EXPLAINABLE ARTIFICIAL INTELLIGENCE IN THE DIAGNOSIS OF CARDIOVASCULAR DISEASES IN SMALL SAMPLES

Natalia Gusarova, Ivan Tomilov, Danil Zmievskii, Vladimir Shilonosov, Tatiana Polevaya and Aleksandra Vatian
ITMO University,
Saint Petersburg, Russia

ABSTRACT

In this work several DNN-based solutions aimed at classifying different of CVDs in conditions of small training samples are proposed. Using GradCAM technology, the main areas of attention of DNNs in these conditions are indentified, and then experimentally being compared with the areas of attention of cardiologists when making a clinical decision. It is shown that even on small samples it is possible to achieve the efficiency of SOTA solutions obtained on large datasets, which should increase the confidence of medical staff, including less qualified ones, in DNN as a means of supporting rapid diagnosis of CVD in acute cases. In turn, combining the results of the physician and AI tools can improve the quality of diagnosis and, therefore, the patient's chances for a speedy and successful cure.

KEYWORDS

Cardiovascular Disease, Explainable Artificial Intelligence, XAI, Small Samples

1. INTRODUCTION

1.1 Nature of the Problem

Cardiovascular disease is the cause of death worldwide, and this trend is only increasing in the 21st century. Although AI-based decision support applications in such areas have large enough datasets for training, which are the basis for a quite high level of accuracy, however, practitioners are in no hurry to use such tools. The reason is largely due to the fact that machine learning methods based on deep neural networks (DNNs), which are the de facto standard for use in medical applications, are fundamentally “opaque”. Without being able to follow the decision-making process performed by the DNN, the doctor, who is personally responsible for the patient, has no reason to fully trust its results and cannot fully take them into account when making his decisions.

The problem is even more exacerbated when it comes to life-threatening conditions and rapidly developing diseases such as myocardial infarction. In this case, on the one hand, the level of responsibility of the doctor for the clinical decision for a particular patient increases. On the other hand, the DNN should quickly retrain on a small set of data, which can be obtained in a short time of the development of the disease in this patient. Under these conditions, the performance accuracy of the DNN expectedly falls, which further reduces the doctor's confidence in its results. As a rule, in such cases, the results formed by DNN, which could act as a source of "second opinion" for the attending physician, are completely disavowed, which negatively affects the diagnosis and further course of the disease.

The way out of this collision may be to involve the concept of explainable artificial intelligence (XAI) [Vilone, Speith]. Namely, it is necessary to transfer the process of work of the DNN to the state of a “transparent box”, that is, to enable the doctor to track and interpret all stages of this process. For example, when diagnosing a heart attack, a sign that is traditional for clinical practice is elevation of the ST segment of the ECG. If, when analyzing this ECG, the DNN pays the most attention to the ST segment, then for the doctor this will be evidence that the network is working correctly, and its results can be trusted.

1.2 Previous Works

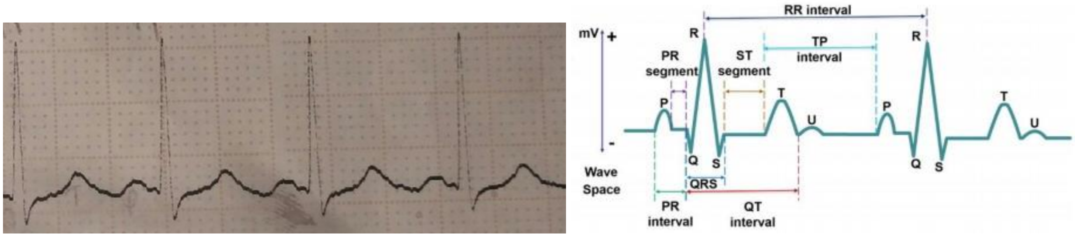
The literature [Barz, Olson, Brigato] presents a number of ways to improve the efficiency of DNN in the context of small samples, including: transfer learning [Xu], data augmentation [Pan], GAN’s [Li], Bayesian neural networks [Madan], etc. However, as a detailed analysis shows, to support the operational clinical decisions, which is typical for CVD, two approaches seem to be the most promising.

The first one is to reduce the dimension of the model. According to [Brigato], model complexity is a critical factor when only a few samples per class are available. Here, measures such as the transition to lower-dimensional DNNs while maintaining the basic architecture, dropout regularization, and others that help reduce the number of adjustable model parameters can be applied.

