluates the use of vibroarthrography as a method of cartilage lesion detection. 47 patients with diagnosed cartilage lesions, and 51 healthy control group patients have been enrolled in this study. The cartilage in the study group was evaluated intraoperatively by experienced orthopaedic surgeon. Signal acquisition was performed in open and closed kinematic chain based on 10 knee joint movements from 0-90 degrees. By using EEMD-DFA algorithms, reducing classifier inputs using ANOVA and then classifying using artificial neural networks (ANN), a classification accuracy of almost 93% was achieved. A sensitivity of 0.93 and a specificity of 0.93 with an AUC of 0.942 were obtained for the multilayer perceptron network. These results allow to apply this testing protocol in a clinical setting in the future.

1. INTRODUCTION

In the last decades, with the increase in life expectancy of the population, there has been a significant increase in the prevalence of musculoskeletal diseases, including osteoarthritis (OA) (Glyn-Jones et al., 2015; Hunter & Bierma-Zeinstra, 2019). This disease, characterized...
by gradual breakdown/damage of joint cartilage, is one of the leading causes of pain and disability among the elderly, significantly affecting their quality of life. The knee joint is the largest and at the same time one of the most sensitive and most frequently damaged joints of the human body, making it also prone to the development of secondary degenerative changes on the basis of untreated trauma (Karpiński, 2022; Karpiński et al., 2023). Due to its crucial role in daily physical activity and its high susceptibility to stress and injury, it is particularly often affected by this condition. Early diagnosis and monitoring of the progress of OA are key to preventing its development and minimizing its effects (Krakowski et al., 2020; Loeser et al., 2012).

Traditional diagnostic methods such as magnetic resonance imaging (MRI) and X-ray imaging, while effective in assessing the severity of disease, often fail to detect subtle changes in cartilage tissue in their early stages (Krakowski et al., 2021a; 2021b). In this context, there is growing interest in methods that allow more accurate assessment of the mechanical properties of cartilage and early detection of degenerative changes before they become apparent on traditional diagnostic imaging studies. An alternative to typical methods of assessing articular cartilage damage may be vibroacoustic diagnostics, which involves evaluating the sounds and vibrations generated by moving knee joints (Bączkowicz & Majoreczyk, 2014; Kręcisz & Bączkowicz, 2018; Łysiak et al., 2020).

Vibroacoustic machine diagnostics is a method of assessing the condition of machines by analyzing their vibrations and sounds (Delvecchio et al., 2018; Jedliński et al., 2015). It uses special sensors to collect signals that are analyzed to detect early signs of damage. The technique makes it possible to plan maintenance, minimize downtime and reduce costs associated with failures, while increasing safety and extending the life of machines (Dąbrowski & Dziurdź, 2016; Jedliński et al., 2022). The method can find application in the area of musculoskeletal diagnostics in assessing knee joints, among other applications. Vibroacoustic diagnosis of knee joints, using advanced techniques to analyze signals generated by moving joint surfaces, opens up new possibilities in this area. This method, based on the recording and analysis of vibroacoustic signals emitted by the joint during movement, together with classification methods based on machine learning, allows non-visionary assessment of the condition of articular cartilage (Gong et al., 2021). Previous studies have shown that changes in the mechanical parameters of cartilage, caused by degenerative processes, affect the characteristics of the vibroacoustic signals generated. Analysis of these changes can therefore provide valuable diagnostic information, enabling not only early diagnosis of OA, but also monitoring of the progress of the disease and the effectiveness of the treatment methods used.

However, the recording and processing of vibro-arthrography (VAG) signals in knee joint diagnosis faces difficulties related to the complexity of the signals, their nonlinearity and non-stationarity, the presence of external interference, interference from other joints, the lack of standards in signal recording and analysis, and challenges in interpreting the obtained results (Balajee & Venkatesan, 2021; Kręcisz et al., 2022). To effectively analyze VAG signals, advanced processing techniques and methods based on machine learning are needed, which requires large data sets and significant computational resources (Balajee et al., 2023; Patankar et al., 2023). The development of VAG signal processing methods requires interdisciplinary collaboration to improve the diagnosis of knee joint pathology.

The use of Ensemble Empirical Mode Decomposition (EEMD) algorithms combined with Detrended Fluctuation Analysis (DFA) can enable effective analysis of nonlinear and
non-stationary vibroacoustic signals (Wu, 2015). EEMD adaptively decomposes the signal into simple components, making it easier to analyze complex signals without assuming linearity or stationarity. Then, DFA is used to study long-term correlations and fractal scaling in these signals, which is useful in identifying their complex behavior (Chen et al., 2002; Hu et al., 2001). The combination of these techniques offers improved signal analysis, identification of hidden patterns, noise reduction and versatility for applications in various fields such as mechanical engineering, environmental sciences and healthcare, contributing to more accurate diagnostics and the development of predictive models (Aranchayanont et al., 2016; Nalband et al., 2016; Wu et al., 2014).

