

Prediction of the fact and degree of coronary artery disease using the processing of clinical and instrumental data by artificial intelligence

T. P. Abdualimov¹, A. G. Obrezan^{1,2}

¹ Sogaz International Medical Center,
1/6, Cheboksarskiy per., St. Petersburg, 191186, Russian Federation

² St. Petersburg State University,
7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation

For citation: Abdualimov T. P., Obrezan A. G. Prediction of the fact and degree of coronary artery disease using the processing of clinical and instrumental data by artificial intelligence. *Vestnik of Saint Petersburg University. Medicine*, 2021, vol. 16, issue 3, pp. 153–158.

<https://doi.org/10.21638/spbu11.2021.302>

Aim of the study was to analyze the possibility of using neural network analysis to predict the severity of coronary bed lesion. The study was also designated to determine the performance and accuracy of the trained neural network model receiving input as the structured data and ECG images with the parameters and leads positioning differ from the training sample, and also to compare the efficiency of detecting transient myocardial ischemia with traditional diagnostic methods, such as 24-hour Holter monitoring, treadmill test. Neural network analysis of the available clinical, laboratory and instrumentation data allow to configure the network parameters for further prediction of coronary artery disease. The results obtained in the form of an AUC score allow to consider this method to be effective in the coronary artery disease diagnosis using recorded ECG tape with parameters and lead positioning differ from initial training sample. The efficiency of transient myocardial ischemia detection on the training sample of the trained neural network is higher in comparison with traditional diagnostic methods, such as 24-hour Holter monitoring, treadmill test.

Keywords: coronary arteries, neural networks, deep learning, ECG.

Introduction

According to the 5th edition of the European Cardiovascular Disease Statistics (published in 2017 by the European Society of Cardiology (ESC)), cardiovascular disease (CVD) is the first cause of mortality and morbidity in Europe. In 2015, more than 85 million people on the continent suffer from cardiovascular disease (48 % of men and 52 % of women), resulting in 3.9 million deaths (45 % of all causes of death). In the European Union (EU), 49 million people suffered from cardiovascular diseases with more than 1.8 million deaths (European Cardiovascular Disease Statistics 2017). These findings support that cardiovascular diseases, despite significant advances in diagnosis and treatment, are still the most common cause of morbidity and mortality in Europe. Early accurate diagnosis and prediction assessment are key factors to improve and optimize CVD outcomes.

The development of mathematical methods and hardware allows to use much more progressive algorithms than standard ones — artificial neural networks. The operating

principle based on fuzzy logic. Neural networks are machine-learning method, modeled on the way human processing information [1]. Working with a wealth of input data led to the widespread use of neural networks in numerous areas of medicine [2].

Each neuron (node) in the network stores a numerical value, and all of them form the network architecture. The connection between neurons governs using weight matrices. The last layer encodes the desired results. A neural network containing more than one hidden layer is defined as “deep”, which allows to use multiple level abstractions [3].

Machine learning methods include technologies and algorithms that could improve the work of cardiologist as the amount of data available in clinical practice has significantly increased.

Potentially, it could facilitate and make each stage of patient management more effective [4].

Materials and methods

The study included 130 patients, who underwent elective or emergency coronary catheterization.

The indications for coronary angiography were verified according to the recommendations of the European Society of Cardiology (EOC). The study was carried out in accordance with Good Clinical Practice and Declaration of Helsinki principles. Inclusion and exclusion criteria were defined.

Key inclusion criteria:

- signing the informed consent prior to the study, including the statistical processing of medical history data;
- aged over 18 years;
- indications (elective or emergency) for coronary catheterization;
- recorded electrocardiography (25 mm/s) one day before / or less before the performed coronary catheterization.

Key exclusion criteria:

- ECG identification of arrhythmias as the atrial fibrillation, AV nodal reentrant tachycardia, ventricular tachycardia while recording;
- previous stenting and / or coronary artery bypass grafting;
- pronounced disturbances on the recorded ECG;
- registration of ECG more than 24 hours before coronary catheterization;
- any surgical or medical state that, according to the researcher, could significantly interfere the work of machine learning algorithm in relation to the accuracy of the results.

The doctor, conducting the study, analyzed the medical record data (complaints, anamnesis, objective, laboratory and instrumental data) and download these results into a machine learning database in a binary format.

