Maciej GOŁGOWSKI[1], Stanisław OSOWSKI[1,2]

Military University of Technology, Faculty of Electronic Engineering (1),
Warsaw University of Technology, Faculty of Electrical Engineering (2)

doi:10.15199/48.2025.03.47

Deep CNN ensemble for anomaly detection in ECG

Abstract. The paper proposes an ensemble composed of CNN networks for the detection of anomalies in ECG waveforms. The approach is composed of two stages: transformation of ECG signals into images and association of the images with the appropriate class of anomaly using the ensemble of CNN classifiers. The experiments have been performed on the publicly available database Complex Physiologic Signals PhysioNet and directed to recognize three types of ECG signals. The results of numerical experiments are presented and discussed.

Streszczenie. Artykuł przedstawia automatyczną metodę rozpoznania anomalii w sygnałach EKG. Jako anomalie rozważane są: arytmia i zastoinowa niewydolność serca na tle sygnałów reprezentujących stan normalny. Proponowane rozwiązanie stosuje ciągłą transformacje falkową, której wyniki w postaci obrazu podawane są na zespół głębokich klasyfikatorów CNN, odpowiedzialnych za rozpoznanie klasy. Badania eksperymentalne proponowanego systemu przeprowadzono na zbiorze danych EKG zaczerpniętych z bazy Complex Physiologic Signals PhysioNet. Wyniki eksperymentów wykazały wysoką efektywność zaproponowanego rozwiązania. (**Głęboki zespół CNN do wykrywania anomalii w ECG**)

Keywords: anomaly detection of ECG, wavelet transform, classification, CNN ensemble of classifiers.
Słowa kluczowe: wykrywanie anomalii, transformacja falkowa, ECG, klasyfikacja, CNN, zespół klasyfikatorów

Introduction

An electrocardiogram (ECG) is an electrical signal measuring the heart's activity. The ECG waveform is not smooth, with many sudden transitions. It is composed of such elements as waves, intervals, segments, and the QRS complex. The waves represent the positive or negative deflection from baseline and include the P wave, Q wave, R wave, S wave, T wave, and U wave [1,2]. The most important in recognition of different anomalies is the QRC complex.

Different signal-processing approaches have been proposed for this task. Most of them apply mathematical models, which can represent the most characteristic features of the ECG waveforms. The typical methods are based on approximation, filtering, Fourier and wavelet decomposition, application of artificial neural networks, support vector machine, and nowadays deep neural structures [3-6].

The most efficient tool seems to be now the application of the wavelet transformation to the ECG signal, since this decomposition may represent very well different details of the signals in the time domain.

The numerical results reported in the published papers depend on the database used in experiments. The declared accuracy and sensitivity values change from 95% up to 100%, depending on the applied method and the used database [3-6].

This paper will present the new approach to ECG wavelet analysis based on the application of continuous wavelet transformation cooperating with convolutional neural networks (CNN) organized in an ensemble, which is responsible for a final recognition of anomaly class.

The database used in experiments

The experiments will use the publicly available database “Research Resource for Complex Physiologic Signals PhysioNet” [7]. The 162 ECG time recordings contain three classes of people with different heart phenomena that will be tested. They represent class 1 corresponding to cardiac arrhythmia (ARR) in which the heartbeats are irregular, too fast, or too slow, class 2 composed of samples with congestive heart failure (CHG) (the heart is unable to pump sufficiently to maintain blood flow to meet the body's needs) and class 3 of normal sinus rhythms (NSR). All of them represent recordings taken from the MIT-BIH Arrhythmia Database, MIT-BIH Normal Sinus Rhythm Database, and BIDMC Congestive Heart Failure Database [7]. The recordings of the ECG are sampled at 128 Hz. The obtained database was composed of 6144 ARR samples, 1920 samples representing CHG, and 2304 normal samples. The typical representations of these three types of waveforms are presented in Fig. 1.