Machine learning (ML) can be used to analyze acoustic signals generated by the moving knee joints of patients with osteoarthritis (OA), opening up new possibilities in the diagnosis and monitoring of the disease (Wang et al., 2021). The process involves collecting acoustic signals, digitally processing them, including extraction of characteristic features (e.g., frequency, amplitude), and then using ML algorithms to classify and predict joint conditions (Karpiński et al., 2022a, 2022b). This method offers a number of advantages, such as non-invasiveness, the possibility of early diagnosis, and monitoring of disease progression. At the same time, challenges include ensuring high data quality and the need to collect a large number of independent signals for training. The development of this technology has the potential to significantly improve the quality of life for people with OA by offering more precise and personalized medicine (Cai et al., 2013).

This article aims to present the results of a study on vibroacoustic diagnosis of knee joints using filtering with EEMD-DFA algorithms and input reduction using ANOVA. Classification results using MLP and RBF artificial neural networks are presented. The paper also discusses how changes in the mechanical properties of cartilage, caused by degenerative processes, affect the vibroacoustic signals generated, and how these changes can be used to develop new diagnostic and therapeutic methods for the treatment of OA.

2. MATERIALS AND METHODS

2.1. Group characteristics and recording of vibroacoustic signals

The study included 148 knee joints, of which 101 were healthy joints while 47 were joints with confirmed osteoarthritis confirmed by physical examination and intraoperatively during surgical procedures. The basis for the analyses carried out were acoustic signals collected during tests conducted on a group of volunteers in the conditions of the orthopedic department, as well as in the laboratories of the Lublin University of Technology. The study received a positive opinion from the Bioethics Committee of the Medical University of Lublin, approval number KE-0254/261/2019. The measurement system was built on the basis of Arduino Mega2560 R3. Signals were collected using a CM01B piezoelectric contact microphone, which was connected to the analog inputs of the Arduino module. The microphones were fixed with double-sided adhesive tape in the central part of the patella. This location was chosen for the small amount of soft tissue that could adversely affect the recording process. To ensure patient safety, a galvanic barrier was used on the USB connector, and the device was powered by an 11.1V lithium-ion battery. Measurement of the angular position (flexion phase) of the limb was carried out using an EMS22A50-D28-
LT6 encoder from Bourns, placed in the axis of rotation of the knee joint in a Breg T-Scope Knee orthosis. The structure of the system is shown in the block diagram shown in Figure 1.

![Block diagram of the measurement system with one microphone placed on the patella](image)

Vibroacoustic signals were recorded during a sequence of knee joint straightening and bending movements from 90° to 0° and back to 90°. The tests were carried out in both closed kinetic chain (CKC) and open kinetic chain (OKC). The assumption of carrying out the measurement in such sequences of movements was due to significant differences in biomechanics and different loads on the joint depending on the kinetic chain. The procedure for recording vibroacoustic signals in a closed kinetic chain assumed rising from a sitting position, with knee flexion to 90°, to a full standing position (knee flexion angle equal to 0°), and then lowering to a sitting position with knee flexion again to 90°, lasting about 2 seconds. When signals were recorded in an open kinetic chain, the procedure involved straightening the knees from a sitting position with the knees bent at 90° to full extension (0°) and bending back to 90°. As in the closed kinetic chain, the duration of one cycle was about 2 seconds. In both cases, the signals were recorded during 10 full repetitions of the described sequences. The signals thus acquired were subjected to further processing procedures.

### 2.2. Signal processing based on EEMD- DFA methods

In this paper proposes a combination of EEMD- DFA methods to extract frequency ranges relevant to degenerative disease progression. The results of the filtering will be forwarded for further analysis conducted using neural networks to identify the changes occurring. This represents an important step toward the development of expert systems for non-invasively determining the condition of knee joints in terms of the presence of osteoarthritis.

The signal processing method known as EEMD (Ensemble Empirical Mode Decomposition) makes it possible to extract signal bands within a certain frequency range. It was introduced to analyze nonlinear and non-stationary signals. It is derived from the EMD (Empirical Mode Decomposition) method and is an improved variation of it. In analyses based on the classical EMD algorithm, there is a phenomenon of mixing of
frequency components from different extracted ranges, especially in cases where unusual phenomena such as noise or impulse interference are present in the spectrum. During the occurrence of modal mixing in EMD results, aliasing occurs resulting in the extraction of components containing different scales. To prevent this phenomenon, an improved version of the algorithm was introduced by adding a white noise component to the signal. The addition of white noise improves the efficiency of the decomposition and increases its stability. The algorithm thus modified, called EEMD, is one among a whole group of xEMD methods, such as noise-assisted Multivariate EMD (MEMD) (Rehman & Mandic, 2010) or Fast Multivariate EMD (FMEMD) (Liu et al., 2019), among others.

