Visnyk NTUU KPI Seriia – Radiotekhnika Radioaparatobuduvannia, 2024, Iss. 97, pp. 46–57

UDC 616.12-07

Neural Networks Detection of Low-Amplitude Components on ECG Using Modified Wavelet Transform

Mnevets A. V., Ivanushkina N. G.

Electronic Engineering Department, Faculty of Electronics,
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine

E-mail: *amnevec-ee22@lll.kpi.ua, niva-ee@lll.kpi.ua*

This study is devoted to identification of low amplitude components from ECG signals by different time-frequency analysis methods when main power spectrum falls on high-amplitude components. It was also analyzed the problem of choosing correct scale system for determination low-amplitude components on the scalogram by artificial intelligence models. As a result of the study, several modifications of the continuous wavelet transform were proposed. First modification was based on the use of a scaling function and a modified wavelet. Second modification was based on the use of cosine similarity at each iteration of convolution followed by the use of a scaling function. The main idea of the study was to modify the wavelet transform in such a way as to select the components which has the target amplitude and reduce all other components that complicate the neural networks analysis of the interested fragments of the signal. Also, possible procedures for signal restoring were proposed for preserving the effect of using scaling modifications. The testing of the proposed modified algorithms was carried out on the basis of artificially created signals as well as on the basis of real ECG signals with late potentials superimposed on them. Visual analysis of scalograms and signal reconstructions obtained using the modified wavelet transform showed that the modified wavelet transform is capable of extracting low-amplitude components from the signal with much greater spectral power than the transform without modifications. In addition, the ability of common neural network models to distinguish between cardiac cycles with and without late potentials was tested. As a result, it was found that models that were trained on scalograms obtained using a modified wavelet transform train faster and are less susceptible to local minima stucking. The results of classification of signals with and without late potentials based on trained neural network models showed that training using scalograms obtained on the base of a modified wavelet transform allows achieving 99% classification accuracy, which is 1-49% more than that using scalograms obtained on the base of on the classical wavelet transform.

Keywords: electrocardiography; wavelet transform; late atrial potentials; late ventricular potentials; neural networks

DOI: 10.20535/RADAP.2024.97.46-57

Introduction

Modern heart electrical activity analyzing methods make it possible to identify ECG pathologies of various types by detecting abnormalities in different locations and representations. Electrocardiogram (ECG), like other biological signals, can contain diagnostically significant information in time, frequency, phase and other domains. Moreover, sometimes the combined representations like time-frequency are better suited for diagnosing some types of arrhythmias, for example atrial fibrillation [1].

Analysis in different representations of ECG has advantages and disadvantages. For example, algorithms based on time-amplitude analysis are easy to interpret by doctors and often do not require a lot of computing resources. But, noise components of various frequencies, superimposed on the analyzed signal can significantly distort its shape and lead to an incorrect evaluation.

Frequency analysis allows to decompose a signal by its frequency components that increase the resistance to noise, but sometimes it is very difficult to distinguish QRS complex components of close frequencies especially with different amplitudes [2].

The combined time-frequency domain takes into account both frequency and time information simultaneously. The representation of this type of analysis is carried out in the space of three coordinates: time, frequency and some amplitude characteristic, which displays the ECG signal energy at a certain frequency and point in time. The most common visuali-

zation tool for this type of representation is scalogram. Despite the ability to analyze a signal in the frequency and time domains simultaneously, this type of analysis also has its drawbacks.

One of the drawbacks besides the high computational complexity is that the time-frequency representation of the signal carries a large amount of information (coefficients), which can lead to complicate assessment. It is also very difficult to obtain high time and frequency resolution at the same time. Also, it is needs to sacrifice frequency resolution to increase time resolution and vice versa.

For example, in short-time Fourier transform (STFT), the time-frequency composition of the entire signal represents a sequence of short-term frequency spectra by small fragments. It is not difficult to conclude that spectral components on ECG whose duration is less than the window width could be lost in this transformation. On the other hand, with a further decreasing the window width, the frequency resolution simultaneously decreases. Theoretically, this behavior is described by the Heisenberg uncertainty principle, which states that it is impossible to simultaneously achieve high resolution in time and frequency [3]. But in practice, STFT also has a number of other losses. For example, at the edges of window functions with STFT, a leakage effect is observed, which is caused by the blurring of the signal energy across adjacent frequencies when the signal is time limited [4]. To solve this problem, special window functions are used. Another popular approach is to use multiple window functions for analysis in a method called multi-taper STFT.

There are also conversions aimed at minimizing leakage effects and maximizing time and frequency resolution. For example, the Wigner-Ville transformation. This transformation is excellent for analyzing signals that have a small number of frequency components in their composition, but as the frequency composition of the signal increases, for example on ECG signal, the influence of cross terms of the transformation increases also, which leads to distortions in the time and frequency domains. To minimize the problem of cross terms, modifications of the Wigner-Ville transformation can be used [5].

To analyze non-stationary ECG signals that have a lot of low amplitude components and the zone of interest localized over a short period of time, continuous wavelet transform (CWT) is used preferably [6]. This type of transformation gives the ability to represent ECG signal in a time-frequency domain, but at the same time has much less losses compared to STFT. Also, the wavelet transform is less receptive to edge effects than STFT.

But wavelet analysis has some disadvantages. Firstly, the result of the transformation depends significantly on the choice of wavelet and scaling coefficients, which can complicate the interpretation. Secondly, from the principle of the wavelet transform, at small values of scales, the low frequency resolution persists [6]. With a large number of coefficients, the computational complexity of automated detectors increases. Due to the large number of coefficients, the probability of machine learning overfitting and the problems of gradient fading increases also.

Therefore, it is necessary to develop an artificial intelligent method or refine existing neural networks approaches to obtaining time-frequency representations of ECG signal that can improve the quality of processing and pathologies determination in ECG and others biomedical signals using modern algorithms.

1 Formulation of the problem

To assess the quality of proposed transformations a pathological ECG signal was generated and time-frequency conversions was carried out using various methods (Fig. 1):

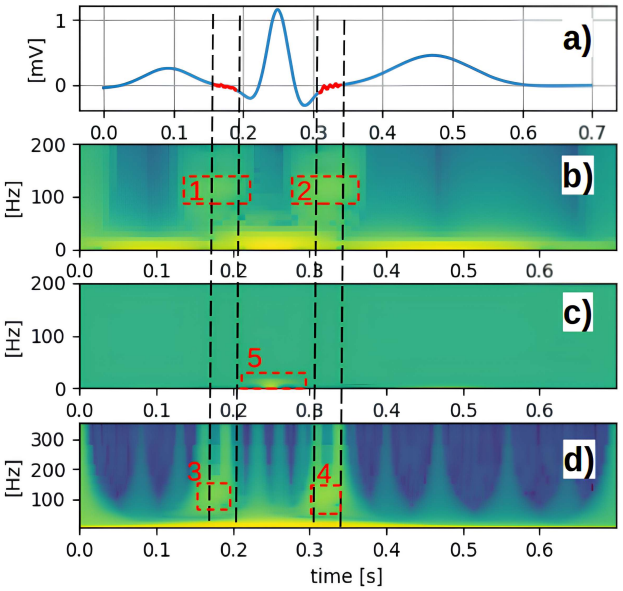

Fig. 1. Time-frequency representation of the QRS complex with late atrial and ventricular potentials; cardiac cycle of pathological ECG signal with late potentials highlighted in red (a); STFT scalogram (b); Wigner-Vile scalogram (c); CWT scalogram (d); approximate localization of late atrial potentials on STFT and CWT (1, 2, 3, 4); manifestation of R peak on Wigner-Ville scalogram (5)

Signals of late ventricular and atrial potentials were superimposed on the artificial QRS complex of the ECG signal, so that their location, amplitude and duration corresponded to the biological aspects of this types of the low amplitude components manifestation.