At the first stage of data collection for each case, structured parameters were entered into a tabular form, as well as the ECG image in jpeg format into the database. Numerous morphometric, objective, laboratory and instrumental data of the patients were used to train neural networks, such as: age, gender, diagnosed acute coronary syndrome (ACS) or

chronic coronary syndrome (CCS), ST segment pathology on the ECG, the presence or absence of concomitant pathology (diabetes mellitus, hypertension, obesity, anemia, previous stroke, atherosclerosis, arrhythmias, dyslipidemias), aggravated heredity, bad habits (smoking, alcohol abuse), stress factors, low physical activity, menopause, increased nutritional intake.

The abovementioned factors were filled in a structured binary form (0, 1) in a tabular format. Registration of ECG on a sample, developed for a neural network training, was carried out using one type of apparatus and the record was transmitted to the machine learning operator in jpeg format. Thus, 22 parameters (key features) were used to develop a neural network learning algorithm.

Neural network was trained on the data obtained from the analysis of coronary angiograms. As “targeted” values were taken: performed stenting or recommended CABG based on coronary catheterization, atherosclerosis, left main coronary artery stenosis, left main coronary artery subocclusion, anterior interventricular artery occlusion, anterior interventricular artery subocclusion, anterior interventricular artery stenosis, circumflex artery occlusion, circumflex artery subocclusion, circumflex artery stenosis, right coronary artery occlusion, right coronary artery subocclusion, right coronary artery stenosis.

The degree of the coronary artery stenosis was filled in the table in numerical form as a percentage, then converted to binary form (1 — stenosis more than 50 %), the rest of the parameters were filled in binary form according to the presence or absence of lesion. The above “targeted” values were predicted by a trained machine learning algorithm.

The algorithm had to solve the coronary arteries classification problem, to predict the absence or presence of lesion. The neural network was used to solve the classification problems, receiving input as the structured data and images and providing output as a multifactorial classification of the main coronary arteries. The ratio for training and testing was 100/30. Predicting and results evaluation were performed on a test sample.

Python libraries were used for designing a neural network architecture (pandas — for tabular data; TensorFlow — for neural networks constructing and their training). Supervised learning method was used on the available data, where the outcomes were known (coronary angiography data), and neural network parameters were adjusted to minimize the error.

The analysis of structured tabular data of the training sample, including 100 patients, was carried out. The average age of the patients was 64 years (from 31 to 89 years), 52 male patients and 48 female patients.

The men median age is less than that of women. In 62 out of 100 patients was verified ACS, ST segment elevation was diagnosed in 19 patients with ACS. Typical anginal pain was observed in 53 out of 62 patients with ACS, as well as in 14 patients without ACS, mainly with chronic coronary artery disease in medical history. Type 2 diabetes had 28 out of 100 patients. In most of cases was observed hypertension. Atherosclerosis was verified in 44 patients. Almost 25 patients had cardiac rhythm disturbances, such as the atrial fibrillation, Ryan high-grade ventricular premature beats. In 27 patients was verified obesity. Anemia of varying severity was observed in 18 cases; 5 patients suffered from acute cerebrovascular accident; 36 patients complained of dyspnea; in 27 cases was observed an increased level of cholesterol. Alcohol abuse was determined in 3 patients, then tobacco smoking in 24 patients; 32 patients in the «training» sample regularly experience

psycho-emotional overload; 32 patients had low physical activity and leading a sedentary lifestyle; 24 patients had increased nutritional intake.

The study was also designated to determine the performance and accuracy of the trained neural network model receiving input as the structured data and ECG images with the parameters and leads positioning differ from the training sample. Patients of the test sample underwent ECG on the devices of various manufacturers. Speed of ECG recordings, amplitude and leads positioning were determined randomly. ECG tapes for each case were converted into digital jpeg format using a portable camera.

Each patient of the test sample underwent 12-channel Holter daily monitoring and treadmill test. The detection results of transient myocardial ischemia were filled in the table for each method separately. (0 — absence of transient ischemia, 1 — presence.)

The assessment of the accuracy of lesions prediction of the main coronary arteries was carried out based on the CT coronary angiography data, the efficiency comparison of transient myocardial ischemia detection was conducted on the basis of prediction calculation of necessity to perform the revascularization of the coronary arteries by the neural network analysis method and the results obtained during 24-hour Holter monitoring, treadmill test.

Results

The neural network simultaneously received input as the ECG images (size of the digital three-dimensional matrix was $200 \times 200 \times 1$) and structured tabular data. As the output, the neural network predicted multilevel values of main coronary arterial lesions in the probabilistic form.

As the neural network layers for images processing were used fully-connected, convolutional, batch normalization and dropout layers, while for structured data processing was used only fully-connected layers. Within the neural network was taken concatenate layer for generalizing images and dataset weight. After concatenate layer, there are two fully-connected layers. An input layer includes 13 neurons for predicting each parameter. «Adam» (adaptive learning rate optimization algorithm calculating an exponential weighted moving average of the gradient and squared gradient) was used as an optimizer, cross entropy as a loss function.

