UDK 004.9

Yuriy V. Khoma[1], Candidate of Technical Sciences, Assistant Professor of the Institute of Computer Technology, Automatics and Metrology, E-mail: yurii.v.khoma@lpnu.ua, Scopus ID: 55675125300,
ORCID: http://orcid.org/0000-0002-4677-5392
Miroslaw Szmajda[2], PhD Hab. Eng, Professor, Director of Automatic Control Institute,
E-mail: m.szmajda@po.edu.pl, Scopus ID: 24767924400, ORCID: http://orcid.org/0000-0003-3467-951X
Mariusz Pelc[3], PhD Hab. Eng, Professor, Senior Lecturer of Department of Computing & Information Systems, E-mail: m.pelc@greenwich.ac.uk, Scopus ID: 9335890000,
ORCID: http://orcid.org/0000-0003-2818-1010
[1] Lviv Polytechnic National University, St. Bandery, 12, Lviv, Ukraine, 76013
[2] Opole University of Technology, Proszkowska, 76 Street, Opole, Poland, 45-758
[3] University of Greenwich, Park Row, London SE10 9LS, UK

DEVELOPMENT OF SCIENTIFIC-METHODOLOGICAL APPROACHES OF MACHINE LEARNING APPLICATION IN BIOSIGNALS PROCESSING

Annotation*. Current state and future perspectives of machine learning usage in the computer bioinformatics systems are analyzed in the article. It is shown that heterogeneousness of data and wide range of bioinformatics tasks influenced the development of the specialized solutions for each separate domain or application. This complicates the possibility to compare effectiveness of certain methods as well as usage of the best variants for the system design for the new tasks. Research results are presented, which are related to the development of principles for the design of the biosignal computer processing systems involving the machine learning. The expediency of separation the system levels within the process of biosignals processing is reasoned, and their functions are outlined. Innovativeness of the suggested approach lies in separation of the function of lower, middle and upper levels from methods with the help of which they are realized, as well as from the implementation variants for these methods based on the hardware and software components. It is highlighted that the middle system level is significantly invariable both in regards to the task to be solved and to the biosignal type. At the same time the upper level is specific as to the task, and the lower level is specific as to the type of biosignal. Distinct outlining of functions for each system level and the inter level interfaces opens prospectives for structuring information during the analysis of the known decisions, which simplifies the analysis and comparison of the effectiveness of these solutions. Design process of the computer system for the specific tasks gets simplified and potentially quickens due to the possibility of transferring the best results between the related tasks. On the basis of the developed three system levels concept the range of tasks related to machine learning application and biosignal processing on all the system levels was studied and analyzed.*

Keywords*: bioinformatics; biosignal computer analysis; machine learning; system levels of the biosignal processing; inter level interfaces*

Introduction. Biosignals carry valuable information, and for more than a hundred years they have been used in medical diagnostics. Development of microelectronics and computer technologies positively influenced the characteristics improvement in medical diagnostics. This favored the growth of domains and tasks, where biosignal analysis discloses new possibilities. Currently various hardware is available for registration and processing of dozens of biosignals aimed to solve of different bioinformatics tasks. Among such tasks the most popular ones include 24-hour clinic monitoring, biometric identification and authentication, control of the psycho-emotional states and reactions of the person, direct interaction of the human with the technical means through the ‘brain-computer’ interface and bioengineering for neuro-rehabilitation [1-5].

Traditional methods of biosignal processing include both performances of the formal and informal procedures. Formal methods are realized by the analog and digital means and foresee aquisition, conditioning, analog-to-digital transformation, and later a number of the specific transformations aimed at detection of the informative (diagnostic) features [6-7]. The final processing stage is complicated and cannot be formalized, as by its nature it is related with the analysis and interpretation of the research result on the basis of the detected informative features and is traditionally executed by a human expert [1].

However, the analysis of even the denoised biosignals containing the relevant information is not always an easy task for a doctor. Some types of biosignals, for example electromyography one, is too complicated for the visual analysis by a human. Biosignals are also variable by nature, which can

© Khoma, Y., Szmajda, M., Pelc, M., 2020
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/deed.uk)

produce subjective divergences of the analysis results executed by different doctors [1]. These factors define the need not only to improve the hardware, but also to develop new methods for the objective biosignal analysis using not only the formal processing algorithms, but also heuristic approaches based on the machine learning technologies providing the intellectual analysis of biosignals [8-9].

Literature review. Today there exists a great interest in the machine learning application in the healthcare [9-10] for the analysis of the electronic medical records [11], medical diagnostics [12] and many other domains of bioinformatics, including affective informatics and biometrics [3; 6]. Tasks related to the biosignals processing treat the machine learning as a technology which is, first of all, applied towards the intellectual analysis and the results interpretation, thus, used on the final stage [1; 6]. However, the recent researches showed that profit from the machine learning algorithms becomes more significant on all the previous stages of the biosignals transformation, where the efficiency of the traditional methods of the digital processing is limited. This also encourages the development of the intellectual methodologies of biosignal processing [13-15].

During the last hundred years there have appeared hundreds of publications containing the description of systems, methodologies of the research execution and analysis of their results, which relate to the dozens of the specific tasks in the domain of bioinformatics. Development of the specialized solutions for obtaining of the biomedical values, absence of the common formalized approaches complicate knowledge and methodologies transfer from one domain into another, especially what concerns the bioinformatics systems. Existing publications contain description of biosignal processing pipelines in the form of the sequence of stages, presentation of the signal and data transformation and analysis methods applied on each stage, as well as implementation examples [6; 14; 16-18].

However, even the publications dedicated to the solution of one and the same task via the processing of the same biosignal contain different number and names of the stages. For example, work [16] used only 4 stages while working with the ECG signal in the biometric systems, and the work [18] has twice as much stage because of the additional details. Sometimes the researches use records from the pre-collected biosignal databases, so they do not contain the stages of acquisition and registration [6], and some publications do not even have the stages related with the biosignal conditioning [19]. Variability of the research results presented in the publications grows due to detailed treatment of stages suggested for biosignal computer processing.

