Heartbeat Sound Classification Based on Neural Network

Dr. Faris Ali Jasim Shaban[1]
Computer Technology Engineering Department, Imam Ja'afar Al_Sadiq University, Baghdad, Iraq
e-mail: faris2005@gmail.com
farsi@sadiq.edu.iq

Abstract:

The classification of heart sounds is essential for the early detection of cardiovascular disorders. Although heart sound classification has made tremendous strides in recent years, the majority of them still rely on conventional segmentation features and shallow structure-based classifiers. As a result, we suggest a novel method for classifying heart sounds that is based on deep residual learning and enhanced Mel-frequency cestrum coefficient characteristics. The improved features of the heart sound signal are first computed after preprocessing. The neural network uses these features as input features after that. The deep residual network extracts even more pathogenic information from the heart sound stream. The heart sound signal is finally categorized into various groups based on the features that the neural network has learned. In-depth studies of various network parameters and connectivity methods are presented in this research. For the dataset used in this study, the proposed approach obtains an accuracy of 94.43%.

Keywords : CNN , RNN ,

1. Introduction:

The heart makes sounds by rhythmically contracting and distilling. The heart, which supplies other organs with oxygen and other nutrients as well as removes waste products from metabolism to support cell maintenance a proper physiological condition, is the body's powerhouse and its most important organ. The four chambers of the heart are the left atrium, left ventricle, right atrium, and right ventricle , The first, second, third, and fourth heart sounds are produced during the cardiac cycle, which happens before the following pulse. Auscultation of the heart sound is a straightforward, important, and reliable technique that has been used for more than 180 years to identify cardiovascular disorders the initial heartbeat, which signals the onset of ventricular systole, is distinguished by its length, intensity, and volume. The second heart sound, which has a shorter length, lower strength, and a lower pitch, signals the start of ventricular diastole. The third heart sound follows the second heart sound. They have a larger wavelength and last for between 0.04 and 0.05 seconds. The majority of kids and around half of adults report hearing it, and it is not always a sign of abnormality [1] . The long wave sound, which lasts for approximately 0.04 seconds before the first heart sound in the fourth heart sound, is heard. The atrial sound, which is a mechanical wave brought on by the atria contracting and the ventricles rapidly filling with blood flow. On the ECG, the majority of healthy persons can detect a little fourth heart sound that is difficult to hear during routine auscultation. The doctor records the four fundamental heart sounds and examines how they differ from the normal heart sound based on the clinical status of the patient , Acoustic devices are used to capture electrocardiogram (ecg (PCG) in order to diagnose and treat individuals., and complex machinery has become a standard component of medical equipment thanks to manufacturing. With the widespread usage of PCG, it has become increasingly common to apply Medical and physiological data may be extracted from PCG data using encoding and machine intelligence techniques.

Since it was first suggested in 2006, CNN has grown to be a mature deep learning framework thanks to the recent advancements in the field of deep learning Cuz to its convolutional, that discovers regional image trends, it has grown to be a popular approach in computer vision. By using the appropriate audio processing techniques, example is translating voice spectral graphs into real biomedical signals, CNN is also gradually being applied to the classification of biomedical signals and the semantic recognition of speech. A class of neural networks known as recurrent neural networks (RNN) is particularly adept at handling sequential data Since it was first suggested in 2006, CNN has benefited from the recent advancements in a well-known computational intelligence framework for pattern recognition, its convolution layers, which discovers local patterns in images, has made it a popular approach in computer vision [2][4].

Related voice recognition methods, such converting physical inputs from people in speaking spectra patternsRnns are also increasingly being used for physiological signal classification & semantic voice commands (RNNs) are a family of neural networks with a focus on handling sequential data. Long term memory (LSTM) and gated recursive units (GRU) are enhanced variants of RNN that offer superior performance in a variety of applications, such as picture annotation, speech recognition, and machine translation. The classification since heart vibrational frequencies are serial data with interpreted correlation, they can be handled effectively. Figure 1[5] depicts the wave representations of either the S(1 to 4)tones throughout the two stages of systole..

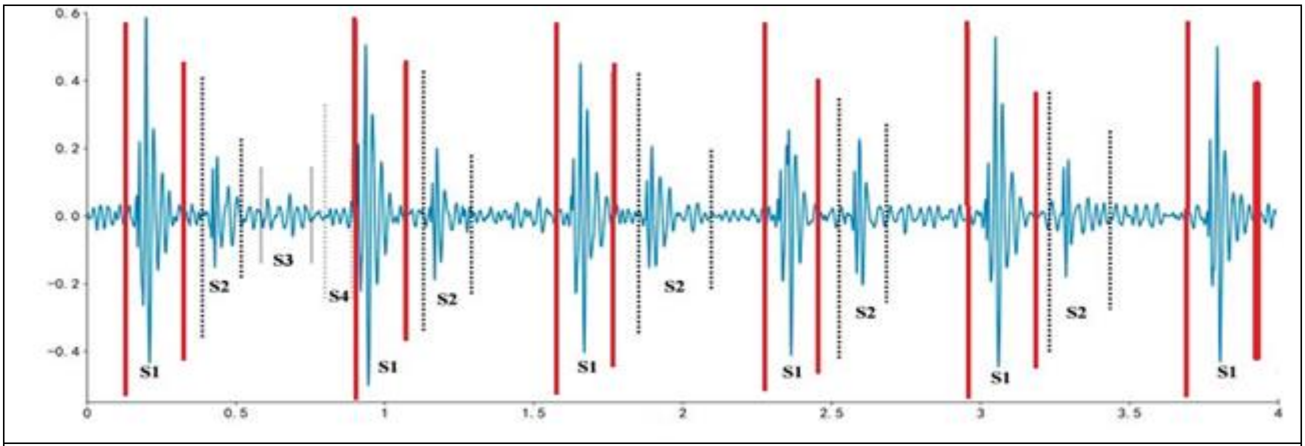

Fig.1 S1 to s4 Oscillations are used to visualize the systole and ventricular systole periods of sound.

