ISSN: 3049-7272

Review Article

Journal of Biomedical Sciences and Biotechnology Research

Artificial Intelligence Can Diagnose any Disease from the Data of an Electrocardiogram

Raul Isea

Fundación Instituto de Estudios IDEA, Hoyo de la Puerta, Baruta, Venezuela

***Corresponding author**
Raul Isea, Fundación Instituto de Estudios IDEA, Hoyo de la Puerta, Baruta, Venezuela.

Received: February 02, 2024; **Accepted:** February 15, 2024; **Published:** February 20, 2024

ABSTRACT
The electrocardiogram is a test that records the electrical activity of the heart, and has recently been shown that it can also detect other non-cardiovascular conditions, such as diabetes, measles, Alzheimer's, arterial hypertension, fatty liver, hiperpotasemia, hypothyroidism, malaria, etc. For this reason, this paper proposes a computational technique to analyze and detect patterns based on the position and length of the segments between the peaks that make up the electrocardiogram signals, using deep learning techniques created by Artificial Intelligence. The program started by evaluating a database of heart arrhythmia signals, which included more than 120 electrocardiograms grouped between signals from normal and arrhythmia patients, employing convolutional neural network (CNN). The program had a prediction accuracy rate of 94.3%.

Keywords: Electrocardiogram, ECG, Artificial Intelligence, CNN, Neural Networks

Introduction

An electrocardiogram (often abbreviated EKG or EKG) is a graphic representation of heart health that has been used for hundreds of years. It is a cost-effective, painless test that allows real-time visualization of results, enabling the diagnosis of ventricular hypertrophy, arrhythmias, sudden cardiac death or other cardiovascular diseases [1].

Remember that the ECG signal (Figure 1) is the result of twelve leads, distributed mainly from the chest, arms and legs, and its design dates back to 1922 thanks to Willem Einthoven (1860-1927), who defined the foundations of interpretation this although Augustus D. Waller (1856-1922) published the first human electrocardiogram [2].

At the same time, new computational methods were developed based on a paradigm created by artificial intelligence (AI) based on the analogy of the human brain. This paradigm enabled machine learning and deep learning through pattern recognition in large volumes of data [3].

In the 1950s, Arthur L Samuel (1901-1990) created an algorithm inspired by the process of neural networks that can recognize patterns in massive data and coined the term machine learning (translated as "machine learning"). This program transforms data into models without imposing assumptions on them, creating a foundation for deep learning that helps automatically discover patterns or abstractions in data to solve problems [3].

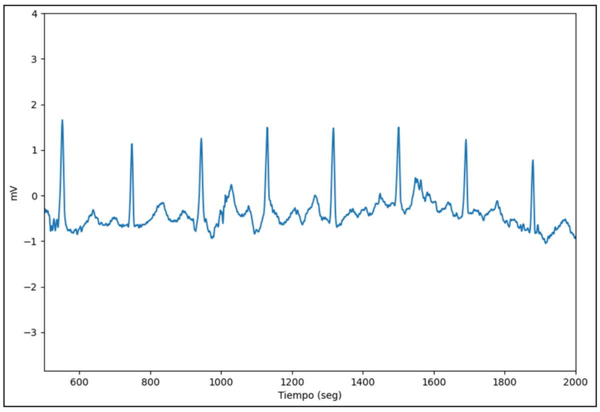

Figure 1: Signal from a 12-derivation electrocardiogram visualized with the Python programming language.

Machine learning can be supervised or unsupervised. Supervised learning is an algorithm that learns from labeled results, so it is usually said to use labeled data. In contrast, unsupervised learning uses only input data without labels and thus adapts to the data without prior knowledge, which is ideal for grouping ("clustering") data and also creating hierarchical cluster análisis [4].

Convolutional Neural Networks (CNN) are among the deep learning algorithms used for time series analysis, although

Citation: Raul Isea. Artificial Intelligence Can Diagnose any Disease from the Data of an Electrocardiogram. J Biomed Sci Biotech Res. 2024. 2(1): 1-4.
DOI: doi.org/10.61440/JBSBR.2024.v2.03

more recently they have been used for image classification and recognition [5,6].

This type of algorithm was popularized by Krizhevsky and coworkers in 2012 with the development of AlexNet for image recognition, and later optimized by Google to reduce code computing two years later [7,8].

Training, optimization, and inference are the three steps that make up code development in CNN. Because the training stage uses the supervised learning paradigm, this stage consumes the most computing time. Next, there is the optimization process to simplify the model obtained in the previous step, and avoid the need to re-train the network. Finally, evaluation with a specific problem is just the inference phase [9,10].