Currently there is a need to solve many different tasks, which are based on the computer processing of dozens of different, sometimes interconnected, biosignals [20-22]. This variety additionally increases because of different conditions (stand-alone or mobile applications) and the specific scenarios of biometric systems usage (not only the one time measurements, but also the continuous monitoring at fixed location [20], remote [23], with the wearable equipment [22; 24]). This directly influences the ways in which the suggested methods are implemented and the constructive design of the separate system modules, for example, variants related to with the cloud computing, ‘stand-alone’, ‘wearable’ application or others [20; 22; 25].

Many of these researches describe the specific solutions in the form of the monolith structure [24-25]. This complicates the possibility to compare effectiveness of separate methods and use the best variants for system design for the new tasks. Thus, there has appeared the need to structure the knowledge in the subject area related to the biosignal computer processing and formalization of the approaches to the design of the bioinformatics systems.

The purpose of the article

The goal of the current research is to present the three system level concept for computer processing of the biosignals. In the basis of the approach proposed lies the idea of the distinct separation of functions inhered to system levels from the methods and means applied for the realization of such functions. Such separation will become the ground for knowledge structuring in the subject area, enable comparing and evaluation of the efficiency for different methods, help to choose better decisions on the stage of bioinformatics systems design for the certain tasks, considering the conditions and scenarios of their usage.

To achieve this goal the following tasks are to be solved:

1. Reason the expediency of system levels separation in the biosignal processing flow.
2. Outline the functions of the lower, middle and the upper levels.
3. Analyze the possibility and provide examples of machine learning algorithms application on each of the system levels.

Main part. Challenges of the modern bioinformatics and perspectives of the machine learning application

Review and analysis of the publications from the recent years allow to point onto the problematic aspects in the bioinformatics, and namely the ones concerning application of the machine learning technologies.

There are many different biosignals, moreover using their synergy (co-processing) due to over-additive effect allows to increase the accuracy and reliability of the results obtained from execution of one definite task. Such an approach (Fig. 1a) in a modern bioinformatics was named as sensor fusion [26]. Example of the practical application of the synergetic approach is the polygraph (lie detector). In order to improve the trustworthiness of the test results the multichannel measurements of the interconnected ECG, EEG, PPG and some other signals are used, which reflect the corresponding psycho-emotional reactions of the person onto the questions asked. In order to realize the synergetic approach it is necessary to maximally unify the measuring transformations, first of all in the input analog circuits of the measuring channel, as the measuring transformations of each biosignal appear to be rather specific [7; 24].

Currently the need of practice sets demand, which is a bit controversial as for the synergy, namely – processing of one type of biosignal oriented onto solution of several different problematic tasks of the bioinformatics (Fig. 1b). For example, via the processing of corresponding electrocardiogram the versatile tasks are solved – medical diagnostics (disease detection), biometrics (person identification) and clinic monitoring (24-hour monitoring of the heart performance).

By contrast with the synergetic approach, the stress of the problem solving in the multi-purpose variant is shifted from the lowest system level (input circuits of the biosignal sensors) to the highest level of the computer processing system, thus, it relates to distinguishing of the informative features, analysis and interpretation of the research results. That is why in order to realize the multipurpose approach it is important to secure the flexibility of the system reconfiguration from solving one task to another. For this purpose the wide application of the machine learning technologies looks promising.

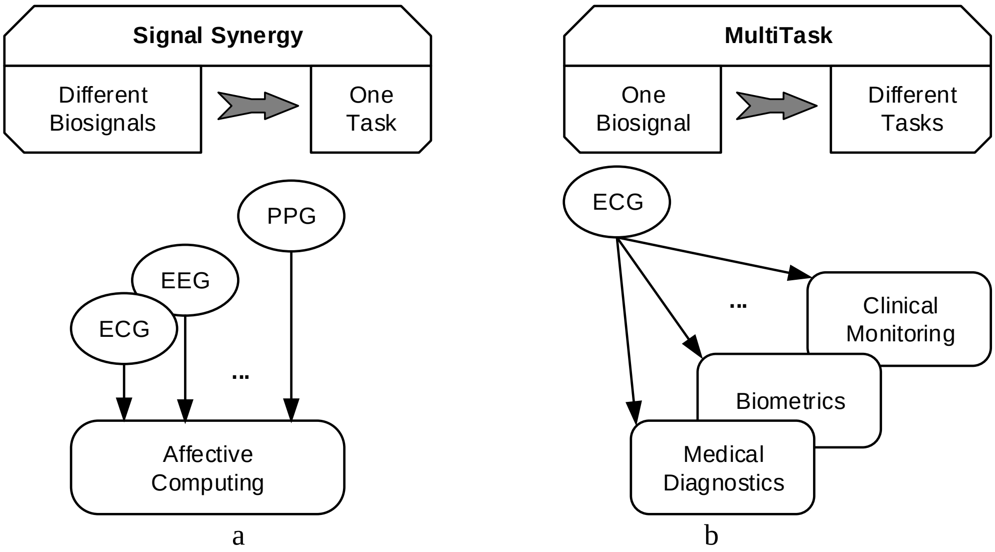

Fig. 1. Controversial approaches to the biosignal processing:
a – synergetic; b – multipurpose

The common peculiarity of all the biosignals include low level, their variability and specific distortions of the biologic origin, which can conditionally be divided into three types – interference, shielding, natural variability of the measuring biological parameter. Frequency band of obstacles and distortions is often overlapping with the frequency band of the informative components of the biosignal [6; 17; 27]. Of course, these factors significantly complicate the flow of the biosignal computer processing, and effectiveness of the traditional methods of the digital signal processing, based on the formal models, is rather limited. That is why the search of the new approaches to provide the invariance of the biosignals measuring transformations to the influence of the disturbing factors continues, first of all the ones based on the machine learning technologies [28-29]. It is important to stress, that the final stage of the biosignal processing related to the analysis and the interpretation of the results is difficult or even impossible to formalize.

However, apart from the mentioned barriers, currently in the domain of the biosignal computer processing has appeared such a powerful driving technology as machine learning, especially deep learning, which relies on the neural networks and provides unbelievable advancement in the similar tasks, which are difficult to formalize. This is confirmed by the fact that the similar algorithms are successfully applied for the analysis in other domains: images, sound/speech, text/language, therefore, to the data of very versatile nature, complexity and structure.