The presence of late potentials on the ECG can be a sign of cardiomyopathy, atrial fibrillation and other life-threatening pathologies. Late potentials occur as a result of conduction disturbances in the myocardium as a result of structural changes, disruption of

ion channels, or myocardial metabolism. Therefore, timely identification of late potentials allows to begin timely treatment and, as a result, improve the patient's quality of life [7].

An artificial signal of late potentials was generated as a superposition of several sinusoids in the frequency range of 80-120 Hz, phase-shifted among themselves and having a spread in amplitude. The sampling frequency of the resulting signal (Fig. 1a) was 1000 Hz, and the duration was 0.7 seconds.

The STFT scalogram was constructed using a Hann window [8] of 100 samples. A high window overlap value (90%) was chosen due to the fact that the fragments of interest have a short duration.

The CWT scalogram was constructed based on the Morlet wavelet with a central frequency parameter ω_0=6 [rad/sample], the number of scale coefficients for analysis was chosen 198 from 3 to 200 inclusive. The Wigner-Ville algorithm was built on the basis of a common approaches for generating a time-frequency representation of a signal [9].

In addition, for more discriminative ability, the coefficients of time-frequency representations on scalograms were provided on a logarithmic scale, since without the use of logarithms, the manifestation of late ventricular and atrial potentials on scalograms were almost indistinguishable.

Analyzing the obtained representations (Fig. 1), it can be observed that on the STFT scalogram (Fig. 1b) manifestations of late ventricular and atrial potentials are observed in the required frequency range approximately (80-120 Hz) (Fig. 1 1, 2). As a result of selecting a large overlap, fuzzy boundaries of late potentials in the time domain were obtained. In addition, these boundaries are difficult to distinguish even on a logarithmic scale.

When analyzing the Wigner-Vile distribution (Fig. 1c), even on a logarithmic scale, only the main component of the QRS complex, (R peak), is clearly observed (Fig. 1 5).

On the scalogram constructed using CWT (Fig. 1d), the manifestation of late atrial and ventricular potentials (Fig. 1 3, 4) has more accurate localization in time. In the frequency domain, late potentials are also more clearly expressed than with STFT.

Thus, it can be noted that the resulting manifestations of late potentials on scalograms are difficult to distinguish even on a logarithmic scale. Therefore, identifying late potentials and other low amplitude features using time-frequency representation is a difficult task even for neural networks.

To solve these types of problems, it is necessary to have some kind of algorithm that, like an optical microscope, can clearly focus on components of the required amplitude and analyze only them. Using this approach, it would be possible to significantly clear the time-frequency representation from the distortions and make this representation much easier for artificial intelligence analysis. To implement this approach, amplitude analysis methods such as Empirical Mode Decomposition [10], Hilbert-Huang Transform [11], and Singular Spectrum Analysis [12] may be suitable. But these methods are based on decomposing signal by modes, where it is not possible to set the preferred amplitude range in which the zone of interest lies. Furthermore, short-term late potentials will be difficult to select as a signal pattern, especially in the presence of noise.

2 Materials and methods

To construct the required algorithm, the continuous wavelet transform was taken as a basis. The process of this transformation can be described by the following formula:

$$CWT\left(s,\tau\right)=\frac{1}{\sqrt{|s|}}\int_{-\infty}^{\infty}x\left(t\right)*\ \psi^{*}\left(\frac{t-\tau}{s}\right)dt, \quad (1)$$

where s, τ – the required wavelet transform coefficients, $x\left(t\right)$ – analyzed signal, ψ^{*} – complex conjugate function to ψ (where ψ – mother wavelet), s – scale factor, τ – time shift parameter, t – time variable.

In this formula $\frac{1}{\sqrt{|s|}}$ is a normalizing factor, which ensures that the norm of the scaled and time-shifted wavelet is kept equal to 1, that is $||\psi_{s,\tau}(t)|| = ||\psi(t)|| = 1$ and parameter s should be bigger than 0.

In this study, the modification of the wavelet transform is performed using the Morlet wavelet. The Morlet wavelet was chosen due to its good localization in the time and frequency domains, flexibility and accuracy in tuning the wavelet frequency, and similarity to late ventricular and atrial potentials. So, the Morlet wavelet can be described by the following formula:

$$\psi\left(t\right)=\pi^{-\frac{1}{4}}e^{i\omega_0 t}e^{-\frac{t^2}{2}}, \quad (2)$$

where $\psi\left(t\right)$ – Morlet wavelet function of time (t), ω_0 – central frequency of wavelet (corresponds to the highest energy concentration in the Morlet wavelet), $e^{i\omega_0 t}$ – complex exponential which generates sine wave, $e^{-\frac{t^2}{2}}$ – function of the Gaussian window that modulates the sine wave, ensuring that the wavelet has a finite length. Member $\pi^{-\frac{1}{4}}$ represents a normalization factor that included to equation to make the Morlet wavelet wavelet energy equals to 1.

To reconstruct the original signal $x\left(t\right)$ from the obtained CWT coefficients $X(s,\tau)$, the following formula can be used:

$$x\left(t\right)=C_{\psi}^{-1}\int_{0}^{\infty}\int_{-\infty}^{\infty}X(s,\tau)\,\frac{1}{\sqrt{|s|}}\widetilde{\psi}\left(\frac{t-\tau}{s}\right)d\tau\,\frac{ds}{s^2}, \quad (3)$$

where $x\left(t\right)$ – the initial signal, C_{ψ}^{-1} – normalization coefficient depends on wavelet ψ, $X\left(s,\tau\right)$ – coefficients,

obtained after CWT of signal $x(t)$, $\frac{1}{\sqrt{|s|}}$ – normalization factor, $\widetilde{\psi}(t)$ is a dual function of $\psi(t)$ which is used for inverse continuous wavelet transform. Integration in this case is carried out according to two parameters, first according to the shift parameter (τ) then according to the scaling parameter (s).

The Morlet wavelet itself is usually used as the dual function [13]. Thus, in the case of Morlet wavelet, the inverse wavelet transform can be performed using the same function as the forward wavelet transform.

For the inverse continuous wavelet transform, the normalization coefficient could be calculated by the following formula:

$$C_\psi = \frac{\int_{-\infty}^{\infty} |\widehat{\psi}(\omega)|^2}{|\omega|} d\omega, \tag{4}$$

where $|\widehat{\psi}(\omega)|$ – Fourier transform of wavelet function $\psi(t)$, ω – angular frequency, related to f by the equation $\omega = 2\pi f$. For inverse continuous wavelet transformation in discrete form, simplification is very often used to calculate the normalization coefficient [14]. Then the formula will take the form:

$$C_\psi = \sum (|\psi(t)|^2). \tag{5}$$

This simplification directly sums the squares of the modulus of the wavelet function in the time domain, instead of integrating its Fourier transform. This formula describes the normalization coefficient in a practical approximation sufficient to restore the initial signal.

Since the feature of the Morlet wavelet is that it has its own central frequency, conversion from scales to frequencies for discrete signals can be done using the following formula:

$$f_a = \frac{\omega_0}{2\pi a}, \tag{6}$$

where a – scale of the wavelet, ω_0 – central frequency of Morlet wavelet in [rad/sec].