The Signal of an ECG

A normal heartbeat is an electrical impulse that originates in the sinus node (SA) and travels along the heart muscle (Figure 2). This heart rhythm consists of a P wave caused by atrial depolarization, a QRS complex caused by ventricular depolarization, a T wave, and finally a U wave caused by ventricular repolarization, although the U wave is not seen in many patients.

Between the waves there are segments that play an important role in the diagnosis of the patient. There is a PR segment (between the P and R waves) that varies from 0.12 seconds to 0.2 seconds; and the time interval between the two R waves, which determine the heart rate. On the other hand, QT intervals, which are usually half the length of the RR interval.

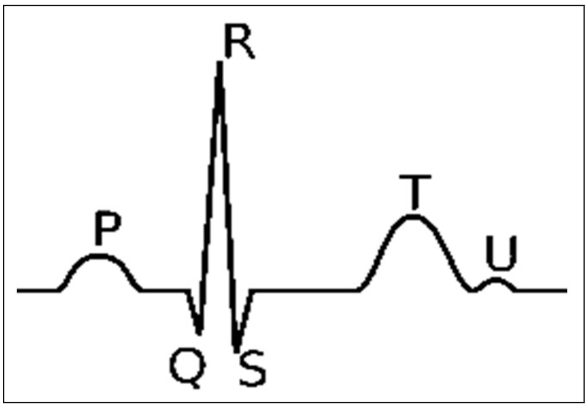

Figure 2: Heart rate, where P, QRS, T and U waves are showed.

The P wave is the first to occur in the cardiac cycle and represents atrial depolarization. The maximum amplitude is 0.25 mV and the duration is typically less than 0.10 s; The Q wave is the first negative wave of the QRS complex that occurs when ventricular depolarization begins. Q waves in peripheral leads usually do not exceed 0.03 seconds, although in lead III they usually exceed 0.04 seconds.

The QRS complex consists of a series of waves that represent ventricular depolarization and have a duration of 0.06 to 0.10 seconds. The T wave is ventricular repolarization and is shorter than the QRS complex. Since the rising part is slower than the falling part, it is usually asymmetric. Finally, the U wave appears behind the T waves and is positive.

Although many heart problems are not always detected in a standard record of just 10 seconds, an electrocardiogram (ECG) is a common tool used by cardiologists to discover a wide variety of cardiovascular abnormalities. Hence it is necessary to carry out a long-term monitoring of the electrical activity of the heart that can last even a couple of days. For this purpose, portable devices have been developed that can measure heart rate at any time of the day and send this data to the Internet based on a two or three-derivation system [11].

Piotr Augustyniak showed that it was possible to reconstruct a twelve-derivative electrocardiogram signal from signals consisting of three groups of derivatives and Atoui et al. improved this calculation using neural networks that could derive a normal 12-lead electrocardiogram from three series of electrocardiograms [12]. It is also important to consider the work of H. Zhu and colleagues, who developed a methodology that made it possible to reconstruct a 12-lead ECG from one of three leads (I, II and V2) [13]. The preferred strategy is a linear segmentation method based on adaptive area segmentation (also known as APSPL) [13].

Non-cardiac Diseases

Recently, diseases other than cardiovascular diseases are also detected in the EKG. For example, a recent study was published in which the EKG was able to diagnose diabetes in 1,262 people after analyzing more than 10,000 heartbeats per person long before the first blood samples were taken [14].

In addition, there is a risk of atrial fibrillation, a type of arrhythmia often associated with cerebral blood flow events, increasing the heart rate, which increases the risk of cardiac clots. Such a study may seem insignificant, but today it must be remembered that tinnitus is a silent and difficult to diagnose disease.

Furthermore, a study with 1024 datasets conducted between December 2020 and December 2021 has also been able to predict hyperpotasemia in patients with kidney problems using a simple electrocardiogram. Without overlooking another study that predicts the risk of developing Alzheimer's disease after more than seven years of clinical follow-up in more than 100,000 patients aged 60 or over [15].

Moreover, anorexia nervosa (AN), which is known to be an eating behavior disorder is characterized by sinus bradycardia and repolarization changes demonstrated in QT prolongation and increased dispersion, can be detected by studies based on analysis of data from an electrocardiogram [16].

For all this, this work aims to develop a computer methodology capable of detecting patterns in cardiac electrical signals, which can be used to detect and possibly diagnose cardiac and non-cardiac diseases.

This article analyzes signals from patients with cardiac arrhythmias. Remember that this is an abnormality in the heart and is divided into bradyarrhythmias (heart rate below 60 beats per minute) and tachyarrhythmias (heart rate above 100 beats per minute, which in turn are divided into supraventricular and ventricular). Bradyarrhythmia is usually asymptomatic and characterized by the absence of cardiac signals [17].