Table 1 provides the comparison of several key aspects, characteristic for the digital processing of signals and machine learning.

However, apart from the potential advantages, usage of the machine learning on the practice also presents the specific demands, which, up till the moment they are solved, create multiple problems for usage of the machine learning in bioinformatics. For example, it relates to synthesis of the neural network architecture adequate to the character of the analyzed data (in our case – biosignals), as well as choice of the optimal hyper-parameters of this network. Another problem is related to the need of big amounts of the relative data, received during the experiment from the real objects. Exactly that data is the basis for adjustment of the machine learning models on the training stage for this or that applied task.

Currently the machine learning is applied mostly on the final stage related to the interpretation of the experiment results. There are many researches directed onto usage of the machine learning in the medical diagnostics to support the decision making system of the functional diagnostic doctor. Similarly there are a lot of publications on the topic of the machine learning application in the biometric systems, affective computing and other domains of the bioinformatics. However, the machine learning in the similar researches performs the classification function based on the distinguished (with the help of the formal procedures) informative features. There is no complex research related to the search of the way to use the machine learning on different stages of the computer processing of biosignals, where the traditional formal methods are insufficient.

Development of the scientific-methodological approaches to application of the machine learning in bioinformatics

The final goal of transformation and processing of biosignals in the computer bioinformatics system is to obtain the necessary information in the form of the certain biomedical parameters and interpret them within the context of information from the knowledge base. There are many different types of biosignals, by processing of which different tasks both in the domain of medical diagnostics and other domains of bioinformatics are solved.

Fig. 2 presents the hypothetical toolset (in the form of certain methods and means) with the help of which the desirable scenarios where one type of the biosignal is applied for the solution of several bioinformatics tasks are realized, or otherwise – one and the same task is solved via application of different biosignals. As was provided above, the synergy, namely merging of several biosignal processing methods, may be used for the formation of the integral biomedical values and increase of the authenticity of the experiment results.

Table 1. Comparison of technologies of the digital signal processing and the machine learning

Technology	Models	Basis	Advantages	Disadvantages
Digital signal processing	Formal	Known dependencies between variables	Effective implementation	Limited effectiveness for biosignals
Machine learning	Fuzzy	Experimental data	Heuristicity	Need of big data

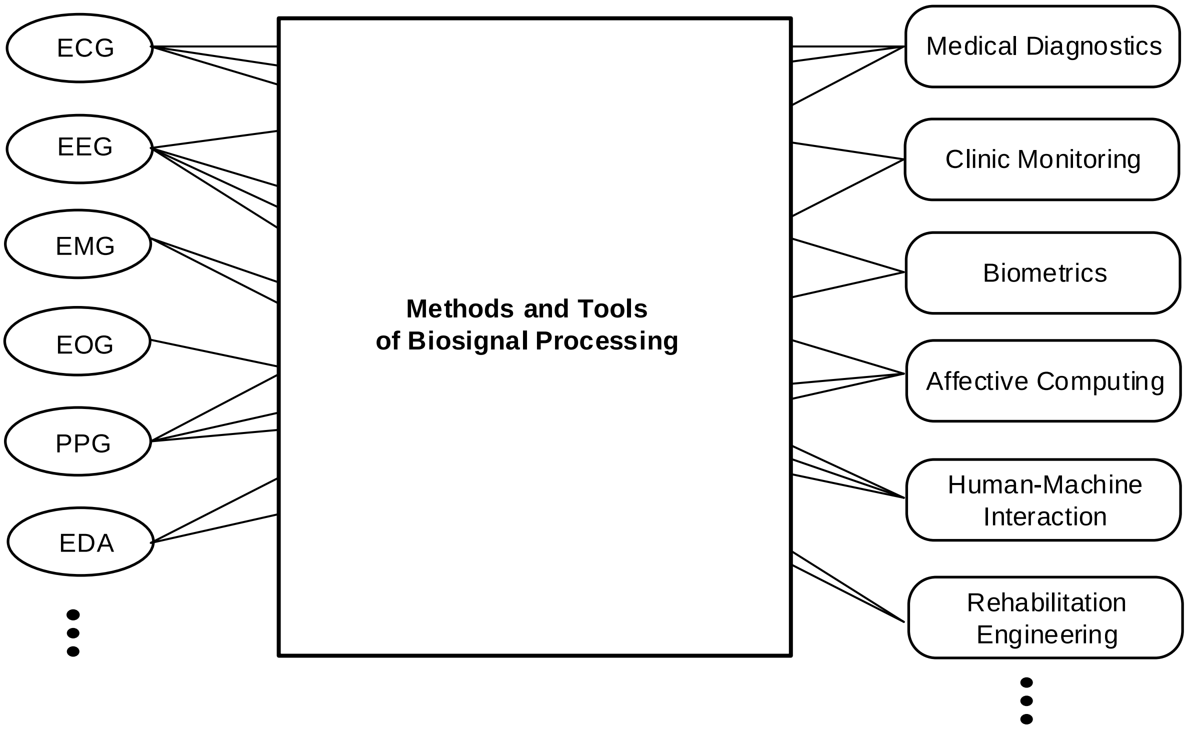

Fig. 2. Relations between biosignals and bioinformatics tasks

From the other side there exists a multipurpose (problem oriented) approach which provides the possibility to solve several different tasks of bioinformatics with the help of the corresponding processing of one biosignal type. Nevertheless, even within one application, which is to be the medical diagnostics, processing of the ECG-records is used to perform the research aimed at detection of different heart performance anomalies [19]. Apart from this, electrocardiogram may be also used to identify the person [6]. Due to the partial overlap of the transformation pipelines and processing of signals and data (Fig. 2) the question of hardware and software aggregation for the effective development of the computer measuring system naturally arises (considering the expenses and time spent necessary for the obtaining of the result).

Analysis of the references allows stating that within the structure of the designed biometric systems the authors outline separate structure elements, aligning their functions with the separate stages of the biosignal processing, where certain signal transformation and analysis methods are used. Yet, unfortunately, there is no generally accepted approaches to the division of these stages and their functions are not distinctly outlined. This complicates the possibility to compare different methods based on their accuracy and computing capacity. That is why, as was stated above, there is a need to structure and systemize the knowledge relating to the biosignal processing. It is important to choose the best solutions during the design of the computer system for the specific task and considering the conditions and certain usage scenarios.