The second approach uses architectures specialized for learning from small samples, such as Siamese neural networks (SSN) [Bromley] and few-shot learning (FSL) [Santoro].The SNN inputs consist of paired signals (ECG in our case), with each signal passed through an identical deep-convolutional subnetwork. The SNN is typically trained using a contrastive loss function, thereby highlighting the difference in input signals. The FSL is based on a similarity function which is formed by pre-training the deep model on a large-scale labeled dataset. From a technological point of view, it is important that in order to represent the similarity function a SNN can be used. Then, using this similarity function, the FS model can determine the closest class for each newly seen item.

Both SSN and FSL, alone and in combinations, are widely used to detect differences in artifacts of various domains [Liu D., Wang, Figueroa-Mata], and in particular in medicine, including clinical language processing [Li (2023), Oniani], high-tech medical imaging [Li (2020), Deepak], sleep staging based on single-channel EEG [You], ECG time series classification [Gupta, Pałczyński].

The main processed artifact in the classification of CVD using DNN is the electrocardiogram (ECG) (Figure 1), which consists of a sequence of QRS complexes [Wagner, Liu F.F., Moody]. From the point of view of the formation of the DNN feature space, CVD can be divided according to the type of manifestation on the ECG into CVD associated with displacements of the main (R) wave, and CVD associated with local changes in other components of the QRS complex. The first type includes such CVDs as supraventricular premature beat, premature ventricular contraction, fusion of ventricular and normal beat, etc. The most important representative of CVD of the second group is myocardial infarction (MI), which can manifest itself on the ECG as an elevation (STE) and (or) depression (STD) of the ST segment. Obviously, the classification of CVD of the second type is a much more difficult task than the first, regardless of the nature and type of classifier used, which must be taken into account when comparing the effectiveness of classifiers.



a b

Figure 1. Fragment of ECG (a) and its characteristic intervals (b)

An analysis of the literature allows us to select as SOTA-level for the classification of CVD by ECG the works [Gupta, Pałczyński, Jeong, Strodthoff], where the achieved level of accuracy was 57–92%, depending on the type of diseases classified (more detailed data are given in section 3.1). It should be emphasized that the same level is typical for cardiologists when making operational decisions on the ECG. According to the results of a meta-analysis [Cook], the pooled accuracy in ECG interpretation performed by medical personnel of various skill levels ranges from 56% for residents to 69% for practicing physicians and 75% for cardiologists.

In this regard, the explainability of AI-based models for the performed ECG diagnosis acquires a special role. The literature [Paralic] notes the lack of explainability of DNN-based methods in relation to the medical field and to ECG interpretation, in particular. For example, the review [Jeyakumar] considers XAI methods for

ECG data within the framework of general approaches to multiple types of data, with an emphasis on the explainability of the ECG as a whole. To do this, explanation-by-example methods are offered as the optimal way of XAI. As a means for generating such examples, it is proposed to use a pair of encoder-decoder [Attia]. A similar problem is solved in [Paralic] – to mark parts of ECG the most important for manual disease diagnosis, and perturbation-based XAI methods turned out to be optimal.

For visual explanations of DNN predictions of CVD different XAI methods are used. For example, [Anand] applied SHapley Additive exPlanations (SHAP) for ECG containing CVD of arrhythmia group. [Strodthoff] employed the Gradient×Input method. However, in ECG analysis performed by DNN, backpropagation methods of XAI dominate, among which GradCAM [Taniguchi, Raza, Ganeshkumar] ranks first.

A number of works have proposed methods for testing the effectiveness of the developed XAI tools. In [Paralic, Cabitza] for this purpose the questionnaire offered to doctors is used. [Strodthoff] highlights areas of the ECG curves that do or don't contribute to the correct assessment of myocardial infarction. [Siddiqui] proposes a framework for comparative evaluating of XAI methods accomplished using expertly annotated ECG signals. The technique uses two metrics - similarity and stability.