But the recalculation formula also could depend on the direct implementation of the wavelet transform. If the Morlet wavelet depends not on time ($\psi(t)$) but on discrete sample ($\psi[n]$), then ω_0 can be expressed in [rad/sample]. Then the formula for converting scales to frequency will be:

$$f_a = \frac{\omega_0}{2\pi a \Delta t}, \tag{7}$$

where a – scale of the wavelet, ω_0 – central frequency of Morlet wavelet in [rad/sample], Δt – time interval between signal counts. The choice of scale system also affects the quality of the wavelet transform.

Fig. 2 shows the distribution of Morlet wavelet spectrum depending on the selected scale system. The x-axis has the logarithmic scale for more detailed display. The first graph (Fig. 2a), obtained using a linear scale system ($s = n$, $n \in [0 : N]$), shows that the spectra overlap significantly at low frequencies and have large gaps in high frequencies. To arrange the spectra of wavelet functions more homogeneous, it can be used a logarithmic scale system (Fig. 2b):

$$s_k = s_{min} * r^k, \; k \in [0, N-1], \tag{8}$$

where s_k – logarithmic scale, s_{min} – minimal scale, r – scaling coefficient.

If the largest (s_{max}) and smallest (s_{min}) scale was known within which we want to scale the wavelet so for the required number (N) of scales the scaling coefficient could be calculated as:

$$r = \left(\frac{s_{max}}{s_{min}}\right)^{\frac{1}{N-1}}. \tag{9}$$

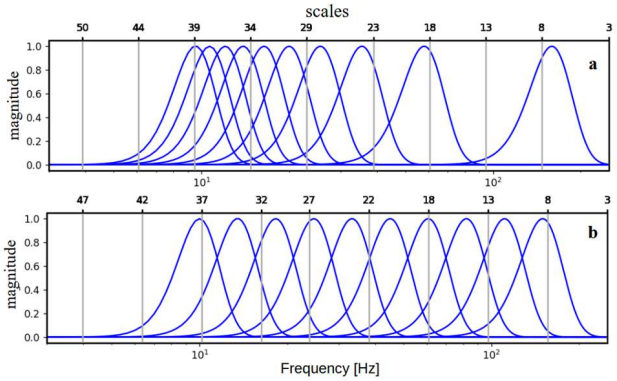

Fig. 2. Distribution of Morlet wavelet spectrum using different scale systems: distribution using a linear scaling system (a); distribution using logarithmic scaling system (b)

With discrete transform, the maximum wavelet frequency is limited by the sampling frequency, and according to the Nyquist-Shannon theorem it is $f_s/2$. At the same time, the minimum frequency is limited by the signal length and depends on wavelet scaling. The maximum wavelet transform frequency is $1/T_{signal} = f_s/N_{samples}$, where f_s – sampling frequency. Also, the minimum and maximum frequency of a signal can be described by the frequency range of the signal itself. Using (6) or (7), it can be expressed maximum and minimum scale, which according to formulas (8), (9) can be turned into a uniformly spaced sequence of logarithmic scales.

To test the proposed algorithms, the MIT-BIH Arrhythmia Database [15] and Paroxysmal Atrial Fibrillation Events Detection from Dynamic ECG Recordings [16] databases were used. The first database was chosen due to the presence of annotated cardiac cycles and other cardiac events. The second database was chosen because it contains 100 ECG recordings of healthy people with a duration more than three minutes and high sampling frequency (about 500 Hz).

Also, to analyze the results and modifications, similarity measurement functions such as Pearson correlation [17] and cosine similarity [18] were used, and binary

classification scores, such as sensitivity, specificity, and overall accuracy were calculated as well [19].

To analyze the features selection quality, based on the proposed methods, common neural network architectures were selected [20], such as ResNet, DenseNet, MobileNet v2, MobileNet v3, ShuffleNet, Resnext 50.

3 Modifications

Since the basis of the CWT (1) function is a cross-correlation, and the coefficients of the continuous wavelet transform express the cross-correlation coefficients between the original signal and the shifted scaled mother wavelet [21], the obtained correlation coefficients were filtered using some correction function (f_{scale}). Then the formula for the direct Wavelet transform with this modification will be:

$$CWT\left(s,\tau\right) = \\ = \int_{-\infty}^{\infty} x\left(t\right) * \psi'^{\left(\frac{t-\tau}{s}\right)} * \ f_{scale}\left(x\left(t\right), \psi'\left(t-\tau\right)/s\right) dt, \tag{10}$$

where $f_{scale}\left(x\left(t\right), \psi'\left(t-\tau\right)/s\right)$ – a scaling function that depends on the current value ($x\left(t\right)$) and the current value of the wavelet ($\psi'\left(t\right)$) depending on the value of the scale factor s and offset τ, ψ' – a modified wavelet that normalizes the amplitude of the scaled wavelet function to a certain target amplitude A, which must be highlighted from the signal.

The main idea is to compare at each iteration the amplitude of each signal component with a modified wavelet ψ' that has a target amplitude A. Using the scaling function f_{scale}, in the case of similar amplitudes, the value of $CWT\ \left(s,\tau\right)$ will be amplified, and in the case of a difference, the coefficients will be reduced or reset to zero.

The modified wavelet $\psi'(t)$ can be described by the following formula:

$$\psi'(t) = A \times \frac{\frac{1}{\sqrt{|s|}}\psi\left(\left(t-\tau\right)/s\right)}{max\left(\left|\frac{1}{\sqrt{|s|}}\psi\left(\left(t-\tau\right)/s\right)\right|\right)}, \tag{11}$$

where A – target amplitude, $\psi\left(\left(t-\tau\right)/s\right)$ – time shifted and scaled mother wavelet.

The following function may be suitable for f_{scale} in (10):

$$f_{scale}\left(a,b\right) = \left(\frac{\min\left(|a|,|b|\right)}{\max\left(|a|,|b|\right)}\right)^{q}, \tag{12}$$

where a, b – values of comparing, q – quality factor, which characterizes the degree of function increasing according to degree of similarity a and b.

As can be seen in Fig. 3, the more the a/b ratio moves away from 1, the smaller scaling factor becomes. Values of the q parameter set the rate at which the scaling factor decreases, making the bandwidth of acceptable amplitude ratios narrower as the quality factor q increases.

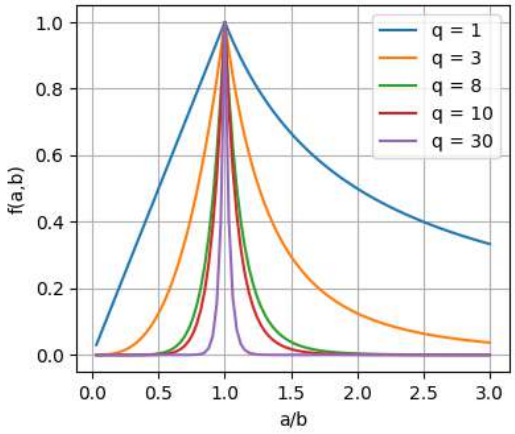

Fig. 3. Dependence of the f_{scale} function on the a/b ratio for different values of the quality factor q

Thus, the closer the amplitude of the compared values, the higher the scaling factor f_{scale} will be. This graph is symmetric about 0, and have the same distribution for negative a/b. To more precisely control the range of target amplitudes, it can be used other types of windows. For example, to isolate components from a signal in the range of 5-20 mV, a target amplitude A can be selected as 10 mV and apply a unit rectangular window in the range a/b in [0.5-2]. In such a case, components beyond 5-20 mV will be suppressed by the modified CWT. Also, to minimize distortions associated with the choice of a rectangular window, other types of windows could be used, for example the Tukey window [8].