Generalizing the known methods of the biomedical researches the author considers it reasonable to distinguish three system levels and define inter level interfaces as provided on Fig. 3. Apart from this, it is offered to separate the system functions from methods (e.g. methods of bioimpedance into voltage transformation, ejection detection, distinguishing of the informative signals, etc.) and also from certain implementation options of these methods (electronic components of the lower level, computing platforms, hardware and machine learning libraries on the middle and the upper levels).

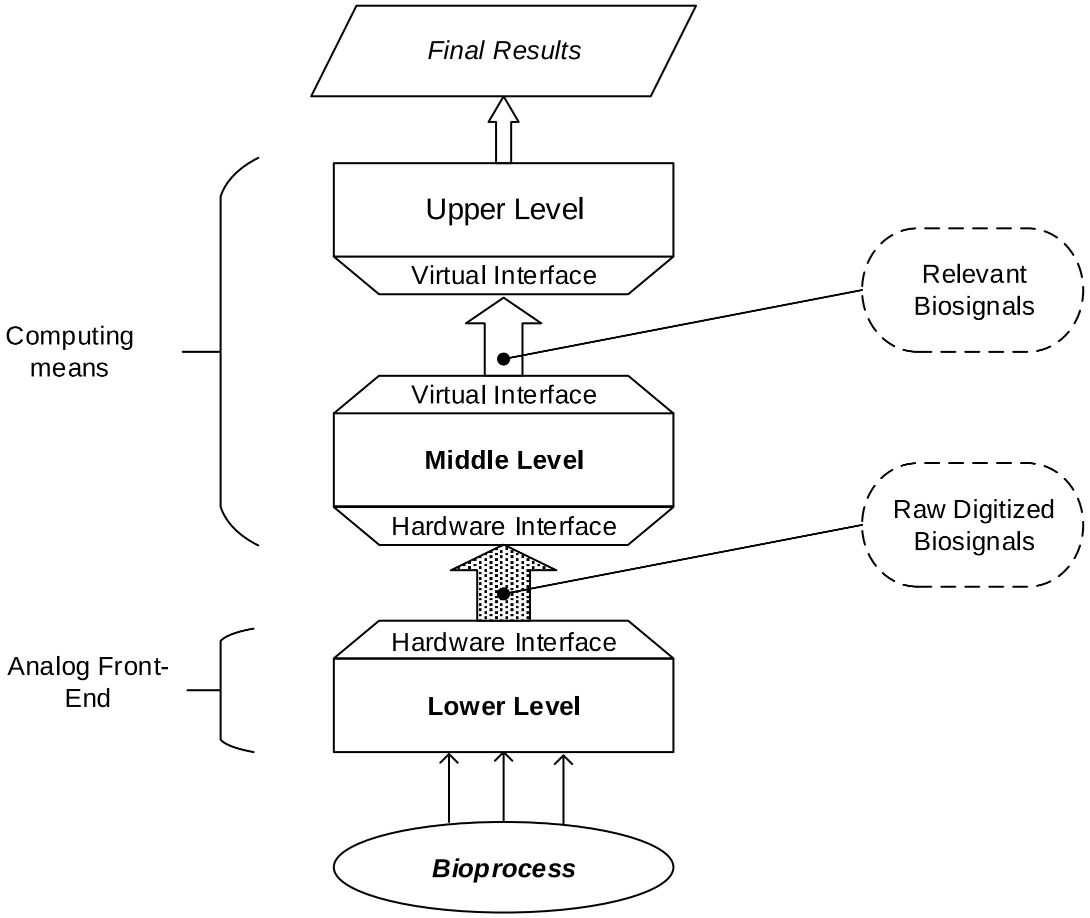

Fig. 3. System levels and the inter level interfaces

The push for distinguishing of the three levels was the following. The lower level or the level of the measuring transformations is tightly related to the nature of the certain biosignal, and thus, is very specific regarding biosignals. Its main designation is to form the electric biosignal on the basis of the psycho physiological processes or values and secure its compatibility with the computer system. The upper level of the intellectual analysis is, vice versa, related to the certain task to be solved. Its main designation is to provide the objective and the most trustworthy result based on the analysis of the informative features distinguished from biosignals (descriptors). The main designation of the middle level or the level of the digital conditioning is to maximally clean the biosignal from the distortions of different kind, meanwhile thoroughly storing the informative components. Thus, the middle system level is rather independent, from both the task to be solved and the type of the biosignal. At the same time, the upper level is specific as to the task, and the lower level – specific as to the type of biosignal.

Distinguishing features and attributes of the lower level:

– at the input biological processes and values (spatial distribution of biopotential variable in time, bioimpedance parameters,);

– executes measuring transformations, namely formation of the biosignals, their amplification, filtering and digitalization [17];

– at the output – ‘raw’ digitized biosignals;

– built mostly from the analog components.

Distinguishing features and attributes of the middle level:

– at the input – ‘raw’ digitized biosignals;

– performs digital conditioning of the biosignals;

– at the output – relevant digital biosignals as the numerical series;

– built from the digital computing means.

Distinguishing features and attributes of the upper level:

– at the input – relevant digital biosignals;

– executes heuristic processing of the unique informative parameters of the biosignals;

– at the output – results of the research in the form of biomedical parameters and physiological characteristics of the person;

– built from the digital computing means.

Forming and acquisition of the biosignals from the object, as well as the work of the analog measuring converters, related to the lower system level, operates under the difficult circumstances. Therefore, there is a need to condition the biosignal, which represents the main function of the middle system level [1; 30].

Providing of the biosignals in the digital format allows computationally processing and analyzing this data on the middle and the upper levels. Modern technologies of the digital signal processing enable application of the computational capacities on the middle level to improve the correlation signal/noise (digital filtering, synchronous detection, time-frequency transformations), as well as to compensate distortions of the analog part of the measuring pipeline (equalizer).

The signal, cleaned from the noise and the distortions, undergoes the further analysis on the middle system level in order to, for example:

–find informative segments in the registered signal [6; 31];

– normalize in accordance with the level and time [31];

– detect characteristic points, for example R-peaks, on the electrocardiogram [6; 31];

– disclose anomaly fragments in the records and their correction [27; 29].