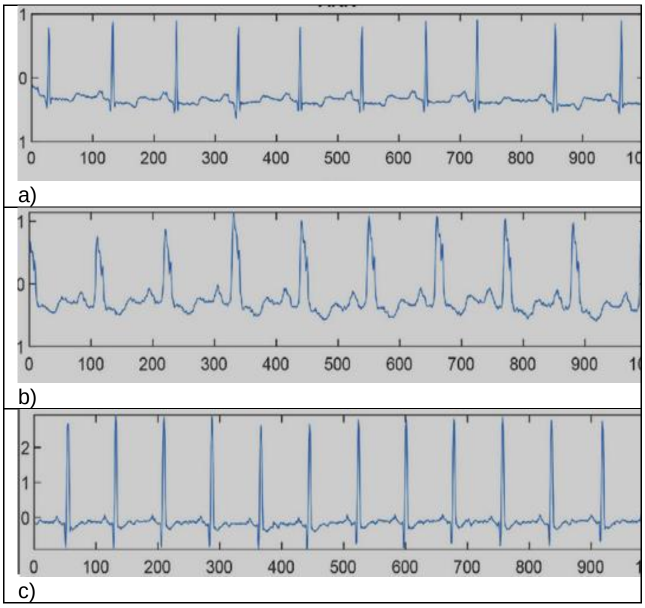

Fig. 1 The examples of the three classes of ECG subjected to the recognition: a) cardiac arrhythmia (ARR), b) congestive heart failure (CHG), and c) normal sinus rhythms (NSR).

Preprocessing steps

The first step in signal processing is to recognize the QRS complex in the ECG recordings. It was done using the „Pan-Tomkins QRS detection algorithm” [2] applying the three-step filtering process. The first step is done by cascading the low-pass and high-pass filters, which are responsible for reducing the signal noise, coming from the muscles, fluctuations of the level of signals, and changes in segment T. The passband of this stage is changing from 5Hz to 15Hz. The applied filters are described in [2].

$$\frac{(1-z^{-6})^2}{(1-z^{-1})^2} \tag{1}$$

$$H_{HP}(z)=\frac{(1-32z^{-16}+z^{-32})}{(1-z^{-1})} \tag{2}$$

In the next step, the differentiation of the slopes of the QRS complex is done using filtration. The differentiation filter is described by the transfer function

(3) $$H(z) = \frac{1}{8}T(-z^{-2} - 2z^{-1} + 2z^{1} + z^{2})$$

The filtered signal is squared to provide the non-negative values of the signal. In the last step, the signal is subjected to integration using the moving window. The output signal of this step is described by the equation

(4) $$y(nT) = [x(nT) - (N-1)T) + x(nT - (N-2)T) + \cdots + x(nT)]$$

where N is the number of samples in the window and T the sampling rate. The result of processing is in the form of a QRS complex of the ECG. Fig. 2 depicts the filtering system.

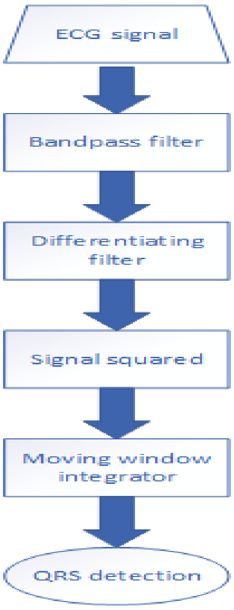

Fig. 2 The cascade of the filtering process of the ECG leading to QRS complex detection.

The ECG signals corresponding to the detected QRS complexes represent the starting point in the recognition problem using convolutional neural networks.

Transformation of signal to an image using CWT

The first step in the application of CNN to class recognition is to convert the signal to the 2-dimensional form of an image. This will be done here using continuous wavelet transformation (CWT) [8-10]. The CWT is the linear transformation of a 1-dimensional signal into a 2-dimensional form represented by the continuous time scale factor *a*, and the continuous time shift *b.* The time scale *a* is inversely proportional to the frequency (the higher the value of *a* the lower the frequency). The general definition of CWT is given in the form [9]

(5) $$W_x(a,b) = |a|^{-1/2}\int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right)dt$$

where $\psi()$ represents the wavelet function used in the analysis of a signal. The CWT applies specially defined wavelet functions, for example, Morlet, Gabor, Gauss, Mexican hat, etc. All of them are explicitly defined.

