Research Article

Virology: Research & Reviews

Can Non-Heart Diseases like Diabetes and Sleep Apnea be Detected by an ECG?

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.*

Submitted: 2023, Nov 13; **Accepted**: 2023, Dec 13; **Published**: 2024, Jan 12

Citations: Isea, R. (2024). Can Non-Heart Diseases like Diabetes and Sleep Apnea be Detected by an ECG. *Virol Res Rev, 1*(1), 01-06.

Abstract
An electrocardiogram (ECG) is frequently used to identify heart problems, but it can also identify apnea, diabetes, and other conditions that are not heart-related. The study uses deep learning methods from artificial intelligence, particularly the Convolutional Neural Network (CNN) approach, to look for patterns in ECG data. To predict illness, two different methodologies are used. The first approach makes use of an ECG scale, while the second employs a gradient-boosting machine (GBM)-based learning technique. The MIT-BIH Arrhythmia database, the Normal Sinus Rhythm database, the BIDMC Congestive Heart Failure database, the Sleep Apnea database, the type 2 diabetes mellitus dataset, and the dataset for healthy volunteers were the sources of the data used in the study. Finally, 92% of predictions were accurate.

Keywords: Electrocardiogram, ECG, Artificial Intelligence, CNN, Apnea, Diabetes Mellitus

1. Introduction
An electrocardiogram is a graphic picture of heart health that has been in use for hundreds of years and is sometimes shortened as ECG or EKG. In order to diagnose ventricular hypertrophy, arrhythmias, heart attacks, or other cardiovascular illnesses, a painless, cost-effective test is used (Verdecchia et al., 1998).

The ECG signal is the result of twelve leads that are primarily distributed from the chest, arms, and legs. Its design dates back to 1922, thanks to Willem Einthoven (1860–1927), who defined the foundations of interpretation. Although Augustus D. Waller (1856–1922) published the first human electrocardiogram [1].

New computing techniques were produced at the same time using a paradigm of artificial intelligence (AI) constructed using the analogy of the human brain. Through pattern identification in massive amounts of data, this paradigm made machine learning and deep learning possible [2].

Convolutional Neural Networks (CNN) are one of the deep learning algorithms used for time series analysis, while they have more recently been used for graphical classification and recognition [3].

With the creation of AlexNet for image recognition in 2012 (KrizhevskyIlya, et al; 2012), this sort of algorithm gained popularity. Google then improved it to minimize code computation two years later [4].

The three stages of code creation in CNN are training, optimization, and inference. The training step requires the most computation time since it employs the supervised learning paradigm. The optimization technique is then used to streamline the model created in the previous stage and prevent the network from needing to be retrained. The inference step of assessment for a particular problem is all that remains [5].

1.1 What an ECG Signal Means
An electrical impulse that begins in the sinus node (SA) and moves across the heart muscle constitutes a typical heartbeat (Figure 1). The P wave, which is brought on by atrial depolarization, the QRS complex, which is brought on by ventricular depolarization, the T wave, and finally the U wave, which is brought on by ventricular repolarization but is not frequently observed in patients, make up this heart rhythm.

There are sections in between the signals that are crucial to the patient's diagnosis. The time gap between the two R waves and the PR segment, which occurs between the P and R waves and ranges from 0.12 to 0.2 seconds, is what defines the heart rate. In contrast, QT intervals are typically half as long as RR intervals.

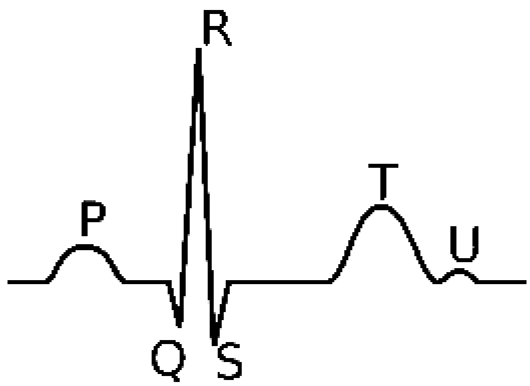

Figure 1: The P, QRS, T, and U waves of the heart are displayed.

The first wave to appear during the cardiac cycle is the P wave, which denotes atrial depolarization. The Q wave, which is the initial negative wave of the QRS complex and appears when ventricular depolarization starts, has a maximum amplitude of 0.25 mV and a normal duration of less than 0.10 s. Although they frequently reach 0.04 seconds in lead III, Q waves in peripheral leads typically do not exceed 0.03 seconds.