However, in order to characterize the noise-assisted variation of the algorithm, it is necessary to mention its prototype, the classic EMD (Empirical Mode Decomposition) method. It involves the decomposition of a signal \( x(t) \) into a number of components with a specific frequency range, called IMF (Intrinsic Mode Function) or modes. To consider a component of \( x(t) \) as an IMF, called \( c_1(t) \), it must satisfy two conditions:

1. the number of extrema and the number of zero crossings must be equal or differ at most by one,
2. the mean value of local envelopes (upper and lower) is equal to zero.

The residue \( r_1(t) \) obtained by subtracting \( c_1(t) \) from the input signal serves as a signal for the decomposition of further IMF functions. The process of dividing the signal into components is repeated until a monotonic residual signal is obtained. So we can write that:

\[
x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t)
\] (1)

The improvement of EMD in the form of an evolution of the algorithm to EEMD involves adding a finite-amplitude white noise component to the signal. The EEMD application can be presented as:

\[
x_i(t) = x(t) + w_i(t),
\] (2)

where \( x(t) \) is the input signal, \( w_i(t) \) is white noise with a length equal to the length of \( x(t) \) and \( i=1, 2, \ldots, M \). \( M \) is the number of ensembles.

The decomposition looks similar to that presented for EMD:

\[
x_i(t) = \sum_{j=1}^{N} c_{ij}(t) + r_i(t),
\] (3)

where \( j=1, 2, \ldots, N \). \( N \) is the number of IMFs represented as \( c_{ij}(t) \) and \( r_i(t) \) denotes the residual signal of the \( i \)-th waveform.

Finally, the ensemble means of the subsequent IMFs are determined.

\[
c_j(t) = \frac{1}{M} \sum_{i=1}^{M} c_{ij}(t).
\] (4)

The EEMD technique was used for acoustic signals recorded using a sensor placed on the patella at its central point. Figures 2 and 3 show the decomposition of VAG signal waveforms for a person diagnosed with patella chondromalacia (OA) and a person from the
The individual components of the decomposition reflect the division of the recorded signals in terms of frequency. IMFs $c_1$, $c_2$ and $c_3$ represent the time courses of VAG signals with high-frequency components, and IMFs with indexes higher than 3 represent signals from the low-frequency range.

Fig. 2. A waveform showing decomposition for a patient diagnosed with knee osteoarthritis

Fig. 3. A waveform showing the decomposition for a person from the control group

The second stage of processing the recorded signals was the use of the Detrended Fluctuation Analysis (DFA) technique. The signals originating from the human body are rhythmic and repetitive by nature. They are characterized by relationships within signals, also known as correlations. DFA is a method for detecting long-range correlations in signals (LRC). There are known cases of using the DFA algorithm and the $\alpha$ exponent value to detect degenerative changes in human gait as Parkinson's disease progresses (Ota et al.,
2011). Another, no less important, property of DFA is the ability to separate the useful part of the signal from artifacts (Wu et al., 2014).

Using the DFA technique, the fractal scaling parameter is determined, which enables the description of small fluctuations corresponding to the dynamics characteristics of the waveform. The waveform \( c_i(t) \), which is the isolated IMF, will be the input signal here. To simplify the notation, let's call it \( q(t) \). DFA analysis begins with determining the average of the tested waveform:

\[
\bar{q} = \frac{1}{N} \sum_{t=1}^{N} q(t)
\]

The cumulative sum is then determined:

\[
y(m) = \sum_{k=1}^{m} [q_k(t) - \bar{q}]
\]

The integrated time series \( y(m) \) are divided into several windows of length \( n \). In each of the windows, the integrated time series has been adjusted using a local trend in the form of a polynomial function \( y_f(m) \). The trend can take different forms, depending on the degree of the polynomial. By subtracting the local trend \( y_f(m) \) from the time series \( y(m) \) in each window, the detrended fluctuation function \( Y(m) \) was obtained.

\[
Y(m) = y(m) - y_f(m)
\]

For a window of length \( n \), the root mean square (rms) fluctuation is then calculated:

\[
F(n) = \sqrt{\frac{1}{N} \sum_{m=1}^{N} [Y(m)]^2}
\]

The calculations are repeated for windows of length \( n \) (the lengths change due to different scales) so that the relationship between \( F(n) \) and \( n \) itself is ensured (Hu et al., 2001). The existence of a power law relationship between \( F(n) \) and the selected window size indicates the presence of a scaling relationship: \( F(n) \sim n^\alpha \). Scaling factor \( \alpha \) reflects the occurrence of correlations, reflecting the occurrence of different signal properties. If \( 0.5 < \alpha < 1 \) we can talk about the occurrence of long-range correlation (LRC), while if the coefficient takes values \( 0 < \alpha < 0.5 \), the signal exhibits anticorrelation properties. If the coefficient takes a value \( 0.5 \) the signal properties are those of white noise one.