Let's summarize. Building AI-based CVD classification tools with small training samples is an open area of research. The main ways here are recognized as reducing the dimension of the model and using FSL-based architectures. The accuracy of existing systems varies greatly depending on the type of CVD and is comparable to that of practitioners. To increase the confidence of doctors in the results of AI systems, it is advisable to use XAI. However, in the subject area under consideration, XAI solves either the problem of improving the usability of ECG analysis in the manual diagnosis of CVD, or the problem of assessing the effectiveness of the AI system itself in diagnosing CVD. Evaluation of the possibilities of combining the results obtained by doctors and the AI system in the development of clinical decisions on CVD is not presented in the literature available to the authors.

1.3 Our Contribution

1. We have proposed DNN-based solutions to classify different CVDs having small training samples, based on two approaches - reducing the dimension of the model and SNN + FSL-complexing, which provide efficiency on small samples that is superior to SOTA-level.

2. Using GradCAM technology, we identified the main areas of attention of DNNs in these conditions and compared them with the areas of attention of cardiologists when making a clinical decision;

3. We have shown that interpretations of CVD by typical ECG performed by specialized DNN-based AI tools and medical personnel are adequate in terms of the accuracy achieved and comparable in terms of the distribution of areas of attention when making a decision.

2. METHOD AND MATERIALS

To develop a solution that uses the principle to reduce the dimension of the model, we used the ResNet-18 architecture [Jeong] as the initial schema. To adapt the network to work on small samples, the following measures were used: decrease in the number of network parameters due to the transition to a smaller number of layers (from 18 to 10); replacement of the activation function from ReLu to Gelu; combining regularizers L1 and L2; optional use of the dropout regularizer. To develop a few-shot solution, we used the architecture [Gupta] as the initial schema, with the following changes: after the last convolutional layer, a dropout layer was added for regularization in order to identify the most relevant features, and the vector proximity function was changed from L2-distance to cosine similarity as more robust to the direction of embeddings. Both solutions use the Grad-CAM [Selvaraju] technique, which is compatible with any DNN architecture, as a means of implementing attention. F1-measure and accuracy were used as performance evaluation metrics.

To verify the zones of attention allocated by DNN, they were compared with the zones of attention used in the diagnosis of relevant diseases in clinical practice. For this purpose, expert cardiologists (three specialists) were involved, who performed manual marking of their areas of attention on ECG records extracted from the dataset and visualized in the form equivalent to native paper ECG. For all types of classified CVD, zones of attention were manually allocated in the diagnosis of characteristic QRS complexes. The total number of labeled data for cardiovascular diseases was 70 cases. For a comparative assessment of the areas of attention allocated by medical experts and the DNN, the Jaccard measure was used:

$$K_J = \frac{S(A \cap B)}{S(A \cup B)} \ , \qquad (1)$$

where the numerator is the measure of the intersection of the compared objects, and the denominator is the measure of their union, respectively. In the case of comparing the classification by cardiovascular diseases, the length of the corresponding zone along the time axis was used as such a measure.

To improve the comparability of the results, the training and testing of the developed solutions was carried out on the same datasets as the SOTA sources. Characteristics of datasets and specific application in each case are presented in Table 1.

Table 1. Datasets characteristics

Name, source	Characteristics	Where applied	Application specifics
CPSC2018 [Liu F.F.]	6877 records, 8 classes	[Jeong]	Classification into 5 classes out of 8
		Present work	Decimation by 4 times (up to 1741 records), 5 classes used
PTB-XL [Wagner]	21837 records, 5 classes, 71 types of heart diseases	[Pałczyński]	Classification into 5 and(or) 20 classes
		Present work	Classification into 2 classes
ECG5000 [Chen F.]	Length - 140, 5 classes	[Gupta]	for training, classification into 5 classes
MIT-BIH Arrhythmia Database [Moody]	109,000 heart-beats, 47 classes	[Gupta]	for testing, classification into 5 classes
PTB [Bousseljot}	549 records of 12-lead ECG, 9 classes	[Strodthoff]	Classification into 2 classes

3. RESULTS AND DISCUSSION

3.1 Results

Efficiency estimates of the developed solutions in comparison with SOTA models are presented in Table 2 for FCL-based schemes and in Table 3 for ResNet-based schemes. The following designations are used: N – the number of classes, K –- the number of labeled samples for each of the N classes, n/d – no data available; I-AVB – first-degree atrioventricular block, RBBB – right bundle branch block, PVC – premature ventricular contraction, STD – ST-segment depression, STE – ST-segment elevated, MI – myocardial infarction.