The ventricular depolarization waves that make up the QRS complex have a wavelength of 0.06 to 0.10 seconds and are composed of a succession of waves. The QRS complex is longer than the T wave, which represents ventricular repolarization. It is typically asymmetrical because the rising portion rises more slowly than the lowering portion. The U wave, which is positive, then emerges behind the T waves.

Recently, the ECG has also been used to diagnose disorders other than cardiovascular disease. For instance, a recent study found that 1,262 people with diabetes were able to be diagnosed long before the first blood tests were collected by using the ECG to analyze more than 10,000 heartbeats per person [6].

Furthermore, the ability to predict hyperkalemia in individuals with renal difficulties using a straightforward ECG was also demonstrated in a trial with 1024 participants carried out between December 2020 and December 2021 [7]. Without ignoring a different investigation that forecasts the likelihood of Alzheimer's disease in more than 100,000 individuals aged 60 or older after more than seven years of clinical follow-up [8].

Additionally, studies based on the analysis of electrocardiogram data have been able to identify anorexia nervosa (AN), which is known to be an eating behavior disorder, which is characterized by sinus bradycardia and repolarization changes demonstrated in QT prolongation and increased dispersion. In order to identify and diagnose cardiac and non-cardiac disorders, this effort intends to create a computer approach that can recognize patterns in cardiac electrical impulses [9].

2. Computational Methodology

All calculations were performed using the Python programming language, for which several deep learning libraries have been created for data analysis, such as Scikit-learn Keras and PyTorch as well as result visualization libraries such as MatPlotLib and numerical computing libraries such as Numpy and Scipy [10-15].

The data studied are 96 from the MIT-BIH Arrhythmia Database, 30 from the Normal Sinus Rhythm Database, 36 from the BIDMC Congestive Heart Failure Database, 28 from sleep agnea, 22 from type 2 diabetes mellitus, and 48 from healthy people of various [6,16-19]. The use of the arrhythmia database is also justified by the fact that a number of studies have been published in the scientific literature and that subsequent research will compare the findings, for example, to assess the effectiveness and accuracy of the calculation algorithms described in the literature. First of all, the data is not filtered because this could alter the results, changing the patterns of the electrocardiographic signs [20-22]. This is an important point, and as will be shown in future work, filtering information can change results when different diseases are evaluated using the same computational methodology.

There are two methods for predicting disease. The first method is founded on Gradient Boosting Machines (GBMs). It's important to keep in mind that GBM are learning algorithms that continually train new models to produce more precise estimates of response variables. Their goal is to create new base learners that have the highest correlation with the ensemble's associated loss function's negative gradient. The researcher has the option of selecting a loss function, taking into account both already-existing loss functions and employing their own task-specific loss. The second method employs scalograms of preprocessed signals generated with CWT and Morse wavelets [23]. The ECG signal can be subjected to multiresolution analysis by translating it from the time domain to the frequency domain using the CWT and Morse wavelets. The wavelet transform is an effective method for decomposing a signal from a mother wavelet into shifted or scaled forms.

3. Results

The algorithm analyzed 228 ECG signals of which 96 are arrhythmia, 48 signals obtained from healthy volunteers, 22 with type 2 diabetics mellitus, 28 correspond to sleep apnea, and 34 are patients where people have no major arrhythmias [6,17,19]. The following nomenclature was used to group distinct ECGs by illness: CHF for congestive heart failure, ALZ for healthy individuals, AGN for sleep apnea, ATR and ARR for different types of arrhythmias, and NSR for normal sinus rhythm (another name for healthy people). In the first calculation, using Boots Machine's model-based gradient method, we calculated a 35-feature vector. This configuration is used to train and interpret automatic learning models. According to the previously described methodology, the feature-vectors are graphically represented by importance in Figure 2 to visualize the results obtained (the number of feature-vectors will be determined in future studies based on the accuracy of disease prediction using ECG data). In addition, this figure shows that the decision tree obtained from training overlaps with the feature vector represented by the blue bar. This decision tree is only valid for the group of diseases analyzed in the paper. Further studies on this type of outcome will be carried out in the future.