Computational Methodology

Data were obtained from the PhysioNet arrhythmia database located at physionet.org. All calculations are performed using the free and distributed programming language Python, where several deep learning libraries have been created for data analysis, such as Scikit-learn, Keras and PyTorch, as well as result viewing libraries such as MatPlotLib and numerical computing libraries such as Numpy and Scipy [18-23].

The next step was to filter the signal generated by the electrodes touching the skin and the breathing person. (mainly). For this purpose, all data analysis programs use an electronic filter called Butterworth [24]. The signals were then normalized to oscillate between -1 and 1, the default values of the neural network, and then baseline removed using the cubic spline method [25].

The database of arrhythmic ECG signals used in the work was MIT-BIH, from which more than one hundred and twenty normal patients (32 in total) and cardiac arrhythmic (88 ECG) signals are available [26]. A CNN-like network was trained from this data.

The reason for using the arrhythmia database is that several studies have been published in the scientific literature and future work will compare the results, e.g. to determine the efficiency and accuracy of the calculation algorithms presented in the scientific literatura [27-29].

Results

The CNN algorithm was based on 120 ECG signals, of which 96 were arrhythmias and 34 were patients without heart disease. For example, one ECK from a normal patient and another with an arrhythmia is shown without data processing in Figure 3. This classification allowed the data to be labeled, which helped the program distinguish between the two classes. Each electrocardiogram contains more than thirty-five thousand data units, containing approximately 70.3 megabytes of data.

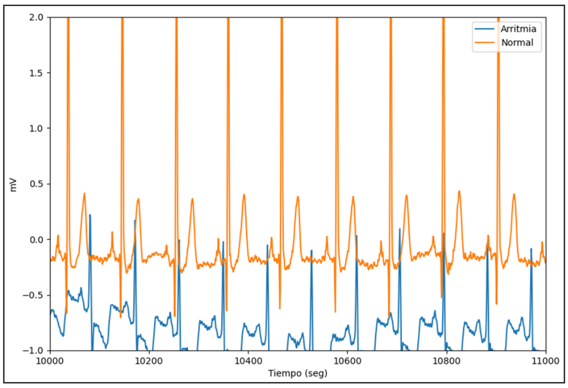

Figure 3: Graphical representation of an ECG of a normal person (orange color) of another with arrhythmia (azul). These signals are the input data of the program that evaluates the deep learning network.

Filtering the signal and then removing the baseline from all signals in the database was a crucial step. A Butterworth type filter is used for filtering and Figure 3 illustrates the generated signal.

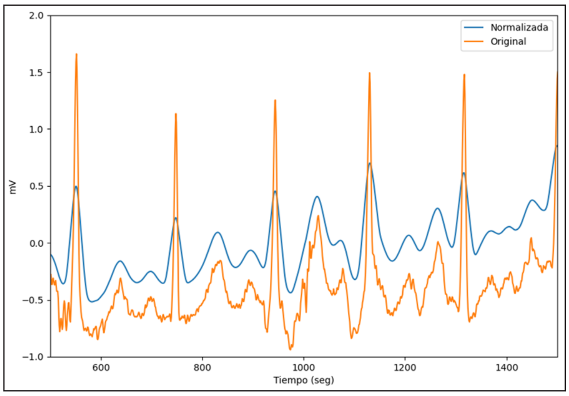

Figure 4: ECG signal from an unfiltered patient (orange), while in blue the filtered and normalized signal is between -1 and 1. This last sign is the one used in the deep learning program.

Filtering the signal and then removing the baseline from all signals in the database was a crucial step. A Butterworth type filter is used for filtering and Figure 3 illustrates the generated signal.

The training process was carried out using routines defined in PyTorch (the code will be published soon in a dedicated computer magazine). If the prediction fails after ECG evaluation, the program counts the accuracy as one, and if it fails, it counts as zero. An overall accuracy of 0.943 (94.3%) was obtained by averaging the number of hits against the training score.

Later, the percentage risk of cardiac arrhythmias will be determined and then validated by various cardiac arrhythmia prediction programs, as the results are not yet agreed upon.

Conclusions

Using artificial intelligence (AI) algorithms to predict and improve patient quality of life is one of the challenges of information technology. Recent scientific publications use the electrical activity of the heart to detect cardiovascular problems and other diseases such as diabetes, Alzheimer's disease and fatty liver long before the first symptoms appear.

This is possible by training a neural network that is specialized for each condition that affects a person. To do this, it starts with diseases caused by cardiac arrhythmias to refine and improve the training of the neural network model.

Finally, it is important to remember that the final diagnoses must be made only by doctors, and the current work can only be a guide to predict the possible risk of a particular disease.

References

1. Verdecchia P, Schillaci G, Borgioni C, Ciucci A, Gattobigio R, et al. Prognostic value of a new electrocardiographic method for diagnosis of left ventricular hypertrophy in essential hyper-tension. J Am Coll Cardiol. 1998. 31: 383-390.
2. Waller AD. A demonstration of electromotive changes accompanying the heart's beat. J Physiol (Lond). 1887. 8: 229-234.
3. Habehh H, Gohel S. Machine Learning in Healthcare. Curr Genomics. 2021. 22: 291-300.