To restore the signal with saving the influence of the scaling function, it can be used the standard formula of the inverse wavelet transform (3). But since the inverse transform uses an unmodified wavelet, and the direct modified transform used a scaled wavelet ψ', the influence of the modified wavelet can be corrected using the following formula:

$$x\left(t\right) = C_{\psi}^{-1} \int_{0}^{\infty} \int_{-\infty}^{\infty} X\left(s,\tau\right) \frac{1}{\sqrt{|s|}} \widetilde{\psi}\left(\frac{t-\tau}{s}\right) \frac{max\left(\left|\frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right)\right|\right)}{A} \psi'\left(\frac{t-\tau}{s}\right) d\tau \frac{ds}{s^2}, \tag{13}$$

where $x(t)$ – the initial signal, C_ψ^{-1} – normalization coefficient, $X(s,\tau)$ – coefficients obtained after modified CWT of signal x(t), $\frac{1}{\sqrt{|s|}}$ – normalization factor, $\psi(t)$ – mother wavelet, which uses for direct wavelet transformation, $\psi'(t)$ – modified wavelet using (11).

For testing proposed CWT modification, two artificial tested signals were created. The tested signal (Fig. 4c) was created as subsequence of three heterogeneous sinusoidal signals. The first component of this signal was a 30 Hz sine wave at unity amplitude with high frequency noise, making the sine wave distorted and spread in the frequency domain. The second component was a sinusoid with a smoothly varying frequency of 1-10 Hz with unity amplitude. This makes this signal fragment non-stationary in the frequency domain. The third component was sinusoid with a frequency of 30 Hz and low amplitude, about 0.025. The sampling rate of tested signal was 500 Hz and duration about 1.7 seconds. The component with low amplitude is 40 times smaller than other signal components. Thus, the main power spectrum of the scalogram fall on the high-amplitude components of the signal, and the low-amplitude component was indistinguishable.

Using the modified CWT (10), it is possible to significantly increase the scale resolution of the low-amplitude component on the scalogram (Fig. 4a) and significantly increase its amplitude during inverse transformation (13) (Fig. 4d). For the modified transformation obtained in Fig. 4a, the Morlet wavelet was chosen $w_0 = 6$ rad/sample, a linear system of 50 scaling coefficients within the range of 10 to 200 scales was used, with a target amplitude of $A = 0.025$ in (11) and a quality factor of $q = 5$ for (12).

When analyzing the reconstructed signal (Fig. 4d) as well as the CWT scalogram (Fig. 4a), it was noticed that high-amplitude components was not sufficiently suppressed and some fluctuations are observed on the graph (Fig. 4d) within 0.5-1.25 seconds. These fluctuations can be caused by the fact that, over a certain time range, the analyzing modified wavelet function can coincide in amplitude and partially in shape with the analyzed signal. To smooth out these fluctuations, a modification that will be described further may be useful.

The other signal for testing (Fig. 4e) was an artificial cardiac cycle with LVP, LAP, where the late potentials are a superposition of several sinusoids in the frequency range of 80-120 Hz and amplitude about 35 uV, like in Fig. 1a. The sampling rate of this signal was 500 Hz and duration about 0.8 seconds. For modified CWT (10) the linear 50 scales system was used within the range of 3 to 200 scales, with a target amplitude of $A = 0.035$ and quality factor of $q = 5$.

Analyzing the coefficients of the modified CWT (Fig. 4b), obtained from the test signal of the cardiac cycle with late potentials (Fig. 4e), it can be noticed that although there are distinguishable peaks in the frequency range of 80-100 Hz, the energy of other components with higher amplitudes is not sufficiently suppressed. A similar result can be observed when analyzing the reconstructed signal (Fig. 4f). Although there are some peaks and fluctuations within time of late potentials, the selection of late potentials from the signal is insufficient.

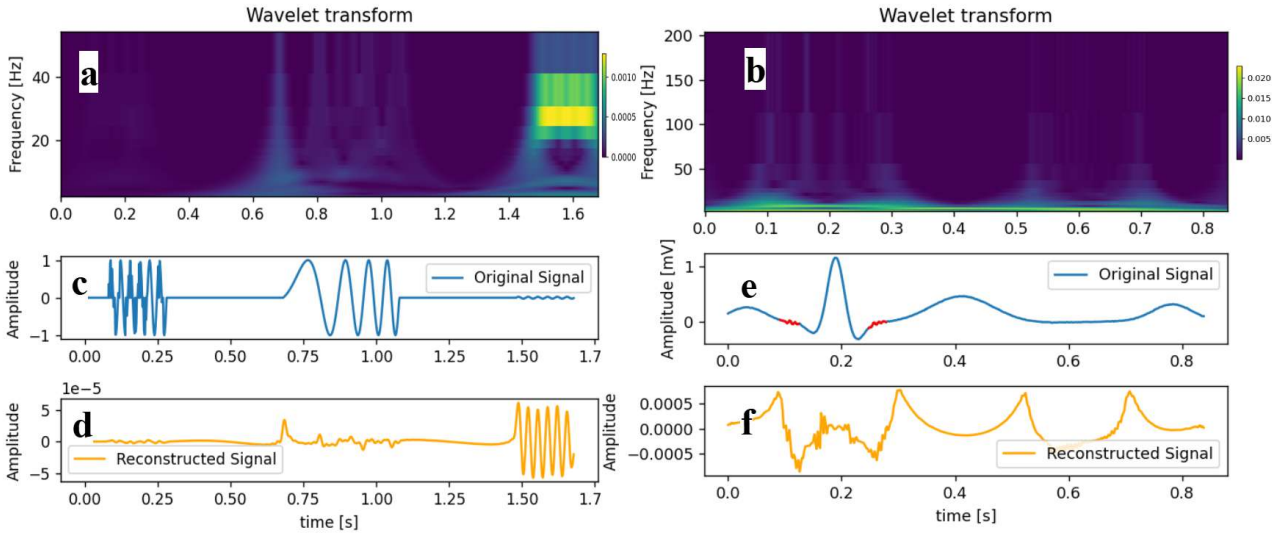

Fig. 4. Results of continuous wavelet transform and signal reconstruction, obtained using a modified wavelet transform algorithm (10); scalogram of CWT coefficients for heterogeneous sinusoidal signal (a); scalogram for signal of ECG cardiac cycle with LVP and LAP (b); original heterogeneous sinusoidal signal (c); reconstruction obtained by modified inverse transform (13) for heterogeneous sinusoidal signal (d); original signal of cardiac cycle with LVP and LAP (e); reconstruction obtained by modified inverse transform (13) for signal of cardiac cycle with LVP and LAP (f)

This difference in the results between the two test signals can be explained by the fact that the sinusoidal signal (Fig. 4c) is very similar in shape to the analyzing Morlet wavelet. In addition, the Morlet wavelet and fragments of the test signal with sinusoids (Fig. 4c) oscillate around 0, which gives a greater coincidence of amplitudes and, as a result, a greater value of the modified wavelet transform coefficients. On the other hand, although the late potentials are close in shape, frequency and amplitude to the analyzing wavelet, on the test signal (Fig. 4e) they are shifted in amplitude relative to 0. This key difference from a signal with sinusoids leads to the fact that the scaling function inaccurate calculates scaling coefficients and, as a result, does not effectively identify components with the target amplitude. In this case, an approach is needed that is less sensitive to fluctuations in the amplitude of the analyzed signal.