In the first step, the analyzed ECG signals are converted into sections of a length equal to 1024. In CWT decomposition different scale factors have been tried: from 16 to 128. Fig. 2 presents the typical results of the application of Mexican hat to convert three types of ECG signals (ARR, NSR, and CHG) to the image form [17]. The wavelet transformation is applied for scale *a* equal to 64. The resulting image is of the size 64x1024. The horizontal axis represents the shift (from 1 to 1024) and the vertical one the scale changing from 1 to 64)

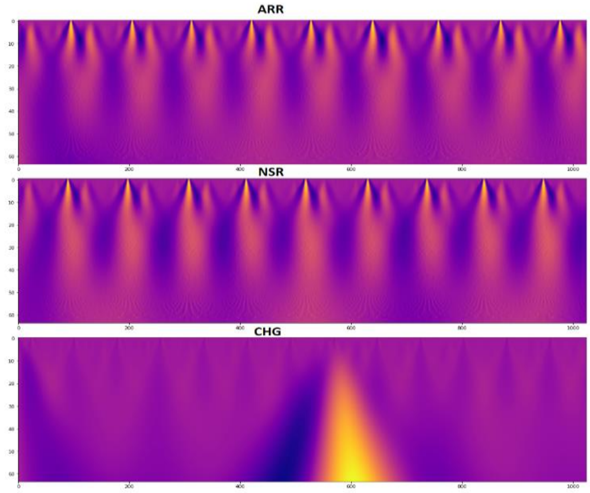

Fig. 3 The resulting images of the CWT transformation of three representative ECG signals. The upper image corresponds to cardiac arrhythmia (ARR), the middle one – to normal sinus rhythm (NSR), and the bottom one – to congestive heart failure (CHG).

Analyzing the CWT representations of these three signals some differences can be observed. They are very well visible for CHG versus the other two classes. The differences between NSR and ARR are less expressive, but also visible in the details of the intensity level of colors. They are better seen at lower resolution after the compression of the image. Fig. 4 presents the NSR and ARR images after a four-times reduction of the resolution.

The differences in the images corresponding to different types of ECG waves give a good perspective for recognizing the anomalies in the ECG signal.

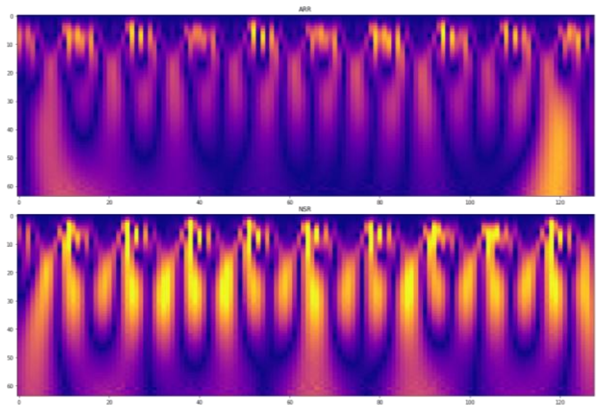

Fig 4 The CWT images of NSR and ARR signals after four times reduction of the resolution of original images.

The compressed images resulting from CWT operation represent the input attributes to the classification system composed of convolutional neural networks. In experiments, we used compressed images due to some limitations of the computational resources.

Convolutional Neural Network

The convolutional neural network is a multilayer deep neural structure designed to analyze the two-dimensional data representing images [14-15]. In common architecture, it implements at the same time the automatic generation of the diagnostic features of images and the final recognition and classification stage based on softmax operation.

The generation of features is done here by many locally connected layers of neurons. The input images are processed sequentially using such operations as convolution, nonlinear ReLU activation, pooling, and normalization [14]. The succeeding layers of locally connected neurons generate a set of images of smaller and smaller dimensions organized in the form of tensors. Some loss of information following the reduction of the size of images is compensated by increasing their population (the depths of tensors).

Finally, the set of images in the last locally connected layer is flattened to the vectorial form. The elements of this vector are treated as the numerical descriptors (features) of the input image. They represent the attributes of the final softnet classifier, applying the softmax function. The softmax generates the probability of class membership of the actual input images[14,15].