In the research that is the subject of this article, after decomposition using EEMD, DFA analysis was performed for each of the obtained IMF functions. For DFA, the range of window length \( n \) from 50 to 500 with an increment of 10 was assumed. Based on the scaling factor results obtained \( \alpha \) for each of the IMFs, it was decided to leave the components that exhibit LRC features for further analysis (\( \alpha > 0.5 \)). Each decomposed waveform was reconstructed based on components showing features of long-range correlation. Figures 4 and 5 show the relationship between the mean fluctuation and the length of the k window in the closed kinematic chain (CKC) (Figure 4) for a person diagnosed with knee chondromalacia (OA) and a healthy person (HC) (Figure 5).
The vibroacoustic signals prepared in this way were further processed, and discriminants were determined for the reconstructed signals.

2.3. Determination and selection of the most informationally important signal measures

For vibroacoustic signals pre-processed and filtered using EEMD-DFA algorithms, based on literature analysis (Karpiński et al., 2021a, 2021b), 15 measures/discriminant of vibroacoustic signals were determined separately for each of the adopted test protocols, i.e. open and closed kinetic chain. These discriminants are: Variance (Var), Standard Deviation (StDev), Root Mean Square (RMS), Form Factor (FF), Crest Factor (CF), FM4, Kurtosis (Kur), Peak to Peak Value (PPV), M6A, Impulse Factor (IF), Straightened Mean (SM), Mean (M), M8A, Skewness (SK) and Peak Value (PV). In order to reduce the number of classifier inputs, it was decided to reduce the inputs using analysis of variance, ANOVA, which is a statistical method used to examine observations that depend on one or many factors acting simultaneously.
Analysis of variance, known as ANOVA, is a statistical method used to examine observations that depend on one or more factors. The results obtained using this method explain the probability with which selected factors can account for differences between the observed group means. Testing using one-way ANOVA allows you to determine whether one independent variable influences the results of one dependent variable. It has greater possibilities than the Student's t-test because it can also be used when the independent variable has many levels. It involves comparing the inter-group variance and the intra-group variance.

In the case of data in the form of recorded VAG signals, there were large values of inter-group variance within the designated measures/signal discriminant, which helped in the selection of parameters for classifying the condition of the knee joints. Calculations of the significance of the indicators used were based on one-way analysis of variance for each predictor variable, grouped by class. The features were then ranked using value \( p = 0.05 \). For each predictor variable, the hypothesis was tested that the predictor values grouped by response class came from a population with the same mean, against the alternative hypothesis that the population means were not the same. The results match \( \log(p) \).

Based on the results obtained, measures whose significance was greater than 10 were selected for further analysis (Fig. 6): Variance, Standard Deviation, RMS, Form Factor, Crest Factor, FM4 and Peak-to-Peak Value. Due to the similar method of calculating the FM4 parameter and Kurtosis, it was decided to use only the first one in further analyses. Finally, the 7 most informationally significant measures were adopted for each chain, which gave a total of 14 classifier inputs. These indicators are briefly described below.

![Feature importance scores using ANOVA algorithm](image)

Fig. 6. Feature importance scores using ANOVA algorithm

Variance is a statistical measure that describes the dispersion of results around the arithmetic mean. The parameter describes how diverse the data set is in terms of concentration around the average or dispersion of values in the set. Results oscillating around zero mean sets with very similar values. An indicator that provides similar information is the standard deviation. It also gave good results in terms of significance, but it was not used for further analyzes due to the similarity of properties in terms of variance. For a set of values \( A \) with the number of observations \( N \), the variance is expressed as:
where $\mu$ we interpret as mean:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} A_i$$  \hspace{1cm} (10)

The second parameter adopted was RMS, i.e. mean square or effective value. It is a statistical measure of signal that allows one to estimate the order of magnitude of a series of numerical data. It is especially useful when data differs in sign, producing meaningful results unlike the arithmetic mean. The RMS value calculated for set A is calculated as follows:

$$A_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |A_i|^2}$$  \hspace{1cm} (11)

Another parameter of relevance for further analysis was the form factor (FF). It is defined as the ratio of the RMS value to the average value of the absolute value of a given waveform:

$$FF = \frac{A_{RMS}}{\bar{A}}$$  \hspace{1cm} (12)

where $\bar{A}$ is marked as mean.