Table 2. Efficiency estimates of the developed solutions in comparison with SOTA models for FCL-based schemes

Source	Arcitecture	Dataset	CVD type	N	k	Accuracy	F1-score
[Gupta]	FSL +SSN	ECG5000	normal / arrytm	5	1	0.848	0.852
					2	0.905	0.907
					3	0.920	0.921
					4	0.920	0.922
					5	0.923	0.924
[Pałczyński]	FSL + SVM with linear kernel	PTB-XL	normal / not normal	2	5	0.912	0.903
			normal / arrytm / MI	5	5	0.788	0.717
			normal / arrytm / MI	20	5	0.653	0.320
Present work	FSL +SSN	PTB-XL	normal / MI	2	1	0.712	0.703
					2	0.691	0.705
					3	0.829	0.831
					4	0.832	0.834
					5	0.832	0.834

Table 3. Efficiency estimates of the developed solutions in comparison with SOTA models for ResNet-based schemes

Source	Arcitecture	Dataset	CVD type	N	Accuracy	F1-score
[Jeong]	ResNet-18	CPSC2018 (full dataset)	I-AVB	5	0.837	0.803
			RBBB		0.860	0.850
			PVC		0.613	0.636
			STD		0.822	0.760
			STE		0.571	0.525
[Strodthoff]	ResNet-based	PTB	normal / MI	2	n/d	0.932
Present work	ResNet-10	CPSC2018 (decimation by 4 times)	I-AVB / normal	4	0.853	0.834
			RBBB / normal		0.890	0.852
			PVC		0.702	0.647
		PTB-XL	normal / MI	2	0.715	0.683

Figure 2 shows an example of the implementation of XAI in the developed solutions (a solution based on the ResNet-10 architecture, the input is characteristic ECG leads, labeled by experts as acute myocardial infarction). The attention zones of the DNN are marked with vertical stripes, and the intensity of the color corresponds to the degree of attention of the neural network. A fragment of such an ECG, marked up by an expert, and the corresponding DNN attention zones are shown in Figure 3, a, b.

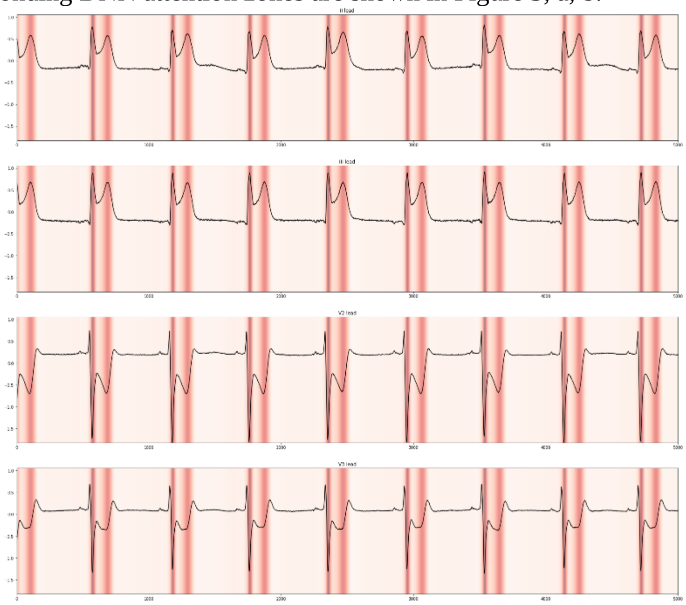
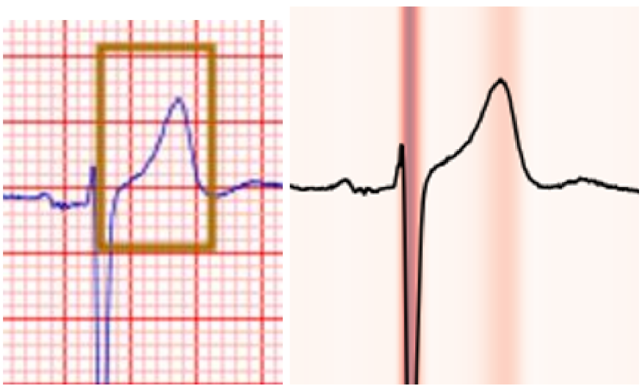


Figure 2. DNN attention zones in classification of ECG with MI



a b

Figure 3. Comparison of attention zones of an expert (a) and of DNN (b) in the classification of ECG with MI

The mean value and standard deviation of the Jaccard measure (1), averaged for both proposed solutions, were: mean = 0,461, std = 0.076 respectively.