To minimize distortions associated with the amplitude shifting of the analyzing wavelet and analyzed signal, it could possible to replace the procedure of calculating cross-correlation with a cosine similarity in the wavelet transform:

$$CWT_{cos}(s,\tau) = \frac{\int_{-\infty}^{\infty} x(t) * \psi'^{*}\left(\frac{t-\tau}{s}\right) dt}{\sqrt{\int_{-\infty}^{\infty} |x(t)|^2 \, dt} \sqrt{\int_{-\infty}^{\infty} \left|\psi'\left(\frac{t-\tau}{s}\right)\right|^2 dt}}, \quad (14)$$

where $x(t)$ – analyzed signal, $\psi'(t)$ – modified wavelet (11), $\psi'^{*}(t)$ – complex conjugate of modified wavelet.

Thus, the result of the wavelet transform will be a measure of cosine similarity between the analyzed signal and the modified shifted and scaled mother wavelet. Then the scaling function will be applied as follows:

$$CWT_{modif}(s,\ \tau) = \frac{\int_{-\infty}^{\infty} x(t) * \psi'^{*}\left(\frac{t-\tau}{s}\right) dt}{\sqrt{\int_{-\infty}^{\infty} |x(t)|^2 \, dt} \sqrt{\int_{-\infty}^{\infty} \left|\psi'\left(\frac{t-\tau}{s}\right)\right|^2 dt}} \times \times \int_{-\infty}^{\infty} f_{scale}\left(x(t) - corect(t), A\right) dt, \quad (15)$$

where x(t) – analyzed signal, $\psi'(t)$ – modified wavelet (11), $\psi'^{*}(t)$ – complex conjugate of modified wavelet, f_{scale} – scale function (12), A – target amplitude of the selecting components, $corect(t)$ – correction function that can adjust the offsets of interested signal fragments relative to 0.

In (15), compared to (10), the integral value of the scaling function is multiplied by the result of the cosine similarity. This procedure is intended to simplify calculations and to smooth out the scaling factor. It also eliminates conversion noise caused by rapid fluctuations of the signal under study and the characteristics of the scanning wavelet. For simplify demonstration of some operations of this modified transform using a fragment of artificial non-stationary signal, the calculation scheme was presented (Fig. 5)

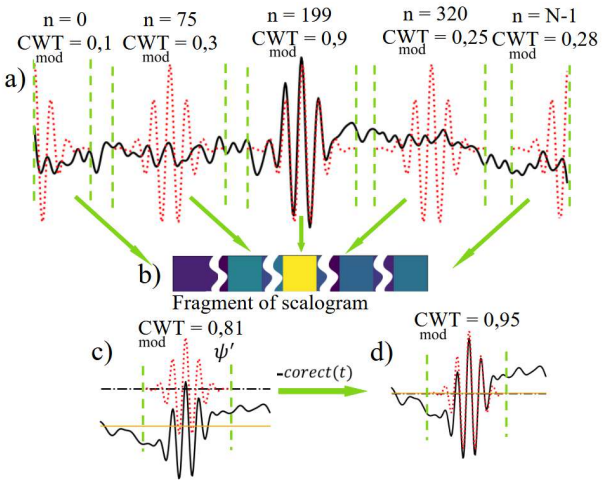

Fig. 5. Calculation peculiarities scheme of modified CWT; coefficients calculation based on the modified wavelet ψ' (a); an example of building a scalogram based on calculated coefficients (b); example of modified CWT coefficient calculation without correction function (c), example of modified CWT coefficient calculation using the correction function (d).

Since the Morlet wavelet is modeled by a Gaussian envelope (3), and based on the fact that 99.7% of the entire area of the Gaussian curve is concentrated in the range of $\pm 3\sigma$, it is possible to determine the length of the wavelet in seconds or samples, cut it and compare it with a fragment of the analyzed signal with the corresponding duration. The proposed algorithm for calculating the modified CWT for a discrete signal is based on this principle.

Fig. 5a shows how the scanning wavelet shifts along the analyzed signal, and at each iteration CWT_{modif} is calculated started from certain sample n.

Based on the obtained coefficients, a fragment of the scalogram (Fig. 5b) is constructed for a given scales. Also Fig. 5c, d shows how the correction function affects the result of the coefficient calculating. On Fig. 5c the low-frequency baseline drift distorts the signal fragment and makes it shifted relative to zero. Fig. 5d shows the result of the correction function applying. The correction function eliminates the influence of baseline drift allowing the signal and scanning wavelet be compared more accurately. In the simplest case, the correction function removes the mean from the compared signal fragment on each iteration, making its values symmetrical relative to 0.

Restoring a signal while maintaining amplitude selection transformation can be a computationally difficult task that requires taking into account the influence of all scaling and correction functions and may be the subject of additional research. This study proposes

a simplified recovery procedure:

$$x(t) = C_{\psi}^{-1} \times \\ \times \int_0^{\infty} \int_{-\infty}^{\infty} CWT_{modif}(s,\tau) \psi' \left(\frac{t-\tau}{s} \right) CWT(s,\tau) d\tau \frac{ds}{s^2}, \tag{16}$$

where $CWT_{modif}(s,\tau)$ – coefficients of modified wavelet transform (15), ψ' – modified wavelet (11), $CWT(s,\tau)$ – coefficients of continuous wavelet transform (1), C_{ψ}^{-1} normalization coefficient depends on wavelet ψ.

The main principle of reconstruction (16) is that since the obtained coefficients of the modified CWT lie in the range from 0 to 1, which makes them compared to the scaling coefficients for classical CWT. In the case of maximum similarity of amplitudes and shape, the energy of the classical CWT will be transferred in full, but if the amplitude and shape are different, the energy of the classical CWT will be reduced by the modified CWT coefficients.

4 Results and discussion

To evaluate the results of the modified wavelet, transform on the base of the algorithm described above, testing was carried out using an artificial and real ECG signal with late ventricular and atrial potentials as a target low amplitude components.

The artificial signal for testing (Fig. 6b) was generated as QRS complex with added LVP, LAP, where the late potentials are a superposition of several sinusoids in the frequency range of 80-120 Hz and amplitude about 35 uV, like in Fig. 1a. and Fig. 4e. The duration of artificial signal about 0.6 sec and sampling rate is 1000 Hz.

The real signal with sampling frequency of 500 Hz was chosen from MIT-BIH Arrhythmia Database. The presence of R peaks annotated by doctors in this database reduces the averaged cardiac cycle construction error associated with incorrectly detected R peaks by automated algorithms. Using the real signal of healthy person from this database, the late ventricular potentials was added in 0.035 second after each R peak. Using this modified signal, the averaged QRS complex with VLP was generated (Fig. 6e). The late potentials were generated as a hyperposition of action potentials which was simulated using parallel conductance model [22] for a number of cells.

For the artificial signal, a modified CWT with a Morlet wavelet was applied with $\omega_0 = 6$ [rad/sample], the target amplitude A was chosen as 0.035 and quality factor $q = 4$, and the subtraction of the average from signal part was chosen as a correction function at each iteration. The scale system was chosen logarithmic (9) for the frequency range 40-200 Hz, with a number of 40 scales. For the real signal, exactly the same parameters were chosen, except for the target amplitude A, which was 0.025.