The Crest Factor (CF) shows the ratio of the peak value to the RMS value. Provides information on how extreme the waveform peaks are. High Crest Factor values in terms of VAG signals (vibroarthrography signals) often indicate degeneration of the cartilage of the knee joints.

The Peak to Peak Value (PPV) parameter is also useful in terms of classifying the condition of the knee joints. This is a measure that determines the range between the highest and lowest amplitude values.

The last measure selected for classification is FM4, known as the difference signal kurtosis. Describes how peaked or flat the difference signal amplitude is. The results of kurtosis and FM4 gave results that were very similar but slightly better in favor of FM4, so it was selected for further research. The tan parameter is calculated using the equation:

$$FM4 = \frac{1}{N} \sum_{i=1}^{N} (d_i - \bar{d})^4 \left[ \frac{1}{N} \sum_{i=1}^{N} (d_i - \bar{d})^2 \right]^{-2}$$  \hspace{1cm} (13)

where $d$ is difference signal, $\bar{d}$ is the mean of this signal.

In the next stage, the measures selected above were used as inputs to classifiers.

### 2.4. Classification and evaluation of classifiers

Machine learning (ML) is revolutionizing medical diagnostics by analyzing medical images for precise identification of abnormalities, predicting disease risk based on clinical and genetic data, personalizing therapy, and being supported by virtual therapy assistants. These technologies facilitate faster and more accurate diagnosis, enable early disease
detection and optimized treatment, leading to better healthcare and outcomes for patients. Each of these applications contributes to the development of precision medicine that is more personalized, efficient and evidence-based. With machine learning, it is possible to better understand each patient’s unique characteristics and tailor a treatment plan to them, which can significantly improve quality of life and patient outcomes. The use of neural networks in medical diagnostics offers promising opportunities to improve diagnostic accuracy, reduce the time required for diagnosis, and personalize therapeutic approaches, which together contribute to better disease management and improved treatment outcomes (Machrowska et al., 2020; Machrowska et al., 2020a).

In the present study, neural networks with radial RBF basis functions and multilayer perceptron neural networks were used to solve the classification task (Falkowicz & Kulisz, 2024; Szabelski et al., 2022). All calculations related to classification were carried out using Statistica 13.3 software. As inputs to the classifiers, respectively, 7 indicators each assumed to be the most informationally relevant, determined separately for measurements carried out in an open and closed kinetic chain. A total of 14 independent parameters of vibroacoustic signals were used. Classification involved assigning cases to one of two groups OA - subjects with osteoarthritis and HC - healthy subjects without any changes in the articular cartilage.

A confusion matrix, ROC curves and their key parameters were used to compare the results of the obtained classifiers, as well as measures resistant to group differentiation such as Matthews correlation coefficient and F1 Score (Chicco & Jurman, 2020).

3. RESULTS

This section of the paper presents the results of using artificial neural networks (ANNs) of the multilayer perceptron (MLP) and radial basis function (RBF) types to classify vibroacoustic signals in the context of diagnosing health conditions. The aim of the study was to assign the analyzed signals to one of two groups: healthy controls (HC) and osteoarthrosis (OA) patients. The classification was based on selected measures characterizing vibroacoustic signals that may indicate differences in the mechanisms leading to joint dysfunction. Table 1 shows the accuracy parameters in the learning, test and validation sets.

<table>
<thead>
<tr>
<th>Network name</th>
<th>Accuracy of learning (%)</th>
<th>Accuracy of testing (%)</th>
<th>Accuracy of validation (%)</th>
<th>Learning algorithm</th>
<th>Error function</th>
<th>Activation (hidden)</th>
<th>Activation (output)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP 14-14-2</td>
<td>96.15</td>
<td>77.27</td>
<td>86.36</td>
<td>BFGS 38</td>
<td>Entropy</td>
<td>Logistic</td>
<td>Softmax</td>
</tr>
<tr>
<td>MLP 14-47-2</td>
<td>96.15</td>
<td>86.36</td>
<td>86.36</td>
<td>BFGS 79</td>
<td>SOS</td>
<td>Logistic</td>
<td>Tanh</td>
</tr>
<tr>
<td>RBF 14-18-2</td>
<td>80.77</td>
<td>77.27</td>
<td>86.36</td>
<td>RBFT</td>
<td>SOS</td>
<td>Gauss</td>
<td>Linear</td>
</tr>
<tr>
<td>RBF 14-19-2</td>
<td>81.73</td>
<td>72.73</td>
<td>81.82</td>
<td>RBFT</td>
<td>SOS</td>
<td>Gauss</td>
<td>Linear</td>
</tr>
</tbody>
</table>

The highest accuracy in the learning set of 96.15% was observed for MLP-type classifiers. For networks with radial basis functions, the values obtained were lower oscillating around 81%. In the test set, the highest accuracy was obtained for the MLP 14-47-2 network and was 86.36%, while the lowest value of 72.73% was observed for the RBF
14-19-2 network. In the validation set, an accuracy value of 86.36% was obtained for 3 of the 4 classifiers under consideration, while the fourth classifier under consideration achieved 81.82%. The learning algorithm for the MLP classifiers was BFGS while for the RBF network it was the RBFT algorithm.

