ISSN 1870-4069

Phonocardiogram Classification Using Neural Networks for Anomaly Heart Detection

Juan Eduardo Tovar Díaz, Said Polanco Martagón, Marco Aurelio Nuño Maganda, Yahir Hernández Mier, Mario Enrique García Luna

Universidad Politécnica de Victoria, Mexico

{1830013,spolancom,mnunom,yhernandezm,2439004}@upv.edu.mx

Abstract. Heart diseases have been one of the leading health problems over time, being one of the leading causes of death among the population, affecting people of all ages and social classes. Socioeconomic status, lifestyle, and lack of awareness about symptoms have contributed to its prevalence. In Mexico, in 2022, out of approximately 650,000 deaths, 200,000 were caused by heart disease. In response to this issue, the present research demonstrates the development of a Feedforward Neural Network (FNN) model for detecting cardiac anomalies by analyzing time-domain and frequency-domain features extracted from pre-processed heart sound recordings from the PhysioNet Challenge 2016: Heart Sound Classification dataset. The model demonstrated its effectiveness by achieving a classification accuracy of 99.66% for heart sounds. Amplitude change, pitch shifting, noise addition, noise removal were applied as data augmentation techniques. In addition, SMOTE data balancing techniques were applied to increase the diversity of the dataset and improve the model's performance. Results show that the proposed model can identify complex patterns in heart sound recordings through the extracted features, with an accuracy of almost 99.66%. Although the model performed well, it is important to acknowledge the potential influence of bias on the results, even after applying data augmentation techniques and data balancing techniques.

Keywords: Feedforward Neural Network (FNN), heart sound, data augmentation.

1 Introduction

Heart diseases have been one of the significant problems that medicine has faced over time, as they have been a leading cause of death among the population, affecting people of all ages and social conditions. Factors such as socioeconomic status, lifestyle, and lack of knowledge about symptoms contribute to the proliferation of this problem. In recent years, especially in Mexico, it has been observed that one of the leading causes of mortality in the population is heart disease [4]. Out of about 650,000 deaths in 2022, heart disease was responsible

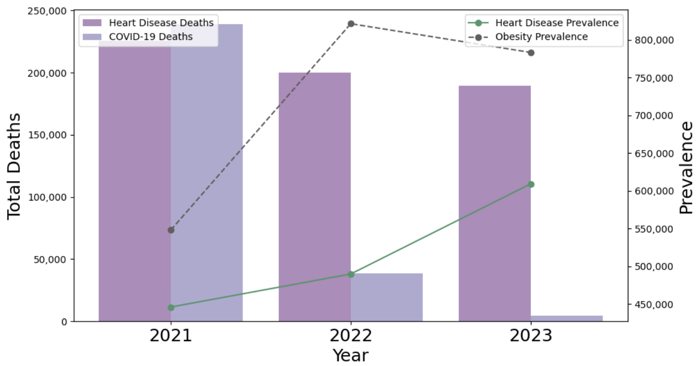


Fig. 1. Comparison of Deaths and Disease prevalence (2021-2023).

for 200,000 of them. Heart diseases significantly impact individuals across all age groups, with a higher incidence observed from the age of 45. To further emphasize the fact that heart diseases are a serious problem, in 2021, at the height of the COVID-19 pandemic, the difference in causes of death between COVID-19 and heart diseases was barely an estimated 13,000 deaths, with COVID-19 totaling nearly 239,000 deaths compared to 226,000 deaths caused by heart diseases [4]. This trend continues in the years 2022 and 2023.

In addition to this issue, the number of people suffering from heart disease has only continued to rise. In Mexico, the number of diagnosed individuals with hypertension increased from 445,993 in 2021 to 548,045 in 2022 [10]. Moreover, the incidence of obesity, a condition linked to hypertension [12], has shown a significant increase. In 2021, 489,731 people were diagnosed with obesity [10], but in 2022, this figure nearly doubled compared to the previous year, reaching 821,255 patients diagnosed with obesity [10]. Although it can be observed that in 2023, the number of patients with obesity decreased to a total of 783,207, those with high blood pressure increased to 609,070 [10]. See figure 1 for a summary of the data.