The final stage in the computer processing of the biosignal is realized on the upper system level. It is the most complicated from the computational point of view, as it supposes execution of number of such operations as [6]:

– extraction of the informative, for example diagnostic, features (feature extraction) from the biosignal;

– feature space reduction;

– pattern recognition;

– interpreting of the results and decision making.

Two latter operations are difficult to formalize and they are usually executed by the means of the machine learning [6; 19].

Notwithstanding that middle and upper system levels are built from the digital computational means, the functions assigned to them significantly differ, which was the reason for their separation. The interface created between these levels is virtual, to the contrast with the hardware analog-to-digital interface between the lower and the middle system levels (Fig. 3).

It is worth marking the possibility to use one method, for example wavelet transformation, on the different system levels: on the middle level to decompose signal to the scalogram or on the upper level to form the descriptors for the diagnostic classifier [6; 32].

Thus, the innovativeness of the suggested approach lies in the separation of the system level functions from the methods with the help of which they are being realized, as well as hardware and software implementation variants for these methods, and also reasoning of the functions for each system level. Application of such an approach will favor the knowledge structuring in the subject area, where the biosignal processing is used, provide the possibility compare different methods and evaluation of their effectiveness, open the choice of the best solutions while designing computer biometric system for the certain task, considering the conditions and use cases.

Outlining and solving of the problematic tasks for biosignal computer processing within the paradigm of the system level

It is worth stating, that the suggested system levels are first related to the functions of the computer processing, and not with the methods or hardware-software means, with the help of which these functions are realized. In other words, there may be multiple variants to implement these functions, which, consequently, are reflected by the big amount of publications. Application of the suggested approach based on the division of the biosignal processing pipeline onto three system levels. Distinct definition of the functions for each level and inter level interfaces opens the perspectives to structure knowledge and information in the current subject area. There is also the possibility to compare different biosignal processing methods, to evaluate their effectiveness and to flexibly design the computer systems to solve the various tasks in the domain of bioinformatics.

Application of the suggested approach has the following advantages both during the analysis of the known solutions, and on the stage of the computer bioinformatics system design:

– possibility to structure methods of the software-hardware processing means and biosignal analysis in the published researches;

– performance of the correlation analysis and effectiveness comparison for the known solutions

– simplification and speed up of the system design process for the specific tasks due to the possibility to transfer the best results from the contiguous tasks;
– application of different types of digital processing and machine learning becomes more flexible and visual on all the structural levels without introduction of the collisions into the principles of the biosignal digital processing;
– separation of the functions from methods and means additionally highlights the possibility and expediency to widely apply the artificial intelligence algorithms on the middle and especially on the lower system levels, where, as shown by the results of the literature review, the lack of corresponding scientific activities is clearly noticeable.

Thus, based on the developed concept of three system levels it is possible to consider and analyze the number of the problematic tasks related to the application of the machine learning for the biosignal processing on all the system levels. Hence, on the lower level the machine learning means allow to realize algorithmic correction of the instrumental errors of the bioimpedance measuring channel [28]. Because of the inertial properties of the electronic components the dynamic errors of the measurements appear. Complicated non-linear dependency of the bioimpedance transform function into the real and imaginary voltage components complicates usage of the formal methods for the correction of the dynamic distortions. At the same time, trained on the non-linear model according to the Monte-Carlo method the neural network equalizer is the effective corrector of the errors arising in the bioimpedance measuring channel.

On the middle level, while executing the biosignal conditioning, in case the formal DSP algorithms are ineffective due to the mutual overlapping of the useful signal spectral components with the obstacles signal spectral components. To solve these issue application of neural network autoencoders makes a lot of sense [41]. What concerns the upper system level, the example of the machine learning usage for processing of one and the same biosignal but for the different tasks (biometric identification, functional diagnostics) is described in the works [1;19]. Usually on the upper level, the operations related to the processing of the information in the weakly formalized tasks, such as pattern recognition, are being performed. It is worth stating that work [19], as well as many others, lacks the functions of the lower and the middle levels, as the filtered biosignals from the open source databases are being analyzed. This also increases the expediency to use the concept of the system levels for the analysis, comparison and consolidation of the practices in the domain of biosignal processing.

Apart from this, the widely spread variant of ‘raw’ biosignal usage from the open source libraries greatly fits into the paradigm of the system levels which is necessary, for example, on the stage of the neural network classifier training [33]. Thus, on the stage of the model training two system levels are used, and on the working stage – three of them. In order to secure the compatibility of the training set with the ‘raw’ biosignal selected from the lower level, it is necessary to apply the programming module, which unifies the data received from the different equipment (discretization frequency, resolution, record duration) and enable their introduction to the second system level. The author realized such an approach, and the results of the research were published in the work [33].

It is worth stating that the complication in the structural components design and the efficiency of their work may be different depending on the type of the signal and the aim of the research. Analysis of the references revealed that the best processed hardware and software means on acquisition, transformation and analysis of the electric biopotentials first of all are for the electrocardiograms and encephalograms, as exactly this equipment is widely used in the medical diagnostics [34-35]. From the other side, rheographs, frequency analyzers, as the instrumental provisioning of the bioimpedance measurement, are significantly more complicated equipment, as obtaining of the relevant information requires additional specific transformations [36-37]. Thus, in the paradigm of the system levels the improvement of the analog part of the bioimpedance measurement channel is really essential, as well as the efficiency increase of the specialized algorithms for the digital signal processing.

Conclusions. Biosignals carry the valuable information about the health condition, anatomic peculiarities and psychophysical reactions of the person. In order to obtain this information the biosignals are being measured and processed, using the corresponding instrumental provisioning. Development of the information technologies, increase of the computational efficiency and lowering of the price on the hardware-software means opened the possibilities to apply the machine learning for the intellectual analysis of the biomedical data. The biosignals application domain has also significantly grown, which defines the necessity not only to improve the equipment but also to develop the new approaches to their objective analysis.