4. Lopez C, Tucker S, Salameh T, Tucker C. An unsupervised machine learning method for discovering patient clusters based on generic signatures. J Biomed Inform. 2018. 85: 30-39.
5. Sarvamangala DR, Kulharni RV. Convolutional neural networks in medical image under-standing: a survey. Evol Intell. 2022. 15: 1-22.
6. Sun X, Yin Y, Yang Q, Huo T. Artificial intelligence in cardiovascular diseases: diagnostic and therapeutic perspectives. Eur J Med Res. 2023. 28: 242.
7. KrizhevskyIlya A, SutskeverGeoffrey I, Hinton E. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems. 2012. 25.
8. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, et al. Going deeper with convolutions. InProceedings of the IEEE conference on computer vision and pattern recognition. 2015. 1-9.
9. Akbilgic O, Butler L, Karabayir I, Chang PP, Kitzman DW, et al. ECG-AI: electrocardio-graphic artificial intelligence model for prediction of heart failure, European Heart Journal - Digi-tal Health. 2023. 2: 626-634.
10. Vaz JM y Balaji S. Convolutional neural networks (CNNs): concepts and applications in pharmacogenomics. Mol Divers. 2021. 25: 1569-1584.
11. Bansal A, Joshi R. Portable out-of-hospital electrocardiography: A review of current tech-nologies. J Arrhythm. 2018. 34:129-138.
12. Augustyniak P. On the Equivalence of the 12-Lead ECG and the VCG Representations of the Cardiac Electrical Activity; Proceedings of the 10th International Conference on System-Modelling-Control; Zakopane. 2001. 51-56.
13. Zhu H, Pan Y, Cheng KT, Huan R. A lightweight piecewise linear synthesis method for standard 12-lead ECG signals based on adaptive region segmentation. PLoS ONE. 2018. 13: e0206170.
14. Kulkarni AR, Patel AA, Pipal KV, Jaiswal SG, Jaisinghani MT, et al. Machine-learning algorithm to non-invasively detect diabetes and pre-diabetes from electrocardiogram. BMJ Inno-vations. 2022. 6: bmjinnov-2021.
15. Amini M, Pedram M, Moradi A, Ouchani M. Diagnosis of Alzheimer's Disease by Time-Dependent Power Spectrum Descriptors and Convolutional Neural Network Using EEG Signal. Comput Math Methods Med. 2021. 5511922.
16. Upeguia CV, Gómez J. Alteraciones electrocardiográficas en anorexia nervosa: revisión crítica de la literatura. Rev colomb psiquiatr. 2014. 44: Bogotá.
17. Gaztañaga L, Francis E Marchlinski, Brian P Betensky. Mecanismos de las arritmias car-diacas. Cardiología. 2012. 65: 174-185.
18. Scikit-learn Team. Scikit-learn: Machine Learning in Python, Pedregosa JMLR. 2011. 12: 2825-2830.
19. Chollet F. Keras. 2015. Available at: https://github.com/fchollet/keras.
20. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Sys-tems. 2019. 32: 8024-8035.
21. Hunter JD. Matplotlib: A 2D Graphics Environment. Computing in Science & Engineer-ing. 2007. 9: 90-95.
22. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, et al. Array program-ming with NumPy. Nature. 2020. 585, pp: 357-362.
23. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, et al. SciPy 1.0: Funda-mental Algorithms for Scientific Computing in Python. Nature Methods. 2020. 17: 261-272.
24. Ochoa B, Barboza A, Estrada F, Díaz F, et al. Sistema de Adquisición y Procesamiento de Señales Electrocardiográficas. Revista de Sistemas, Cibernética e Informática. 2010. 7.
25. Meyer CR, Keiser HN. Electrocardiogram baseline noise estmations and removal using cubic splines and state-space. Computation Techniques. Computers and Biomedical Research. 1977. 459-470.
26. Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol. 2001. 20: 45-50.
27. Wang T, Lu C, Sun Y, Yang M, Liu C Ou C. Automatic ECG classification using continu-ous wavelet and convolutional neural network. Entropy (Basel). 2021. 23: 119.
28. Wang X, Zhu T, Xia M, Liu Y, Wang Y, et al. Predicting the prognosis of patients in the coronary care unit: a novel multi-category machine learning model using XGBoost. Front Cardi-ovasc Med. 2022. 9: 764629.
29. Zheng J, Zhang J, Danioko S, Yao H, Guo H, et al. Scientific Data. 2020. 7: 48.

Copyright: © 2024 Raul Isea. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.