Various ways have been presented to provide a diagnosis. The most commonly used method is the analysis of electrocardiogram, which interprets the heart's electrical activity [2]. There are other more invasive methods, such as magnetic resonance imaging (MRI) and echocardiograms. However, auscultation, a technique involving using a stethoscope to listen to the sounds produced by the heartbeats [14], is a fundamental and accessible method in diagnosing heart conditions. AI has emerged as a promising tool, showcasing its versatility in various medical applications such as disease diagnosis, treatment optimization, and patient care management. Among the most notable models are neural networks, particularly feedforward neural networks (FNN), which offer an effective method for extracting and classifying information. This research uses feedforward neural networks (FNN) to analyze and classify heart sounds

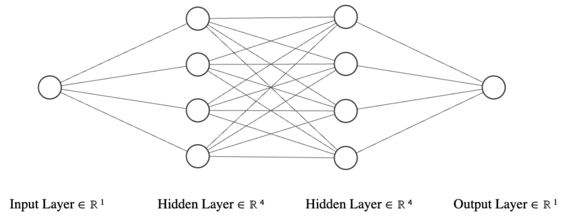


Fig. 2. Structure of a Feedforward Neural Network.

based on extracted features from the PhysioNet Challenge 2016: Heart Sound Classification database. Aiming to enhance the early detection of cardiac problems, the study provides an efficient and accurate tool for an early diagnosis.

2 Theoretical Framework

2.1 Neural Network

A neural network [13] is a network of interconnected nodes designed to generate an output. Biological studies of the nervous system heavily inspire these elements. In short, neural networks are a model of artificial intelligence designed to mimic the functioning of the human brain, assisting machines in learning patterns and making decisions. The elements are represented as "neurons"(N), which, when grouped, form layers to which a numerical value is assigned.

2.2 Feedforward Neural Network

They are artificial neural networks (ANN) in which the connections between neurons do not form a cycle. Feedforward neural networks were the first type of artificial neural network invented. They are called feedforward because information travels only forward through the network (without loops), first through the Input Nodes/Input Layer(IL), then through the Hidden Nodes/Hidden Layers(HL), and finally through the Output Nodes/Output Layer(OL) [16]. Figure 2 shows the structure of a Feedforward Neural Network (FNN).

2.3 Dataset

In this investigation, the PhysioNet/CinC Classifying Heart Sounds Challenge 2016 training dataset was implemented and has been extensively employed in numerous related studies. This dataset tries to represent real-life cases containing phonocardiograms in which the heartbeat is clear and recordings where noise is present. The recordings have a duration of 5 to 120 seconds and were obtained from both healthy and sick patients, including children and adults. This dataset resulted from a combination of nine different databases, and the equipment used to obtain the samples varied across the different databases, as did the environment for each one [8].

2.4 Data Augmentation

It is a set of techniques that generate new information based on limited information. It provides a large amount of data to machine learning models for their training. It also reduces the likelihood that the model will suffer from overfitting while training with the obtained data. Additionally, it helps improve the accuracy and performance of the created models [9].

2.5 Data Balancing

Data balancing encompasses strategies aimed at alleviating class imbalance between different categories. The most commonly employed SMOTE approach is among the many ways to alleviate the issue. This method generates synthetic data from the existing data of the minority class [1].

2.6 Time Domain Features

The simplest way to analyze an audio signal is through time, as it is a time series. Every audio signal evolves, and by visualizing them, we can observe certain key features that help predict and analyze similar signals [17].

Figure 3 shows an example of the main features extracted from the audio signal a0007.wav.