1.2 The aim of the research:

Moreover, as some background noise is unavoidably captured as cardiac tones are being acquired, this significantly lower the model's classification accuracy [6][7]. Consequently, it is essential to feature until supplying the neural net with the source cardiac beat input for learning. Many feature extraction techniques, such as discrete wavelet transform and mill-frequency spectral coefficients, are frequently utilized in heart sound categorization applications (MFCC). In this study, the neural network's input tensor is composed of depending on the first- and 2nd differential correlations MFCC. This feature extraction the human brain is capable of extracting physiological and medical data from the ecg signals data using this approach., increasing classification accuracy while reducing the impact of noise on the outcomes Deep learning techniques avoid the issues of manual [5][8].

Interference, when it comes to complicated processes and weak generalization compared to conventional heart sound categorization algorithms, Heart sound classification was integrated by MFSC and CNN [9].

The model's performance is constrained by the dearth of substantial, trustworthy open-heart sound data sets. The report offers three of the most popular heart tones datasets sample to address this issue. It aids in significantly enhancing the learning model's performance. To better represent the properties of the cardiovascular issues in our study, we used enhanced MFCC as input features. To further reduce gradient latency and deterioration during training, we employ a residual network. The driving force behind our work is outlined in Figure 2 [10].

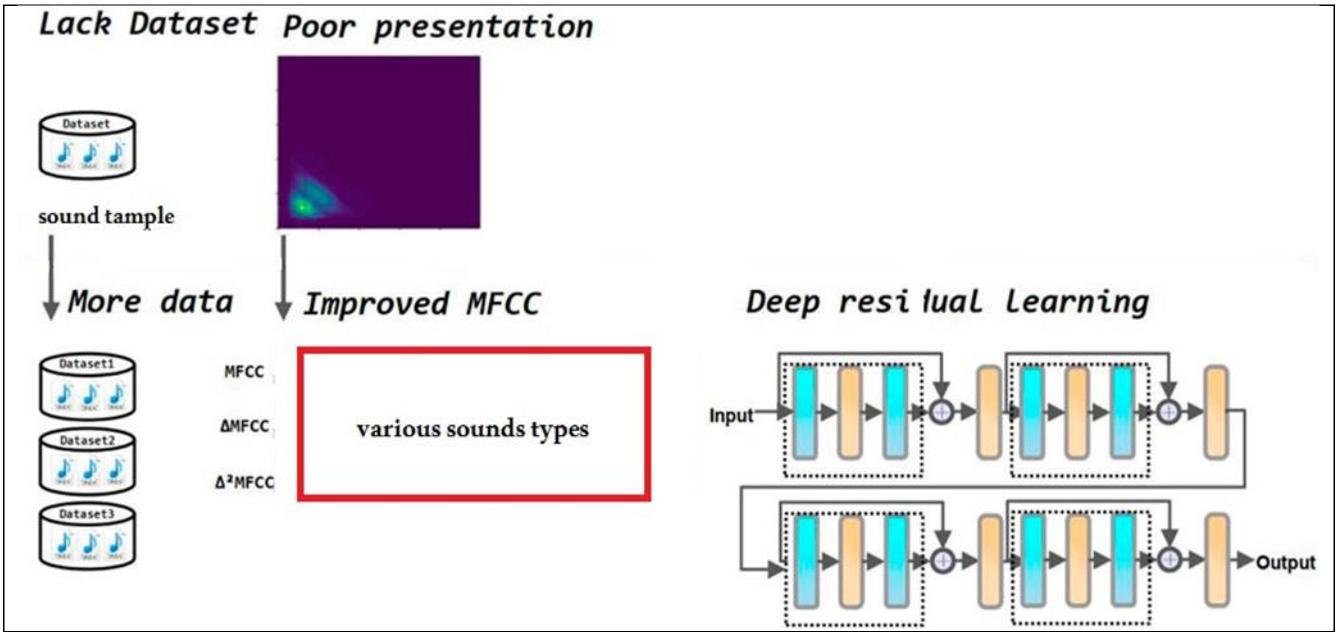

Fig.2 proposed method

1.3 Related work:

Nowadays, one of the most useful clinical diagnostics methods to treat cardiac disorders is ecg signals auscultatory technology. It has the advantages of being not invasive, effective, and convenient, and it can provide information about the heart's physiology and pathology. However, because clinical diagnostic conditions are so complex, there is so many of sound pollution, and inexperienced doctors are frequently distracted by the noise of the outside area, which leads to an inaccurate diagnosis. Doctor Marcus used angiography to observe the test of cardiac disease in the 1970s, dispelling long held misconceptions about cardiac troubles. In 1929, Dr. Werner in Germany used a catheter method to send drugs to the cardiac, opening new ways to the use of physical models to study and notice cardiovascular troubles [10][11].

