Advance Publication by J-STAGE

Journal of Biomechanical Science and Engineering

DOI:10.1299/jbse.20-00017

Received date: 8 January, 2020

Accepted date: 31 March, 2020

J-STAGE Advance Publication date: 9 April, 2020
Knee osteoarthritis detection based on the combination of empirical mode decomposition and wavelet analysis

Rui GONG*, Kazunori HASE**, Hiroaki GOTO***, Keisuke YOSHIOKA*** and Susumu OTA****

*Graduate School of Systems Design, Tokyo Metropolitan University
1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, Japan
**Faculty of Systems Design, Tokyo Metropolitan University
1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, Japan
E-mail: kazunori.hase@tmu.ac.jp
***livet Inc.
#301, 4-9-1 Myojincho, Hachioji, Tokyo 192-0046, Japan
****The Faculty of Rehabilitation and Care, Seijoh University
2-172 Fukinodai, Tokai, Aichi 476-0014, Japan

Abstract
The early-stage of knee osteoarthritis (OA) is usually asymptomatic. However, timely detection of osteoarthritis can prevent further cartilage degeneration via appropriate exercise prescription and behavioral change. In this article, a noninvasive method to diagnose the OA of a knee recording the knee vibroarthrographic (VAG) signals over the mid-patella during the standing movement is proposed. A method that combines empirical mode decomposition (EMD) and wavelet transform is developed to analyze the nonstationary VAG signals. The least squares support vector machine algorithm (LSSVM) that is a type of support vector machine is used to classify the knee joint VAG signals (26 normal and 25 abnormal) collected from healthy subjects and patients suffering from the knee OA using the Kellgren and Lawrence grading system III and IV (KLGS III and IV). The LSSVM classifier achieves an accuracy of 86.67% in differentiating the normal and abnormal subjects that proves the effectiveness of the autocorrelation function features and continuous wavelet transform (CWT) features. Therefore, the VAG signals can be clinically significant for the classification of healthy and OA subjects.

Keywords: Osteoarthritis (OA), Vibroarthrographic (VAG), Empirical Mode Decomposition (EMD), Wavelet transform, Least Squares Support Vector Machine Algorithm (LSSVM)

1. Introduction

Osteoarthritis (OA) represents a progressive degradation process and usually occurs in the knee joint because this joint supports almost the entire weight of a human body, and also provides motion during walking (Vigorita, 2008). Therefore, the knee is the most vulnerable, both to the injuries and OA in the human body. The knee OA is an idiopathic disease characterized by a degeneration of articular cartilage. A breakdown of the cartilage matrix can lead to the development of fibrillation and fissures, appearance of gross ulcerations, and disappearance of the full thickness surface of the joint. These processes are accompanied by changes in the bone due to osteophyte formation and thickening of the subchondral plate (Martel-Pelletier, 1999). The pathological manifestations of the joints are osteophyte formation and narrowing of the joint space (Fife et al., 1991; Prescott et al., 2009). The knee OA mainly occurs in the elderly (Shamir et al., 2009). About 30% of people over the age of 65 suffer from knee OA (Bączkowicz, Majorczyk, & Kręcisz, 2015). Moreover, the incidence of knee OA is higher in females than in males, and its possibility increases with age. However, the knee OA is not limited to the elderly; a significant proportion of OA cases is associated with trauma and obesity; a large number of OA cases are idiopathic (Coggon et al., 2001).
The most common symptom of knee OA is joint pain. The pain tends to worsen with activity, specifically following a period of rest that has been known as the gelling phenomenon (Sinusas, 2012). Early detection of knee OA helps to control the disease via surgery or physical therapy (Frank, Rangayyan, & Bell, 1990). At present, knee OA is mainly diagnosed via general clinical examination including the analysis using the X-ray, computed tomography, magnetic resonance imaging (MRI), arthroscopy, and vibroarthrographic (VAG). The detection of the knee OA using the X-ray images is the most common method, and it is based on the Kellgren–Lawrence (KL) classification grades. The KL system is a validated method for classifying the individual joints into one of the five classes, where the Class 0 represents normal joint, and Class 4 denotes the most severe radiographic disease (Kellegren, J. H., 1957). Since the differences before Grade 3 are extremely limited, doctors can have different assessments of the OA and hence, draw different conclusions regarding its presence and severity. On the other hand, the MRI technology can be susceptible to the detection of the articular cartilage surface and soft tissue. The cartilage MRI diagnosis remains dependent on visual assessments, so there is a risk that small early OA lesions may not be detected (Cashman, Kitney, Gariba, & Carter, 2002). Further, arthroscopy is a rod-shaped optical instrument with a diameter of about 5 mm for observing the internal structure of joints. It represents an endoscope used to diagnose and treat joint diseases. Arthroscopy is equipped with a lens at the end of a thin tube. The diagnosis is conducted inserting the thin tube into a joint, and then the structure inside the joint is displayed on a monitor; thus, the cartilage inside the joint can be directly observed. Although arthroscopy improves the accuracy of the OA diagnosis, it should be avoided since it is invasive and can cause complications (Polly et al., 1988).