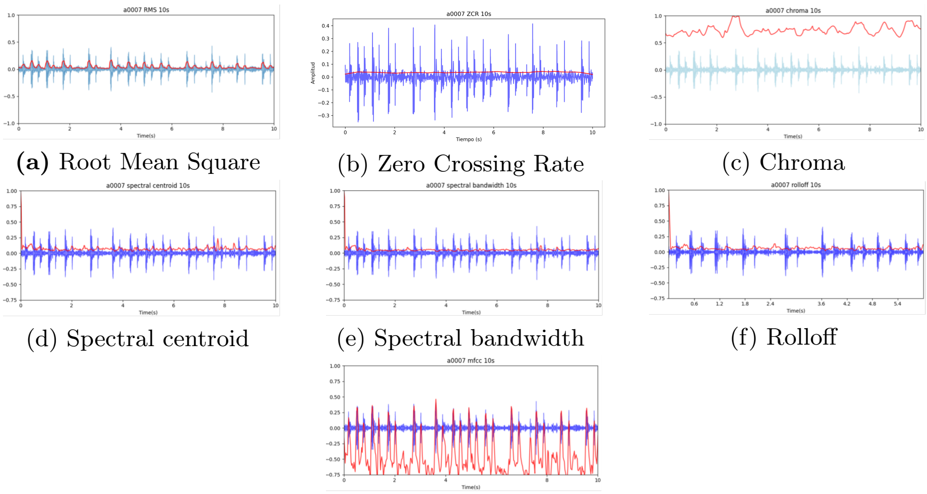
Among these key features, the following will be utilized in this research.

- **RMS (Root Mean Square).** The volume or intensity of a sound is one of the most important characteristics of the human auditory system. Mathematically, volume is defined as the signal's magnitude's root mean square (RMS) value. This feature is used for speech discrimination and music, speech segmentation, and classification of acoustic scenes [9]. See figure 3.a.
- **ZCR (Zero Crossing Rate).** The ZCR can be defined as the zero-crossing rate of an audio signal over a specific time. Mathematically, it is the number of times a signal changes from positive to negative and vice versa, the number of times a signal crosses zero [17]. See figure 3.b.

2.7 Frequency Domain Features

Temporal domain graphs show the variations of the signal over time. The characteristics of the signal are extracted using the Fourier Transform [9] to analyze the signals in frequency.

- **Chroma.** Chromatic characteristics are an audio representation where the entire frequency spectrum of the audio signal is grouped into 12 bins representing the musical octave's 12 semitones (Chroma). This mapping can be obtained through the audio signal's short-time Fourier transform (STFT). These features are particularly useful for identifying the similarity between different interpretations of a musical piece or audio signal [17]. See figure 3.c.



(a) Root Mean Square (b) Zero Crossing Rate (c) Chroma

(d) Spectral centroid (e) Spectral bandwidth (f) Rolloff

(g) MFCC

Fig. 3. Features extracted from the audio a0007.wav. In each subplot, the blue waveform corresponds to the original audio signal, while the red curve represents the extracted feature over time: (a) Root Mean Square (RMS), (b) Zero Crossing Rate (ZCR), (c) Chroma, (d) Spectral Centroid, (e) Spectral Bandwidth, (f) Rolloff, and (g) Mel Frequency Cepstral Coefficients (MFCCs).

- **Spectral centroid.** Indicates where the centroid of the spectrum is located. It describes the "brightness" of an audio signal, calculated by considering the spectrum as a distribution where the values are the frequencies and the probabilities of observing them are the normalized amplitudes. It is used to measure music's timbre and for musical classification [17]. See figure 3.d.
- **Spectral bandwidth.** It is a second-order statistical value that determines low-bandwidth sounds compared to high-frequency sounds [17]. See figure 3.e.
- **Rolloff.** It is the frequency below a certain percentage of the total frequencies of the audio signal, for example, 95% [17]. See figure 3.f.
- **Mel Frequency Cepstral Coefficients (MFCC).** They are the central representation of an audio signal. The MFCCs represent an audio clip's short-term power spectrum based on the power spectrum's discrete cosine transform on a non-linear mel scale. In the MFCCs, frequency bands are equally spaced on the mel scale, closely mimicking the human auditory system [15]. See figure 3.g.

3 State of the Art

The literature review on heart sound classification using feedforward neural networks (FNN) reveals various methodologies proposed for the same purpose.

Krishnan, Balasubramanian, and Umapathy [6] propose the creation of four neural network models: three one-dimensional convolutional neural networks (1D-CNN) and one feedforward neural network (FNN) aimed at classifying unsegmented sounds from a portion of the PhysioNet Challenge 2016: Heart Sound Classification dataset. These sounds are subsequently divided into 6-second sections and used for training the models, with the results considered for the feedforward neural network (FNN), which achieved an accuracy of 82.52%.