In Johns H University, the first cardiac defibrillators were used in clinical settings in the 1980s, and the first telemetry sys were created so that physicians could monitor the signs and data of cardiac disease patients from another place (distance). More recently, with the advanced technology ways, devices like ECG heart sound wave analyzer and advanced electronic stethoscope wich used in clinical settings, but within to the inescapable parameters in the process, these devices have not been without their drawbacks. Current methods for digitally denoising heart sound signals include wavelet decomposition, empirical modal decomposition, and digital filters. More precise and efficient cardiac sound analysis techniques are anticipated to be accomplished in the future with the development of AI or big data, and various advanced technologies [9][11].

One of the main factors determining the outcomes is the dataset, and heart sound categorization is no exception. The more fitting of the samples can be proposed, and the ability to generate of the model can be updated, generally speaking, the massive data types collection and the more extensive the cardiac wave's sound data. A study by Milani et al. (2022) found that the significant, reliable open source heart tones dataset makes it difficult to use DL technologies for heart wave's audio classification process. In this study, more thorough, less noisy heart sound reconstructions were made using the special heart waves tones dataset from Liu et al (2016), and Pascal cardiac waves sound data-set from Gomes et al (2013), and Yaseen cardiac waves sound data-set Son & Kwon in (2018). and more trustworthy dataset of heart sounds. Both positive and negative sample imbalances may have an impact on the model's performance. It is presumptive that there is an imbalanced distribution of (+) and (-) samples in the feature area. The CNN trained to learn the networks of relationship sample at this case. In most feature space regions, it forecasts that having more samples will result in reduced loss. The model eventually fails as a result, and the projected parameter are consistently focused close To sample with additional samples as well. In other words, the model performs very well Applied to the special training groups, but inappropriately and incorrectly for the validation process. the sample potential to be generalized is greatly diminished. Deep learning approaches have been employed by numerous researchers to address heart sound categorization issues. During MFCC information collection, the effect of (DCT) on classifier was investigated. Kui et al. in 2021. The M-F-C-C method of extraction procedure, which skips the(D-C-T) stage, has an intermediate state called MFSC. Given that CNN is simply a non-linear modification of the data whereas D-C-T is simply a change in specific, This procedure has the effect of removing any pathophysiological details from the heart problems., making MFSC a viable option for DL to classify the sounds of a heartbeat [11].

2. The method:

Nowadays, one of the most useful clinical diagnostics methods to treat cardiac disorders is ecg signals auscultator technology. Given data integration, that down samples and filtering, and trims the source heart sounds, is the first step. The next phase is creating a feature, which involves taking input feature vectors and fusing conventional, 1st, and 2nd order MFCC. The third phase involves building using a deep convolution learning algorithm and instructing it with extracted features. Lastly, the trained model is used to predict the test samples, and accuracy is measured. The paper's workflow is depicted in Figure 3. The three ways the methodology is novel are as follows: 1) Utilizing the reliable heart tones sample (dataset) from three separate sources, which significantly boosts the deep learning model's performance. 2) Using enhanced MFCC as features for inputs data to better capture Both static and dynamic properties of the heartbeat's audible pulse. Improving gradient degradation and vanishing during learning by utilizing a residual human heart [12].

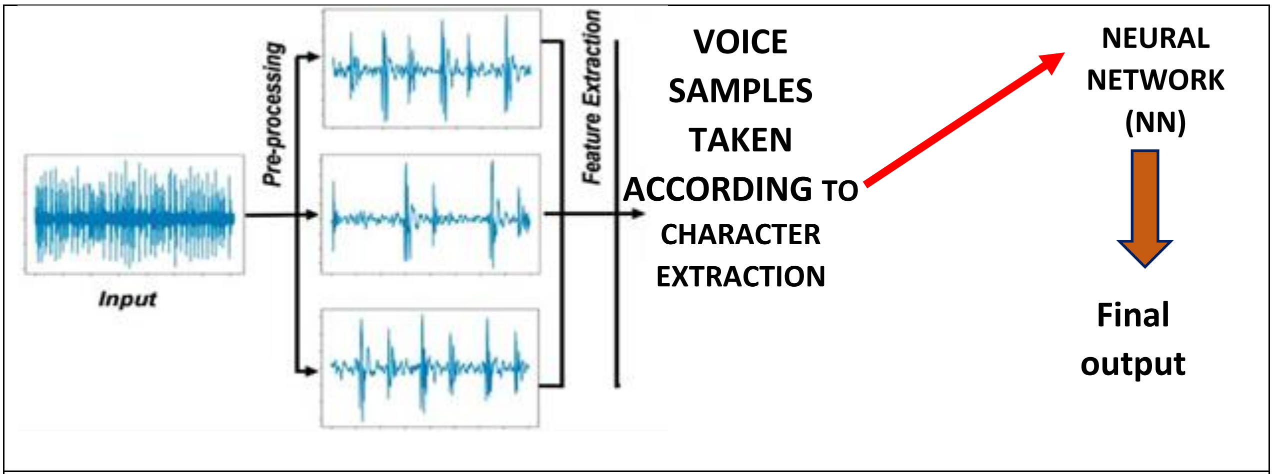

Fig.3 our proposed system chart

2.1 Collection fusion:

The datasets chosen for this article have varied label categorization requirements ,The labels must be harmonized and all files must go through data pre-processing before being fed into the neural network.utilizing all available heart sound datasets from various sources aids in further improving the generalization of the model. This study classifies the classification of the combined cardiac pulse evidence onto 3 groups: normal, pathological, &noise, depending on the traits of the label types of the dataset [13].