The traditional diagnosis of joint disease is based on listening to the sound and feeling the vibrations (Kernohan et al., 1990). Due to the recent development of advanced technologies, such as digital signal processing and machine learning, the measurement and analysis accuracy of VAG signals has dramatically improved (Ota et al., 2019). There has been a new technology based on acceleration instead of palpation, where the diagnosis is conducted analyzing the VAG signals collected during the knee flexion and extension motion. Compared with the healthy patients, those with articular cartilage, meniscus, and ligament have the VAG signal with unique characteristics in time series. However, these characteristics are surrounded by noise. Even when the noise is reduced, a part of the signal is still lost. Since the generation mechanism of the VAG signals is complex and random, the range of the signal-to-noise ratio of these signals cannot be determined a priori (Wu, Krishnan, & Rangayyan, 2010). Therefore, we choose the empirical mode decomposition (EMD) for signal preprocessing. For the time–frequency analysis of the VAG signal, different from commonly used traditional Fourier transform (FT) and short-time-Fourier-transform (STFT), in this work, we use the wavelet transform. Namely, the FT decomposes a signal into its frequency components, so the information in the time domain is lost, while the STFT provides the frequency components in the local time intervals of fixed duration. However, using both of them, it is very difficult to analyze signals that contain non-periodic and fast transient features, which is why we use the continuous wavelet transform (CWT) that can provide good time resolution and relatively poor frequency resolution at high frequencies and good frequency resolution and relatively poor time resolution at low frequencies (Percival, 2010).

In this paper, a combined method of the EMD and CWT is proposed to improve the diagnostic accuracy of knee OA.

2. Methods
2.1 VAG signal acquisition

As shown in Fig. 1, the VAG signal can be recorded using an accelerometer. The procedure is as follows: the sensor is attached to the subject’s tibia where 15cm under from the patella-center using medical double-sided adhesive tapes; the subject is required to hold arms tightly and legs open from shoulder width sitting on a height of 420 mm chair without armrests, stand up as steady as possible in 2 seconds (Ota, Susumu, et al., 2016). The purpose of holding arms tightly is to reduce the effects of the upper limb shake when the subject is standing up. The accelerometer sensor measures the acceleration and deceleration amplitude of the knee joint during the course of flexion-extension. Moreover, during the measurement of the VAG signal, an electronic goniometer is used to measure the bending angle of the knee joint.

We collected the raw signal by an accelerometer sensor (BW21SG2, Fuji Ceramics), a conversion connector (BNCP-ZR, Fuji Ceramics), and pre-amplifier (CA201, Fuji Ceramics) with a sampling rate of 25 kHz. The signal was digitized...
with a resolution of 12 bit per sample using a PC oscilloscope (picoscope 4424, Pico Technology) and the picoscope 6 software (Pico Technology). Each subject is measured 4 times, and the available data closest to 2 seconds is selected.

Fig. 1 The knee joint VAG data acquisition setup. Channel 2 records the VAG signal at the middle position of the patella; Channel 1 records the goniometer voltage (angle) information. (a) sitting position with arms held tightly (b) standing position with arms held tightly. (c) shows the front view of the position of accelerometer

2.2 VAG signal preprocessing and analysis

In clinical applications, the use of VAG signals for computer-assisted diagnosis and analysis of knee joint pathology signal requires exceptional high quality and accuracy. However, since the sensor is attached to the skin on the patella surface, it can acquire several different types of signals, including the mechanomyogram (MMG) and electromyogram (EMG) signals, ambient interference, and random noise (Cai et al., 2012). Due to the participation of several muscles in the sitting–standing movement, the EMG and MMG signals are readily induced by the flexion and extension movements (Islam, Sundaraj, Ahmad, & Ahamed, 2013).