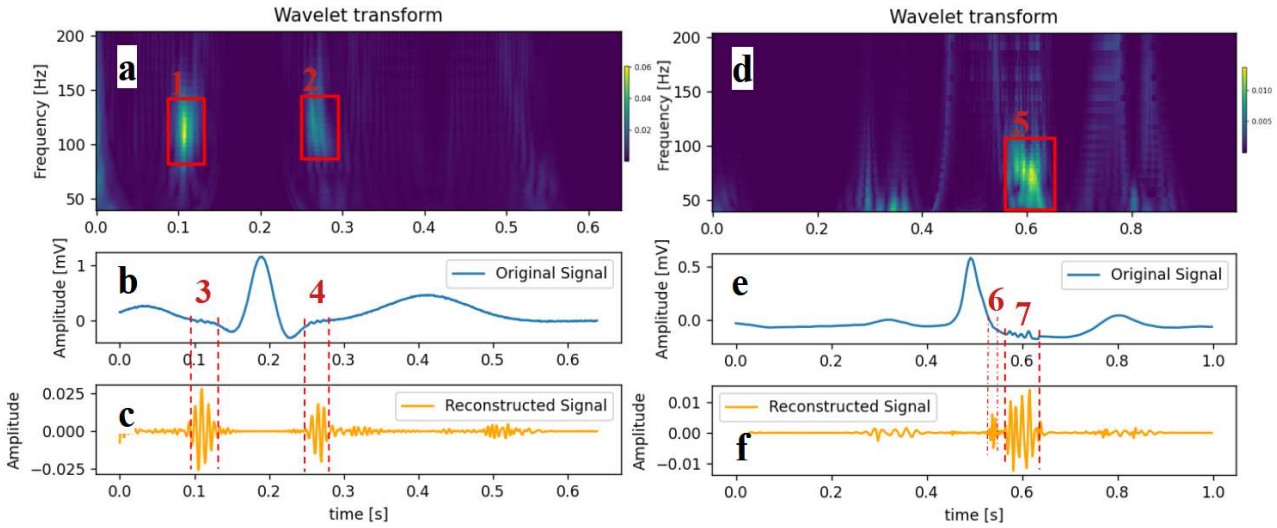

Fig. 6. Modified CWT (15) and reconstruction for signals with late potentials; modified CWT scalogram for artificial signal with VLP and ALP (a); artificial signal with VLP and ALP (b); reconstruction of artificial signal (c); modified CWT scalogram for real QRS wave signal with VLP (d); real QRS wave signal with VLP (e); reconstruction for real signal (f); manifestation of VLP and ALP on a scalogram (1, 2); manifestation of VLP and ALP on artificial signal and corresponding fragments on reconstruction (3, 4); manifestation of VLP on a scalogram (5); manifestation of VLP on real QRS wave signal and corresponding fragments on reconstruction (6, 7)

Analyzing the scalograms (Fig. 6 a, d), obtained using a modified wavelet transform (15), it can be seen that in the frequency and time ranges of the presented late potentials there are highlighted peaks of spectral power (Fig. 6 1, 2, 5). These peaks do not merge with neighboring coefficients, and high-amplitude signal components do not distort the time-frequency representation in target location. This peculiarity gives the opportunity to analyze late potentials in the time-frequency domain more accurately.

Analyzing the signal reconstructions (Fig. 6 c, f), it can be seen the clear signal bursts that correspond to the time frame of the low-amplitude components (Fig. 6 3, 4, 7), and the high-amplitude components are significantly suppressed. For comparison, Fig. 7 shows a scalogram of a classical CWT for a similar signal built according to the same parameters as the modified CWT. In this case, it can be seen that the greatest spectral power is concentrated around the high-amplitude high-frequency component of the R peak. In addition, the localization of late potentials on the scalogram is complicated and noisy due to the coefficients of other high-amplitude signal components. Proposed modified CWT allows to distinguish those oscillations that could not be detected by classical wavelet transform. For example, in Fig. 6f fluctuations are observed in the zone of late potentials (Fig. 6 6), which are indistinct in the analyzed signal and scalogram but are very clearly visible in the reconstruction. These fluctuations were not artificially imposed as in (Fig. 6 7), so it may belong to the patient whose ECG was used for this comparison.

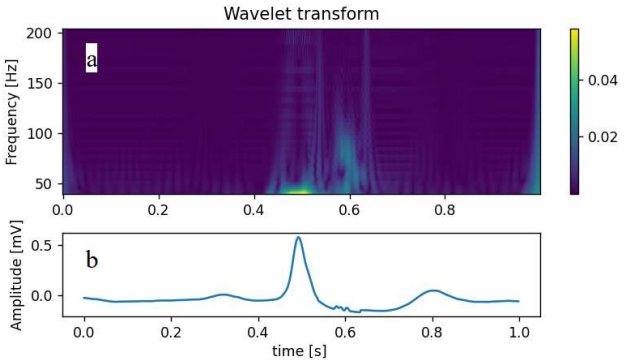

Fig. 7. The real cardiac cycle of ECG signal with LVP and CWT scalogram that was obtained using the traditional CWT with Morlet wavelet; CWT scalogram (a); real QRS complex with LVP (b).

To visually analyze the effect of noise on the artificial signal (Fig. 6b), noise with a signal to noise ratio (SNR) of 35 dB was added. This noise appears in the reconstruction (Fig. 6c) when the signal amplitude fluctuates close to zero, for example in the range of 0.4-0.6 seconds, and also partially distorts the waveform in the range of late potentials. This disadvantage can be eliminated by using more detailed scaling, different correction functions or could be solved during the further research. But in real high-resolution cardiography, the noise level on the average cardiac cycle is lower, so the influence of noise was less observed. When assessing the Pearson correlation coefficient between the late potentials on reconstructions and late potentials on the averaged cardiac cycle, it was obtained a value of 0.6. Such distortion may be caused by noise, and may also appeared in the result of scaling functions usage. Other reason could be the inverse transformation inaccuracies. Improving the quality of reconstruction when using proposed modified wavelet transform is also could be the subject of further research.

To recognize two classes (ECG with LVP and ECG without LVP) based on the constructed scalograms, the Paroxysmal Atrial Fibrillation Events Detection from Dynamic ECG Recordings database [16] was used. About 100 records with a normal 12 leads ECG were selected from this database. The generated late potentials of 0.1 mV amplitude were superimposed on each lead as in Fig. 6e. Then the average cardiac cycle was obtained using an automatic cardiac cycle detector [23].

For each averaged cardiac cycle with and without VLP, a scalogram was calculated using a modified CWT, as well as a classic CWT. For the continuous wavelet transform, the Morlet wavelet with ω_0=6[rad/sample] was chosen. A logarithmic scale system was calculated for the frequency range of 40-200 Hz and the number of scales was 40. For the modified CWT, the target amplitude A was chosen as 0.06, the quality factor $q = 2$, and the subtraction of the average value from signal part was chosen as a correction function at each iteration. Thus, one thousand scalograms for averaged cardiac cycles with LVP and one thousand scalograms without LVP was calculated using the modified CWT. And also, the same amount scalograms with and without LVP was calculated using classic CWT. Based on these data, various common neural networks models were trained. Train to test split ratio was 85%/15%. For all types of tested neural networks optimizer Adam [24] was chosen. The learning rate was set as 0.01, batch size was chosen as 70 scalograms per batch, the number of epochs was set as 10.

Analyzing the graphs of the classification accuracy (Fig. 8 a-f), it can be concluded that for all presented neural network architectures, the classification accuracy at the first training epoch when using the modified CWT is significantly higher than the accuracy using the classical CWT. The convergence rate of architectures which trains using a modified CWT was much higher. Despite the fact that the classification accuracy of both methods was approximately equal at 10th epoch, at epochs 1-4 the quality of classification was generally higher with the modified CWT. In addition, there are models that stucked in a local minimum (Fig. 8 d, f) and until tenth epoch could not improve the classification accuracy using the classic CWT. At the same time, models that were trained on

modified CWT did not get stuck in the local minima and reaches much higher classification accuracy. These results indicates that the modified CWT is capable of much more accurately identifying the required features from the signal under study.