On the other hand, Chowdhury, Poudel, and Hu [3] propose a more sophisticated approach than the previous one, using a feedforward neural network to classify sounds from the same dataset. However, they added a smaller dataset composed of 18 audio files. These audio files are segmented using the zero-crossing rate and the Shannon energy envelope, allowing the detection of cardiac cycle beats and segmenting them by each cycle.

After that, the features constituting the Mel spectrogram and Mel-frequency cepstral coefficients are extracted from each audio file and used for training a feedforward neural network (FNN), which achieved an accuracy of 97.10%. Khan, Abid, and Khan [5] use the PhysioNet Challenge 2016: Heart Sound Classification dataset in two different ways: first, with unsegmented audio, and second, with segmented audio. They use the zero-crossing rate for segmentation to determine when a cardiac cycle starts and ends. Subsequently, they use Hidden Markov Models (HMM) and a modified HMM to improve the accuracy when determining the duration of a cycle.

With both unsegmented and segmented information, Mel-frequency cepstral coefficients are extracted, and training is conducted using different types of classifiers, including a Feedforward Neural Network (FNN), which achieved a precision of 79.3% with segmented audio and 80.9% with unsegmented audio. In some studies, although modifications to the neural network's architecture are introduced, the purpose and the data used remain constant.

Li et al. [7] used a Convolutional Neural Network to analyze and classify information from the segmented audio of the PhysioNet Challenge 2016: Heart Sound Classification dataset. Features from the Mel-frequency cepstral coefficients were extracted from the segmented audio, which was then used to train the model, achieving an accuracy of 96.48%. Similarly, Norman, Ting, Salleh, and Ombao [11] propose three models of convolutional neural networks: a one-dimensional Convolutional Neural Network (1D-CNN), which receives the raw audio directly and learns its features; a two-dimensional Convolutional Neural Network (2D-CNN), which learns from features extracted through Mel-frequency cepstral coefficients; and a combination of both models (TF-ECNN).

Focusing on the 1DCNN model, it achieved results already present in the current theoretical framework, reaching 87.23% in accuracy, 87.57% in sensitivity, 85.84% in specificity, and 86.7% in modified accuracy. In Table 1, a summary of the reviewed studies from the current state of the art can be observed.

Table 1. Comparison of the studies reviewed in the state of the art.

Art	Model used	Dataset	Features	Segmentation	Accuracy
[6]	FNN	Part of the PhysioNet Challenge 2016: Heart Sound Classification	Raw audio signal	Sections of 6 seconds	82.52%
[3]	FNN	PhysioNet Challenge 2016 + additional dataset of 18 audio recordings	MFCC	Segmentation by heartbeats cycles using ZCR and Shannon Envelope	97.10%
[5]	FNN	PhysioNet Challenge2016	MFCC	Segmentation by heartbeats cycles using ZCR and HMMS.	79.3% (segmented) 80.9% (Unsegmented)
[7]	CNN	PhysioNet Challenge 2016 + additional dataset of 45 audio recordings	MFCC	Segmentation by 5-second cycles.	96.48%
[11]	1D-CNN	PhysioNet Challenge2016	Raw audio signal	Unsegmented	87.23%

Compared to the studies present in the state of the art, this work differs by the wide range of features extracted (ZCR, RMS, Chroma, Spectral Centroid, Spectral Bandwidth, Rolloff, and MFCC) from each of the audio files in the dataset, making the obtained information more significant and representative for each audio. Furthermore, using a Feedforward Neural Network (FNN) demonstrates that it is possible to classify complex and delicate information without needing a more complex and heavy architecture such as a Convolutional Neural Network. This implements a tool based on this type of model (FNN) that is feasible for use in real-life cases, as a lightweight model allows it to be implemented on various devices.

4 Methodology

4.1 Dataset Modification

The audio files from the dataset are divided into six directories, labeled from a to f (6 folders), these folders were simplified in 2 categorized into "normal" and "abnormal" with a total of 665 abnormal audios and 2575 normal audios respectively. As can be observed, audio files exhibit a significant imbalance, considering the total number of audio files per class, with the "normal" labeled audio files being more dominant.