2.2 Electronic filtering:

Due to technology restrictions and the impact of the surrounding environment, numerous noises will unavoidably be gathered in the heart sound audio. In this article, the heart sound audio was filtered to lessen the effect of noise during neural network training. This study removes the elevated sighs by passing the cardiac sound audio through a third 400 Hz Cutoff limited filters. in the heart sound signal in order to preserve the low frequency components of the heart sounds that contain significant physiological information [13].

2.3 Downward sampling:

To 2000 hz, all sound waves are compressed. in order to simplify the figure's and also the presupposes difficulty in terms of computing that cardiac tones data from various sources can produce feature maps of the same size during the feature engineering process [10][13].

2.4 Sound clipping:

This study compresses the audio into units of 2 s to make the most of the duration of the current heartbeat sounds and united sounds given the large variability in length amongst heart sound audios. However, due to the difficulty in expressing the diseased aspects of Too short a heart in the heart sound and the strong temporal association of aberrant immune in acoustics of heartbeat, this work excluded heart sound samples under 2 seconds [15].

2.5 Enhancement engineering:

Most of the time, deep learning models are unable to fully learn from random data, and it is important to hard-code heart sound features using feature engineering. This study used a better feature extraction approach to get a pathologic picture of cardiovascular illnesses that was accurate. The perception of frequency by the human ear is logarithmic. The Changes with the low frequency levels affect it, whilst changes in the high frequency levels have no effect on it. The performance of the model was impacted by the feature engineering's usage of linearly distributed spectrum diagrams. The nonlinear link among sound frequency as well as the natural ear, which is reflected in MFCC, can be used to derive pathogenic aspects from heart sound. The following equation represents the MFCC calculation [11][14].

\

$$\mathrm{Mel}(f)=2595\lg(1+f/700) \qquad 1$$

To close the difference between the low and high -freq components of the signal, the heart tones sound waves is subjected to a HPF(high pass) in this study [15]. This is how the signal x[n] specifically operates. {{where $\alpha \cong 1.$}}

$$y[n]=x[n]-\alpha x[n-1] \quad 2$$

It is important to execute a Fourier transform on the audio samples waves in order to acquire the distribution of frequency levels component in the mp3 waves of the heartbeat. As the Fourier transform requires a stable input data, the audio samples must first be framed and frame. In framing, the original signal is divided into numerous manageable time-based blocks, each of which is referred to as a frame [16]. This is how the Hamming function of window W(n) is displayed.{{ where the α value is 0.46}}

$$W(n)=(1-\alpha)-\alpha\cos(2\pi n/(N-1)),0\leq n\leq N-1 \quad 3$$

This work employed discrete of Fourier transformation on the data after windowing and framing to convert the time-domain data into a f-hz domain data and produce the spectrum that equal Xk [17].

$$X(k)=\sum_{n=0}^{n-1} x(n)e\hat{}{-j2\pi nk/N},0\leq n,k\leq N-1 \quad 4$$

According to Eq. 5, the power P(k) = signal spectrum as the square of its modulus. The heart sound signal's energy properties are more properly expressed by the power spectrum, which keeps some amplitude components while discarding the heart sound signal's phase characteristics [18].

$$P(k)=\frac{1}{N} X(k)|\hat{}2 \quad 5$$

2.6 Extracting features dynamically:

The various heart wave data also contains very important features, can be exploited to further increase classification value. MFCC represents the static data of the heart wave signal, This work extracts the 1st difference coefficient D(n) and the 2nd difference coefficient D.2.(n) to reflect the dynamic information of the heart sound signal. The following is a description of the calculating formulas [20].

$$D(n)=\frac{1}{\sqrt{\sum_{i=-k}^{i=k} i\hat{}2}} \sum_{i=-k}^{i=k} i \cdot C(n+i) \quad 6$$

$$D2(n)=\frac{1}{\sqrt{2\sum_{i=-k}^{i=k} i\hat{}2}} \sum_{i=-k}^{i=k} i \cdot C(n+i) \quad 7$$

Where C (n + I is a frame of MFCC coefficient and the value of k is assumed to be 2. These are depicted in 2D in Figure 5, where MFCC is the outcome of Equation 6; MFCC is the outcome of Equation 10; and 2MFCC is the outcome of Equation 11. These are all (199,13) in size, and we combined them to create a (198,39.2) feature for the CNN input [21].

Figure 5 depicts the network structure in this article.