Analyzing the classification results based on the sensitivity, specificity and overall accuracy of the trained models (Table 1), it can be seen that the models that were trained on CWT data in most cases showed higher classification accuracy than the algorithms that were trained on the basis of the classic CWT. For the neural networks MobileNet v3 and ResNet 50, ten epochs were not enough to increase the classification accuracy beyond 50% for the classic CWT. Besides that, for the modified CWT the classification accuracy was much higher and reached 96%.

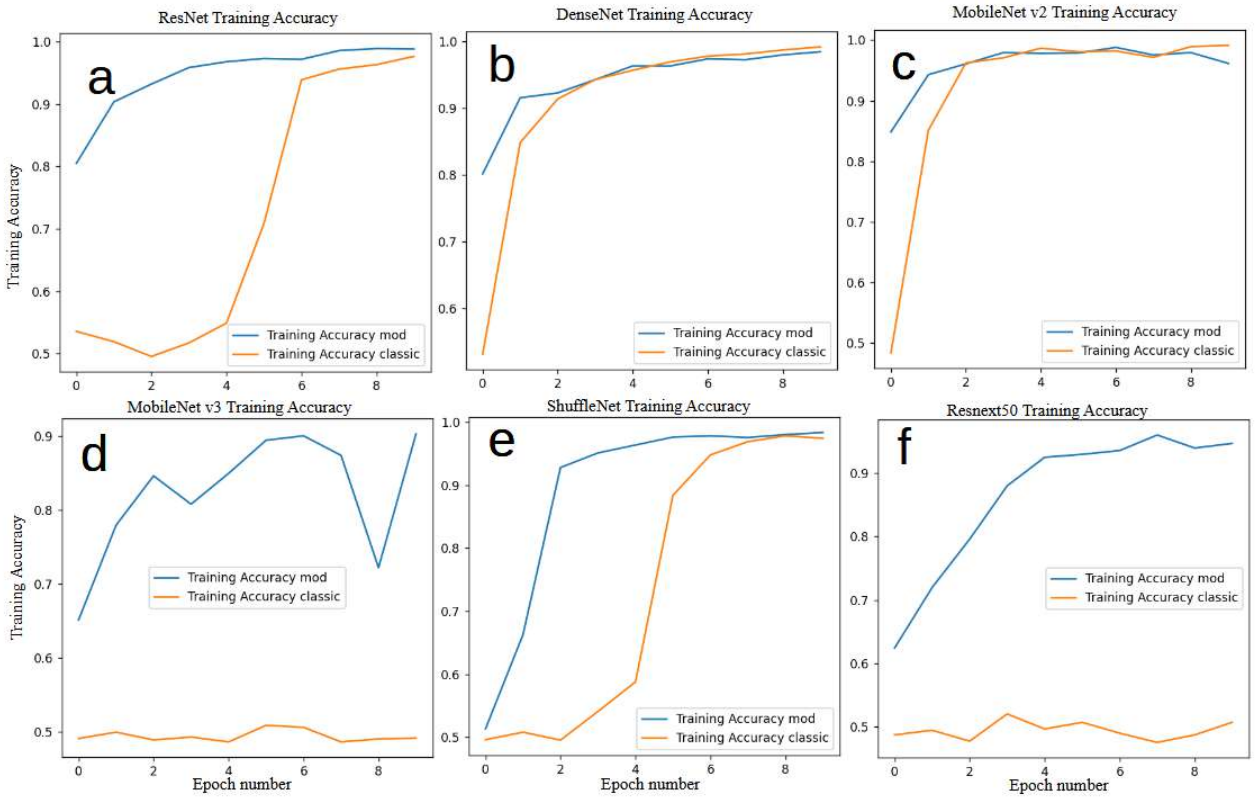

Fig. 8. Accuracy comparing of training process of different neural networks for classical and modified CWT algorithm; accuracy of ResNet network by each epoch (a); accuracy of DenseNet network (b); accuracy of MobileNet v2 network (c); accuracy of MobileNet v3 network (d); accuracy of ShuffleNet network (e); accuracy of Resnext50 network (f)

Табл. 1 Classification quality comparison of common neural network architectures trained on data generated by modified and normal CWT

Model name	Sensitivity		Specificity		Accuracy	
	norm cwt	mod cwt	norm cwt	mod cwt	norm cwt	mod cwt
ResNet	0.82	0.9	1	0.98	0.97	0.99
DenseNet	0.97	0.96	0.99	0.98	0.98	0.98
MobileNet v2	0.96	0.91	0.99	0.99	0.98	0.95
MobileNet v3	1	1	0	0.64	0.5	0.83
ShuffleNet	0.93	0.97	0.99	0.99	0.96	0.98
Resnext 50	0	0.98	1	0.94	0.5	0.96

Conclusion

This study analyzed methods of time-frequency analysis like Short-time Fourier transform (STFT), Wigner-Vile transform, Continuous wavelet transform (CWT) and also deep learning models for low components detection in ECG.

For increasing low components detection ability, the CWT modifications were proposed. Such modifications as scaling function implementation, wavelet function

modification, as well as the use of cosine similarity in the process of constructing CWT coefficients have shown their effectiveness in the task of identifying low-amplitude components from artificial and real signals.

Based on the analysis of the issue of choosing a scaling system for CWT, a logarithmic scale system was chosen and described for more detailed localization of low-amplitude components.

The modifications that were proposed for the wavelet transform procedure showed the ability to identify signal components with the target amplitudes and reduce other non-target components.

The results of testing various models of neural networks as well as visual analysis showed that the proposed modified wavelet transform algorithm is capable of identifying low-amplitude components from the time-frequency representation of the signal, where almost the entire power spectrum is occupied by other high-amplitude components. For example, the proposed modifications can achieve a classification accuracy of 99% and increase the accuracy of neural networks training obtained with scalograms based on the classical wavelet transform for 1-49%.

The proposed method also helps to increase the speed and stability of neural network training.

The suggested modifications of the wavelet transform can be widely used in medicine to identify low-amplitude components of the ECG and other biological signals, as well as in other industries.

References

[1] Hu Y., Zhao Y., Liu J., Pang J., Zhang C., and Li P. (2020). An effective frequency-domain feature of atrial fibrillation based on time–frequency analysis. *BMC Med. Inform. Decis. Making*, Vol. 20, No. 1, DOI: 10.1186/s12911-020-01337-1.

[2] Ceschi R., Gautier J.-L. (2017). *Fourier Transform. CHAPTER 2 In: Fourier Analysis*. Wiley, pp. 39-96, DOI: 10.1002/9781119388944.ch2.

[3] Borisagar K. R., Thanki R. M., and Sedani B. S. (2018). *Fourier Transform, Short-Time Fourier Transform, and Wavelet Transform. In: Speech Enhancement Techniques for Digital Hearing Aids*. Springer International Publishing, pp. 63-74. DOI: 10.1007/978-3-319-96821-6_4.

[4] Lyon D. (2009). The Discrete Fourier Transform, Part 4: Spectral Leakage. *J. Object Technol.*, Vol. 8, No. 7, p. 23. DOI: 10.5381/jot.2009.8.7.c2.

[5] Athanassoulis A. G., Mauser N. J., and Paul T. (2009). Coarse-scale representations and smoothed Wigner transforms. *J. de Mathématiques Pures et Appliquées*, Vol. 91, No. 3, pp. 296-338. DOI: 10.1016/j.matpur.2009.01.001.