4.2 Data Augmentation

Data augmentation was performed on both classes for better data balancing. The following techniques were used to generate more audio from existing audio.

- **Amplitude change.** This amplitude change is a multiplication of the audio signal by a random value between the intervals of 0.1 and 10, with a step of 0.1.
- **Pitch shifting.** The pitch change is performed by randomly selecting from a maximum number of semitone steps of 2. The pitch can be altered in both negative and positive directions. For example, if the maximum number of steps is 2, the selected value can range between -2 and 2.
- **Noise Addition.** Gaussian noise is added to the audio at the same length as the original signal. This is done by calculating the audio's energy and normalizing the noise to the desired noise level relative to the audio's energy. The desired noise level is a random value ranging from 0 to 0.5, with a step of 0.001. This approach ensures that the noise level does not exceed appropriate limits, preventing the creation of unhelpful audio samples.
- **Noise Removal.** Bandpass filter with a low cutoff frequency of 25 Hz and a high cutoff frequency of 400 Hz, applied with a second-order filter design.

Each audio file underwent an amplitude change and a combination of amplitude changes with one of the remaining modifications. The combinations were as follows:

- Amplitude change (AC)
- Amplitude change and pitch shifting (ACPS)
- Amplitude change and noise addition (ACNA)
- Amplitude change and noise removal (ACNR)

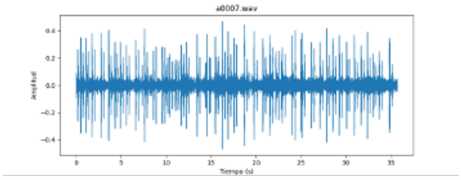
The audio classified as "normal" underwent these data augmentation techniques four times, while the "abnormal" audio was applied eight times. Figure 4.a shows the original audio signal, while figure 4.b shows the signal with the amplitude change. Figure 4.c displays the signal with both amplitude and pitch-shifting techniques. Figure 4.d shows the amplitude change with noise addition, and Figure 4.e illustrates the amplitude change with noise removal.

The newly generated audio files initiated feature extraction from the time and frequency domains. The extracted features were as follows, RMS, ZCR, Chroma, centroid, bandwidth, roll-off, MFCC

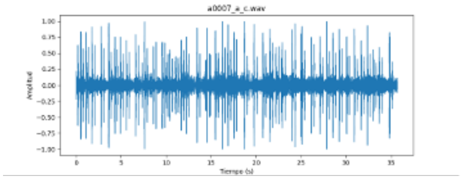
This process was carried out using the Librosa library. The mean and variance were calculated for each extracted feature, and the data was stored and categorized in a CSV file. This CSV file will be used to train the feedforward neural network.

4.3 Data Balancing

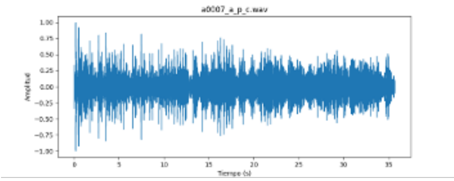
SMOTE data balancing technique is applied with a random seed of 42 to balance the amount of data labeled as "abnormal" with those labeled as "normal." A



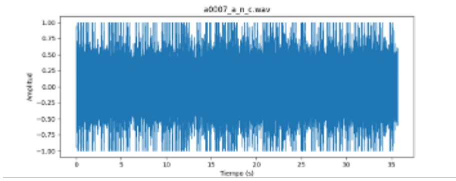
(a) Audio original



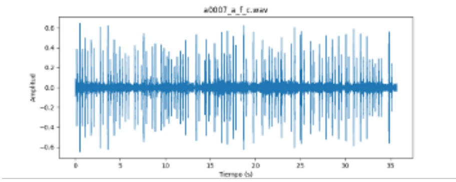
(b) Audio with AC.



(c) Audio with ACPS.



(d) Audio with ACNA.



(e) Audio with ACNR.

Fig. 4. Data augmentation of audio a0007.wav.