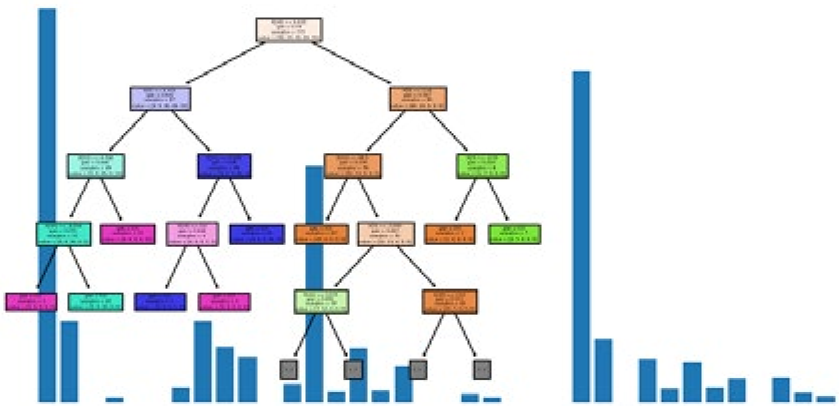

Figure 2: The feature-vector obtained from the Gradient Boots machine's model is shown in the blue bar. Overlap the tree decision (see text for more details).

Figure 3 depicts the feature values obtained with the second methodology (scalogram). As can be seen, the maximum values correspond to the before method (basically, the numbers 1, 2, 13, and 25). In reality, keep in mind that these features were determined using two very distinct computational approaches.

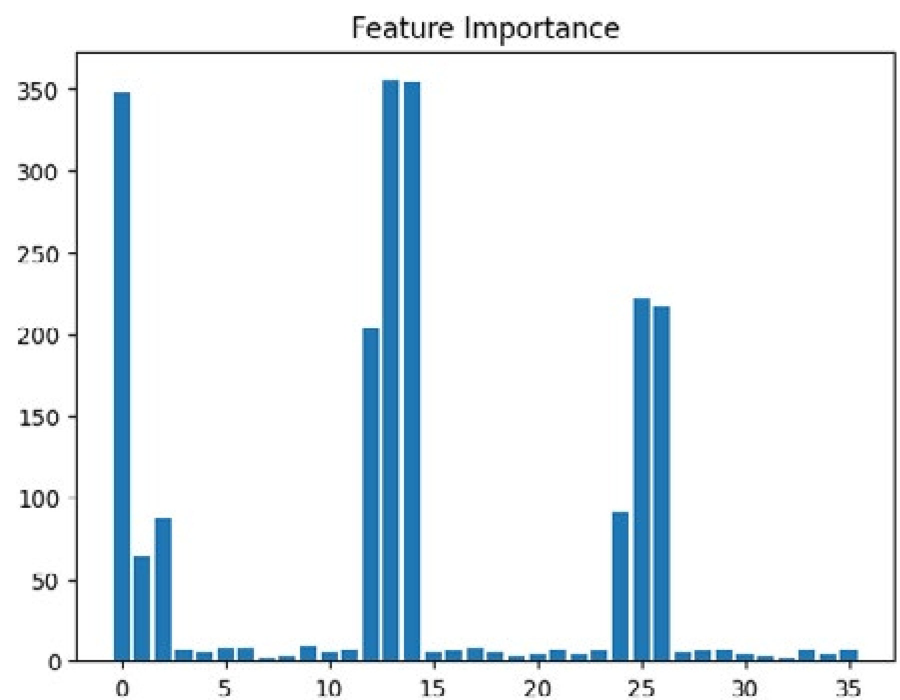

Figure 3: Plot the 35-feature Vector Generated by the Second Methodology (Scalogram).

It is interesting to remember that calculating scalograms has the advantage of allowing you to visualize patterns based on the type of sickness. A person's disease can be predicted in this way. All of this is feasible because the absolute value of a signal's continuous wave transformation (also known as CWT) is the outcome of a scalogram.

Moreover, the scalograms display the frequency content and temporal position of distinct ECG features, and the CWT may eliminate undesired frequency components from the data. One advantage of using scalograms is that they can be used to denoise the ECG signal (see figure 4).