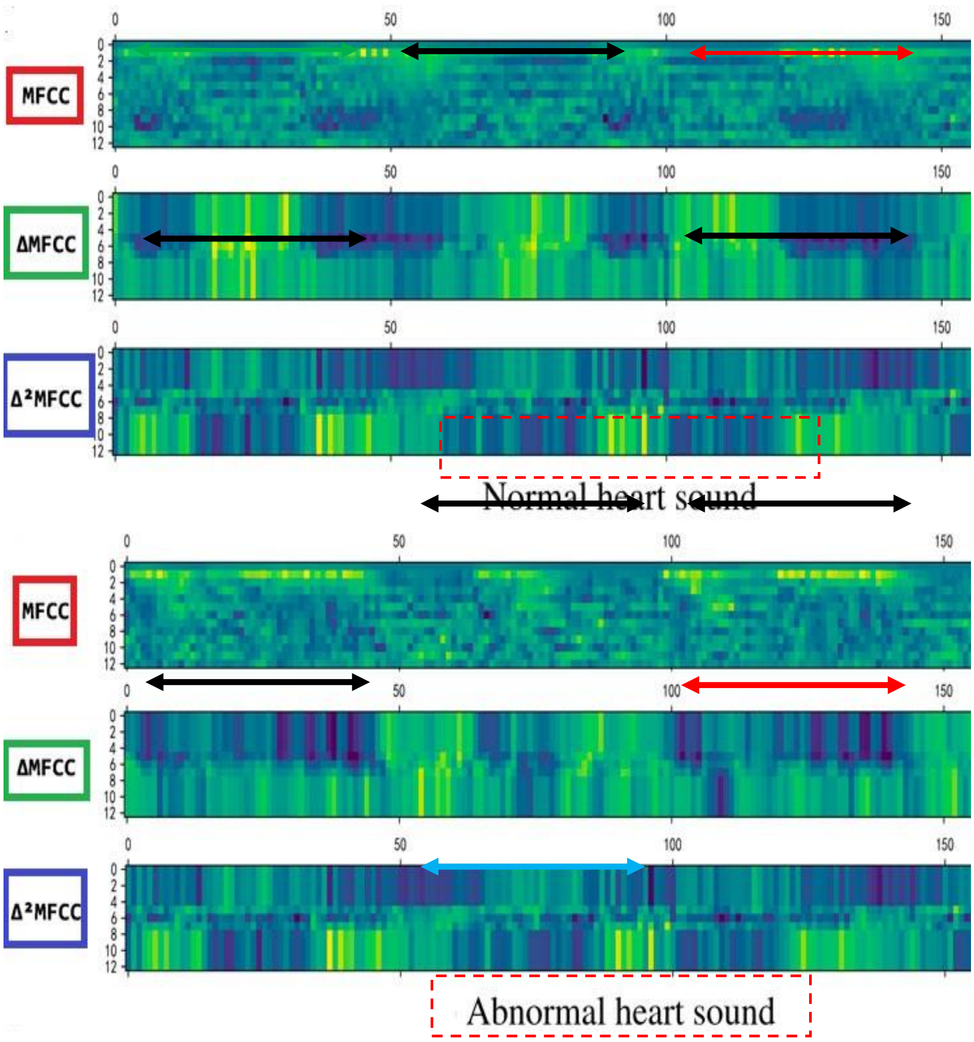

Fig.4 **(A)** Normal heart sound. **(B)** Abnormal heart sound samples

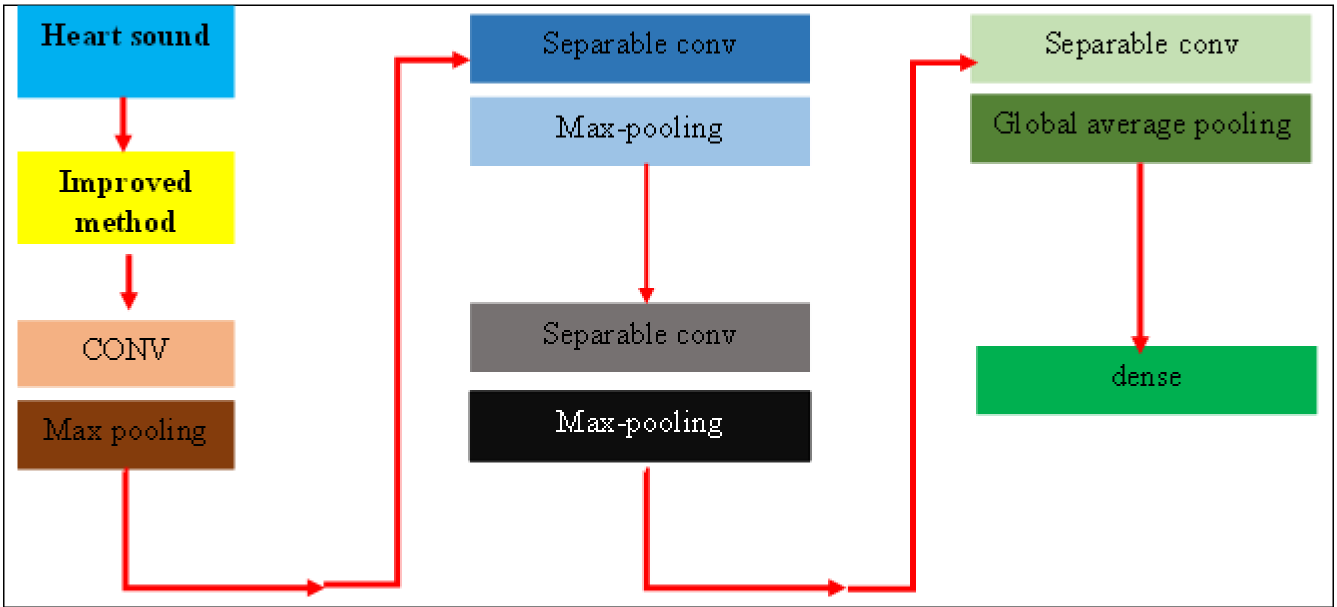

Fig.5 the method net

Large one of heart tones spectroscopy produced from conventional ANN can teach a neural network useful features. CNN not only shares many similarities with conventional fully connected neural networks but also has various variations and optimizations based on them. The fundamental idea behind how convolutional neural networks operate is to transform the source data into a 2D matrix representation [17]. CNN offers a straightforward structure, practical functionality, and trainability. Figure 6 illustrates the CNN convolution calculation algorithm's basic operation.

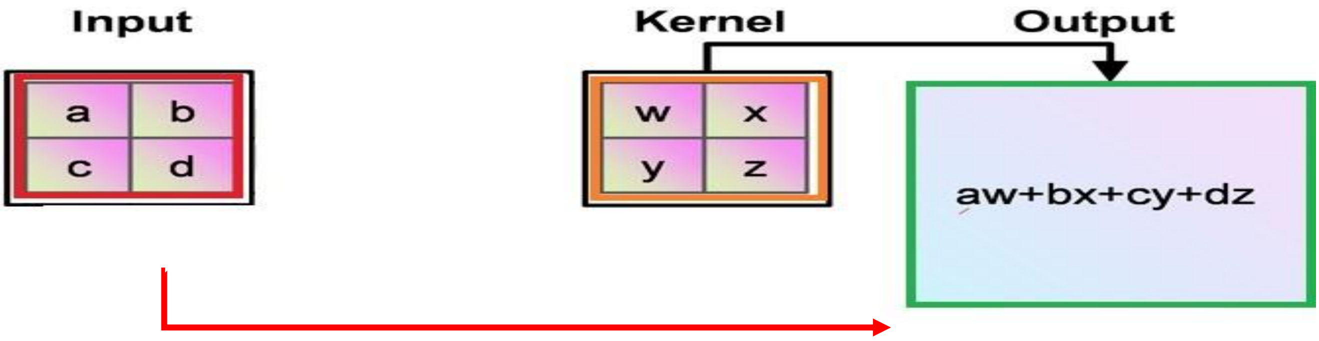

Fig. 6 CNN convolution algorithm

Figure 7 illustrates the fundamental concept of the residual structure produced in this study. Typically, a layer of the lattice can be represented wich {y = H(x)}, while the remaining mass of the lattice is given by:
: Since{ H(x) ={ F(x)+x}}and {F(x) = H(x)}, where (y equal to x) is observed one and H(x) is the expected one, the residual value, or F(x), is H(x), which is why the network is so named. The amount of information the deep network gathers as it spreads out reduces layer level by level as the network gets deeper. With this processing, the information appears to be accumulating in a layer-by-layer fashion. This is excellent for preventing data loss, which is quite helpful. The complete network just requires that portion of the input data and output data difference, which simplifies the experiment objective and complexity, the remaining block safeguards the value of the data by transfer the input information directly near the output level [22][25].

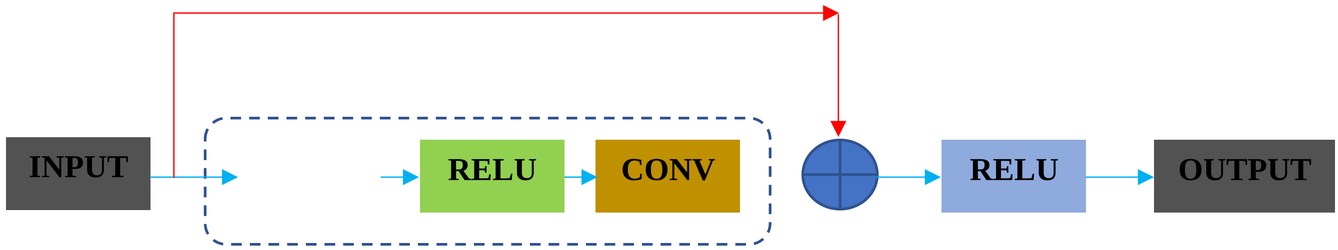

Fig.7 the structure of Residual

2.7 Dataset:

The heart speech datasets from the Physio Net Challenge 2016 competition, the Kaggle platform heart voice dataset, and the Yassin heart voice dataset are all used in this study. Table 1 lists the specifics of this data set [26] .

Table 1: PhysioNet dataset.

THE FILE	NORMAL	ABNORMAL
Training 1	292	113
Training 2	52	426
Training 3	11	57
Training 4	54	44
Training 5	74	90
Training 6	86	85
Training 7	57	66

For more information, see Table 2 [26]. The sound range in this dataset is(1 to 30) seconds.

Table 2: CinC dataset.

NORMAL	MURMUR	EXTRAHS	AIRCRAFT
31	34	19	40
320	95	no	no
351	133	19	40

3. Results:

The code attached to our research. We conducted the test on a specific database according to models and audio samples that were mentioned in the theoretical section.

Our code has been worked on according to the proposed algorithm and is under study, using Python code as a programming language (Fig. 8), as it is considered the most popular in the current day, and therefore we will include a sample of shapes and how to extract features from audio models and templates (Fig. 9).

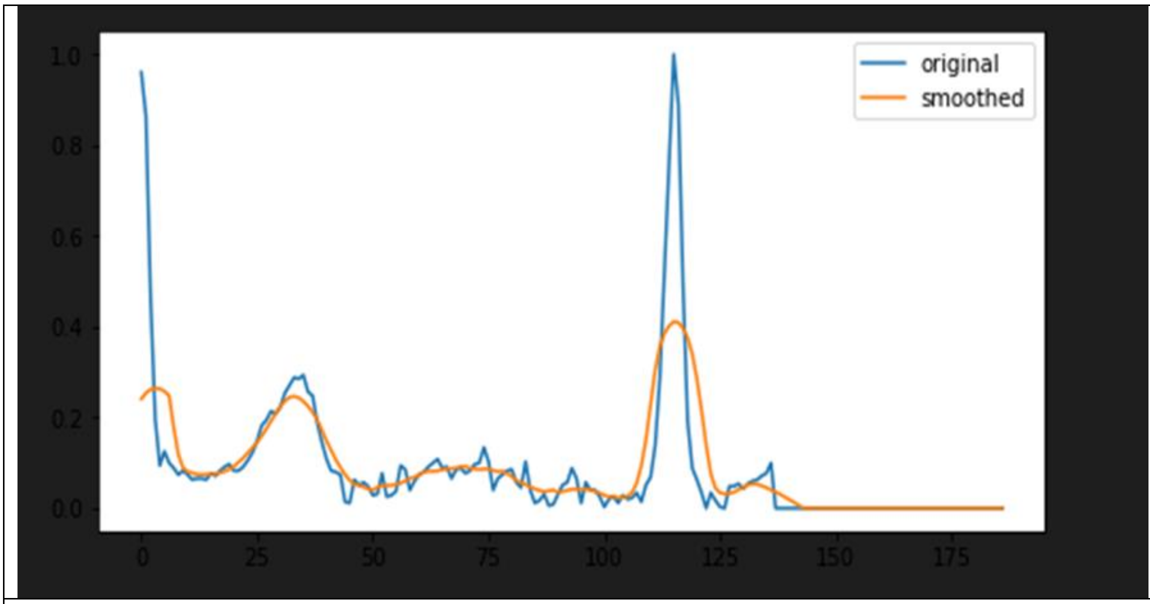

Fig.8 A sample of the output showing how to process each audio template and perform the smoothing process on the audio wave according to the proposed algorithm

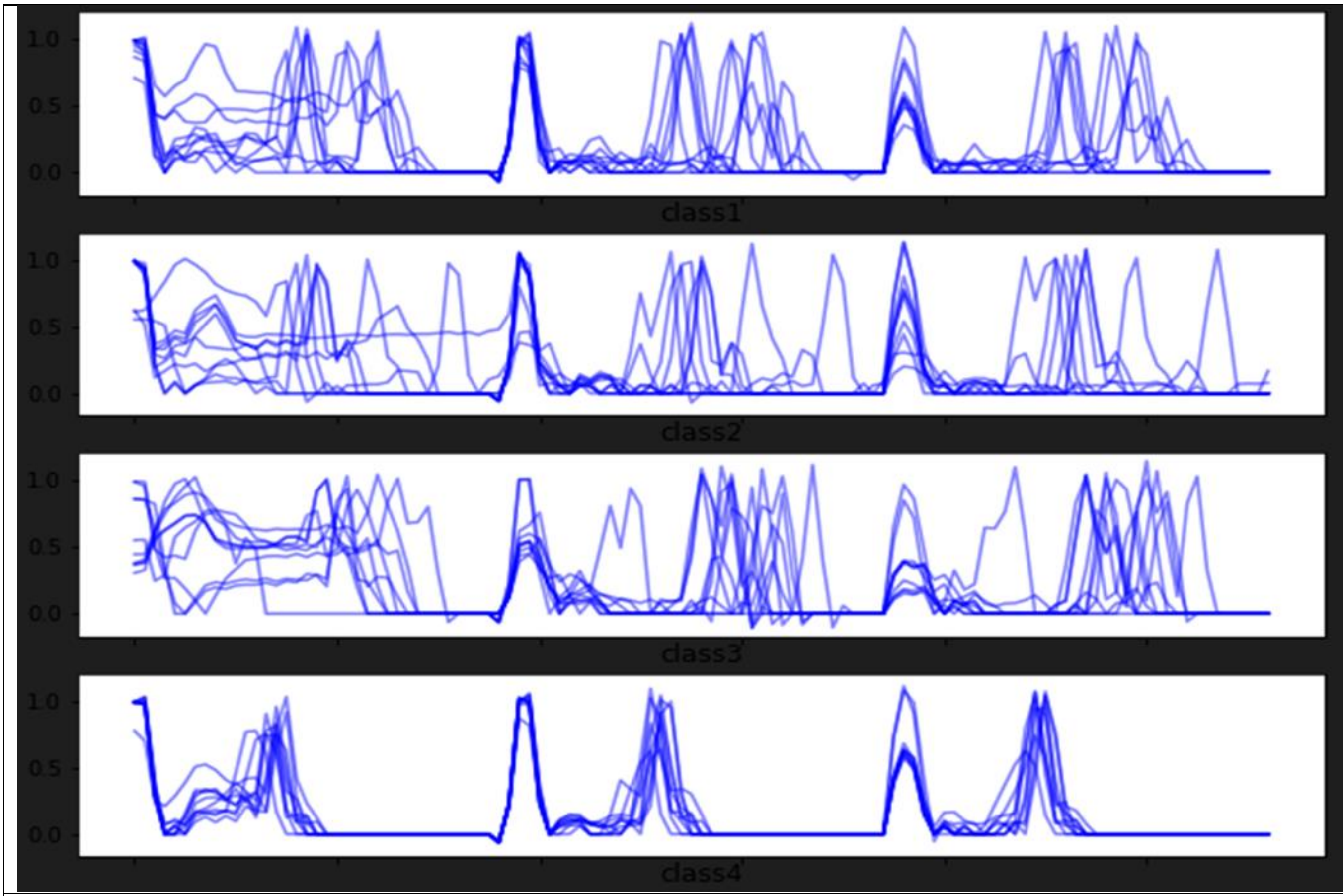

Fig.9 Applying the studied algorithm to different audio samples according to different levels

By comparing the results that appeared together through the code and the study, we found that the proposed method is one of the best methods used in order to classify heartbeats, for two main reasons, which we will mention.

First: Most of the proposed methods depend on the method of comparing the audio models with previous trainings of the networks in order to know the type and the specific type of sound. As for our proposed method, it depends on determining the type through the advantage of extracting features on the audio models and templates of the heartbeat, and in this case it is more accurate than others.

Second: Most of the methods and algorithms that work to identify the heartbeat take the sample as it is without modification and perform processing operations on it. In our proposed method, we pre-process the sample and then extract the features.

4. Conclusion:

For no dependable heart sound datasets, we integrated data sets from three different platforms in our study, which gave the neural network a strong foundation for training.
Experiments demonstrate that using these data as input to the CNN system can significantly increase the performance of the type. And we employed an improved feature and data extraction approach based on the investigated method ,
The suggested approach quickens the CNN training process and improves updating, which successfully notice effects of disappearance on the recognition of the medical signals and achieves a high value rate based on generated data set, which is bigger than the case Of the various artistic movements.

References:

[1] Yang T.-c. I., Hsieh H. (2016). "Classification of acoustic physiological signals based on deep learning neural networks with augmented features," in *2016 computing in Cardiology conference (CinC)* (IEEE), 569–572.

[2] Yu K., Jia L., Chen Y., Xu W., et al. (2013). Deep learning: Yesterday, today, and tomorrow. *J. Comput. Res. Dev.* 50, 1799–1804.

[3] Zeinali Y., Niaki S. T. A. (2022). Heart sound classification using signal processing and machine learning algorithms. *Mach. Learn. Appl.* 7, 100206. doi:10.1016/j.mlwa.2021.100206

[4] Zhang W., Han J. (2017). "Towards heart sound classification without segmentation using convolutional neural network," in *2017 computing in Cardiology (CinC)* (IEEE), 1–4.

[5] World Health Organization, "Cardiovascular diseases (CVDs)," May 2019,

[6] E. J. Benjamin, S. S. Virani, C. W. Callaway et al., "Heart disease and stroke statistics-2018 update: a report from the American Heart Association," *Circulation*, vol. 137, no. 12, 2018.

[7] Nation Heart Lung and Blood Institute, "Arrhythmia," May 2019,

[8] R. J. Martis, U. R. Acharya, and H. Adeli, "Current methods in electrocardiogram characterization," *Computers in Biology and Medicine*, vol. 48, no. 3, pp. 133–149, 2014.

[9] M. M. A. Hadhoud, M. I. Eladawy, and A. Farag, "Computer aided diagnosis of cardiac arrhythmias," in *2006 International Conference on Computer Engineering and Systems*, pp. 262–265, Cairo, Egypt, 2006.

[10] Y. Chen, S. Wang, C.-H. Shen, and F. Choy, "Intelligent identification of childhood musical murmurs," *Journal of Healthcare Engineering*, vol. 3, pp. 125–140, 2012.

[11] L. H. Cherif, S. M. Debbal, and F. Bereksi-Reguig, "Choice of the wavelet analyzing in the phonocardiogram signal analysis using the discrete and the packet wavelet transform," *Expert Systems with Applications*, vol. 37, no. 2, pp. 913–918, 2010.

[12] S. Mallat, "Group invariant scattering," *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.

[13] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1872–1886, 2013.

[14] S. Mallat, "Understanding deep convolutional networks," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, article 20150203, 2016.

[15] R. Leonarduzzi, H. Liu, and Y. Wang, "Scattering transform and sparse linear classifiers for art authentication," *Signal Processing*, vol. 150, pp. 11–19, 2018.

[16] J. Bruna and S. Mallat, "Classification with scattering operators," in *Computer Vision and Pattern Recognition*, pp. 1561–1566, Providence, RI, USA, June 2011.

[17] J. Andén and S. Mallat, "Multiscale scattering for audio classification," in *International Society for Music Information Retrieval Conference*, pp. 657–662, Miami, Florida, USA, 2011.

[18] J. Andén and S. Mallat, "Deep scattering spectrum," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4114–4128, 2014.

[19] R. Mark and G. Moody, "MIT-BIH Arrhythmia Database," May 2019, http://ecg.mit.edu/dbinfo.html.

[20] M. M. A. Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Information Sciences*, vol. 345, pp. 340–354, 2016.

[21] K. Luo, J. Q. Li, Z. G. Wang, and A. Cuschieri, "Patient-specific deep architectural model for ECG classification," *Journal of Healthcare Engineering*, vol. 2017, Article ID 4108720, 13 pages, 2017.

[22] U. R. Acharya, S. L. Oh, Y. Hagiwara et al., "A deep convolutional neural network model to classify heartbeats," *Computers in Biology and Medicine*, vol. 89, pp. 389–396, 2017.

[23] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, 2016.

[24] N. M. M. Nascimento, L. B. Marinho, S. A. Peixoto, J. P. do Vale Madeiro, V. H. C. de Albuquerque, and P. P. R. Filho, "Heart arrhythmia classification based on statistical moments and structural co-occurrence," *Circuits Systems and Signal Processing*, vol. 39, no. 2, pp. 631–650, 2020.

[25] Y. Kaya and H. Pehlivan, "Feature selection using genetic algorithms for premature ventricular contraction classification," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, pp. 1229–1232, Bursa, Turkey, November 2015.

[26] V. Mondjar-Guerra, J. Novo, J. Rouco, M. G. Penedo, and M. Ortega, "Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers," *Biomedical Signal Processing and Control*, vol. 47, pp. 41–48, 2019.