In recent years, one can observe the real boost in the domain of bioinformatics, which is, among others, represented by hundreds of publications dedicated to the computer processing of dozens of biosignals used for the solutions for the versatile tasks. The specificity of these tasks influenced the development of the specialized solutions in the hardware manufacturing, as well as complicated the knowledge transfer among the experts in different domains and became the factor holding back the bioinformatics development. Thus, there arose the necessity to structure knowledge, especially in issues concerning approaches, methods and means of the computer processing and analysis of biosignals.

In the work, the expediency to use the concept of the three system levels was suggested and proved, and these levels relate not so much to the structure of the bioinformatics system, but to the functions, which are to be performed on each stage. Thus, the lower level is assigned with the function of biosignal formation, acquisition and digitalization. That is why the ‘raw’ signal obtained from the hardware analog-to-digital interface is passed to the middle level, where it is conditioned using different digital signal processing algorithms and various machine learning models. The relevant biosignal received at the output via the virtual interface is passed onto the upper system level for the intellectual analysis and interpretation of the research results mostly on the basis of the machine learning technology.

The examples of usage and advantages of the current concept application are provided. Taking into consideration the multi variability of the known solutions in the design of the biometric systems, as well as new, sometimes controversial demands to their functionality (for example synergy of biosignals versus the multipurpose approach) the separation of the system levels from the methods and implementation allows to structure knowledge and formalize the process of system design for the specific tasks.

One of the main questions arising before the application of the machine learning is the need in big amounts of data received during the experiment. Right on the basis of this data the machine learning models are fitted on the training stage. There are open source databases with different types of biosignal, which may be introduced on the middle level. However, this creates the need in the program converter which task is to provide compatibility by converting data into common format. The other example presenting application of the offered approach is the possibility to use the machine learning on all the system levels (anomalies detection and correction, correction of the sensor errors), and not only, as traditional, on the upper level in the role of the classifier based on top of informative features. The next step in the development of the current competency may become the standardization of the demands to the inter level system interfaces.

Apart from the above provided hierarchy principle in the form of the three system levels, the further development of the scientific-methodological bases is worth directing onto the development of the module principle for the realization of different biosignal processing functions. On the lower system level these would be hardware module, and on the middle level and upper levels – program modules, realizing the above listed functions of biosignal conditioning, as well as detection of informative classification features, etc.

References

1. Rangayyan, R. M. (2002). “Biomedical Signal Analysis: A Case-Study Approach”. *John Willey and Sons Inc.* 552 p.

2. Jain, A. K., Flynn, P. & Ross, A. A. (2008). “Handbook of Biometrics”. *Springer*, 556 p.

3. Luneski, A., Konstantinidis, E. & Bamidis, P. (2010). “Affective Medicine: a review of Affective Computing efforts in Medical Informatics*”, Methods of Information in Medicine,* 49 (3). pp. 207-218.

4. Ramadan, R. A. & Vasilakos, A. V. (2017). “Brain Computer Interface: Control Signals Review”. *Neurocomputing*, Volume 223, (5) pp. 26-44.

5. Reaz, M. B. I., Hussain, M. S. & Mohd-Yasin, F. (2006). “Techniques of EMG signal analysis: detection, processing, classification and applications”. *Biol. Proced.* Online, 8(1), pp. 11-35.

6. Lugovaya, T. S. (2005). “Biometric human identification based on electrocardiogram”. Saint-Petersburg, Russian Federation, *Publ.* “*LETI”.*

7. Khoma, Y. V., Stadnyk, B. I., Mykyichuk, M. M. & Frish, S. (2018). “Metody i zasoby vymiryuvannya ta komp’yuternoho opratsyuvannya biosyhnaliv”. [Methods and assets of biosignal measuring and computer processing]. *Measuring Equipment and Metrology*. Vol. 79(3), pp. 5-16 (in Ukrainian).

8. Seonwoo, M., Byunghan, L. & Sungroh, Y. (2017). “Deep learning in bioinformatics”. *Briefings in Bioinformatics*, Vol. 18(5), pp. 851-869.

9. Miotto, R., Wang, F., Wang, S., Jiang, X. & Dudley, J. T. (2018). “Deep learning for healthcare: review, opportunities and challenges”. *Briefings in Bioinformatics*, Vol. 19(6), pp. 1236-1246.

10. Arsirii, O. O. & Manikaeva, O. S. (2019). “Models and methods of intellectual analysis for medical-sociological monitoring’s data based on the neural network with a competitive layer”. *Applied Aspects of Information Technology*. Vol. 2, No. 3, pp. 173-185. DOI: 10.15276/aait.03.2019.1.

11. Jang, H. & Cho, K. (2019). “Applications of deep learning for the analysis of medical data”. *Arch. Pharm*. Res. 42, pp. 492-504.

12. Duke, J., Ryan, P., Suchard, M., Hripcsak, G., Jin, P., Reich, C., Schwalm, M.-S., Khoma, Y., Wu, Y., Xu, H., Shah, N., Banda, J. & Schuemie, M. (2017). “Risk of Angioedema Associated with Levetiracetam Use: Findings of the Observational Health Data Sciences and Informatics”. *Epilepsia* Vol. 58(8), pp. e101-e106.

13. Pinto, J. R., Cardoso J. S. & Lourenco, A. (2018). “Evolution, Current Challenges, and Future Possibilities in ECG Biometrics”. *IEEE Access*, Vol. 6, *IEEE Access* 6: pp. 34746-34776.

14. Rajkomar, A., Dean, J. & Kohane, I. (2019). “Machine learning in medicine”. *New England Journal of Medicine*, 380(14), pp. 1347-1358.

15. Zanini, R. A., Colombini E. L. & de Castro, M. C. F. (2019). “Parkinson’s Disease EMG Signal Prediction Using Neural Networks”. *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy, pp. 2446-2453.

16. Q. Gui, Z. Jin & W. Xu (2014). “Exploring EEG-based biometrics for user identification and authentication”. *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1-6.

17. Wieclaw, L., Khoma, Y., Falat, P., Sabodashko, D. & Herasymenko, V. (2017). “Biometric Identification from Raw ECG Signal Using Deep Learning Techniques”. *The 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*. 21-23 September 2017, Bucharest, Romania, pp.129-133.

18. (2012). Gacek, A., Pedrycz W. (Eds.). “ECG Signal Processing, Classification and Interpretation*”. A Comprehensive Framework of Computational Intelligence*. *Verlag: Springer*, London, 278 p.