Table 2. PhysioNet Challenge2016: Heart sound classification Dataset

	Class					
	Original dataset		After data augmenting		After data balancing	
	Normal	Abnormal	Normal	Abnormal	Normal	Abnormal
Total	2575	665	12,441	5,985	12,441	12,441

new CSV file is then created where the data balancing technique has already been applied, and this file is used to train the feedforward neural network.

Table 2 shows the number of samples of each class before and after data augmentation and the bedore and after data balancing.

The test data is located within a specific folder of the PhysioNet Challenge 2016: Heart sound classification Dataset called "validation", which contains 150 normal audios and 151 abnormal audios. The same features were extracted from each of the audios as in the training data for the subsequent creation of a test CSV file.

5 Experimental Results

Four tests were conducted, training different structures of a feedforward neural network, where each test was repeated 10 times through 250 epochs with a learning rate of 0.0001.

Table 3. Neural Network Architectures and optimizers.

Test	Network Structure	Optimizer	Avg training time (min)
1	Sequential (Hidden Layer: 64 N, ReLU) → (Hidden Layer: 32 Neurons, ReLU) → (Output Layer: 2 Neurons, Softmax)	Adam	5 min.
2	Sequential (Hidden Layer 128 N, ReLU) → (Hidden Layer: 64 Neurons, ReLU) → (Output Layer: 2 Neurons, Softmax)	Adam	5 min. 30 sec.
3	Sequential (Hidden Layer: 256 N, ReLU) → (Hidden Layer: 128 Neurons, ReLU) → (Output Layer: 2 Neurons, Softmax)	RMSprop	7 min.
4	Sequential (Hidden Layer: 256 Neurons, ReLU) → (BN) → (D: 0.5) → (Hidden Layer: 128 Neurons, ReLU) → (BN) → (Dropout: 0.5) → (Hidden Layer: 64 N, ReLU) → (Output Layer: 2 Neurons, Softmax)	SGD	8 min.

Table 4. Metrics of the models trained

Test	Accuracy (Acc)	Normal Acc	Abnormal Acc	Min Acc	Max Acc	Mean Acc	F1-score	Spec	Sens
1	.93	.93	.93	.897	.93	.919	.93	.93	.927
2	.98	.99	.97	.93	.98	.969	.983	.97	.99
3	**.9966**	**1**	**.9933**	**.93**	**1**	**.989**	**.996**	**.993**	**1**
4	.906	.966	.846	.88	.906	.894	.915	.84	.96

The following equipment specifications used in the training of the models:

- **CPU:** Intel i3-12100
- **RAM:** 16 GB 3200 MHZ DDR4
- **GPU:** Nvidia GeForce RTX 3060 12GB VRAM

Table 3 shows the structure and optimizers of the models trained. Table 4 details the results obtained from each configuration. As shown, the neural network in test 3 achieves an accuracy of 0.9966 in the test subset, with a mean of 0.989. This contrasts with the current state of the art, showcasing a broader range of extracted features from each audio file, enhancing the depth and representativeness of the information obtained. Figure 6 shows the confusion matrices of the best result of each configuration.

Figure 5 shows the accuracy distribution of all the models trained for each test performed in this study.

6 Conclusions and Future Work

In this work, a feedforward neural network is used to classify the unsegmented heart sounds from the PhysioNet challenge, i.e., each heart cycle of each

