Automatic Diagnosis of Cardiovascular Diseases Through Analysis of Heart Sound Signals Using CNN Models: A Survey

Aya Bakr
Department of Electrical Engineering
Faculty of Engineering, Aswan University
Aswan, Egypt

Samia Heshmat
Department of Electrical Engineering
Faculty of Engineering, Aswan University
Aswan, Egypt
samia.heshmat@aswu.edu.eg

***Abstract*— Heart diseases often cause changes in heart sounds and murmurs before other symptoms appear, making auscultation a crucial first step in diagnosing cardiovascular conditions. However, heart sound analysis has not been widely adopted due to uncertainties about the origins of these sounds and the lack of reliable quantitative methods for analyzing them. Since heart sound signals contain much more information than the human ear or traditional visual inspection methods can detect, automated classification is essential for early detection, especially in primary healthcare settings. This paper explores the use of deep Convolutional Neural Networks (CNNs) for classifying heart sounds as normal or abnormal. It provides a detailed analysis of CNN-based approaches, highlighting their strengths in feature extraction and classification accuracy compared to conventional methods. The paper also discusses key challenges, including model generalization, data quality, and integration with other diagnostic tools. By reviewing recent advancements, this study emphasizes the potential of CNNs to improve early diagnosis and enhance patient outcomes in cardiovascular health.**

Keywords---component; Convolutional neural networks(CNN); phonocardiograms (PCGs); and heart sound classification.

I. INTRODUCTION

Cardiovascular diseases (CVDs) are among the leading causes of mortality worldwide, making early detection crucial for effective treatment. Traditionally, physicians use a stethoscope placed at specific cardiac auscultation points to assess heart sounds. While advanced imaging techniques such as echocardiography and computed tomography (CT) offer more precise diagnoses, they are costly and time-consuming, limiting their suitability for large-scale screenings, particularly in resource-limited settings. Heart sounds, generated by myocardial contractions, provide valuable clinical insights into cardiovascular hemodynamics, aiding in both the prevention and early diagnosis of cardiovascular diseases (CVDs) [1].

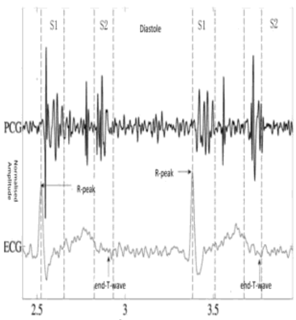

Fig. 1. The process of phonocardiography with a simultaneous electrocardiogram recording [5].

Structural abnormalities in heart valves often go unnoticed in the early stages, as symptoms may not be immediately apparent. These abnormalities can lead to vessel narrowing, altered blood flow, and abnormal arterial-venous connections, generating murmurs. Consequently, the automated classification of heart sounds is essential for the early detection of CVDs. Recent advances in artificial intelligence (AI) have facilitated the development of automated heart sound analysis techniques, improving diagnostic accuracy and efficiency [2,3].

Phonocardiography (PCG) is a well-established method for recording heart sounds, capturing signals generated during both the systolic and diastolic phases of the cardiac cycle. These sounds contain critical physiological information about the atria, ventricles, and major blood vessels, reflecting their functional state [4]. The two primary fundamental heart sounds (FHSs), the first heart sound (S1) and the second heart sound (S2), correspond to key cardiac events. S1 occurs during isovolumetric ventricular contraction when the mitral and tricuspid valves close, while S2 marks the onset of diastole with the closure of the aortic and pulmonary valves. Accurate segmentation of FHSs is vital for identifying the sequential cardiac states: S1, systole, S2, and diastole. Figure 1 illustrates a PCG process synchronized with an electrocardiogram (ECG), where the QRS complex helps determine the locations of S1 and S2. Extracting features from these heart sounds provides essential diagnostic information for evaluating cardiac health. Given their clinical importance, automated heart sound classification has gained significant attention in recent years. Research in this field has primarily focused on two approaches: traditional methods and deep learning-based methods. With the rise of big medical data and AI advancements, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have shown great promise in heart sound