3.2 Discussion

As can be seen from Figure 2, the attention of the neural network is stably localized on the characteristic elements of the ECG - this is the peak of the R wave and the ST segment. In both cases, the first "burst" of attention of the neural network falls on the peak of the R wave, which is the main indicator of heart rate, i.e. the nature of the work of the heart muscle as a whole. At the same time, the second "burst" of attention of the neural network is localized in that area of the ST segment, which determines a specific type of pathology. Thus, we can say that the distribution of attention zones of the neural network corresponds to the physiological mechanism of the manifestation of certain pathologies of the cardiovascular system on the ECG.

Small std value of the Jaccard measure with a sufficiently high mean value indicate that the tactics of expert doctors and the "tactics" of DNN work are stable and quite comparable. This conclusion is well confirmed by comparison of Figures 3 (a, b): DNN steadily localizes attention to two ECG zones corresponding to meaningful signs of the disease; the same zones are included in the area of attention of the doctor-expert; at the same time, the expert doctor, considering these zones as the main ones, also distributes his attention over the entire ST-segment of the ECG.

An analysis of Tables 2 and 3 shows that the effectiveness indicators of the developed solutions are generally consistent or even somewhat higher compared to SOTA level, even with a 4-fold reduction in the training set. We emphasize that this level is comparable or even higher than the level of medical personnel of various skill levels when deciphering the ECG.

At the same time, in all studies, the effectiveness indicators depend on the type of CVD, in particular, when classifying MI and its potential manifestations, such as STD and STE, they are usually less than when classifying diseases with a purely arrhythmic manifestation on the ECG. In this case, the stability of the coincidence of the "opinions" of the expert and the DNN is important, which can increase the confidence of the experts in the DNN data, even at relatively small values of the F-measure, and vice versa. For example, in [Strodthoff], with a high value of the F1-measure (Table 3), on 4 out of 12 ECG leads, network attention varies, and it identifies dominant signs against the presence of MI, and on 8 – for it. At the same time, in our work, at moderate values of the F1-measure (Tables 2, 3), this agreement is quite high (std = 0.076).

Finally, comparing the results of Figure 3 and Tables 2, 3 gives grounds to treat DNNs and expert doctors as equally powerful classifiers, but taking into account different aspects of the manifestation of CVD on the ECG. This, in turn, opens the way for formal ensembling of such decisions, which has proven to be effective in various decision-making problems [Ali].

4. CONCLUSION

We have proposed DNN-based solutions aimed at classifying different of CVDs in conditions of small training samples. The effectiveness of the developed solutions depends on the CVDs classified, but in all cases is at the level of the best DNN-based SOTA models, as well as with the best levels of healthcare experts..

Using GradCAM technology, we identified the main areas of attention of DNNs in these conditions, experimentally compared them with the areas of attention of cardiologists when making a clinical decision, and obtained a match at std = 0.076. Such a coincidence, even at relatively low levels of F-measures (F1=0.834) can serve as a basis for increasing physicians' confidence in using DNN as a support tool for making clinical decisions in CVD, and vice versa.

Thus, we have shown that even on small samples it is possible to achieve the efficiency of SOTA solutions obtained on large datasets, which should increase the confidence of medical staff, including less qualified ones, in DNN as a means of supporting rapid diagnosis of CVD in acute cases. In turn, combining the results of the physician and AI tools can improve the quality of diagnosis and, therefore, the patient's chances for a speedy and successful cure.

Our research also revealed a number of limitations of the considered approaches. As noted in Section 1.2, diagnosis of various diseases based on ECG can be associated with both relatively simple (albeit somewhat non-linear) signs characteristic of disorders such as arrhythmias, and rather subtle and complex signs, as in the case of myocardial infarction and its subtypes. . The doctor, analyzing the ECG, necessarily draws on a fairly extensive context, both in the form of previous experience and in the form of additional signs. At the same time, the neural network model, which is trained based on the detection of ECG similarity, relies only on the features it has identified. Due to the specifics of training based on minimizing the l2-distance between ECG embeddings, the complexity of the identified features and the resolution associated with the adjusted parameters may be insufficient. Partially, this problem is solved by adding mechanisms for evaluating the attention of the model, however, additional research is required on the influence of architecture, advanced regularization methods, the chosen method for generating embeddings and calculating the distance between them, as well as other parameters on the properties listed earlier.