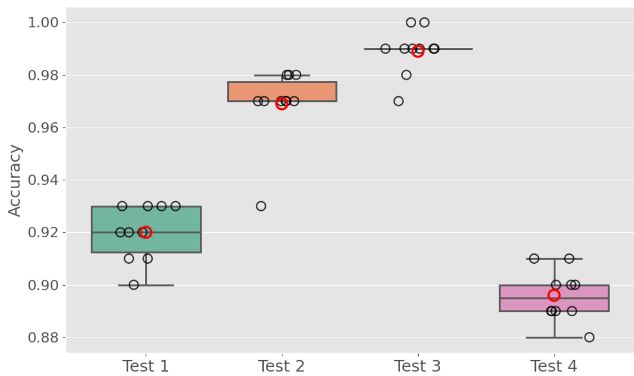
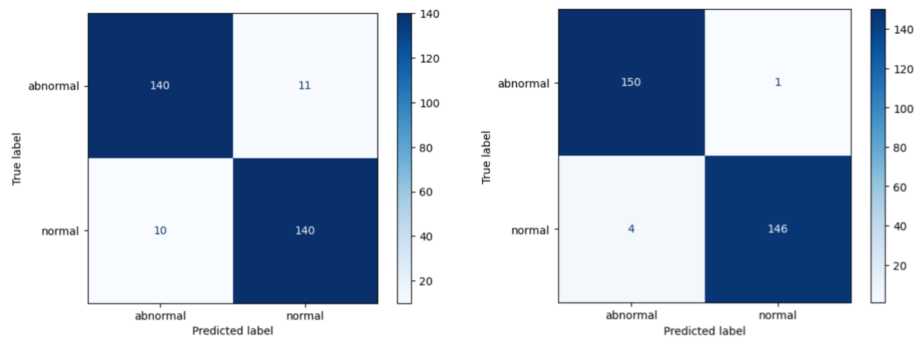
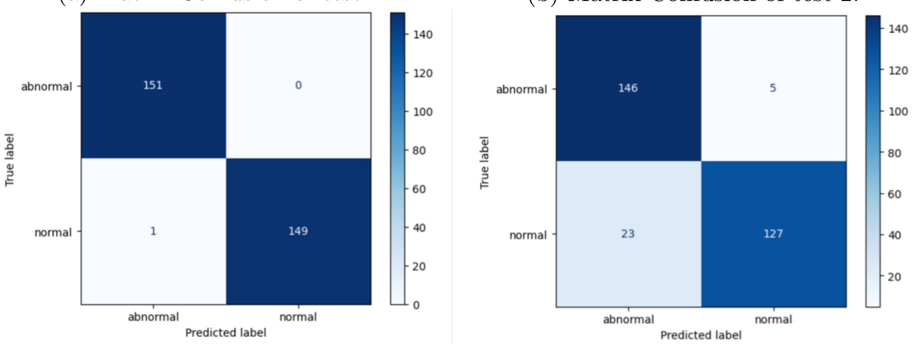


Fig. 5. Distribution of Accuracy Obtained by the Four Neural Network Models Across the 10 Performed Tests.



(a) Matrix Confusion of test 1.

(b) Matrix Confusion of test 2.



(c) Matrix Confusion of test 3.

(d) Matrix Confusion of test 4.

Fig. 6. Confusion matrices of the tests performed.

record has not been divided to train the model. Then, according to experiments, feedforward neural networks have demonstrated efficacy in identifying abnormalities within these sounds.

Furthermore, it is shown that feedforward neural networks are nearly on par with more complex and heavier models, such as convolutional neural networks (CNN), making their implementation on smaller and more accessible devices feasible for real-world applications.

The results obtained in the study demonstrate the potential of feedforward neural networks (FNN) for heart sound classification, as well as their ability to classify complex patterns, specially when using data augmentation techniques and extracting a wide range of features extracted from each phonocardiogram in the dataset.

Data augmentation techniques allowed us to enrich samples from the dataset, while data balancing techniques allowed us to balance the minority class results in higher accuracy for the proposed model, exhibiting the capabilities inherent in feedforward neural networks.

This result could help healthcare practitioners to properly access and use non-complex, lightweight models, for efficient and extensive application in practical circumstances such as electronic stethoscope applications, medical mobile apps, or Internet of Things medical applications.

Currently, the proposed technique has some limitations when used in a real environment, since the model could yield false positives because in this case, the input audio is not segmented into single duration chunks.