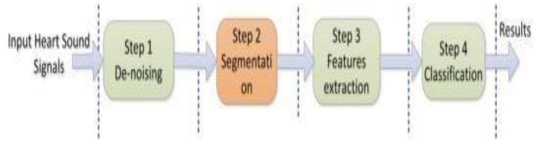

Fig. 2. The heart sound classification process.

classification. However, challenges remain in ensuring model robustness, generalizability, and integration with clinical workflows. Addressing these challenges is critical to enhancing early diagnosis and improving patient outcomes. The remainder of this paper is organized as follows: Section 2 details the heart sound classification process. Section 3 discusses traditional and deep learning-based classification methods. Section 4 focuses on CNN-based approaches for heart sound analysis. Finally, Section 5 presents the conclusion.

II. Heart Sound Classification Process

The automatic classification of heart sounds typically involves four key stages: denoising, segmentation, feature extraction, and classification, as illustrated in Figure 2. Each step plays a crucial role in ensuring accurate identification of normal and abnormal heart sounds.

A. Denoising

Heart sound recordings are often affected by various types of noise, including friction between the stethoscope and skin, electromagnetic interference, and physiological sounds such as breathing and lung noises [6]. These unwanted signals can overlap with heart sounds, making effective noise reduction essential for improving segmentation, feature extraction, and classification accuracy. Several denoising techniques have been developed to filter out these interferences, including wavelet denoising, empirical mode decomposition, and digital filtering [7]. A promising research direction involves designing a wavelet basis function specifically tailored for heart sound signals, leveraging prior knowledge about their characteristics [8].

B. Segmentation

Segmentation divides heart sound signals into four primary components: the first heart sound (S1), systole, the second heart sound (S2), and diastole. These segments contain valuable diagnostic information, but variations in heart cycle duration, the number of heart sounds, and the presence of murmurs can make accurate segmentation challenging. To address these issues, various segmentation techniques have been proposed, including envelope-based methods [9,10], electrocardiogram (ECG)-assisted approaches [11], probabilistic models [12–15], feature-based methods [16], and time-frequency analysis techniques [17]. Most segmentation algorithms assume that the diastolic phase is longer than systole; however, this assumption is not always valid, particularly for abnormal heart conditions such as those in infants or patients with cardiac disorders [18]. Among available methods, ECG-assisted segmentation, which leverages the relationship between the QRS complex and heart sounds, has demonstrated superior performance. However, these approaches require advanced hardware and software resources. Additionally, many publicly available heart sound databases do not include synchronized ECG signals, limiting their practical application.

C. Feature Extraction

Feature extraction transforms raw heart sound signals into meaningful, low-dimensional representations for analysis. Various techniques have been explored, ranging from handcrafted features to machine learning-based approaches. Commonly used features include Mel-frequency cepstral coefficients (MFCCs) [19,20] and heart sound spectra (spectrograms) derived from short-time Fourier transform (STFT) and discrete wavelet transform (DWT) coefficients [21]. Additionally, time-domain, frequency- domain, and time-frequency features extracted from the S1 and S2 components provide critical diagnostic insights [22]. A key challenge in STFT-based feature extraction is balancing time and frequency resolution, as window size directly affects both. Wavelet transformation, on the other hand, offers superior time-frequency resolution and has proven to be more effective in capturing essential features of heart sound signals [21].

D. Classification

Classification is the final step in the heart sound analysis process, categorizing phonocardiogram signals as either normal or abnormal. Traditional classification techniques include Gaussian mixture models, support vector machines, random forests, and hidden Markov models, which rely on extracted features to detect patterns associated with heart conditions. Recently, deep learning approaches, particularly deep convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have gained popularity in heart sound classification. These models automatically learn hierarchical features from raw data, improving classification accuracy and enabling early detection of cardiovascular abnormalities.

III. Analysis Methods of Heart Sound Signals

Heart sound classification methods can be broadly categorized into two main types: traditional methods and deep learning-based techniques. The following sections provide a detailed discussion of each.