2.2.1 Empirical mode decomposition

EMD is a method for the analysis of nonlinear and nonstationary signals (Huang, Shen, & Long, 1999). The EMD decomposes a given signal into a set of intrinsic mode functions (IMFs) using a sifting process according to the algorithm mentioned in Huang et al., (1998). Accordingly, signal \( x(t) \) is adaptively decomposed into the IMFs \( c_i(t) \), \( i = 1, 2, \ldots, k \) and residue \( r(t) \), as follows:

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

(1)

where \( k \) denotes the number of IMF functions. The residue \( r(t) \) reflects the average trend of a signal \( x(t) \) or a constant value. The upper envelope \( E_u(t) \) and lower envelope \( E_l(t) \) are obtained using the local maxima and minima of the signal \( x(t) \), respectively. The mean of the upper and lower envelope is expressed as

\[
m_1(t) = (E_u(t) + E_l(t))/2
\]  

(2)

By subtracting the mean from the original signal we have

\[
h_1(t) = x(t) - m_1(t)
\]  

(3)

Then, it is checked whether \( h_1(t) \) satisfies the conditions for the IMFs. The above steps are repeated until \( h_1(t) \) is an IMF. The first IMF after \( n \) iterations is expressed as
\[ c_i(t) = h_{i(t-1)}(t) - m_{in}(t) \quad (4) \]

The first residue \( r_1(t) \) can be calculated as
\[ r_1(t) = x(t) - c_1(t) \quad (5) \]

Repeat the whole algorithm with \( r_1(t), r_2(t), \ldots \) until the residue is a monotonic function. After \( k \) iterations \( x(t) \) is decomposed according to Eq. (1).

### 2.2.2 Reconstructed signal

The value and noise of the raw signal are distributed in the IMFs; thus, it is necessary to remove the noise to reconstruct the signal. In this paper, the mean squared error (MSE) and peak signal-to-noise ratio (PSNR) measures are used to select an IMF and then reassemble the signal.

The MSE assesses the quality of an intrinsic mode function \( c_i(t) \), and is expressed as
\[ MSE(c_i(t)) = \frac{1}{N} \sum_{n=1}^{N} (x(t) - c_i(t))^2 \quad (6) \]

The PSNR (in dB) is defined using the MSE as
\[ PSNR(c_i(t)) = 20 \cdot \log_{10} \left( \frac{PPV(c_i(t))}{\text{MSE}(c_i(t))} \right) \quad (7) \]

where \( PPV(c_i(t)) \) of an intrinsic mode function \( c_i(t) \) is defined as the difference between its maximum positive peak and minimum negative peak that is given by
\[ PPV(c_i(t)) = \text{Max}(c_i(t)) - \text{Min}(c_i(t)) \quad (8) \]

Generally, the filters perform denoising process through the subtraction operation. For instance, a median filter of a time series will smoothen the signal and lose some of them. The disadvantage is that it is difficult to analytically treat the effect of a median filter. This denoising method is not suitable for time series such as VAG signals when the high-frequency features should be classified. Thus, in this work, addition is chosen to reconstruct the signal. The two smallest MSE values and the largest PSNR values of all the IMFs, if they are the same, then select the next one. These three IMFs are chosen. They are reconstructed, and the signal \( y(t) \) is obtained.

### 2.3 Feature extraction

#### 2.3.1 Feature extraction using CWT

To observe the frequency characteristics, a CWT is performed on the reconstructed signal. The CWT transforms the reconstructed signal \( y(t) \) into a function \( D(s, p) \) of two real variables \( s \) and \( p \) that is given by
\[ D(s, p) = \int_{-\infty}^{\infty} y(t) \psi_{s,p}(t) \, dt \quad (9) \]

Function \( \psi_{s,p}(t) \) in Eq. (9) is referred to as the wavelet function, and is defined by translating along the time axis and stretching a so-called “mother wavelet” function \( \Psi(t) \) that is expressed as
\[ \psi_{s,p}(t) = \frac{1}{\sqrt{s}} \psi \left( \frac{t-p}{s} \right) \quad (10) \]

where \( s \) is the scale, and \( p \) is the shift of the wavelet in the time domain. The wavelets can be real or complex functions. For the application studied in this paper, a real wavelet is used (D. B. Percival and A. T. Walden, 2010). The Morlet wavelet is given by
\[ \Psi(t) = e^{-\frac{t^2}{2\sigma^2}} \cos(2\pi f_0 t) \quad (11) \]
where \( \sigma \) denotes the approximate bandwidth and \( f_0 \) is the center frequency. Although the device supports frequencies up to 2000 Hz, the sample rate of the signal is 25000 Hz, so we decided to divide the frequency domain to 250 scales for 100Hz per scale.

The CWT coefficients \( X \) are represented using a matrix with rows and columns and times and scales. In order to facilitate the following calculations, \( X \) needs to be normalized in the range \((0,1)\); the normalized matrix is denoted as \( Y \), which is expressed as

\[
Y = \frac{X - X_{\min}}{X_{\max} - X_{\min}}
\]

where \( X_{\min} \) and \( X_{\max} \) denote the minimum and maximum value of \( X \), respectively. Thus, the CWT coefficients \( Y \) in the intervals of 100 Hz below the frequency of 2500 Hz are summarized into a feature.