Acknowledgments. This study was funded by the Secretaria de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) through a master's degree scholarship.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ahsan, M. M., Siddique, Z.: Machine learning-based heart disease diagnosis: A systematic literature review (2021), https://arxiv.org/abs/2112.06459
2. Brady, W. J., Lipinski, M. J., Darby, A. E., Bond, M. C., Charlton, N. P., Hudson, K. B., Williamson, K.: Electrocardiogram in Clinical Medicine. Wiley (2020), https://books.google.es/books?id=MCP5DwAAQBAJ
3. Chowdhury, T. H., Poudel, K. N., Hu, Y.: Time-frequency analysis, denoising, compression, segmentation, and classification of pcg signals. IEEE Access, vol. 8, pp. 160882–160890 (2020) doi: 10.1109/ACCESS.2020.3020806
4. INEGI: Estadísticas de defunciones registradas (edr) (2023), https://www.inegi.org.mx/programas/edr/, Último acceso: 30 de Junio del 2024
5. Khan, F. A., Abid, A., Khan, M. S.: Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. Physiological Measurement, vol. 41, pp. 055006 (2020) doi: 10.1088/1361-6579/ab8770
6. Krishnan, P. T., Balasubramanian, P., Umapathy, S.: Automated heart sound classification system from unsegmented phonocardiogram (pcg) using deep neural

network. Physical and Engineering Sciences in Medicine, vol. 43, pp. 505–515 (2020) doi: 10.1007/s13246-020-00851-w

7. Li, F., Liu, M., Zhao, Y., Kong, L., Dong, L., Liu, X., Hui, M.: Feature extraction and classification of heart sound using 1d convolutional neural networks. EURASIP Journal on Advances in Signal Processing, vol. 2019, pp. 59 (2019) doi: 10.1186/s13634-019-0651-3
8. Liu, C., Springer, D., Li, Q., Moody, B., Juan, R. A., Chorro, F. J., Castells, F., Roig, J. M., Silva, I., Johnson, A. E. W., Syed, Z., Schmidt, S. E., Papadaniil, C. D., Hadjileontiadis, L., Naseri, H., Moukadem, A., Dieterlen, A., Brandt, C., Tang, H., Samieinasab, M., Samieinasab, M. R., Sameni, R., Mark, R. G., Clifford, G. D.: An open access database for the evaluation of heart sound algorithms. Physiological Measurement, vol. 37, pp. 2181 (2016) doi: 10.1088/0967-3334/37/12/2181
9. Mushtaq, Z., Su, S.-F., Tran, Q.-V.: Spectral images based environmental sound classification using cnn with meaningful data augmentation. Applied Acoustics, vol. 172, pp. 107581 (2021) doi: https://doi.org/10.1016/j.apacoust.2020.107581
10. de México, G.: 20 principales causas de enfermedad nacional (2023), https://epidemiologia.salud.gob.mx/anuario/html/principales_nacional.html, Último acceso: 1 de Agosto de 2024
11. Noman, F., Ting, C. M., Salleh, S. H., Ombao, H.: Short-segment heart sound classification using an ensemble of deep convolutional neural networks. In: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1318–1322 (2019) doi: 10.1109/ICASSP.2019.8682668
12. Organization, W. H.: Hipertensión (2023), https://www.who.int/es/news-room/fact-sheets/detail/hypertension, accedido: 2 de Agosto de 2024
13. Picton, P.: What is a Neural Network?, pp. 1–12. Macmillan Education UK (1994), https://doi.org/10.1007/978-1-349-13530-1_1
14. Reyna, M. A., Kiarashi, Y., Elola, A., Oliveira, J., Renna, F., Gu, A., Alday, E. A. P., Sadr, N., Sharma, A., Kpodonu, J., Mattos, S., Coimbra, M. T., Sameni, R., Rad, A. B., Clifford, G. D.: Heart murmur detection from phonocardiogram recordings: The george b. moody physionet challenge 2022. PLOS Digital Health, vol. 2, pp. e0000324– (9 2023)
15. Sandhya, P., Spoorthy, V., Koolagudi, S. G., Sobhana, N. V.: Spectral features for emotional speaker recognition. In: 2020 third international conference on advances in electronics, computers and communications (ICAECC). pp. 1–6. IEEE (2020)
16. Sharkawy, A.-N.: Principle of neural network and its main types. Journal of Advances in Applied Computational Mathematics, vol. 7, pp. 8–19 (2020)
17. Sharma, G., Umapathy, K., Krishnan, S.: Trends in audio signal feature extraction methods. Applied Acoustics, vol. 158, pp. 107020 (2020)