Training was performed on 100 “epochs” (one epoch = one forward pass and one backward pass of all the training examples). «Batch» size (number of training examples per one iteration)-8, validation sample size is 0.1 Parameters and structure of the neural network was selected empirically. AUC (area under the ROC curve) was chosen as a starting metric for assessing the model quality.

Prediction of the main coronary arterial lesions was carried out. On a test sample including 30 patients, the AUC score was 0.87, where the accuracy was 96 %, precision was 76 %, recall was 71 % and the f1 score — 74,1 %.

It was carried out the efficiency comparison of transient myocardial ischemia detection based on the prediction calculation of necessity to perform revascularization of the coronary arteries by the neural network analysis method and the results obtained during 24-hour Holter monitoring, treadmill test. Results of the neural network analysis method: accuracy 93 %, precision 60 %, recall 100 %, AUC score 96 %, f1 score 75 %, 24-hour Holter monitoring: accuracy 87 %, precision 33 %, recall 33 %, AUC score 63 %, f1 score 33 %, treadmill test.: accuracy 70 %, precision 12 %, recall 33 %, AUC score 54 %, f1 score 18 %.

Discussion

In modern medicine, decision-making is a complex process based on the availability of objective and reliable evidence, access to knowledge, as well as on the correct interpretation of the available data, considering the risk-benefit balance for the patient.

Prediction of the likelihood of a cardiovascular event for a relatively long period of time has become an important aspect of the patient's treatment. Annually, numerous models of assessing cardiovascular risk are designed by the researchers, the majority of which are based on the data obtained from clinical and registry-based studies.

However, only a small part of such models is used in everyday clinical practice. Due to this, qualitatively different approach is of a high importance.

In the present study, a relatively small amount of data was used: 22 parameters, ECG images from 130 patients. As a result, a neural network architecture was designed with a small number of parameters.

The result of multilevel classification of suggested lesions for each major coronary artery by the chosen neural network architecture confirmed the effectiveness of the model (AUC = 0.87). The designed neural network analysis model allow to sufficiently predict the lesion of the main coronary arteries on the basis of structured data and ECG images significantly differ by lead positioning, recording method from ECG images in test samples.

The accuracy of detecting transient myocardial ischemia, determined by neural network analysis and obtained in order to predict the need for coronary revascularization, is higher than that of traditional diagnostic methods, such as 24-hour Holter monitoring and treadmill test. The obtained results suggest a potential practical application of the neural network analysis method in clinical practice.

Conclusion

Neural network analysis of the available clinical, laboratory and instrumentation data allow to configure the network parameters for further prediction of coronary artery disease. The trained neural network predicts the main coronary arterial lesions with precise accuracy — 76 %, AUC score 0.87. The available high accuracy of the trained neural network model, observed in the test sample, based on the input structured data and recorded ECG with the parameters and lead positioning differ from initial training sample, allow to consider this method to be effective in the preliminary diagnosis of the main coronary arteries disease regardless the ECG machine model.

The detection efficiency of the transient myocardial ischemia in the abovementioned test sample is higher in the trained neural network, than in traditional diagnostic methods, such as 24-hour Holter monitoring and treadmill test.

References

1. Zaharchuk G., Gong E., Wintermark M., Rubin D., Langlotz C. P. Deep learning in neuroradiology. *AJNR Am. J. Neuroradiol.*, 2018, vol. 39, pp. 1776–1784.
2. Hyun-Jong J., Kyung-Ok Ch. Applications of deep learning for the analysis of medical data. *Archives of Pharmacal Research*, 2019, vol. 42, pp. 492–504.

3. Litjens G., Kooi T., Bejnordi B.E., Adiyoso Setio A.A., Ciompi F., Ghafoorian M., van der Laak J.A.W.M., van Ginneken B., Sánchez C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.*, 2017, vol. 42, pp. 60–88.
4. Johnson K.W., Torres Soto J., Glicksberg B.S., Shameer Kh., Miotto R., Ali M., Ashley E., Dudley J.T. Artificial intelligence in cardiology. *J. Am. Coll. Cardiol.*, 2018, vol. 71, pp. 2668–2679.

Received: May 31, 2021
Accepted: June 24, 2021

Authors' information:

Timur P. Abdualimov — MD; abdualimov@gmail.com

Andrey G. Obrezan — MD, Dr. Sci. in Medicine; obrezan1@yandex.ru