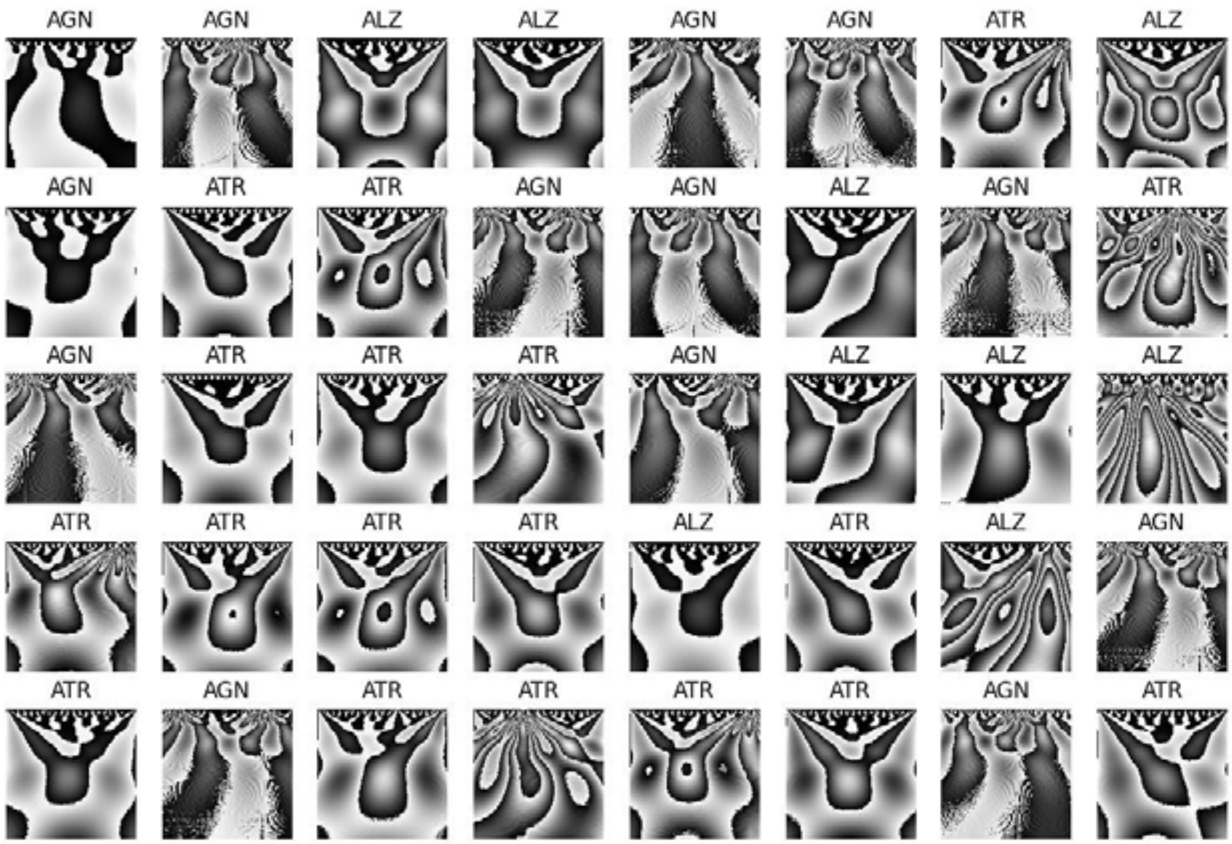

Figure 4: Some scalograms were obtained with the second methodology (nomenclature inside the text).

In addition, a multi-label confusion matrix was generated in accordance with Heydarian et al., and its diagonal represents the genuine positive rate while the remainder displays the false positive rate [24]. The accuracy, recall, and F-measures are calculated using these methods. The F-measure can be interpreted as a weighted harmonic mean of precision and recall; it reaches its best value at 1 and its worst score at 0. Precision is intuitively the ability of the classifier not to label as positive a sample that is negative (where the best value is 1 and the worst value is 0).

In the case of first method, the values obtained were

	precision	recall	f1-score
0	0.84	0.96	0.90
1	0.80	0.50	0.62
2	1.00	1.00	1.00
3	1.00	0.67	0.80
4	0.92	1.00	0.96
5	1.00	1.00	1.00
6	1.00	1.00	1.00
accuracy			0.92
macro avg	0.94	0.88	0.90
weighted avg	0.92	0.92	0.91

where the numbers 0, 1, 2, 3, 4, 5, and 6 correspond to the letters ARR, CHF, NSR, AGN, ALZ, ATR, and DIA, respectively. The lowest prediction value is for CHF and continues with ALZ, while the remainder is very close to 1.0. The percentages in the second method are pretty comparable with the first method (don´t show it). The final accuracy score was 0.92 (92.0%) for both methods of calculation.