19. (2019). Hannun, A.Y., Rajpurkar, P., Haghpanahi, M. (et al.). “Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network”. *Nat Med 25*, pp. 65-69.

20. Crone, B. (2011). “Multiphysiological Parameter Patient Monitoring”. *Analog Devices, Inc.* Technical Article, MS-2126.

21. Von Luhmann, A., Wabnitz, H., Sander, T. & Muller, K.-R. (2017). “M3BA A Mobile, Modular, Multimodal Biosignal Acquisition Architecture for Miniaturized EEG-NIRS-Based Hybrid BCI and Monitoring”. *IEEE Transactions on Biomedical Engineering*, Vol. 64(6), pp. 1199-1210.

22. Ahn, J. W.; Ku, Y. & Kim, H. C. (2019). “A Novel Wearable EEG and ECG Recording System for Stress Assessment”. *Sensors* 19, 1991, pp.1-14.

23. Baby, B., Manikandan, M. S. & Soman, K. P. (2011). “Automated Cardiac Event Change Detection for Continuous Remote Patient Monitoring Devices”. *In Proc.: 1st Int. Conf. on Wireless Techn. for Humanitarian Relief*, pp. 225-232.

24. Ng, K. A. & Chan, P. K. (2005). “A CMOS Analog Front-End IC for Portable EEG/ECG Monitoring Applications”. *IEEE Transactions on Circuits and Systems*, Vol. 52, (11), pp. 2335-2347.

25. Celik, N., Manivannan, N. & Balachandran, W. (2016). “Evaluation of a Behind-the-Ear ECG Device for Smartphone based Integrated Multiple Smart Sensor System in Health Applications”. *(IJACSA) Intern. J. of Advanced Comp. Sci&Appl.*, Vol. 7(7), pp. 409-418.

26. Sasiadek, J. Z. (2002). “Sensor Fusion”. *Annual Reviews in Contro*, pp. 203-228.

27. Khoma, V., Pelc, M., Khoma, Y. & Sabodashko, D. (2018). “Outlier Correction in ECG-Based Human Identification”. In: Hunek W., Paszkiel S. (eds) (2018). “Biomedical Engineering and Neuroscience”. *BCI Advances in Intelligent Systems and Computing*, Vol. 720. pp. 11-22. *Springer*, Cham.

28. Wrzuszczak, M. & Khoma, Y. (2016). “Wykorzystanie sztucznych sieci neuronowych do zmniejszenia błędów przetworników impedancji”. *XLVIII Międzyuczelniana Konferencja Metrologów MKM`, AGH*, Kraków, 5-7 września 2016 (in Polish).

29. Khoma, V. V., Khoma, Y. V., Sabodashko, D. V. & Khoma, P. P. (2019). “Avtoenkodery dla opratsyuvannya promakhiv syhnaliv EKG u systemi biometrychnoyi avtentyfikatsiyi” [Autoencoder for ECG signal outlier processing in system of biometric authentication], *Artificial Intelligence*, No. 1-2, pp. 101-110 (in Ukrainian).

30. Urigüen, J. A. & Garcia-Zapirain, B. (2015). “EEG artifact removal-state-of-the-art and guidelines”. *J Neural Eng.* 12 (3). pp. 2006-2015.

31. Dudykevych, V. B., Khoma, V. V., Chekurin, V. F., Khoma, Y. V. & Sabodashko, D. V. (2019). “Normalizatsiya syhnaliv EKG dlya zastosuvannya v systemakh biometrychnoyi identyfikatsiyi”. [ECG signals normalization for systems of biometric identification]. *Scientific notes of Taurida National V. I. Vernadsky University*,

Issue: Technical Science: Vol. 30 (69), Part 1, No. 4. pp. 49-56 (in Ukrainian).

32. Khoma, Y. V. (2019). "Klasyfikatsiya vibroartrohrafichnykh syhnaliv z vykorystannyam khvylʹkovoho peretvorennya i tekhnolohiy mashynnoho navchannya". [Classification of vibroartographic signals based on wavelet transformation and machine learning techniques]. *Journal of Lviv Polytechnic National University "Information Systems and Networks"*, Vol. 5, pp. 40-52 (in Ukrainian).

33. Pelc, M., Khoma, Y. & Khoma, V. (2019). "ECG Signal as Robust and Reliable Biometric Marker: Datasets and Algorithms Comparison" *Sensors*, 19(10), 2350, pp. 1-8.

34. (2018) "ADAS1000/ADAS1000-1/ADAS1000-2 Low Power, Five Electrode Electrocardiogram". *Analog Front End. Data Sheet. Analog Devices*. 85 p.

35. (2019). "ADS1299-x Low-Noise, 4-, 6-, 8-Channel, 24-Bit, Analog-to-Digital Converter for EEG and Biopotential Measurements". *Texas Instruments Incorporated*. 83 p.

36. Nikolayev, D. V., Smirnov, A. V., Bobrinskaya, I. G. & Rudnev, S. G. (2009). "Bioimpedansnyy analiz sostava tela cheloveka" [Bioelectric impedance analysis of human body composition]. Moscow, Russian Federation, *Nauka*, 392 p. (in Russian).

37. Smajda, M., Khoma, V., Khoma, Y. & Otenko, V. (2019). "Zastosowanie technologii cyfrowego przetwarzania sygnałów w nowoczesnych układach reograficznych". *Przeglad Elektrotechniczny*. R. 95, Nr. 11, pp. 231-237 (in Polish).