[6] Addison P. S. (2005). Wavelet transforms and the ECG: a review. *Physiol. Meas.*, Vol. 26, No. 5, pp. R155-R199. DOI: 10.1088/0967-3334/26/5/r01.

[7] Saito D. et al. (2022). Atrial late potentials are associated with atrial fibrillation recurrence after catheter ablation. *J. Arrhythmia*, Vol. 38, Iss. 6, pp. 991-996. DOI: 10.1002/joa3.12786.

[8] Prabhu K. M. M. (2018). Review of Window Functions. Chapter in: *Window Functions and Their Applications in Signal Processing*. CRC Press, pp. 87-127. DOI: 10.1201/9781315216386-3.

[9] O'Toole J. M. and Boashash B. (2013). Fast and memory-efficient algorithms for computing quadratic time–frequency distributions. *Appl. Comput. Harmon. Anal.*, Vol. 35, No. 2, pp. 350-358. DOI: 10.1016/j.acha.2013.01.003.

[10] Tanaka T. and Mandic D. P. (2007). Complex Empirical Mode Decomposition. *IEEE Signal Process. Lett.*, Vol. 14, No. 2, pp. 101-104. DOI: 10.1109/lsp.2006.882107.

[11] Huang N., Wu Z., and Long S. (2008). Hilbert-Huang transform. *Scholarpedia*, Vol. 3, No. 7, p. 2544. DOI: 10.4249/scholarpedia.2544.

[12] Hassani H. (2021). Singular Spectrum Analysis: Methodology and Comparison. *J. Data Sci.*, Vol. 5, No. 2, pp. 239-257. DOI: 10.6339/jds.2007.05(2).396.

[13] Li S., Ma S., and Wang S. (2023). Optimal Complex Morlet Wavelet Parameters for Quantitative Time-Frequency Analysis of Molecular Vibration. *Appl. Sci.*, Vol. 13, No. 4, p. 2734. DOI: 10.3390/app13042734.

[14] Vrhel M. J., Lee C., and Unser M. A. (1995). Fractal dimension estimation using the fast continuous wavelet transform. In *SPIE's 1995 Int. Symp. Opt. Sci., Eng., Instrum.*, SPIE. DOI: 10.1117/12.217603.

[15] Moody G. B. and Mark R. G. (1990). The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it. In *[1990] Comput. Cardiol.*, IEEE Computer Society Press. DOI: 10.1109/cic.1990.144205.

[16] Wang X., Ma C., Zhang X., Gao H., Clifford G. D., and Liu C. (2021). Paroxysmal Atrial Fibrillation Events Detection from Dynamic ECG Recordings (version 1.0.0). *PhysioNet*. DOI: 10.13026/ksya-qw89.

[17] Benesty J., Chen Jingdong, and Huang Yiteng (2008). On the Importance of the Pearson Correlation Coefficient in Noise Reduction. *IEEE Trans. Audio, Speech, Lang. Process.*, Vol. 16, No. 4, pp. 757-765. DOI: 10.1109/tasl.2008.919072.

[18] Chou E. P. and Hsu S.-M. (2018). Cosine similarity as a sample size-free measure to quantify phase clustering within a single neurophysiological signal. *J. Neurosci. Methods*, Vol. 295, pp. 111-120. DOI: 10.1016/j.jneumeth.2017.12.007.

[19] Stojanovic M. et al. (2014). Understanding sensitivity, specificity and predictive values. *Vojnosanit. Pregl.*, Vol. 71, No. 11, pp. 1062-1065. DOI: 10.2298/vsp1411062s.

[20] Albardi F., Kabir H. M. D., Bhuiyan M. M. I., Kebria P. M., Khosravi A., and Nahavandi S. (2021). A Comprehensive Study on Torchvision Pre-trained Models for Fine-grained Inter-species Classification. In *2021 IEEE Int. Conf. Syst., Man, Cybern. (SMC)*. DOI: 10.1109/smc52423.2021.9659161.

[21] Lee H. K. and Choi Y.-S. (2019). Application of Continuous Wavelet Transform and Convolutional Neural Network in Decoding Motor Imagery Brain-Computer Interface. *Entropy*, Vol. 21, No. 12, p. 1199. DOI: 10.3390/e21121199.

[22] Fagan X., Ivanko K., and Ivanushkina N. (2020). Detection of Ventricular Late Potentials in Electrocardiograms Using Machine Learning. In *Advances in Computer Science for Engineering and Education III*, Springer International Publishing, pp. 487-497. DOI: 10.1007/978-3-030-55506-1_44.

[23] Kalidas V. and Tamil L. (2017). Real-time QRS detector using Stationary Wavelet Transform for Automated ECG Analysis. *2017 IEEE 17th Int. Conf. Bioinf. Bioeng. (BIBE)*, IEEE. DOI: 10.1109/bibe.2017.00-12.

[24] Kingma D. P. and Ba J. (2014). Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs.LG]*. DOI: 10.48550/arXiv.1412.6980.

Виявлення нейронними мережами низькоамплітудних компонентів на ЕКГ за допомогою модифікованого вейвлет-перетворення

Мневець А. В., Іванушкіна Н. Г.

Дане дослідження присвячено ідентифікації низькоамплітудних компонентів ЕКГ-сигналів різними методами частотно-часового аналізу, коли основний спектр потужності припадає на високоамплітудні компоненти. Також було проаналізовано проблему вибору системи масштабів для визначення низькоамплітудних компонентів на скейлограмі за допомогою моделей штучного інтелекту. Як результат дослідження, були запропоновані кілька модифікацій безперервного вейвлет-перетворення. Перша модифікація базується на використанні масштабувальної функції та модифікованого вейвлета. Друга модифікація базується на використанні косинусної подібності на кожній ітерації згортки з подальшим застосуванням масштабувальної функції. Основна ідея дослідження полягає в тому, щоб модифікувати вейвлет-перетворення таким чином, щоб виділити компоненти з цільовою амплітудою та зменшити всі інші компоненти, які ускладнюють аналіз цікавлячих фрагментів сигналу нейронними мережами. Також були запропоновані можливі процедури відновлення сигналу для збереження ефекту використання масштабувальних модифікацій. Тестування запропонованих модифікованих алгоритмів було проведено на основі штучно створених сигналів, а також на основі реальних сигналів ЕКГ з накладеними на них пізніми потенціалами. Візуальний аналіз скейлограм та відновлених сигналів, отриманих за допомогою модифікованого вейвлет-перетворення, показав, що модифіковане вейвлет-перетворення здатне виділяти низькоамплітудні компоненти зі сигналу з набагато більшою спектральною потужністю, ніж перетворення без модифікацій. Крім того, була перевірена здатність загальних моделей нейронних мереж розрізняти серцеві цикли з пізніми потенціалами та без них. У результаті було виявлено, що моделі, які тренувалися на скейлограмах, отриманих за допомогою модифікованого вейвлет-перетворення, навчаються швидше і менш схильні до застрягання в локальних мінімумах. Результати класифікації сигналів з пізніми потенціалами та без них на основі натренованих моделей нейронних мереж показали, що навчання за допомогою скейлограм, отриманих на основі модифікованого вейвлет-перетворення, дозволяє досягти 99% точності класифікації, що на 1-49% більше, ніж при використанні скейлограм, отриманих на основі класичного вейвлет-перетворення.

Ключові слова: електрокардіографія; вейвлет-перетворення; пізні потенціали передсердь; пізні потенціали шлуночків; нейронні мережі