The number of output neurons is equal to the number of recognized classes. The softnet contains the feedforward layer composed of *M* summing units (*M* – number of classes) implementing the operation

(6) $$u_i(\mathbf{x}) = \sum_j w_{ij} x_j + w_0$$

for *i*=1, 2, …, *M*). The signals u_i are attributes of the softmax functions, which are responsible for estimating the probability of *i*th class membership

(7) $$\text{softmax}\,(\mathbf{u})_i = \frac{\exp(u_i)}{\sum_{j=1}^{M} \exp(u_j)}$$

The highest value of the softmax signal indicates the class to which the input image belongs. The CNN network is subjected to the learning stage. In the experiments, we have applied the transfer learning approach, in which the pre-trained network with the actual number of output neurons (in the case of 3 classes it is 3) is learned on the proper learning set created for ECG recognition. The learning algorithm was based on ADAM [13] and implemented in Python.

To increase the efficiency of the classification system we have applied the ensemble of CNN classifiers. Different architectures of CNN provide some independence in their operations, an important condition for the proper performance of the ensemble. Each CNN was trained on the randomly selected 70% of the learning data and then tested on the same testing part separated from the available data. The results of testing are subjected to the majority voting defining the final verdict of class membership.

Results of numerical experiments

The database used in experiments [7] was created in Beth Israel in Boston and contains 162 recordings of ECG waveforms representing three classes:
Class 1 – 96 recordings of arrhythmia (ARR),
Class 2 – 30 recordings of congestive heart failure (CHG)
Class 3 – 36 recordings of normal sinus rhythm (NSR).
The recordings have been sampled with 128Hz and the total number of samples is equal to 65536 in each recording.

The original ECG signals are split into 10-second interval frames [17]. Such an approach is suggested by the authors of the database since it resembles the way the medical experts do. The CWT transformation is applied to each frame producing the set of images, which are then used as input attributes to the CNN classification system.

The total number of images used in experiments was equal to 10368. The whole set was split randomly into learning parts (70%) and testing (the remaining 30%). The experiments were repeated 10 times with the random selection of the learning and testing part. The average of these 10 trials is presented as the final results of experiments. The paper will limit the presentation of results only to the testing data not taking part in t learning stage.

The ensemble system was carefully selected from the set of available pre-trained architectures of CNN. The introductory experiments performed in Matlab [11] have allowed us to choose the final composition of ensemble members providing the best operation of the system. The ensemble was arranged from the following CNN units [11]:
alexnet, mobilenetv2, resnet50, efficientnetb0, squeezenet, googlenet, shufflenet and inceptionresnetv2.
All of them have been learned on different (random selection) learning data and tested on the same dataset not used in learning. The testing results of using these 8 units are fused by the majority voting to provide the final verdict for class recognition.

Different types of wavelet functions have been tried in experiments [9,10,17]. The statistical results have been assessed for three mother wavelets: Gauss 8, Mexican hat, and Morlet, and the application of the scale *a* equal to 64, which was found the most efficient in the introductory experiments. The best results of class recognition have been obtained with the application of a Mexican hat. The confusion matrix for the testing data not taking part in learning is presented for this wavelet function in Table 1.

Table 1 Confusion matrix for testing data corresponding to Mexican hat and application of an ensemble of 8 CNN units.

	ARR	CHG	NSR
ARR	1655	23	35
CHG	8	540	19
NSR	11	15	637

Based on this matrix the values of different quality measures (accuracy, sensitivity, precision, and F1 [12,13]) have been calculated and are depicted in Table 2. In the case of the accuracy ACC, the mean for all recognized classes is given. The sensitivity, precision, and F1 are presented for each class separately.

Table 2 The results of numerical experiments in recognition of three classes of ECG. The scale in CWT was equal to 64.

Wavelet function		Mexican hat
ACC		96.23%
Sensitivity	ARR	96.61%
	CHG	95.24%
	NSR	96.08%
Precision	ARR	98.86%
	CHG	93.43%
	NSR	92.19%
F1	ARR	97.67%
	CHG	94.33%
	NSR	94.10%

The best average accuracy (96.23%) is due to the application of an ensemble. It is better than the corresponding value of the system composed of only a single CNN classifier. The worst result of a single classifier (92.3%) was due to Alexnet, The best one was obtained by Resnet50 (ACC=95.81%).