A. Traditional Methods

- *Stethoscope*

For over two centuries, cardiac auscultation has been a cornerstone of clinical assessment, offering valuable insights into heart health. It remains a widely used, low-cost diagnostic tool for detecting abnormalities and guiding further medical evaluation. However, the diagnostic accuracy of auscultation has declined due to reduced exposure to rheumatic valvular diseases and the growing reliance on advanced imaging techniques like Doppler echocardiography. Consequently, improper teaching and inconsistent application have led to subjective and often inaccurate assessments [23,24]. Despite its limitations, auscultation remains a practical method for detecting heart abnormalities. However, it is highly dependent on the examiner's experience and auditory acuity [25]. Moreover, acoustic stethoscopes may alter heart sounds within the clinically relevant frequency range, making detection more

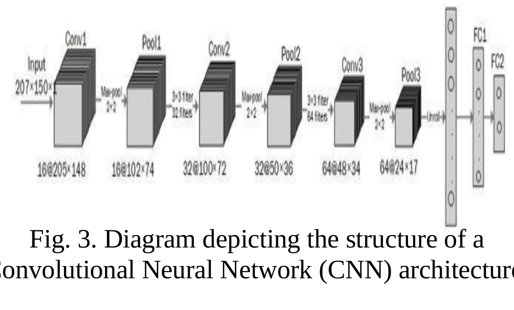

Fig. 3. Diagram depicting the structure of a Convolutional Neural Network (CNN) architecture.

challenging [26]. Additionally, significant variability exists between practitioners in interpreting heart sounds, further limiting its diagnostic reliability [27].

- *Intelligent Phonocardiography – Acoustic Cardiography*

Intelligent phonocardiography was introduced to enhance the quantitative and qualitative analysis of heart sounds. The earliest phonocardiograms, developed by Einthoven and Geluc in 1894 [28,29], evolved into more advanced optical-amplified recordings by Otto Frank in 1904 [30]. These techniques have historically been valuable for identifying abnormal heart sounds, including split sounds, opening snaps, gallop rhythms, and murmurs. With the advent of modern digital signal processing, heart sound analysis has significantly improved, enabling more precise diagnostic capabilities [31]. Studies have successfully applied discrete wavelet transformation to assess the severity of aortic and mitral valve diseases [32]. Additionally, computer- aided diagnostic systems now assist in detecting conditions such as atrial fibrillation, aortic and mitral regurgitation, pulmonary stenosis, ventricular septal defects, and congenital heart diseases [33,34]. These technologies also facilitate the diagnosis of sleep apnea, constrictive pericarditis, and left ventricular hypertrophy, offering cost-effective and efficient monitoring solutions for heart failure patients in both clinical and home settings [35].

- *Digital Stethoscope*

The introduction of digital stethoscopes has transformed computer-aided auscultation. These devices consist of three key components: data acquisition, preprocessing, and signal processing. By converting acoustic signals into electronic data, digital stethoscopes enhance auscultation through amplification and digital processing. The ability to store, analyze, and visualize heart sounds on computers further expands their diagnostic applications [36]. Many digital stethoscopes now include Bluetooth capabilities, allowing heart sounds to be transmitted wirelessly for remote diagnosis, facilitating advancements in telemedicine [37]. By reducing ambient noise and minimizing friction, digital stethoscopes provide clearer heart and lung sound recordings, enabling more precise diagnoses. Additionally, automated acoustic interpretation has the potential to revolutionize cardiovascular diagnostics and enhance bedside clinical education [38].