4. Conclusions

Using artificial intelligence (AI) algorithms to anticipate and improve patient quality of life is one of information technology's problems. Recent scientific articles employ heart electrical activity to diagnose cardiovascular problems and other disorders like diabetes, and they also try to test for Alzheimer's disease before symptoms show. Finally, keep in mind that only doctors can make a final diagnosis, and the existing research can only be used as a guide to forecast the risk of a specific condition [25-27].

References

1. Waller, A. D. (1887). A demonstration on man of electromotive changes accompanying the heart's beat. The Journal of physiology, 8(5), 229.
2. Habehh, H., & Gohel, S. (2021). Machine learning in healthcare. Current Genomics, 22(4), 291.-300.
3. Sarvamangala, D. R., & Kulkarni, R. V. (2022). Convolutional neural networks in medical image understanding: a survey. Evolutionary intelligence, 15(1), 1-22.
4. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).
5. Vaz, J. M., & Balaji, S. (2021). Convolutional neural networks (CNNs): Concepts and applications in pharmacogenomics. Molecular diversity, 25(3), 1569-1584.
6. Novak, V., & Quispe, R. Cerebromicrovascular Disease in Elderly with Diabetes.
7. Xu, D., Zhou, B., Zhang, J., Li, C., Guan, C., Liu, Y., ... & Xu, Y. (2023). Prediction of hyperkalemia in ESRD patients by identification of multiple leads and multiple features on ECG. Renal Failure, 45(1), 2212800.
8. Amini, M., Pedram, M. M., Moradi, A., & Ouchani, M. (2021). Diagnosis of Alzheimer's disease by time-dependent power spectrum descriptors and convolutional neural network using EEG signal. Computational and Mathematical Methods in Medicine, 2021.
9. Upegui, C. V., & Gómez, J. (2015). Alteraciones electrocardiográficas en anorexia nervosa: revisión crítica de la literatura. Revista Colombiana de Psiquiatría, 44(1), 33-40.
10. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.
11. Chollet, F,, et. Al. (2015). Keras.
12. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems, 32.
13. Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in science & engineering, 9(03), 90-95.
14. Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585(7825), 357-362.
15. Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & Van Mulbregt, P. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature methods, 17(3), 261-272.
16. Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH arrhythmia database. IEEE engineering in medicine and biology magazine, 20(3), 45-50.
17. Penzel, T., Moody, G. B., Mark, R. G., Goldberger, A. L., & Peter, J. H. (2000, September). The apnea-ECG database. In Computers in Cardiology 2000. Vol. 27 (Cat. 00CH37163) (pp. 255-258). IEEE.
18. Baim, D. S., Colucci, W. S., Monrad, E. S., Smith, H. S., Wright, R. F., Lanoue, A., ... & Braunwald, E. (1986). Survival of patients with severe congestive heart failure treated with oral milrinone. Journal of the American College of Cardiology, 7(3), 661-670.
19. Schumann, A., & Bär, K. J. (2022). Autonomic aging–A dataset to quantify changes of cardiovascular autonomic function during healthy aging. Scientific Data, 9(1), 95.
20. Wang, T., Lu, C., Sun, Y., Yang, M., Liu, C., & Ou, C. (2021). Automatic ECG classification using continuous wavelet transform and convolutional neural network. Entropy, 23(1), 119.
21. Wang, X., Zhu, T., Xia, M., Liu, Y., Wang, Y., Wang, X., ... & Lai, D. (2022). Predicting the prognosis of patients in the coronary care unit: A novel multi-category machine learning model using XGBoost. Frontiers in Cardiovascular Medicine, 9, 764629.
22. Zheng, J., Zhang, J., Danioko, S., Yao, H., Guo, H., & Rakovski, C. (2020). A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Scientific data, 7(1), 48.
23. Byeon, Y. H., Pan, S. B., & Kwak, K. C. (2019). Intelligent deep models based on scalograms of electrocardiogram signals for biometrics. Sensors, 19(4), 935.
24. Heydarian, M., Doyle, T. E., & Samavi, R. (2022). MLCM: Multi-label confusion matrix. IEEE Access, 10, 19083-19095.
25. Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. circulation, 101(23), e215-e220.
26. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

27. Van Rossum, G., & Drake, F. L. (1995). Python reference manual (Vol. 111, pp. 1-52). Amsterdam: Centrum voor Wiskunde en Informatica.

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.*