The BFGS algorithm is a second-order nonlinear optimization method used to find local minima of a function without calculating second derivatives (Hessian). Instead, it uses an iterative update of the approximate inverse of the Hessian. It is more efficient than first-order methods, but can be expensive to use with a very large number of variables. BFGS finds application in machine learning. The error function in three of the four classifiers considered was the sum of squares of the SOS in the last case it was the entropy function. The activation function of the hidden layer in the case of MPL classifiers was the logistic function while in the case of RBF it was the Gaussian function. The hidden layer activation function for RBF classifiers was a linear function, while for MLP it was softmax and tanh functions, respectively.

The confusion matrix and detailed classification results in each group obtained with the selected classifiers are shown in Table 2. Overall, the highest level classification accuracy was obtained by the 14-47-2 multilayer perceptron network, correctly assigning 97.03% of cases in the HC group and 85.11% in the OA group. In the case of classification using the RBF network, the highest accuracy obtained was 81.08% of the assignment of improved cases with 88.12% in the HC group and 65.96% in the OA group.

<table>
<thead>
<tr>
<th>Network name</th>
<th>HC</th>
<th>OA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP 14-14-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>101.00</td>
<td>47.00</td>
<td>148.00</td>
</tr>
<tr>
<td>Correct</td>
<td>96.00</td>
<td>40.00</td>
<td>136.00</td>
</tr>
<tr>
<td>Incorrect</td>
<td>5.00</td>
<td>7.00</td>
<td>12.00</td>
</tr>
<tr>
<td>Correct (%)</td>
<td>95.05</td>
<td>85.11</td>
<td>91.89</td>
</tr>
<tr>
<td>Incorrect (%)</td>
<td>4.95</td>
<td>14.89</td>
<td>8.11</td>
</tr>
<tr>
<td>MLP 14-47-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>101.00</td>
<td>47.00</td>
<td>148.00</td>
</tr>
<tr>
<td>Correct</td>
<td>98.00</td>
<td>40.00</td>
<td>138.00</td>
</tr>
<tr>
<td>Incorrect</td>
<td>3.00</td>
<td>7.00</td>
<td>10.00</td>
</tr>
<tr>
<td>Correct (%)</td>
<td>97.03</td>
<td>85.11</td>
<td>93.24</td>
</tr>
<tr>
<td>Incorrect (%)</td>
<td>2.97</td>
<td>14.89</td>
<td>6.76</td>
</tr>
<tr>
<td>RBF 14-18-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>101.00</td>
<td>47.00</td>
<td>148.00</td>
</tr>
<tr>
<td>Correct</td>
<td>89.00</td>
<td>31.00</td>
<td>120.00</td>
</tr>
<tr>
<td>Incorrect</td>
<td>12.00</td>
<td>16.00</td>
<td>28.00</td>
</tr>
<tr>
<td>Correct (%)</td>
<td>88.12</td>
<td>65.96</td>
<td>81.08</td>
</tr>
<tr>
<td>Incorrect (%)</td>
<td>11.88</td>
<td>34.04</td>
<td>18.92</td>
</tr>
<tr>
<td>RBF 14-19-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>101.00</td>
<td>47.00</td>
<td>148.00</td>
</tr>
<tr>
<td>Correct</td>
<td>89.00</td>
<td>30.00</td>
<td>119.00</td>
</tr>
<tr>
<td>Incorrect</td>
<td>12.00</td>
<td>17.00</td>
<td>29.00</td>
</tr>
<tr>
<td>Correct (%)</td>
<td>88.12</td>
<td>63.83</td>
<td>80.41</td>
</tr>
<tr>
<td>Incorrect (%)</td>
<td>11.88</td>
<td>36.17</td>
<td>19.59</td>
</tr>
</tbody>
</table>

Details of the classifiers’ parameters are shown in Table 3, while graphical representations of the ROC curves are shown in Fig. 7. Analysis of the parameters that allow comparison of the classifiers shows that the MLP 14-47-2 classifier coped best with solving the presented classification problem, obtaining the highest values of F1 score = 0.889,
MCC=0.842 and the largest area under the ROC curve AUC=0.942 with a sensitivity of 0.930 and specificity of 0.933. The values obtained with the RBF-type network were significantly lower.

Tab. 3. Details of classification parameters using the various methods

<table>
<thead>
<tr>
<th>Network name</th>
<th>Accuracy (%)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>ROC threshold</th>
<th>Precision</th>
<th>Recall</th>
<th>F1 score</th>
<th>MCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP 14-14-2</td>
<td>91.89</td>
<td>0.889</td>
<td>0.932</td>
<td>0.929</td>
<td>0.718</td>
<td>0.851</td>
<td>0.889</td>
<td>0.870</td>
<td>0.811</td>
</tr>
<tr>
<td>MLP 14-47-2</td>
<td>93.24</td>
<td>0.930</td>
<td>0.933</td>
<td>0.942</td>
<td>0.669</td>
<td>0.851</td>
<td>0.930</td>
<td>0.889</td>
<td>0.842</td>
</tr>
<tr>
<td>RBF 14-18-2</td>
<td>81.08</td>
<td>0.721</td>
<td>0.848</td>
<td>0.860</td>
<td>0.579</td>
<td>0.660</td>
<td>0.721</td>
<td>0.689</td>
<td>0.554</td>
</tr>
<tr>
<td>RBF 14-19-2</td>
<td>80.41</td>
<td>0.714</td>
<td>0.840</td>
<td>0.847</td>
<td>0.622</td>
<td>0.638</td>
<td>0.714</td>
<td>0.674</td>
<td>0.536</td>
</tr>
</tbody>
</table>

Fig. 7. Comparison of ROC curves for all classification methods

Analysis of the graph containing the ROC curves confirms the observations presented above and shows that MLP-type neural networks performed significantly better in solving the presented classification problem.

4. DISCUSSION

In recent decades, osteoarthritis (OA) has become one of the leading causes of discomfort and mobility limitations in the elderly population and is beginning to occur with increasing frequency in the younger population, leading to significant challenges in both healthcare and economics. The knee joint, which is one of the most commonly affected joints in OA, is a key area of research to improve diagnosis and treatment. Traditional methods of diagnosing OA, such as physical examination, radiography and MRI, while effective, have their
limitations, including high cost and the availability gap. Vibroarthrography, as a technique based on the analysis of sounds and vibrations generated by moving joint surfaces, has shown promise in initially identifying subtle changes in joint structure before they become apparent in more traditional diagnostic imaging studies. Table 4 shows the classification parameters presented in the actual literature.

Tab. 4. Comparison diagnostic results of proposed method with other related works

<table>
<thead>
<tr>
<th>Authors</th>
<th>Classification methods</th>
<th>Accuracy (%)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin et al.</td>
<td>Statistical analysis</td>
<td>81.52</td>
<td>0.895</td>
<td>0.675</td>
<td>0.807</td>
<td>(Lin et al., 2014)</td>
</tr>
<tr>
<td>Nalband et al.</td>
<td>LS-SVM</td>
<td>83.14</td>
<td>0.981</td>
<td>0.622</td>
<td>0.671</td>
<td>(Nalband et al., 2017)</td>
</tr>
<tr>
<td>Shidore et al.</td>
<td>SVM</td>
<td>87.69</td>
<td>0.857</td>
<td>0.838</td>
<td>0.926</td>
<td>(Shidore et al., 2021)</td>
</tr>
<tr>
<td>Goossens et al.</td>
<td>XGBoost</td>
<td>87.70</td>
<td>0.881</td>
<td>0.833</td>
<td>0.810</td>
<td>(Goossens et al., 2023)</td>
</tr>
<tr>
<td>Yang et al.</td>
<td>Bayesian decision rule</td>
<td>88.00</td>
<td>0.714</td>
<td>0.979</td>
<td>0.957</td>
<td>(Yang et al., 2014)</td>
</tr>
<tr>
<td>Cai et al.</td>
<td>Dynamic weighted classifier Fusion system</td>
<td>88.76</td>
<td>0.737</td>
<td>1.000</td>
<td>0.952</td>
<td>(Cai et al., 2013)</td>
</tr>
<tr>
<td>Karpiński</td>
<td>MLP</td>
<td>90.00</td>
<td>0.917</td>
<td>0.885</td>
<td>0.941</td>
<td>(Karpiński, 2022)</td>
</tr>
<tr>
<td>Rangayyan and Wu</td>
<td>RBF</td>
<td>98.89</td>
<td>0.921</td>
<td>0.882</td>
<td>0.917</td>
<td>(Rangayyan &amp; Wu, 2009)</td>
</tr>
<tr>
<td>Mu et al.</td>
<td>Strict 2-surface proximal classifier</td>
<td>91.01</td>
<td>0.947</td>
<td>0.882</td>
<td>0.950</td>
<td>(Mu et al., 2008)</td>
</tr>
<tr>
<td>Kim et al.</td>
<td>Back propagation neural network</td>
<td>95.40</td>
<td>0.920</td>
<td>0.987</td>
<td>N/A</td>
<td>(Kim et al., 2009)</td>
</tr>
<tr>
<td>Karpiński et al.</td>
<td>MLP</td>
<td>96.32</td>
<td>0.957</td>
<td>0.967</td>
<td>0.996</td>
<td>(Karpiński et al., 2022a)</td>
</tr>
<tr>
<td>Karpiński et al.</td>
<td>MLP, RBF</td>
<td>98.53</td>
<td>0.958</td>
<td>1.000</td>
<td>1.000</td>
<td>(Karpiński et al., 2022b)</td>
</tr>
<tr>
<td>Balajee i wsp.</td>
<td>LS-SVM</td>
<td>98.67</td>
<td>0.934</td>
<td>0.937</td>
<td>N/A</td>
<td>(Balajee et al., 2023)</td>
</tr>
<tr>
<td>Rangayyan et al.</td>
<td>RBF</td>
<td>100</td>
<td>1</td>
<td>1</td>
<td>0.961</td>
<td>(Rangayyan et al., 2013)</td>
</tr>
</tbody>
</table>