Conclusions

The paper has presented the application of wavelet transformation and the ensemble of convolutional neural networks to recognize the anomaly in ECG signals. The set of ECG signals subjected to the recognition was composed of a cardiac arrhythmia, congestive heart failure, and normal sinus rhythms. The CWT was used to transform the signals into the image forms, which represent the input attributes to the CNN classifiers forming the ensemble.

Different types of wavelet functions used in the conversion of signal to the image form have been tried. The best results have been obtained through the application of the Mexican hat. The highest average accuracy of an ensemble in three class recognition reached the value of 96.23%. It was better than the best individual unit included in the ensemble.

The application of an ensemble composed of many different CNN architectures has an advantage over the classical approaches to the construction of the ensembles.

The procedure described in the paper is general and can be applied to different types of signals, not necessarily the ECG. Thus, the proposed system can find applications in different tasks of engineering, not limited only to the medical processes.

***Authors**: prof. dr hab. inż. Stanisław Osowski, Politechnika Warszawska, Instytut Elektrotechniki Teoretycznej i Systemów Informacyjno-Pomiarowych, 00-661 Warszawa, Koszykowa 75 oraz Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, 00-908 Warszawa, Kaliskiego 2, E-mail: stanislaw.osowski@pw.edu.pl; dr inż. Maciej Gołgowski, Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, 00-908 Warszawa, Kaliskiego 2, E-mail: maciej.golgowski@wat.edu.pl.*

REFERENCES

[1] Lenis G., Pilia N., Loewe A., Schulze W. H., Dössel O., Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG, a simulation study, *Comput. & Mathematical Methods in Medicine*, 2017, https,//doi.org/10.1155/2017/9295029

[2] Pan J; Tompkins W. J., A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering. BME-32*(3): 1985, 230–236. doi:10.1109/TBME.1985.325532

[3] Singh R., Mehta R., Rajpal N., Efficient wavelet families for ECG classification using neural classifiers, *Procedia Computer Science*, vol. 132. Pp. 11-21, 2018

[4] Rai H. M., Trivedi A., Shukla S., ECG signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier, *Measurement*, vol. 46, 3238-3246, 2013

[5] Patil D. D., Singh R. P., ECG classification using wavelet transform and wavelet network classifier, *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, vol. 668, pp 289-303, 2018

[6] Arumugam M., Sangaiah A. K., Arrhythmia identification and classification using wavelet centered methodology in ECG signals, *Concurrency and Computation, Practice and Experience*, 2019, http,//doi.org/10.1002/cpe.5553

[7] PhysioNet - The Research Resource for Complex Physiologic Signals, https,//physionet.org/

[8] Daubechies I., *Ten lectures on wavelets*, SIAM, Philadelphia, 1992

[9] S. Mallat, A wavelet tour of signal processing: The Sparse Way, Academic Press, London, 2011.

[10] Teolis A., *Computational signal processing with wavelet*s, Springer Science& Business Media,2012

[11] *Matlab user manual*, MathWorks, 2022, Natick, USA.

[12] Tan P. N., Steinbach M., Kumar V., *Introduction to data mining*, 2014, Pearson Education Inc., Boston

[13] Osowski S., Szmurlo R., Matematyczne modele uczenia maszynowego w językach Matlab i Python, OWPW, 2023. Warszawa

[14] Goodfellow I., Bengio Y., Courville A, *Deep learning*, MIT Press, Massachusetts, 2016

[15] Chollet F., *Deep learning with Python*, Manning Publications Co. 2017

[16] Golgowski M., Osowski S., Anomaly detection in ECG using wavelet transformation, Int. on-line Conf. Computational Problems of Electrical Engineering, Warsaw, 2020, doi: 10.1109/CPEE50798.2020.9238709

[17] Golgowski M., Osowski S., Classical versus deep learning methods for anomaly detection in ECG using wavelet transformation, Przegląd Elektrotechniczny, R. 97 nr 6/2021, doi:10.15199/48.2021.06.13.