Received. 11.02.2020
Received after revision 14.02.2020
Accepted 19.02.2020

УДК 005.8

[1]**Хома, Юрій Володимирович**, кандидат техніч. наук, доцент інституту комп'ютерних технологій, автоматики і метрології, E-mail: yurii.v.khoma@lpnu.ua, Scopus ID: 55675125300,
ORCID: http://orcid.org/0000-0002-4677-5392
[2]**Мірослав Шмайда**, кандидат техніч. наук, професор, директор інституту автоматики,
E-mail: m.szmajda@po.edu.pl, Scopus ID: 24767924400, ORCID: http://orcid.org/0000-0003-3467-951X
[3]**Маріуш Пелц,** кандидат техніч. наук, професор кафедри комп'ютерних і інформаційних систем,
E-mail: m.pelc@greenwich.ac.uk, Scopus ID: 9335890000, ORCID: http://orcid.org/0000-0003-2818-1010
[1]Національний університет «Львівська політехніка», вул. С. Бандери, 12, м. Львів, Україна, 79013
[2]Опольський політехнічний університет, вул. Прушковська, 76, м. Ополе, Республіка Польща, 45758
[3]Грінвічський університет, вул. Парк Роу, м. Лондон SE10 9LS, Великобританія

РОЗВИТОК НАУКОВО-МЕТОДОЛОГІЧНИХ ЗАСАД ЗАСТОСУВАННЯ МАШИННОГО НАВЧАННЯ В ОПРАЦЮВАННІ БІОСИГНАЛІВ

Анотація: *У статті проаналізовано стан і перспективи використання машинного навчання комп'ютерних біоінформатичних системах. Показано, що різнорідність біомедичних даних і широке коло завдань біоінформатики позначилися на виробленні спеціалізованих рішень для кожної окремої сфери чи задачі. Це ускладнює можливість зіставлення ефективності окремих методів та використання кращих варіантів для проектування систем під нові завдання. Представлени результати досліджень, що стосуються розвитку принципів побудови систем комп'ютерного опрацювання біосигналів із використанням технологій машинного навчання. Обґрунтувано доцільність виділення системних рівнів в процесі опрацювання біосигналів та окреслено їх функції. Інноватиність запропонованого підходу полягає у розмежуванні функцій нижнього, середнього і верхнього рівнів від методів, за допомогою яких вони реалізуються, а також варіантів імплементації цих методів на апаратних і програмних засобах. Підкреслюється, що середній системний рівень є значною мірою інваріантним, як до вирішуваної задачі, так і до виду біосигналу. В той же час, верхній рівень є специфічним щодо завдання, а нижній рівень - специфічним щодо виду біосигналу. Чітке окреслення функцій кожного системного рівня та міжрівневих інтерфейсів відкриває перспективи структуризації відомостей під час аналізу відомих рішень, спрощує порівняльний аналіз та зіставлення ефективності цих рішень. Спрощується і потенційно пришвидшується процес проектування комп'ютерної системи під конкретні завдання за рахунок можливості перенесення кращих результатів із суміжних задач. На основі розробленої концепції трьох системних рівнів розглянуто і проаналізувано коло проблемних завдань, пов'язаних із застосуванням машинного навчання опрацювання біосигналів на всіх системних рівнях.*

Ключові слова: *біоінформатика; комп'ютерний аналіз біосигналів; машинне навчання; системні рівні опрацювання біосигналів; міжрівневі інтерфейси*

УДК 004.9

[1]**Хома, Юрий Владимирович**, кандидат технич. наук, доцент института компьютерных технологий, автоматики и метрологии, E-mail: yurii.v.khoma@lpnu.ua, Scopus ID: 55675125300,
ORCID: http://orcid.org/0000-0002-4677-5392
[2]**Міроcлав Шмайда,** кандидат технич. наук, профессор, директор института автоматики,
E-mail: m.szmajda@po.edu.pl, Scopus ID: 24767924400, ORCID: http://orcid.org/0000-0003-3467-951X
[3]**Маріуш Пелц,** кандидат технич. наук, профессор кафедры компьютерных и информационных систем, E-mail: m.pelc@greenwich.ac.uk, Scopus ID: 9335890000,
ORCID: http://orcid.org/0000-0003-2818-1010
[1]Национальный университет «Львовская Политехника», ул. С. Бандеры, 12, г. Львов, Украина, 79013
[2]Опольский политехнический университет, ул. Прушковская, 76, г. Ополе, Республика Польша, 45758
[3]Гринвичский университет, ул. Парк Роу, г. Лондон SE10 9LS, Великобритания

РАЗВИТИЕ НАУЧНО-МЕТОДИЧЕСКИХ ПОДХОДОВ МАШИННОГО ОБУЧЕНИЯ В ОБРАБОТКЕ БИОСИГНАЛОВ

Аннотация: *В статье проанализированы состояние и перспективы использования машинного обучения компьютерных биоинформатичних системах. Показано, что разнородность биомедицинских данных и широкий круг задач биоинформатики сказались на выработке специализированных решений для каждой отдельной сферы или задачи. Это затрудняет возможность сопоставления эффективности отдельных методов и использование лучших вариантов для проектирования систем под новые задачи. Представлены результаты исследований, касающихся развития принципов построения систем компьютерной обработки биосигналов с использованием технологий машинного обучения. Обосновано целесообразность выделения системных уровней в процессе обработки биосигналов и обозначены их функции. Инновативность предложенного подхода заключается в разграничении функций нижнего, среднего и верхнего уровней от методов, с помощью которых они реализуются, а также вариантов имплементации этих методов на аппаратных и программных средствах. Подчеркивается, что средний системный уровень в значительной мере инвариантен, как к решаемой задаче, так и к виду биосигналов. В то же время, верхний уровень является специфическим о касательно задачи, а нижний уровень - специфическим относительно вида биосигналов. Четкое определение функций каждого системного уровня и межуровневых интерфейсов открывает перспективы структуризации данных при анализе известных решений, упрощает сравнительный анализ и сопоставление эффективности этих решений. Упрощается и потенциально ускоряется процесс проектирования компьютерной системы под конкретные задачи за счет возможности переноса лучших результатов из смежных задач. На основе разработанной концепции трех системных уровней были рассмотрены и проанализуваны проблемные задачи, возникающие при применении машинного обучения для обработки биосигналов на всех системных уровнях.*

Ключевые слова: *биоинформатика; компьютерный анализ биосигналов; машинное обучение; системные уровне проработки биосигналов; межуровневые интерфейсы*

Yuriy Khoma, Candidate of Technical Sciences,
Assistant Professor
Research field: Artificial intelligence, bioinformatics, sensors and electrical measurements

Miroslaw Szmajda, PhD Hab. Eng, Professor,
Research field: Advanced Methods of Digital Signal Processing, Bioinformatics and Bioengineering

Mariusz Pelc, PhD Hab. Eng, Professor,
Research field: Computing in Mathematics, Natural Science, Engineering and Medicine