2.3.2 Feature extraction in the spatiotemporal domain

To detect any difference in a reconstructed signal from the abnormal and normal, the features need to be extracted in the spatiotemporal domain that include the skewness and kurtosis of the autocorrelation function (ACF) values (Box & Jenkins, 1990). The ACF \( A(\tau) \) of the reconstructed signal \( y(t) \) is calculated according to the following formulas:

\[
A(\tau) = \frac{1}{N-\tau} \sum_{i=0}^{N-\tau-1} y(i)y(i+\tau)
\]

where \( N \) is the number of samples.

\[
kurtosis = \frac{\mu^4}{\sigma^4}
\]

where \( \mu^4 \) denotes the fourth moment about the mean \( \mu \) of \( A(\tau) \), and \( \sigma \) denotes the standard deviation of \( A(\tau) \).

\[
skewness = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{y(i) - \mu}{\sigma} \right)^3
\]

2.4 Least squares support vector machine (LSSVM) classifier

The LSSVM method introduces the idea of a square sum of errors into the objective function of the standard SVM, where a training data consisting of \( n \) samples is considered.

\[
Train = \{(u_i, v_i) | i = 1, 2, \cdots, n\}, \quad \text{where} \quad u_i \in \mathbb{R}^d \quad \text{and} \quad v_i \in \mathbb{R}
\]

\( \Phi(.) \) is employed to project the input-sample space to the feature space \( \Phi(u) = \{\phi(u_1), \cdots, \phi(u_n)\} \) and construct the following optimal linear decision function (Gencoglu & Uyar, 2009):

\[
y(u) = \omega^T \phi(u) + \beta
\]

Using the structural risk formula, we have:

\[
R = \frac{1}{2} ||\omega||^2 + C R_{\text{emp}}
\]

The weight vector \( \omega \) and offset \( \beta \) can be calculated using Eq. (16). In Eq. (17), \( C \) and \( R_{\text{emp}} \) denote the penalty factor and loss function, respectively. In the LSSVM, the loss function is always equal to the quadratic loss function, which is expressed as

\[
R_{\text{emp}} = \sum_{i=1}^{n} \xi_i^2
\]

where \( \xi_i \) denotes the deviation degree of a misclassification sample from the ideal sample. Then, obtaining the solutions of \( \omega \) and \( \beta \) is equivalent to solving the optimization problem given by

\[
\min R = \frac{1}{2} ||\omega||^2 + C \sum_{i=1}^{n} \xi_i^2
\]
s.t. \( v_i = \omega^T \varphi(u_i) + \beta + \xi_i, \{i = 1, 2, \ldots, n\} \) (20)

By introducing the Lagrangian multiplier \( g = \{g_1, g_2, \ldots, g_n\} \) into Eqs. (19) and (20), we get

\[
L(\omega, \beta, \xi_i, \alpha) = \frac{1}{2} ||\omega||^2 + C \sum_{i=1}^{n} \xi_i^2 - \sum_{i=1}^{n} (g_i(\omega^T \varphi(u_i) + \beta + \xi_i - v_i))
\] (21)

Next, we calculate the derivative of every factor in Eq. (21). To eliminate \( \omega \) and \( \xi \), we introduce the equations \( \omega = \sum_{i=1}^{n} g_i \varphi(u_i) \) and \( \xi_i = \frac{g_i}{2c} \) into Eq. (21). Then, it can be written that

\[
v_i = \sum_{j=1}^{n} (a_j \cdot < \varphi(u_j), \varphi(u_i) >) + \beta + \frac{g_i}{2c}
\] (22)

Suppose that there exists a kernel function \( K(u_i, u_j) = < \varphi(u_i), \varphi(u_j) > \) that satisfies the Mercer condition (Suykens, Lukas, Van Dooren, De Moor, & Vandewalle, 1999). Then, Eq. (22) can be re-written as

\[
v_i = \sum_{j=1}^{n} (g_j K(u_i, u_j)) + \beta + \frac{g_i}{2c}
\] (23)

Solving Eq. (23), the model parameters \( \beta \) and \( [g_1, g_2, \ldots, g_n]^T \) are obtained. Then, the decision function of the LSSVM is expressed as

\[
v(u) = \text{sgn} \left[ \sum_{i=1}^{n} \left( a_i K(u_i, u_i) \right) + \beta \right]
\] (24)

As it is well-known, different types of kernel function \( K(u_i, u_j) \), such as the polynomial kernel function, sigmoid kernel function, linear kernel function, and Gaussian Radial Basis Function (GRBF), lead to different LSSVM performances. In this work, the Gaussian RBF is used because there is only one parameter compared with the other kernel so that it will work well in practice and is relatively easy to calibrate (Land et al., 2011). The Gaussian RBF kernel is given by

\[
K(u_i, u_j) = e^{-\frac{|u_i - u_j|^2}{2\epsilon^2}}
\] (25)

Since \(|u_i - u_j|\) is always a positive value, kernel achieves a maximum output when \(|u_i - u_j| = 0\), and then, approach to zero as \(|u_i - u_j|\) increases. The speed of this approach depends on \( \epsilon \), it is faster at smaller \( \epsilon \) and vice versa. Thus, the factor \( \epsilon \) should be optimized, and we use a heuristic algorithm to determine its optimal value.