B. Deep Learning for Heart Analysis: Advantages and Model Types

Cardiac auscultation is a simple, non-invasive diagnostic technique widely used by healthcare professionals. However, its effectiveness heavily depends on the examiner's expertise, leading to significant variability in detecting pathological heart sounds [39]. Traditional auscultation methods struggle to manage the vast amount of heart sound data generated through long-term monitoring, highlighting the need for more advanced analytical techniques. Deep learning has emerged as a powerful tool for phonocardiogram analysis, particularly in handling large datasets with complex temporal and spectral patterns. Given the high sampling frequency and rich information content of heart sound signals, deep learning models can uncover subtle abnormalities that may be challenging to detect through conventional methods. These models have already demonstrated success in applications such as image classification, speech recognition, and medical diagnostics [40–44]. Compared to traditional imaging techniques, deep learning-based heart sound analysis offers a cost-effective, scalable, and accessible diagnostic alternative, particularly benefiting underserved regions with limited access to specialized healthcare. Among deep learning architectures, CNNs and recurrent neural networks (RNNs) are widely used for heart sound classification. CNNs excel in processing grid-like data, such as spectrograms, by applying convolutional filters to extract spatial patterns. RNNs, on the other hand, are well-suited for sequential data, capturing temporal dependencies in heart sound signals. Despite the challenges posed by the high sampling rates and large data volumes of phonocardiogram signals, modern deep learning models—including CNNs, RNNs, and transformers—have significantly improved heart sound analysis, enabling accurate classification, long-term monitoring, and early detection of cardiac abnormalities. Here's a refined version of your section with improved readability, conciseness, and a more natural academic tone while preserving all key details, citations, and subsection divisions.

IV. Convolutional Neural Network-Based Methods for Heart Sound Signal Analysis

Deep convolutional neural networks (CNNs) have demonstrated exceptional performance in various classification tasks [45,46]. These networks process input signals through multiple layers, including convolutional transformations, non-linear activation functions, and pooling operations. This layered approach enables CNNs to extract essential features while filtering out irrelevant variations, such as temporal shifts in signal characteristics [47–49]. In heart sound analysis, CNNs are commonly used both as direct classifiers and as feature extractors for traditional classifiers like support vector machines [50]. Phonocardiogram (PCG) signals can be transformed into spectrograms using different scales and transformations, converting temporal data into spatial representations. These spectrograms highlight distinctive features of murmurs, facilitating more accurate classification. Given their effectiveness in processing spatial data, CNNs have become the architecture of choice for heart sound classification tasks [51]. CNNs consist of multiple layers, including convolution, batch normalization, pooling, and fully connected layers. Convolutional layers extract features using filters, offering advantages such as parameter sharing and sparse connections. Pooling layers help reduce overfitting by downsampling intermediate outputs, with common methods including max-pooling and global average pooling. Batch normalization stabilizes training by normalizing activations, while fully connected layers handle the final classification task. The output from the last convolution or pooling layer is flattened into a vector before passing through fully connected layers. Activation functions like ReLU are used to prevent vanishing gradient issues and improve sparsity. The choice of activation function in the final layer depends on the classification task. A typical CNN architecture for heart sound classification is shown in Figure 3.

A. CNN-Based Classification Using MelSpectrum and Log-MelSpectrum Features

A 2023 study by Luca Mesin et al. evaluated the effectiveness of two Short-Time Fourier Transform (STFT)-based features—MelSpectrum and Log-MelSpectrum—in

CNN-based heart sound classification. This was the first theoretical comparison of these features in CNN models, including their robustness to additive and multiplicative noise. Results showed that Log-MelSpectrum was more effective at

suppressing additive noise, making it better suited for datasets with domain differences. Using the PhysioNet/CinC Heart Sound Classification Challenge datasets, the study demonstrated that Log-MelSpectrum features yielded higher accuracy and more consistent results when applied to a modified VGG16 CNN model.

- *Research Gap*

While this study provides valuable insights, a direct comparison between MelSpectrum and Log- MelSpectrum in CNN-based heart sound classification remains limited. Further research is needed to analyze their impact on key performance metrics such as accuracy and precision, particularly in noisy environments. Additionally, the relationship between spectral features and CNN model complexity has not been thoroughly explored. Many existing studies rely on small datasets, which may introduce biases. A broader investigation using diverse datasets is necessary to assess these features' generalizability across various heart conditions. Moreover, the potential of transfer learning for CNNs trained on heart sound data with these features remains underexplored, which could help address the limitations of small datasets.

B. