

Article

Cross-Domain Transfer Learning for PCG Diagnosis Algorithm

Kuo-Kun Tseng¹, Chao Wang¹, Yu-Feng Huang^{2,*}, Guan-Rong Chen¹, Kai-Leung Yung³ and Wai-Hung Ip³

¹ School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China; kkseng@hit.edu.cn (K.-K.T.); wangchao257@huawei.com (C.W.); 190110527@stu.hit.edu.cn (G.-R.C.)

² School of Journalism and Communication, Xiamen University, Xiamen 361005, China

³ Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China; kl.yung@polyu.edu.hk (K.-L.Y.); wh.ip@polyu.edu.hk (W.-H.I.)

* Correspondence: biznews@xmu.edu.cn

Abstract: Cardiechema is a way to reflect cardiovascular disease where the doctor uses a stethoscope to help determine the heart condition with a sound map. In this paper, phonocardiogram (PCG) is used as a diagnostic signal, and a deep learning diagnostic framework is proposed. By improving the architecture and modules, a new transfer learning and boosting architecture is mainly employed. In addition, a segmentation method is designed to improve on the existing signal segmentation methods, such as R wave to R wave interval segmentation and fixed segmentation. For the evaluation, the final diagnostic architecture achieved a sustainable performance with a public PCG database.

Keywords: transfer learning; phonocardiogram; biosignal diagnosis



Citation: Tseng, K.-K.; Wang, C.; Huang, Y.-F.; Chen, G.-R.; Yung, K.-L.; Ip, W.-H. Cross-Domain Transfer Learning for PCG Diagnosis Algorithm. *Biosensors* **2021**, *11*, 127. <https://doi.org/10.3390/bios11040127>

Received: 1 March 2021

Accepted: 2 April 2021

Published: 20 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The development of machine learning has enabled great progress in the diagnosis of diseases using biosignals, and the diagnostic performance is constantly improving. The phonocardiogram (PCG) signal is a time-varying heart sound signal with periodicity. The PCG and ECG are the two most important biosensor signals for cardiovascular disease.

The main information in the PCG is within the frequency range of 0–800 Hz, the frequency range of human hearing is mainly between 20–20,000 Hz, and the most sensitive auscultation frequency is 1000–5000 Hz, so there is much important information in the PCG, but it is more difficult to classify compared to ECG.

The important pathological information contained in the PCG signal is not well used. Therefore, the design of an effective algorithm for the PCG signal that can better reveal the information than previous works is an important issue for automatic diagnosis of PCG.

However, mainstream algorithms have begun to move from conventional machine learning algorithms to deep learning algorithms. Compared with traditional machine learning, the main difference of deep learning lies in multiple hidden layers, and feature extraction from each layer is not involved in the design, but is self-learned from the data using a common learning process. In this way, scholars can focus more on improving the adaptability and classification accuracy of the model without spending a lot of effort on feature mining.

The conventional PCG is largely dependent on the doctor's experience, and its reliability is not high. The main object of this work is to propose a deep learning framework for automatic diagnosis with a phonocardiogram. In summary, the contributions of this work are as follows:

- (1) A new segmentation method is proposed to improve existing signal segmentation methods such as RR (R wave to R wave) interval segmentation and fixed segmentation.
- (2) A large kernel is proposed to improve the convolution performance for the PCG waveform.

- (3) Migration transfer learning is designed for this PCG framework, and can enable PCG migration from ECG to PCG.
- (4) This paper proposes a boosting model based on Transfer Learning LKnet (MFS-LKNet Boosting).

The composition of this paper follows: Section 2 is related work, Section 3 is the architecture and algorithm, Section 4 is our expert results, Section 5 discusses some issues of this work, and the final section is our conclusion and references.

2. Related Work

The automatic diagnosis of heart sound is similar to that of the ECG. It is also divided into three parts: preprocessing, feature extraction, and classification recognition. For PCG diagnosis, many scholars have researched phonocardiograms, from noise interference removal, feature point recognition, feature parameter detection, waveform cluster analysis, waveform category judgment, and final diagnostic methods. The following describes the development status of PCG from four aspects: denoising, segmentation, feature extraction, and classification.

2.1. Noise Reduction

The PCG signal collected is generally composed of the heart sound and noise. The noise is generally divided into ambient noise and the noise of the instrument itself. Ambient noise is the noise in the environment in which the heart sounds are collected, and also contains respiratory sound. The noise generated when the instrument rubs against the skin is also an environmental noise. The instrument itself can also generate noise when collecting heart sounds due to the characteristics of its internal electronics. Ambient noise can be avoided artificially, instrument noise can also be reduced by improved devices, but it is still impossible to remove 100%, so a certain denoising step is required to make the useful components more prominent. PCG denoising essentially involves looking for a suitable filter to separate the wanted signal from the noise to provide a denoising effect. At present, the commonly used heart sound denoising methods include the empirical mode decomposition (EMD) denoising method, Hilbert yellow transform denoising method, and the wavelet denoising method [1–3].

2.2. Segmentation

The beat cycle segmentation is a very important part of the PCG automatic diagnosis research based on the feature extraction machine learning model. There are four main types of methods for heart sound segmentation [4]. (1) Envelope-based method—usually uses various techniques to construct an envelope of a heart sound signal, thereby performing heart sound segmentation. Typical methods for constructing envelopes include Shannon energy [5,6], Hilbert transform [7], extraction of heart sound signature waveforms [8], and square energy [9]. (2) Feature-based approach—this method requires the calculation of heart sound features and then segmentation decisions. Typical features include amplitude–frequency features [10,11], phase features [12], periodic component features [13], complexity-based features [14], multilevel wavelet coefficient features [15], and higher-order statistics such as kurtosis [16]. (3) The machine learning-based approach uses neural networks or other nonlinear classifiers that utilize extracted features, for example, a multilayer perceptron neural network [17], K-means clustering [18], dynamic clustering [19], and unsupervised learning methods [20]. (4) Hidden Markov Model (HMM) were originally proposed by Gamero and Watrous [21] in the context of heart sounds. Ricke et al. [22] subsequently developed this method. Then, Gill et al. incorporated timing duration and homomorphic filtering into the HMM method to improve the segmentation accuracy [23,24], and Schmidt et al. extended HMM to a hidden semi-Markov model (HSMM) by modeling the expected duration of heart sounds [25]. Springer et al. [26] further developed the HSMM method by using additional input features, logistic regression,

and an improved Viterbi algorithm, which solved the boundary conditions and improved the overall segmentation of short-term heart sound recording.

2.3. Feature Extraction

The features that the PCG can extract include time-domain features such as mean, variance, area, and amplitude, and various frequency domain features, such as the early-time short-time Fourier transform (STFT) [27,28]. The works use STFT to construct a time-varying transfer function and a time-varying coherence function that relate the heart and chest heart sounds; Ref. [28] used STFT to identify congenital heart disease such as atrial septal defects, ventricular septal defects, and arteriovenous catheterization. STFT reduces the signal stability requirements, is easy to implement, and is easy to calculate. However, its time–frequency window is fixed in the entire time–frequency plane, which cannot meet the requirements of different window functions for different components. Wavelet transform can realize the time-varying spectrum analysis of signals [29], and can analyze the details of the signal in more detail, which is the most widely used heart sound analysis method. In addition, empirical mode decomposition, Mel frequency cepstral coefficients (MFCC), linear predictive cepstral coefficients, etc., can also be used for PCG feature extraction [30–32].

2.4. Diagnostic Classification

Models for ECG classification can be used for PCG, such as SVM, and BP neural network. Jiang et al. [33] used normalized autoregressive power spectral density (NAR-PSD) curves and support vector machine (SVM) techniques to classify cardiac noise. They obtained 459 heart sound signals from six healthy volunteers and 30 patients, including 196 normal and 263 abnormal sound cases, and finally achieved high classification performance. Leung [34] used a neural network classifier to classify heart sounds; and Turkoglu [35] designed an expert system based on a BP neural network to successfully diagnose heart valve disease.

Deep learning studies for PCG diagnosis have brought more possibilities and higher accuracy to PCG classification diagnosis. Potes et al. [36] used the feature-based AdaBoost model classifier and Convolutional Neural Network (CNN) model classifier to classify on the open PCG database, and obtained an average accuracy of 0.79 and 0.82, respectively, with a rate that reached 0.86. Sujadevi et al. [37] demonstrated the feasibility of RNN and its variants LSTM and GRU in the diagnosis of PCG classification.

3. Architecture and Algorithm

The entire diagnostic system includes the following modules: noise reduction and spike removal, signal segmentation, feature extraction, and model classification. Noise has power frequency interference, myoelectric interference, and baseline drift. The signal segmentation method is based on the ECG segmentation method, and is further divided into several short waves, which are the four short-wave states of S1, Systole, S2, and Diastole. Feature extraction based on a traditional machine learning model classification includes 36 time domain features such as mean and variance of the PCG interval and amplitude, 36 frequency domain features based on Hamming window and discrete time Fourier transform, and 52 MFCC coefficient features. For these features, the three classification models of Support Vector Machines (SVM), Random Forest (RF) and Gradient Boosting Decision Tree (GBDT) are used here. The features based on the deep learning neural network are extracted by homomorphic filtering, Hilbert transform, and power spectral density estimation, and the classification model introduces the concept of migration learning. The entire classification system structure is shown in Figure 1.

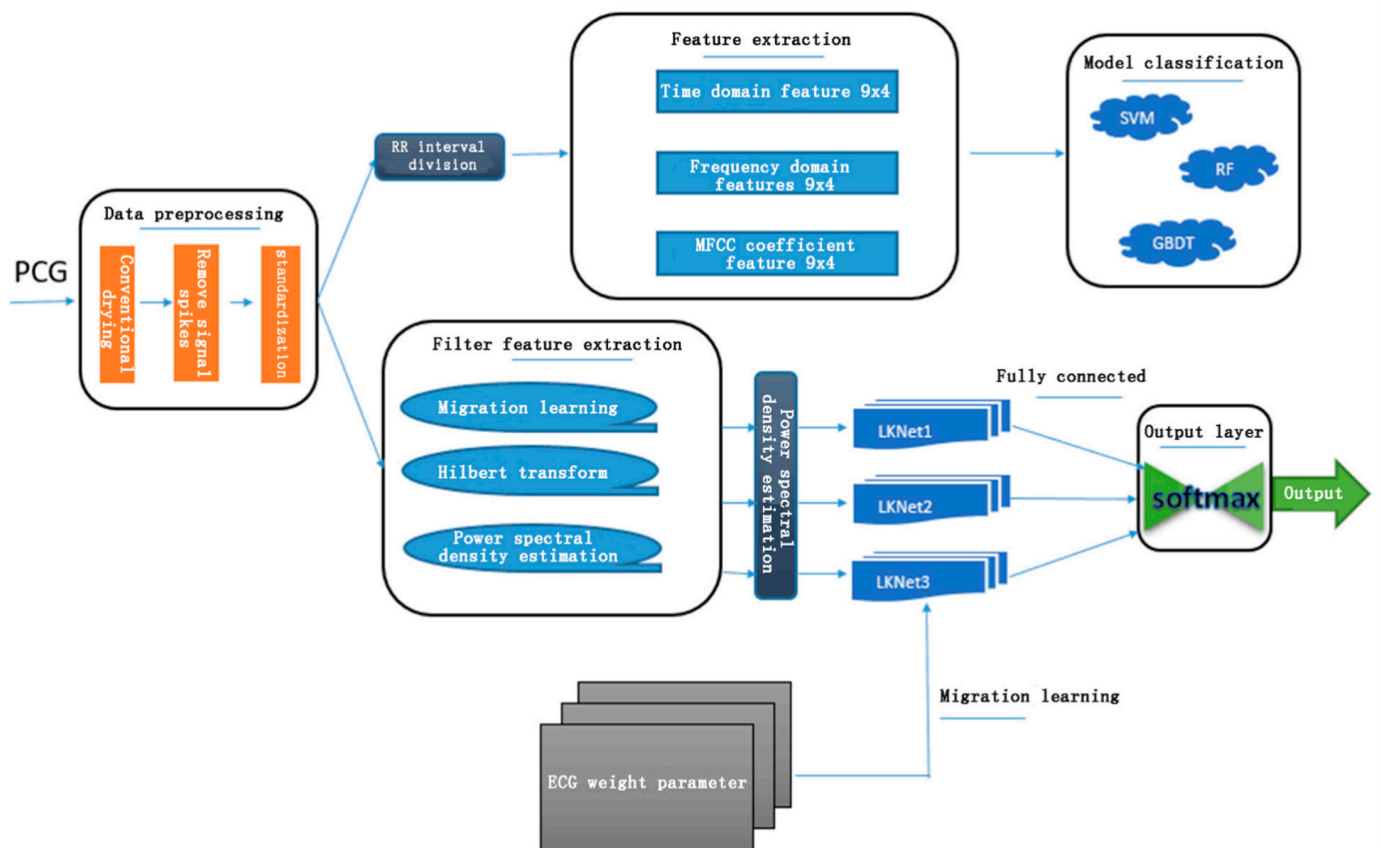


Figure 1. Framework for phonocardiogram (PCG) algorithms.

3.1. PCG Noise Reduction and Removal of Spikes

There is also some noise in the PCG signal, which is inevitable during the acquisition process. Like the ECG signal, the PCG noise can be removed to some extent by the Butterworth filter. Since the frequency of the PCG signal in the data set is 0–2000 Hz, most single-sample PCG signals are also longer in length, so downsampling is first performed to reduce the frequency to below 1000 Hz. The Butterworth band-stop filter is then used to remove the low frequencies in the signal, i.e., to remove signals below 25 Hz and signals above 400 Hz, which ensures that important information is substantially preserved while effectively removing most of the noise.

The PCG signal may have some signal spikes that interfere with the model classification. The algorithm for removing signal spikes mentioned by Schmidt in [25] is mainly for PCG signals. The peak signal in the data set used in this experiment appears less, but since the time complexity of this algorithm for removing spikes is small, the actual experiment does not take up significant time, and it is also reserved.

3.2. PCG Signal Segmentation

There are many mature PCG signal segmentation methods. Chen et al. [38] mainly use K-means cluster analysis for PCG segmentation [39] and use neural networks to segment Fundamental Heart Sounds (FHS) between S1 and S2. Because of the PCG dataset used in this chapter, some samples contain ECG data collected at the same time, and the segmentation effect of ECG signals is very good and the accuracy is extremely high. Therefore, the semihierarchical Markov chain segmentation method proposed in [26] is adopted. After the ECG signal is segmented, the corresponding point is migrated to the corresponding PCG signal as a training set. Figure 2 is a diagram of the heart sound signal and the ECG signal corresponding to the same heart beat information.

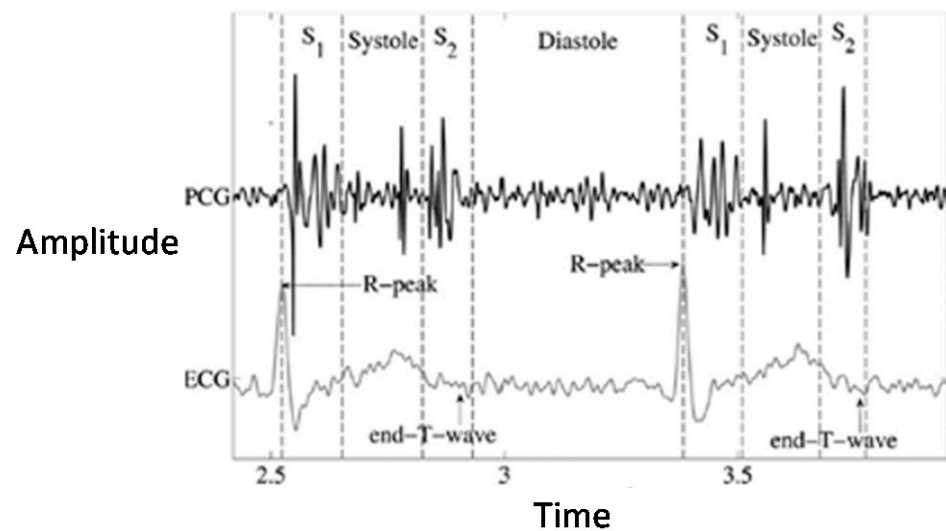


Figure 2. Waveform of PCG.

The specific steps are as follows:

- (1) The signal is divided into 500 ms per window.
- (2) Find the maximum absolute amplitude (MAA) in each window.
- (3) If at least one MAA exceeds three times the median of all window MAAs, perform the following steps. If not, go to step (4).
 - a. Select the window with the largest MAA value.
 - b. In the selected window, the position of the MAA point is defined as the top of the noise spike.
 - c. The beginning of the noise spike is defined as the last zero crossing point before the MAA point.
 - d. The end of the peak is defined as the first zero crossing point after the maximum point.
 - e. The defined noise spike is set to zero.
 - f. Go back to step (2) and continue.
- (4) The program is complete.

In the data set used in the experiment in this paper, the occurrence of spike signals is less frequent, and because the time complexity of the spike removal algorithm is still small, it is also reserved.

3.3. Filter Feature Extraction

The deep learning algorithm itself has its own feature extraction capability, which can directly feed the continuous state samples that are related before and after. After the ECG signal is preprocessed, some continuous transformations can be performed to convert the signal into other states that can be interpreted, such as spectral transform. Through these feature transformations, it is possible to capture the feature information implied by the original information into other spaces or at other levels, while maintaining the continuity of the features themselves.

Through the deformation of the heart sound signal, this paper mainly extracts three types of continuous feature models for deep learning. The methods adopted are homomorphic filtering, Hilbert transform, and power spectral density (PSD) estimation.

- a. Homomorphic filtering is often used for image processing and is a kind of frequency domain filtering. However, homomorphic filtering has its own advantages. Frequency domain filtering can flexibly solve the additive noise problem, but it cannot reduce multiplicative or convolution noise. In order to separate the additive combination signal, the linear filtering method is often used. The nonadditive signal combina-

tion commonly uses the homomorphic filtering technique to convert the nonlinear problem into a linear problem processing, that is, first to nonlinear (multiplicative or convolution). The mixed signal is a mathematical operation that is transformed into additive. Then, it is processed by the linear filtering method, and finally the inverse transform operation is performed to restore the processed image.

In image processing, it is often found that the dynamic range is large but the details of the dark area are unclear. We hope to enhance the dark area details without losing the details of the bright area. In general, we can image $f(x, y)$ modeled as the product of the illumination intensity $i(x, y)$ and the reflection intensity $r(x, y)$:

$$f(x, y) = i(x, y)r(x, y) \tag{1}$$

In general, the illumination of natural pictures is uniform, so i should be a low-frequency component, and the reflection of light by different objects has abrupt changes, so r is a high-frequency component. Now take the logarithm of the two sides and perform the Fourier transform to get the frequency domain of the linear combination:

$$\ln f(x, y) = \ln(i(x, y)) + \ln(r(x, y)) \tag{2}$$

$$FFT(\ln f(x, y)) = FFT(\ln(i(x, y))) + FFT(\ln(r(x, y))) \tag{3}$$

After the operation is completed, the inverse transformation is performed, and the processed image is obtained by exponentiation. Using this for heart sounds can reduce the low-frequency detail and pull up the high frequency.

- b. Hilbert transform is a common method in signal processing. It has a real-valued function $x(t)$ whose Hilbert transform is denoted as $\hat{x}(t)$ or $H[x(t)]$ as follows:

$$x(t) = H[x(t)] = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \tag{4}$$

In contrast to the concept of convolution, it can be found that the expression of the Hilbert transform given above is actually the result of convolving the original signal with a signal. Therefore, the Hilbert transform can be thought of as passing the original signal through a filter, or a system whose impulse response is $H(t)$. Performing a Fourier transform on $H(j\omega)$, one can get the following:

$$H(j\omega) = -j\text{sgn}(\omega) \tag{5}$$

- c. Power spectral density estimation:

$$P_X(e^{j\omega}) = \sum_{m=-\infty}^{+\infty} r_X(m)e^{-j\omega m} \tag{6}$$

where $r_x(m)$ is the autocorrelation function of the random signal. The power spectrum reflects that the power of the signal is distributed in the frequency domain with frequency ω , so $P_x(e^{j\omega})$ is also called the power density:

$$\hat{P}_{AC}(\omega) = \sum_{m=-M}^M \hat{r}(m)e^{-j\omega m}, \quad |M| \leq N - 1 \tag{7}$$

Because the power spectrum obtained by this method is obtained indirectly through the autocorrelation function, it is called an indirect method, also known as an Auto-Correlation (AC) method. When M is small, this formula is not very large, so this method was a commonly used spectral estimation method before the FFT was introduced (that is, before the periodogram method was widely used).

3.4. Transfer Learning LKNet

This paper proposes a transfer learning architecture with a large kernel network (LKNet) as shown in Figure 3. LKNet is a large-size convolution network based on a residual network, which is called transfer learning LKNet. Because ECG data are more common and has cardiac activity related data with PCG, this algorithm uses a transfer learning method to migrate knowledge from the ECG to the PCG network.

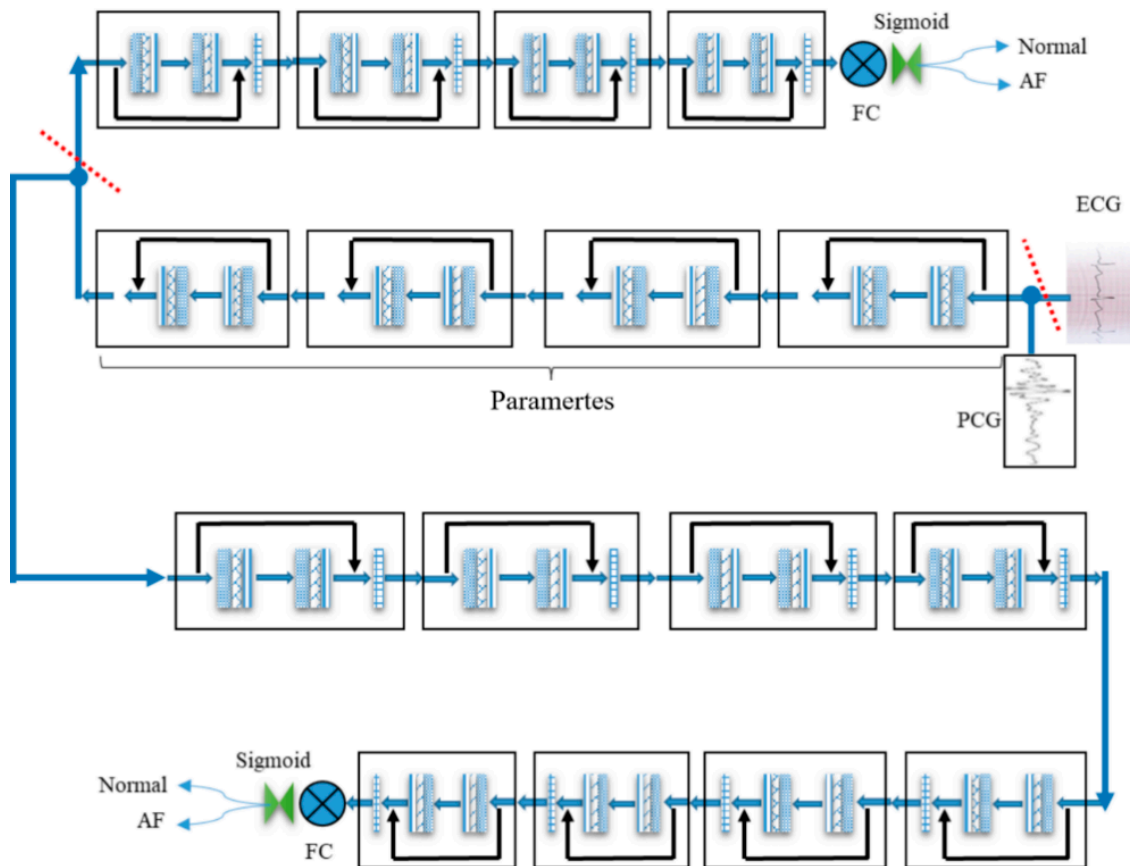


Figure 3. The transfer learning of large kernel network LKNet processes from ECG to PCG.

Transfer learning is to move the knowledge of one domain (i.e., the source domain) to another domain (i.e., the target domain), so that the target domain can achieve better learning results. Usually, the amount of data in the source domain is sufficient, and the amount of data in the target domain is less. This scenario is very suitable for migration learning. For example, we have to classify a task, but the data in this task are not sufficient (target domain), although there is a large amount of relevant training data (source domain); these training data are different from the test data of feature distribution in the classification task that needs to be performed. In this case, if the appropriate transfer learning method can be adopted, the sample is not sufficiently improved.

Transfer learning is widely used in the field of image classification. Usually, a super-large neural network that has been trained is regarded as a source domain, and a model network specific to small sample image data is used as a target domain. At this time, transfer learning is more effective; for example, ResNet101 trained in the ImageNet dataset is used to classify the small sample cat and dog image data collected by itself. Generally speaking, the more source domain samples there are, the fewer target domains there are relative to the source domain. The more similar the samples are, the more obvious the improvement in accuracy is compared to the migration of the target field samples.

The PCG is stored by conversion into audio, and the ECG is stored by conversion into an electrical signal. These two signals, however, are strictly different signals in different

fields, which mean that the similarity is not high enough to be reminiscent of migration learning. Nevertheless, under the existing conditions, through the analysis of the database, the ECG data are relatively clean, the classification effect is very good, and the ECG data sample size is several times that of the PCG, and then the combination of their collection source and the target are the characteristics of the heartbeat. In order to explore the commonality between ECG and PCG signals in deep learning, and the generalization of the improved model LKNet, the migration learning experiment was conducted.

As shown in Figure 3, the specific process follows:

- (1) From the Normal, Atrial Fibrillation (AF) Abnormal, Noise, and Other types of data samples in the ECG database, Noise and Others are removed, and the retained Normal and AF are exactly the same as the sample categories in the PCG database.
- (2) The ECG data are first input to Full Connected (FC), than it is trained by LKNet, and eight blocks are used to obtain a source network with an accuracy of 90% or more.
- (3) The last four of the eight blocks in the source network are removed, then the first four are fixed, and eight blocks are added to form a new network for training PCG data.
- (4) Figure 3 shows the process for the PCG migration from ECG data.

3.5. Boosting LKNet

In order to further improve the accuracy, we can fuse multiple LKNets to boost the transfer learning algorithm described above. The base classifier of the boosting algorithm model is generally a weak classifier. Since a single weak classifier is generally underfitting to the training sample, the model lifting effect is better. However, deep learning models are generally strong classifiers and have the ability to overfit the training set. The boosting model is very sensitive to the abnormal sample, because the input of each classifier in the boosting model depends on the classification result of the previous classifier, which results in exponential error accumulation. The sample size used for deep learning model training is large and a certain degree of mislabeling is allowed, which will seriously affect the performance of the boosting model. Therefore, boosting is usually not combined with deep learning models.

However, Figure 4 shows the main idea of the proposed model. There is some research on PCG in this section that can be applied to boosting methods. This section proposes a multifeature set and CNN-based boosting model (multifeature sets LKNet Boosting, MFS-LKNet Boosting). There are two reasons for the proposed boosting algorithm.

- (1) The first reason is that PCG has fewer public databases. The database selected this time is relatively large, but there are fewer than 3300 samples, and PCG can extract several continuous features, such as state filtering, Hilbert transform, PSD features, and others. It is helpful for the boosting model to learn the characteristics of PCG signals from different angles.
- (2) The second reason is that LKNet is adopted from CNN, the structure proposed in the previous section, and it has ideal performance for single modules. Therefore, the method of our boosting classifier enhances the predictive ability through multiple loop iterations of weight adjustment. The iterative optimization task is mainly completed by adjusting the weights of LKNet, and only a small portion is carried out by the fusion module.

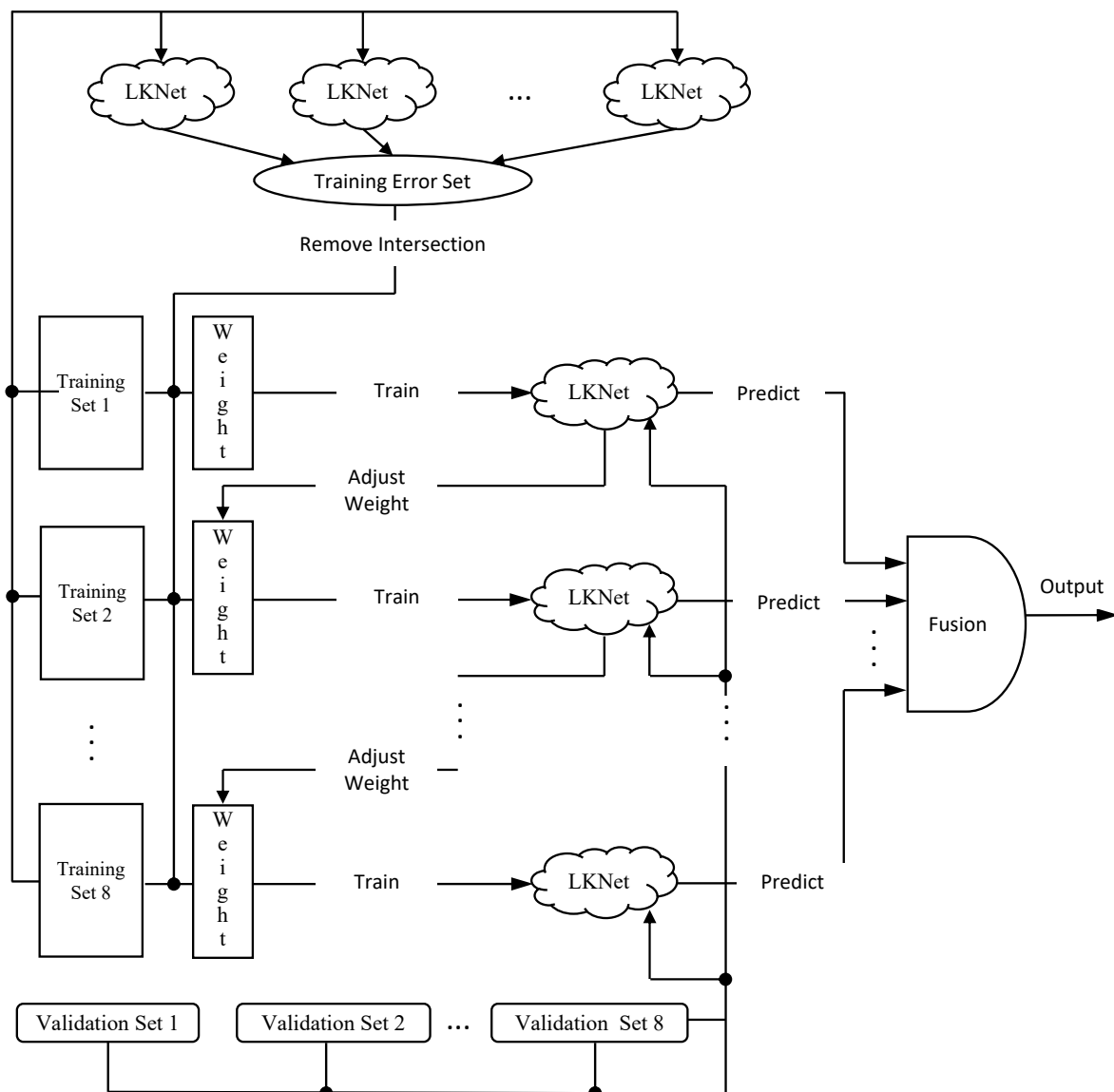


Figure 4. The multifeature sets (MFS)-LKNet Boosting.

4. Experiment

4.1. Data Sources and Evaluation Criteria

The PCG data used in this study are from the world's largest physiological signal library website (www.physionet.org, accessed on 15 April 2021), the public database used for Computing in Cardiology (CinC) challenge 2016. There are two kinds of classification tags, abnormal and normal, and the evaluation standard is the mean of specificity and sensitivity (MAcc).

4.2. Comparison of Different Segmentation

Based on the similarity of data structures between PCG and ECG, and the similarity of the purpose of disease diagnosis based on them, the deep learning model classification part of this work directly uses LKNet. The operation of the sliding window segmentation method previously proposed for the ECG signal does not bring about a significant change in accuracy. The Table 1 shows the results of the experimental classification.

Table 1. Comparison of segmentation methods.

Segmentation Method	Stride (s)	Window (s)	MAcc (%)
Fixed Segmentation	2	2	77.38
	1	2	79.25
Sliding Segmentation	2	5	78.14
	4	10	79.06
Beat Segmentation	-	-	77.21
No Segmentation	-	-	77.64

When the stride and the window are equal, the sliding segmentation becomes a fixed segmentation. The table below shows the results of the window segmentation, fixed segmentation, RR interval segmentation and nonsegmentation. The final results of the three methods of segmentation and the nonsegmentation are almost the same, and the effect of sliding window segmentation is obviously better. Since the length of most PCG signal samples does not exceed 1 min, the window is set to 10 s. The best Mean Accuracy of Percentage (MAcc%) is 79.25 for sliding segmentation with stride 1 and window(s) 2.

4.3. Comparison of Single Classification Method

In this subsection, the classification methods with single features compared in this experimental method are Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting Decision Tree (GBDT), which are the current mainstream machine learning methods. At present, the mainstream deep learning algorithm is the Convolution Neural Network (CNN). Our proposed Large Kernel Network (LKNet) has several versions of improvement, such as LKNet with Homo, Hilbert, Power Spectral Density (PSD), LKNet with ensemble SVM, and LKNet with transfer learning. Among them, transfer learning gives the best performance. The detailed experimental results are shown in Table 2.

Table 2. Comparison of single classification methods.

Models	MAcc (%)
Linear SVM	66.48
RF	66.42
GBDT	65.01
LKNet	79.06
LKNet with Homo, Hilbert, PSD	80.16
LKNet with Ensemble SVM	71.39
LKNet with transfer learning	82.73

It can be seen from Table 2 that, after the PCG signal undergoes homomorphic filtering, Hilbert transform, and PSD estimation, the extracted feature is added to LKNet and the MAcc value is 80.16, which is 1% higher than the original signal input. The addition of transfer learning has increased the model classification results by more than two percentage points. Both PCG signal transformation and transfer learning can bring about an improvement in the model classification accuracy. The traditional machine learning method linear SVM used in this paper has relatively low precision, although many features have been extracted through feature engineering. If we only consider the single module, our transfer learning LKNet has achieved relatively good results.

4.4. Comparison of Fusion Classification

Observing Table 3, Homsi et al. [40] used traditional classifiers based on feature extraction to perform model fusion through Random Forest (RF), LogitBoost (LB) and Cost Sensitive Classifier (CSC), and finally reached a MAcc value of 88.4. The second model compared comes from the fusion model of Potes et al. [36], whose fusion method uses a threshold to select from CNN or AdaBoost classification results and reach the previous high MAcc value of 89%. The last one compared, Tang et al. [41], extracted as many as

515 features, including temporal features in nine categories, namely sign, amplitude feature, energy feature, cepstrum feature, cyclostationary feature, high-order statistical feature, and cross-entropy feature, and then used SVM for fused classification, and reached a MAcc value of 88%. The boosting algorithm based on MFS LKNet proposed in this paper benefits from the good performance of transfer learning LKNet in the previous subsection, and the MAcc value increases by 1% on this basis. Compared with the latest paper with the highest accuracy, there is a 2% improvement, achieving the effect of state-of-the-art.

Table 3. Comparison of fusion methods.

Paper	Model	Sensitivity (%)	Specificity (%)	MAcc (%)
Homs et al. [40]	Fused RF + LB + CSC	94.4	86.9	88.4
Potes et al. [36]	Fused AdaBoost + CNN	96	80	89
Tang et al. [41]	Fused 515 Features SVM	88	87	88
Our paper	MFS-LKNet Boosting	96.34	86.62	92.48

5. Discussion

In this section we discuss two macro issues and two micro issues of this research. The macro issues are algorithm and application contribution, the micro issues are data set and evaluation procedure of this work.

1. What Is the Algorithm Contribution of this Work?

Compared with the transfer learning of images and texts, there is less research on the cross-domain transfer learning of biological signals. In a summary, our algorithm contribution includes proposing sliding window segmentation, LKNet with transfer learning, and LKNet with feature fusion algorithms. For major contribution of cross-domain transfer learning, this research preliminarily proved that two different but similar biological signals can also be transferred learning to improve classification accuracy. In addition, we conducted an in-depth study on the PCG algorithms, which can support PCG researchers to conduct further research with deep learning algorithm.

2. What Is the Application Contribution?

In terms of an application contribution, because heart sound is a diagnostic device commonly used by physicians, this research provides broad practical application value and can assist physicians in making more accurate diagnosis. With the mature development of smart phone, however, our algorithms can be used on mobile devices. In this way, ECG and PCG can be integrated into a separate mobile device, or integrated into a mobile phone. The inference model can be placed on the mobile phone for automated diagnosis, allowing users to collect diagnostic data from time to time, which can make up for the shortcomings of hospital service that are not in a timely manner.

The major value of this research is to explore the possibility of transfer learning between similar heterogeneous multimodal biosignal data. According to our research results, this type of heterogeneous data should be able to learn from each other's experience. We hope that our results can provide some references for the future research.

3. What Is the Data Set?

Our data set is a public ECG/PCG data set of CINC challenge. This PCG data set includes five groups of 3126 heart sound records from 764 subjects, with a sound duration ranging from 5 to 120 s. ECG recordings, collected using the AliveCor device contains 8528 single lead ECG recordings lasting from 9 to 60 s.

As the scope of the selection of input data for our given research problem, our design goal is to realize our algorithm with mobile ECG and PCG equipment. Therefore, we choose the data set by CINC's mobile devices. However, as CINC's ECG and PCG are not collected at the same time, it may be such an impact on the performance that the improvement of transfer learning is not enough. In the future, if data are collected at the same time, the effect of transfer learning should be better.

4. What Is the Evaluation Procedure?

The test procedure is focused on the accurate classification of normal and abnormal PCG. The training and testing sets are each divided into two sets of mutually exclusive populations. The testing set is approximately 10% of the training set, which is a copy of 300 records from the training set. The evaluation is carried on a 10-fold cross-validation with training and testing sets. Based on this, we summarize the overall evaluation process as follows:

Step 1: Divide the data into six training and test sets;

Step 2: First train separately for ECG data set and PCG data set;

Step 3: Transfer the training weight of ECG to the PCG network;

Step 4: Select the test group that is the different as the training set and perform the test;

Step 5: Repeat Step 2 for the cross-validation experiment.

6. Conclusions

In this paper, we successfully propose a PCG diagnosis classification algorithm, and make some improvements, including sliding window segmentation, LKNet with transfer learning, and LKNet with feature fusion algorithms. The experiment is based on a PCG automated diagnostic framework followed by frame steps for signal denoising and spike removal, and sliding window segmentation. For the deep learning feature extraction part, the feature extraction method in the image domain is also used for reference. The MFCC coefficient is also added, based on the machine learning feature extraction. The previously proposed sliding window segmentation method is better applied in this paper. At the same time, it also introduces the concept of transfer learning, designing a set of processes, and successfully applying the characteristic parameters learned from ECG to the PCG diagnostic classification. Finally, this paper proposes a multifeature set with the LKNet-based boosting model, which achieved the best result compared with state-of-the-art methods.

In the future, this research will advance in two directions. On the one hand, the algorithm will continue to be enhanced to improve its accuracy. On the other hand, the algorithm can be applied to other time-series biological signals to enable this research to make a greater contribution.

Author Contributions: K.-K.T.: Substantial contributions to the conception, English paper writing and editing. C.W.: Design of the work, acquisition, analysis, paper writing. Y.-F.H.: Paper editing, final approval of the version to be published. G.-R.C.: Interpretation of the work. K.-L.Y.: Final approval of the version to be published. W.-H.I.: Final approval of the version to be published. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: All patients involved in this study gave their informed consent. Institutional review board approval of our hospital was obtained for this study.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhidong, Z. Research of Independent Component Analysis for Preprocessing Phonocardiograms. *Chin. J. Sens. Actuators* **2003**, *16*, 103–195.
2. Gupta, A.; Misal, A.; Sinha, P.G.; Director, A. Classification of PCG Signals: A Survey. *Int. J. Comput. Appl.* **2014**, *975*, 8887.
3. Adithya, P.C.; Sankar, R.; Moreno, W.A.; Hart, S. Trends in Fetal Monitoring Through Phonocardiography: Challenges and Future Directions. *Biomed. Signal Process. Control* **2017**, *33*, 289–305. [[CrossRef](#)]
4. Liu, C.Y.; Springer, D.; Clifford, G.D. Performance of an Open-Source Heart Sound Segmentation Algorithm on Eight Independent Databases. *Physiol. Meas.* **2017**, *38*, 1730–1745. [[CrossRef](#)] [[PubMed](#)]

5. Liang, H.; Lukkarinen, S.; Hartimo, I. Heart Sound Segmentation Algorithm Based on Heart Sound Envelopogram. In Proceedings of the Computing in Cardiology, Lund, Sweden, 7–10 September 1997; Volume 105–108.
6. Moukadem, A.; Dieterlena, A.; Hueberb, N.; Brandt, C. A robust Heart Sounds Segmentation Module Based on S-Transform Biomed. *Biomed. Signal Process. Control* **2013**, *8*, 273–281. [[CrossRef](#)]
7. Sun, S.; Jiang, Z.; Wang, H.; Fang, Y. Automatic Moment Segmentation and Peak Detection Analysis of Heart Sound Pattern via Short-Time Modified Hilbert Transform. *Comput. Methods Programs Biomed.* **2014**, *114*, 219–230. [[CrossRef](#)]
8. Jiang, Z.; Choi, S. A Cardiac Sound Characteristic Waveform Method for In-Home Heart Disorder Monitoring with Electric Stethoscope. *Expert Syst. Appl.* **2006**, *31*, 286–298. [[CrossRef](#)]
9. Ari, S.; Kumar, P.; Saha, G. A Robust Heart Sound Segmentation Algorithm for Commonly Occurring Heart Valve Diseases. *J. Med. Eng. Technol.* **2008**, *32*, 456–465. [[CrossRef](#)]
10. Naseri, H.; Homaeinezhad, M.R. Detection and Boundary Identification of Phonocardiogram Sounds Using an Expert Frequency-Energy Based Metric. *Ann. Biomed. Eng.* **2013**, *41*, 279–292. [[CrossRef](#)]
11. Kumar, D.; Carvalho, P.; Antunes, M.; Henriques, J.; Eugenio, L.; Schmidt, R.; Habetha, J. Detection of S1 and S2 Heart Sounds by High Frequency Signatures. In Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 1410–1416.
12. Varghees, V.N.; Ramachandran, K.I. A Novel Heart Sound Activity Detection Framework for Automated Heart Sound Analysis. *Biomed. Signal Process. Control* **2014**, *13*, 174–188. [[CrossRef](#)]
13. Pedrosa, J.; Castro, A.; Vinhoza, T.T.V. Automatic Heart Sound Segmentation and Murmur Detection in Pediatric Phonocardiograms. In Proceedings of the Engineering in Medicine & Biology Society, Chicago, IL, USA, 26–30 August 2014; pp. 2294–2297.
14. Nigam, V.; Priemer, R. Accessing Heart Dynamics to Estimate Durations of Heart Sounds. *Physiol. Meas.* **2005**, *26*, 1005–1018. [[CrossRef](#)]
15. Vepa, J.; Tolay, P.; Jain, A. Segmentation of Heart Sounds Using Simplicity Features and Timing Information. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 469–472.
16. Papadaniil, C.D.; Hadjileontiadis, L.J. Efficient Heart Sound Segmentation and Extraction Using Ensemble Empirical Mode Decomposition and Kurtosis Features. *IEEE J. Biomed. Health Inform.* **2014**, *18*, 1138–1152. [[CrossRef](#)] [[PubMed](#)]
17. Sepehri, A.A.; Gharehbaghi, A.; Dutoit, T.; Kocharian, A.; Kiani, A. A Novel Method for Pediatric Heart Sound Segmentation without Using the ECG. *Comput. Methods Progr. Biomed.* **2010**, *99*, 43–48. [[CrossRef](#)] [[PubMed](#)]
18. Gupta, C.N.; Palaniappan, R.; Swaminathan, S.; Krishnan, S.M. Neural Network Classification of Homomorphic Segmented Heart Sounds. *Appl. Soft Comput.* **2007**, *7*, 286–297. [[CrossRef](#)]
19. Tang, H.; Li, T.; Qiu, T.; Park, Y. Segmentation of Heart Sounds Based on Dynamic Clustering. *Biomed. Signal Process. Control* **2012**, *7*, 509–516. [[CrossRef](#)]
20. Rajan, S.; Budd, E.; Stevenson, M.; Doraiswami, R. Unsupervised and Uncued Segmentation of the Fundamental Heart Sounds in Phonocardiograms Using a Time-Scale Representation. In Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 3732–3735.
21. Gamero, L.G.; Watrous, R. Detection of the First and Second Heart Sound Using Probabilistic Models. In Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society, Cancun, Mexico, 17–21 September 2003; pp. 2877–2880.
22. Ricke, A.D.; Povinelli, R.J.; Johnson, M.T. Automatic Segmentation of Heart Sound Signals Using Hidden Markov Models. In Proceedings of the Computers in Cardiology, Lyon, France, 25–28 September 2005; pp. 953–956.
23. Gill, D.; Gavrieli, N.; Intrator, N. Detection and Identification of Heart Sounds Using Homomorphic Envelopogram and Self-Organizing Probabilistic Model. In Proceedings of the Computers in Cardiology, Lyon, France, 25–28 September 2005; pp. 957–960.
24. Sedighian, P.; Subudhi, A.W.; Scalzo, F.; Asgari, S. 2014 Pediatric Heart Sound Segmentation Using Hidden Markov Model. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 26–30 August 2014; pp. 5490–5493.
25. Schmidt, S.E.; Holsthansen, C.; Graff, C.; Toft, E.; Struijk, J.J. Segmentation of Heart Sound Recordings by a Duration-Dependent Hidden Markov Model. *Physiol. Meas.* **2010**, *31*, 513–529. [[CrossRef](#)] [[PubMed](#)]
26. Springer, D.B.; Tarassenko, L.; Clifford, G.D. Logistic Regression-HSMM-Based Heart Sound Segmentation. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 822–832. [[CrossRef](#)]
27. Djebbari, A.; Bereksi, R.F. Short-Time Fourier Transform Analysis of the Phonocardiogram Signal. In Proceedings of the ICECS 2000 7th IEEE International Conference on Electronics, Circuits and Systems (Cat. No. 00EX445), Jounieh, Lebanon, 17–20 December 2000; Volume 2, pp. 844–847.
28. Vikhe, P.S.; Nehe, N.S.; Ahmednagar, L.; Thool, V.R. Heart Sound Abnormality Detection Using Short Time Fourier Transform and Continuous Wavelet Transform. In Proceedings of the 2009 Second International Conference on Emerging Trends in Engineering & Technology, Nagpur, India, 16–18 December 2009.
29. Khadra, L.; Matalgah, M.; El-Asir, B.; Mawagdeh, S. The Wavelet Transform and its Applications to Phonocardiogram Signal Analysis. *Inform. Health Soc. Care* **1991**, *16*, 271–277. [[CrossRef](#)]
30. Jamous, G.; Durand, L.G.; Langlois, Y.E.; Lanthier, T.; Pibarot, P.; Carioto, S. Optimal Time-Window Duration for Computing Time/Frequency Representations of Normal Phonocardiograms in Dogs. *Med. Biol. Eng. Comput.* **1992**, *30*, 503–508. [[CrossRef](#)]

31. Wu, Y.-C. Application of Generalized Window Function in Phonocardiogram Signal Filter. *J. Lanzhou Petrochem. Coll. Vocat. Technol.* **2017**, *17*, 1–29.
32. Gavrovska, A.; Slavković, M.; Reljin, I.; Reljin, B. Application of Wavelet and EMD-Based Denoising to Phonocardiograms. In Proceedings of the International Symposium on Signals, Iasi, Romania, 11–12 July 2013; pp. 1–4.
33. Jiang, Z.; Wang, H. A New Approach on Heart Murmurs Classification with SVM Technique. In Proceedings of the ISITC IEEE Computer Society, Jeonju, Korea, 23–24 November 2007; pp. 240–244.
34. Leung, T.S.; White, P.R.; Collis, W.B.; Brown, E.; Salmon, A.P. Classification of Heart Sounds Using Time-Frequency Method and Artificial Neural Networks. In Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society, Chicago, IL, USA, 23–28 July 2000; pp. 988–991.
35. Turkoglu, I.; Arslan, A. An Intelligent Pattern Recognition System Based on Neural Network and Wavelet Decomposition for Interpretation of Heart Sounds. In Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society, Istanbul, Turkey, 25–28 October 2001; pp. 1547–1550.
36. Potes, C.; Parvaneh, S.; Rahman, A.; Conroy, B. Ensemble of Feature-Based and Deep Learning-Based Classifiers for Detection of Abnormal Heart Sounds. In Proceedings of the Computing in Cardiology Conference, Vancouver, BC, Canada, 11–14 September 2016; pp. 621–624.
37. Sujadevi, V.G.; Soman, K.P.; Vinayakumar, R.; Sankar, A.P. Deep Models for Phonocardiography (PCG) Classification. In Proceedings of the International Conference on Intelligent Communication & Computational Techniques, Jaipur, India, 22–23 December 2017; pp. 211–216.
38. Chen, T.; Kuan, K.; Celi, L.A.; Clifford, G.D. Intelligent Heart Sound Diagnostics on a Cellphone Using a Hands-Free Kit. In *AAAI Artificial Intelligence for Development*; Springer: California, CA, USA, 2010; pp. 26–31.
39. Oskiper, T.; Watrous, R. Detection of the First Heart Sound Using a Time-Delay Neural Network. In Proceedings of the Computers in Cardiology, Memphis, TN, USA, 22–25 September 2002; pp. 537–540.
40. Hernandez, M.; Quintero, N.; Perpiñan, G.; Quintana, A.; Warrick, P. Automatic Heart Sound Recording Classification Using a Nested Set of Ensemble Algorithms. In Proceedings of the Computing in Cardiology Conference, Vancouver, BC, Canada, 11–14 September 2016; pp. 11–14.
41. Tang, H.; Dai, Z.; Jiang, Y.; Li, T.; Liu, C. PCG Classification Using Multidomain Features and SVM Classifier. *Biomed. Res. Int.* **2018**, *2018*, 1–14. [[CrossRef](#)] [[PubMed](#)]