2.5 Classification and ensemble systems

To classify the reconstructed signal as normal or abnormal by an LSSVM classifier, a large set of normal and abnormal signals was used in the training process. The training was performed using the feature vectors having the two features extracted from the ACF of the reconstructed signal: \( \{1\} \) kurtosis of ACF, \( \{2\} \) skewness of ACF with four CWT coefficient features extracted from the frequency bands, \( \{3\} \) CWT coefficient in the range of 100–200 Hz, \( \{4\} \) CWT coefficient in the range of 200–500 Hz, \( \{5\} \) CWT coefficient in the range of 500–1000 Hz, and \( \{6\} \) CWT coefficient in the range of 1000–2500 Hz.

| Feature set | Feature of reconstructed signal |
|-------------|--------------------------------|
| G_a         | \{1\}                        |
| G_b         | \{2\}                        |
| G_c         | \{3\}                        |
| G_d         | \{4\}                        |
| G_e         | \{5\}                        |

The VAG signals were reconstructed using 25 patients suffering from knee OA and 26 young and healthy people. The features were extracted using the presented algorithm. We combined the six features into five sets to verify their
accuracy, as presented in Table 1. In these five sets of features, 10 data were randomly selected for the normal and abnormal signals for the LSSVM training. The remaining data were used for testing.

Since the data amount was not enough, the sensitivity and specificity of each set of features were calculated using 100 simulations. In each LSSVM simulation, training and test data were randomly selected. To evaluate the accuracy of each group of features after mechanical learning, the receiver operating characteristic curve (ROC curve) was used. The ROC analysis was used as it can provide essential information about diagnostic test performance; namely, the closer the apex of the curve toward the upper left corner, the higher the discriminatory ability of the test i.e., the true-positive rate is high, and the false-positive [1-specificity] rate is low (Fan, Upadhye, & Worster, 2006). The area under the ROC curve (AUC) can provide a way to measure the accuracy of a diagnostic test. The larger the area, the more accurate the diagnostic test (Zhu, Zeng, & Wang, 2010).

The AUC of the ROC curve, where \( FP = (1- \text{ specificity}) \) and \( ROC (FP) \) is sensitivity, was calculated using

\[
AUC = \int_0^1 ROC(FP) \, d(FP)
\]  

(26)

The classification accuracy of commonly used AUC ranging in the diagnostic test is summarized in Table 2.

| AUC range       | Classification accuracy |
|-----------------|-------------------------|
| 0.9 < AUC < 1.0 | A (Excellent)           |
| 0.8 < AUC < 0.9 | B (Good)                |
| 0.7 < AUC < 0.8 | C (Worthless)           |
| AUC < 0.7       | D (Not good)            |

Besides, the classification accuracy is mostly expressed as

\[
Accuracy = \frac{TP + TN}{TP + FN + TN + FP}
\]  

(27)

where \( TP \) denotes the true-positive, \( TN \) denotes the true-negative, and \( FN \) denotes the false-negative.

3. Results

We used a median filter to compare the denoised signal with the reconstructed signal, as shown in Fig. 2.
line denotes the reconstructed signal, and the blue line is the signal denoised with the median filter: (a), (c) the knee joint VAG signal of a patient (female) with Grade III knee OA; (b), (d) the knee joint VAG signal of a healthy subject.

From the perspective of the dynamic time warping calculation, the similarity of the signal reduced by the median filter was better than that of the reconstructed signal. The DWT distances of the reconstructed signal of the patient with knee OA and the healthy subjects were 9370 mV and 6725 mV, respectively. The DWT distances of the signal denoised with the median filter of the patient with knee OA and the healthy subject were 7292 mV and 4197 mV, respectively. However, as shown in Fig. 2, obviously, the reconstructed signal retains more peak details to facilitate feature extraction in the next step.

As shown in Fig. 3, there were some differences in the ACF features between the healthy subjects and patients with knee OA. However, a part of them overlapped; thus, it was difficult to classify based on just the ACF features.

![ACF features box plot](image)

Fig. 3 The box plot of ACF features— the red box denotes the ACF kurtosis of the patient (female) with Grade III knee OA; the black box denotes the ACF kurtosis of the healthy subject; the blue box denotes the ACF skewness of the patient (female) with Grade III knee OA; the green box denotes the ACF skewness of the healthy subject.