ACKNOWLEDGEMENT

This work was supported by Russian Science Foundation, Grant № 23-11-00346.

REFERENCES

Ali S., Tirumala S.S., Sarrafzadeh A. (2015) Ensemble learning methods for decision making: Status and future prospects. *14th International Conference on Machine Learning and Cybernetics (ICMLC)*, July 2015. doi: 10.1109/ICMLC.2015.7340924.

Anand A., et al. (2022) Explainable AI decision model for ECG data of cardiac disorders. *Biomedical Signal Processing and Control*. V. 75, May 2022, 103584

Attia Z.I., Friedman P.A. (2023) Explainable AI for ECG-based prediction of cardiac resynchronization therapy outcomes: learning from machine learning? *European Heart Journal*, V. 44, Is. 8, 21 Feb. 2023, Pp. 693–695, https://doi.org/10.1093/eurheartj/ehac733

Barz B., Denzler J. (2020) Deep Learning on Small Datasets without Pre-Training using Cosine Loss. IEEE Winter Conference on Applications of Computer Vision (WACV) 2020. arXiv:1901.09054v2 [cs.LG] 11 Dec 2019

Bousseljot R, Kreiseler D, Schnabel, A. (1995) Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomedizinische Technik*, Band 40, Ergänzungsband 1 (1995) S 317

Brigato L., Iocchi L. (2021) A Close Look at Deep Learning with Small Data. 25th International Conference on Pattern Recognition (ICPR), 2021, pp. 2490-2497, DOI: 10.1109/ICPR48806.2021.9412492

Bromley J., et al. (1994). Signature verification using a "Siamese" time delay neural network. *Advances in Neural Information Processing Systems*. 6: 737–744.

Cabitza F., et al. Painting the black box white: experimental findings from applying XAI to an ECG reading setting. arXiv:2210.15236 [cs.AI]. https://doi.org/10.48550/arXiv.2210.15236

Chen Y., et al. (2015) The UCR time series classification archive, July 2015. www.cs.ucr.edu/∼eamonn/time series data/.

Cook D.A., Oh S.Y., Pusic M.V. (2020). Accuracy of Physicians' Electrocardiogram Interpretations. *JAMA Intern Med.* 2020 Nov; 180(11). P. 1–11.

Deepak S., Ameer P.M. (2021). Brain tumour classification using siamese neural network and neighbourhood analysis in embedded feature space. *IMA*, V.31, Is.3, September 2021, Pp. 1655-1669. https://doi.org/10.1002/ima.22543

Figueroa-Mata G., Mata-Montero E. (2020) Using a Convolutional Siamese Network for Image-Based Plant Species Identification with Small Datasets. *Biomimetics (Basel)*. 2020 Mar; 5(1): 8. doi: 10.3390/biomimetics5010008

Ganeshkumar M., et al. (2021) Explainable Deep Learning-Based Approach for Multilabel Classification of Electrocardiogram. IEEE Transactions on Engineering Management, doi: 10.1109/TEM.2021.3104751.

Gupta P., Bhaskarpandit S., Gupta M. (2021) Similarity Learning based Few Shot Learning for ECG Time Series Classification. *2021 Digital Image Computing: Techniques and Applications (DICTA)*, Gold Coast, Australia, 2021, pp. 1-8, doi: 10.1109/DICTA52665.2021.9647357.

Jeong D.U., Lim K.M. (2021) Convolutional neural network for classification of eight types of arrhythmia using 2D time–frequency feature map from standard 12-lead electrocardiogram. *Sci Rep* 11, 20396. 2021. https://doi.org/10.1038/s41598-021-99975-6

Jeyakumar J.V., et al. How can i explain this to you? An empirical study of deep neural network explanation methods. *Adv. Neural Inf. Process. Syst*. 2020, 33, 4211–4222.