Up to date there is no consensus in examination protocol as well as signal acquisition and signal processing. Simple sensor location can affect results significantly as shown by Madeleine et al (2020). Moreover, there is dependance on results in regard to chosen kinematic chain (Karpiński et al., 2019). On top of those findings, multiple signal processing methods have been proposed in literature (Krishnan et al., 1997; Nalband et al., 2018; Maussavi et al., 1996). Given that, there is no strict protocol of signal processing, and multiple approaches have been proposed in literature. Krishnann et all used adaptive filtering and time frequency analysis to evaluate VAG (Krishnan et al., 2000). What is interesting, the same research group showed that over filtering especially removal of muscle contraction...
artefacts significantly decreased classification accuracy (Krishnan et al., 1997). This phenomenon is understandable in clinical setting, while muscles adopt to joint mobility and overall joint function, therefore muscle contraction noise can provide additional information for the system. What is interesting, VAG have already proved its utility in clinical setting on small group of patients with juvenile arthritis. Researchers showed that joints classified at the beginning of the treatment as diseased at the end were classified as healthy (Semiz et al., 2018). This study shows high potential in clinical setting after resolving issues with signal processing.

EEMDA-DFA which can extract signal bands within a certain frequency range and introduced to analyze nonlinear and non-stationary signals have been implemented in this study. The use of the EEMD-DFA algorithm is beneficial in the analysis of vibroacoustic signals of knee joints due to their complexity and non-stationarity. EEMD allows decomposition of the signal into components of different frequencies, facilitating analysis. DFA, on the other hand, allows the study of long-term correlations in stationary signal components. This adaptive approach allows analysis on different time scales, which is important for variable vibroacoustic signals of knee joints. Proposed algorithms were chosen due to a fact that VAG is nonlinear. This approach have proved to reduce noise ratio and therefore have a positive impact on results. In this study, the approach was found to have a sensitivity of 0.93 and a specificity of 0.93. This shows high diagnostic accuracy and clinical utility. What is important, this study was performed in clinical setting with cartilage status checked visually in every patient by experienced orthopaedic surgeon during surgery. Such method of cartilage evaluation carries little bias and is most accurate approach towards cartilage evaluation. Other research groups as reference used MRI (Kręcisz & Bączkowicz, 2018), which is most accurate available diagnostic modality, however, it was shown in other studies that MRI grossly underestimates cartilage lesion degree (Krakowski et al., 2021), which brings possible bias into those studies. This study carried out on high number of 148 knees showed that our proposed approach have a potential in clinical setting. However higher study and control group is required before it can be lunched as diagnostic modality in orthopaedics.

5. CONCLUSIONS

The proposed method of signal acquisition and processing showed high diagnostic accuracy which can be clinically utilized after confirmation of higher patient group. MLP networks showed highest sensitivity and specificity among analyzed ANN. Due to simple examination protocol, which almost every patient can perform, low cost device, ease of use and mobility, proposed approach can easily be transferred into clinical setting.

Author Contributions

Funding

This research was funded by the Scientific Fund of the Lublin University of Technology: FD/IM/052.

Institutional Review Board Statement:

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Ethics Committee of the Medical University of Lublin consent number KE-0254/261/2019 on 26/09/2019.

Conflicts of Interest

The authors declare no conflicts of interest.

REFERENCES