![CWT coefficients](image)

Fig. 4 The horizontal axis denotes the frequency, and the vertical axis denotes the peak of the CWT coefficients: (a) the patient (female) with Grade III knee OA, (b) the healthy subject.

Using the CWT calculation, we obtained the peak value of the CWT at each frequency range of the flexion and extension actions as shown in Fig. 4. The ratios of the CWT coefficients average values of the 100–200 Hz interval, 200–500 Hz interval, 500–1000 Hz interval, and interval above 1000 Hz to the overall average value were calculated as
features. The ratio of the healthy people was higher in the low-frequency band, while that of patients with knee OA was higher in the high-frequency band.

The classification accuracy rate in the percentage of the Gaussian RBF kernel LSSVM with the $G_a$ feature set was 86.67% that was higher than the accuracy obtained using other feature sets using the same classifier. According to the results presented in Table 3, the AUC obtained using the $G_a$ feature set was 0.9965 with a standard error (SE) of 0.02156 as illustrated in the Fig. 5.

Table 3 Classification results of the different feature sets and ensemble systems on the Gaussian RBF kernel LSSVM.

| Feature set | Accuracy (%) | AUC      | SE      |
|-------------|--------------|----------|---------|
| $G_a$       | 86.67        | 0.9965   | 0.02156 |
| $G_b$       | 76.83        | 0.9744   | 0.02465 |
| $G_c$       | 80.22        | 0.9533   | 0.01993 |
| $G_d$       | 74.19        | 0.8407   | 0.02421 |
| $G_e$       | 77.98        | 0.8935   | 0.0368  |

Fig. 5 The ROC curves provided using the five feature sets with the Gaussian RBF kernel LSSVM classifiers in the ensemble ($p < 0.002$)

In the ensemble experiments, the highest accuracy provided using the feature set was 86.67% ($G_a$), whereas, the lowest was only 74.19% ($G_d$). On the other hand, the highest AUC provided was 0.9965 ($G_a$), whereas, the lowest was only 0.8407 ($G_d$).

Referring to the Table 1 and Fig. 4, the superiority of $G_a$ ensemble in the diagnostic performance was prominent, with the highest overall accuracy, the best AUC value, and the second-lowest SE value; $G_b$ had the second highest AUC (0.9744) and SE values, but the accuracy (76.83%) was not very high. Although $G_c$ had the lowest SE value, the accuracy and AUC values were on the intermediate levels.

According to the analysis, $G_d$ provided the worst result. However, the accuracy was higher than 70%, the AUC value exceeded 0.8, and it was evaluated as B (good). However, the discrimination of the feature sets in the low-frequency band was not very good, and it was worse than that of the feature sets in the high-frequency band. The best result had the accuracy of more than 85%, and the AUC value, higher than 0.99 that was evaluated as A (excellent).

4. Discussion

These accuracy rates of $G_a$ are not lower than those we can find in the published research works, where the values obtained for accuracy are in the range of 84–89%. Moreover, the AUC values of our feature sets are higher, the highest...
value even exceeds 0.99 (Sharma, M., et al, 2019; Athavale, et al, 2020; Nalband, S., et al, 2018).

The reason for the good result of $G_a$ was the CWT coefficients of the three frequencies, the low, medium, and high, were applied as features. It can be seen in Fig. 4 that when only ACF features were missing, the impact on the results was not great. However, when the features of one of the frequency bands were missing, the accuracy and AUC values were greatly reduced, specifically in the high-frequency band.

In view of the excellent performance of the Gaussian RBF kernel LSSVM method, using $G_a$ feature set, we used this method to determine the knee OA level, i.e., to distinguish Grade III and Grade IV OA. The results are presented in Fig. 6, where it can be seen that the sensitivity rate was 48% and the specificity rate was 98%, and the AUC value was only 0.762. The percentages of correctly predicted Grade III and IV knee OA were 4.0% (PV−) and 34.7% (PV+), respectively, which was worse than blind guessing. The reason for such results could be scarce amount of training data, resulting in insufficient machine learning. On the other hand, all the VAG signals corresponded to the late stage of knee OA, so the similarity between Grade III and IV knee OA was too high.

![Fig. 6 The ROC curve analysis for prediction of Grade III and IV knee OA](image)

Although we got a good result, we still need to discuss how we can use the method to diagnose patients in early-stages of knee OA. As the number of patients with early-stages of knee OA rarely, we collected only one VAG signal from Grade II patient. Even though this patient had more pain than Grade III, the X-ray diagnosis report was only Grade II. That’s why this patient was found in early stage. Thus, we put the data into the machine learning test and got the 56.22% possibility of healthy (P<0.01). The result was obviously false negative, because of the signal collected by the sensor is very similar to the healthy.

That’s very difficult to find the cause from overly sparse samples, and mostly the pain in the early stage is inconspicuous, easily overlooked. Nevertheless, we still suspect that it is likely that the frequency range of CWT coefficients is insufficient because we did not study the CWT coefficients of reconstructing signals by EMD below 100Hz and above 2000Hz. In addition, some articles have proposed to use the features of high frequency or ultra-high frequency to improve classification accuracy (Mascaro, B., et al, 2009). This consistent with our finding that the high frequency feature set has a more significant influence on classification.

The reason we don’t use higher frequency in this paper is we will have a heavy calculation burden in reconstruct signal process with the increase of the sampling rate, and there is some serious mode mixing phenomenon in EMD (Wang, Tong, et al, 2012). Next, we are going to try to solve the mode mixing problem by add white noise into the original signal and enhance overall computing speed by a perfect algorithm to extract features from a wider frequency range.

5. Conclusions

In this study, the knee VAG signals that were obtained using an accelerometer and a goniometer were used in the
knee OA detection. The VAG signals were input into a personal computer and reconstructed by the EMD. To obtain the feature set, the autocorrelation function and continuous wavelet transform (CWT) process were applied to the reconstructed signal. Finally, the classification was performed using the LSSVM and the overall evaluation was performed using the ROC curve. The results showed that high accuracy and AUC value were achieved. Consequently, a combined EMD and CWT analysis method of the knee VAG signal that can monitor the knee joint pathology provides an economic alternative to the osteoarthritis patients. However, further research on OA grade classification and its early-stage diagnosis is needed.

References

Athavale, Yashodhan, and Sridhar Krishnan., A telehealth system framework for assessing knee-joint conditions using vibroarthrographic signals., Biomedical Signal Processing and Control, Vol.55, (2020),101580

Bączkowicz, D., Majorczyk, E. and Kręcisz, K., Age-related impairment of quality of joint motion in vibroarthrographic signal analysis, BioMed Research International, (2015), DOI:10.1155/2015/591707.

Box, G. E. and Jenkins, G., Time Series Analysis (1990), Forecasting and Control.

Cai, S., Wu, Y., Xiang, N., Zhong, Z., He, J., Shi, L. and Xu, F., Detrending knee joint vibration signals with a cascade moving average filter, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (2012), DOI:10.1109/EMBC.2012.6346931.

Cashman, P. M., Kitney, R. I., Gariba, M. A. and Carter, M. E., Automated techniques for visualization and mapping of articular cartilage in MR images of the osteoarthritic knee: A base technique for the assessment of microdamage and submicro damage, IEEE Transactions on Nanobioscience, Vol. 99, No. 1 (2002), DOI:10.1109/TNB.2002.806916.

Coggan, D., Reading, I., Croft, P., McLaren, M., Barrett, D. and Cooper, C., Knee osteoarthritis and obesity, International Journal of Obesity, Vol. 25, No. 5 (2001), DOI:10.1038/sj.ijo.0801585.

Percival, D. B. and Walden, A. T., Wavelet Methods for Time Series Analysis (2010), Cambridge University Press.

Fan, J., Upadhye, S., and Worster, A., Understanding receiver operating characteristic (ROC) curves, Canadian Journal of Emergency Medicine, Vol.8, No.1 (2006), pp.19–20.

Fife, R. S., Brandt, K. D., Braunstein, E. M., Katz, B. P., Shelbourne, K. D., Kalasinski, L. A. and Ryan, S., Relationship between arthroscopic evidence of cartilage damage and radiographic evidence of joint space narrowing in early osteoarthritis of the knee, Arthritis & Rheumatism: Official Journal of the American College of Rheumatology, Vol.34, No.4 (1991), pp.377–382.

Frank, C. B., Rangayyan, R. M. and Bell, G. D., Analysis of Knee Joint Sound Signals for Non-Invasive Diagnosis of Cartilage Pathology, IEEE Engineering in Medicine and Biology Magazine, Vol. 9, No. 1 (1990), DOI:10.1109/51.62910.

Gencoglu, M. T. and Uyar, M., Prediction of flashover voltage of insulators using least squares support vector machines, Expert Systems with Applications, Vol. 36, No. 7 (2009), DOI:10.1016/j.eswa.2009.02.021.

Huang, N. E., Shen, Z. and Long, S. R., A new view of nonlinear water waves: the Hilbert spectrum, Annual Review of Fluid Mechanics, Vol. 31, No. 1 (1999), DOI:10.1146/annurev.fluid.31.1.417.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H., 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, Vol. 454, No. 1971 (1998), DOI:10.1098/rspa.1998.0193.