Li D.-C., et al. (2021) A Generative Adversarial Network Structure for Learning with Small Numerical Data Sets. *Appl. Sci*. 2021, 11(22), 10823; https://doi.org/10.3390/app112210823

Li M.D., et al. (2020) Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging. *npj Digit. Med.* 3, 48 (2020). https://doi.org/10.1038/s41746-020-0255-1

Li, Q., He, S. (2023) Similarity matching of medical question based on Siamese network. *BMC Med Inform Decis Mak* 23, 55 (2023). https://doi.org/10.1186/s12911-023-02161-z

Liu D., et al. (2022) Design of Siamese Network for Underwater Target Recognition with Small Sample Size. *Appl. Sci.* 2022, 12(20), 10659; https://doi.org/10.3390/app122010659

Liu F.F., et al. (2018) The China physiological signal challenge 2018: automatic identifcation of the rhythm/morphology abnormalities in 12-lead ECGs. In Te 7th International Conference on Biomedical Engineering and Biotechnology (ICBEB 2018) http://2018.icbeb. org/Challenge.html (2018).

Madan A., Prasad R. (2021) B-Small: A Bayesian Neural Network Approach to Sparse Model-Agnostic Meta-Learning, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 2021, pp. 2730-2734, doi: 10.1109/ICASSP39728.2021.9414437.

Moody GB, Mark RG. (2001) The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 20(3):45-50 (May-June 2001). (PMID: 11446209)

Olson M., Wyner A.J., Berk R. (2018) Modern Neural Networks Generalize on Small Data Sets. *32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada*. 10 p.

Oniani D, et al. (2023) Few-Shot Learning for Clinical Natural Language Processing Using Siamese Neural Networks: Algorithm Development and Validation Study. *JMIR AI* 2023;2:e44293. doi: 10.2196/44293

Pałczyński K., et al. (2021) Study of the Few-Shot Learning for ECG Classification Based on the PTB-XL Dataset. *Sensors* 2022, 22(3), 904; https://doi.org/10.3390/s22030904

Pan Y., et al. Knowledge-based data augmentation of small samples for oil condition prediction. *Reliability Engineering & System Safety*, Volume 217, January 2022, 108114. https://doi.org/10.1016/j.ress.2021.108114

Raza A., et al. (2022) Designing ECG Monitoring Healthcare System with Federated Transfer Learning and Explainable AI. arXiv:2105.12497v2 [cs.LG] 10 Jan 2022

Santoro A., et al. (2016). One-shot Learning with Memory-Augmented Neural Networks. arXiv:1605.06065v1 [cs.LG] 19 May 2016

Selvaraju R.R., et al. (2016) Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. arXiv:1610.02391v4 [cs.CV].

Siddiqui K., et al. (2022) Trust Metrics for Medical Deep Learning Using Explainable-AI Ensemble for Time Series Classification. 2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 18-20 September 2022. DOI: 10.1109/CCECE49351.2022.9918458

Speith T. (2022). A Review of Taxonomies of Explainable Artificial Intelligence (XAI) Methods. *2022 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '22), June 21–24, 2022, Seoul, Republic of Korea. 12 p. https://doi.org/10.1145/3531146.3534639

Taniguchi H., et al. (2021). Explainable Artificial Intelligence Model for Diagnosis of Atrial Fibrillation Using Holter Electrocardiogram Waveforms. Int Heart J 2021;62(3):534-539. doi: 10.1536/ihj.21-094.

The top 10 causes of death (Dec. 2020). https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.

Vilone G., Longo L. (Oct. 2020) Explainable Artificial Intelligence: a Systematic Review. arXiv:2006.00093 [cs.AI]. https://doi.org/10.48550/arXiv.2006.00093

Wagner P., et al. (2020) PTB-XL, a large publicly available electrocardiography dataset (version 1.0.1). *Sci. Data* 2020, 7, 1–5.

Wang Z.-M., et al. (2021) A Few-Shot Learning-Based Siamese Capsule Network for Intrusion Detection with Imbalanced Training Data. *Comput Intell Neurosci*. 2021; 2021: 7126913.

Xu Y., et al.(2022) Iterative self-transfer learning: A general methodology for response time-history prediction based on small dataset, *Journal of Computational Design and Engineering*, Volume 9, Issue 5, October 2022, Pp. 2089–2102, https://doi.org/10.1093/jcde/qwac098

You Y., et al. (2023) A Siamese Network-Based Method for Improving the Performance of Sleep Staging with Single-Channel EEG. *Biomedicines* 2023, 11(2), 327; https://doi.org/10.3390/biomedicines11020327