Islam, M. A., Sundaraj, K., Ahmad, R. B. and Ahamed, N. U., Mechanomyogram for muscle function assessment: a review, PLoS one, Vol. 8, No. 3. (2013), DOI:10.1371/journal.pone.0058902.
Kellegren, J. H. and J. S. L., Radiological assessment of osteoarthritis., Annals of the Rheumatic Diseases, Vol.16, No.4 (1957), pp.494–501.

Kernohan, W. G., Beverland, D. E., McCoy, G. F., Hamilton, A., Watson, P., and Mollan, R.A.B., Vibration arthrometry, Acta Orthopaedica, Vol. 61, No. 1 (1990), DOI:10.3109/17453679008993071.

Land, W. H., Qiao, X., Margolis, D. E., Ford, W. S., Paquette, C. T., Perez-Rogers, J. F., Borgia, J. A., Yang, J. Y. and Deng, Y., Kernelized partial least squares for feature reduction and classification of gene microarray data, BMC Systems Biology, Vol. 5, No .3 (2011), DOI:10.1186/1752-0509-5-S3-S13.

Martel-Pelletier, J., Pathophysiology of osteoarthritis, Osteoarthritis and Cartilage, Vol.96, (1999), pp.371–373.

Land, W. H., Qiao, X., Margolis, D. E., Ford, W. S., Paquette, C. T., Perez-Rogers, J. F., Borgia, J. A., Yang, J. Y. and Deng, Y., Kernelized partial least squares for feature reduction and classification of gene microarray data, BMC Systems Biology, Vol. 5, No .3 (2011), DOI:10.1186/1752-0509-5-S3-S13.

Mascaro, B., Prior, J., Shark, L. K., Selfe, J., Cole, P., & Goodacre, J. Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints. Medical engineering & physics, Vol. 31, NO.8. (2009), pp.1013-1022.

Nalband, S., Valliappan, C. A., Prince, A. A., & Agrawal, A., Time-frequency based feature extraction for the analysis of vibroarthrographic signals, Computers & Electrical Engineering, Vol.69, (2018), pp.720-731.

Ota, S., Ando, A., Tozawa, Y., Nakamura, T., Okamoto, S., Sakai, T., & Hase, K. Preliminary study of optimal measurement location on vibroarthrography for classification of patients with knee osteoarthritis. Journal of physical therapy science, Vol. 28, (2016), No.10, pp.2904-2908.

Ota, S., Fujita, R., Segawa, N., Tanaka, R., Inagawa, S. and Sakai, T., Clinical analytical vibroarthrography method for classifying patients with knee osteoarthritis, Osteoarthritis and Cartilage, Vol.27, (2019), DOI:10.1016/j.joca.2019.02.171.

Percival, D. B., Wavelet Methods for Time Series Analysis (2010), Cambridge University Press.

Prescott, J. W., Pennell, M., Best, T. M., Swanson, M. S., Haq, F., Jackson, R. and Gurcan, M. N., An automated method to segment the femur for osteoarthritis research, In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (2009), DOI:10.1109/IEMBS.2009.5333257.

Shamir, L., Ling, S. M., Scott Jr, W. W., Bos, A., Orlov, N., MacUra, T. J., Eckley, D. M., Ferrucci, L. and Goldberg, I. G., Knee X-ray image analysis method for automated detection of osteoarthritis, IEEE Transactions on Biomedical Engineering, Vol. 56, No. 2 (2009), DOI:10.1109/TBME.2008.2060625.

Sharma, M., Sharma, P., Pachori, R. B., & Gadre., Double Density Dual-Tree Complex Wavelet Transform-Based Features for Automated Screening of Knee-Joint Vibroarthrographic Signals, Machine Intelligence and Signal Analysis, Springer, Singapore, (2019), pp. 279-290.

Sinusas, K., Osteoarthritis: Diagnosis and treatment, American Family Physician, Vol. 85, No. 1 (2012), DOI:10.1136/bmj.1.5222.355-a.

Suykens, J. A. K., Lukas, L., Van Dooren, P., De Moor, B. and Vandewalle, J., Least Squares Support Vector Machine Classifiers: a large scale algorithm, European Conference on Circuit Theory and Design, Vol.99, (1999), pp.839–842.

Vigorita, V. J., Orthopaedic pathology (2008), Lippincott Williams & Wilkins.

Wang, T., Zhang, M., Yu, Q., & Zhang, H. Comparing the applications of EMD and EEMD on time–frequency analysis of seismic signal. Journal of Applied Geophysics, Vol.83, (2012), pp.29-34.

Zhu, W., Zeng, N. and Wang, N., Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations, NESUG proceedings: health care and life sciences, Baltimore, Maryland, Vol.19, (2010), pp.1–9.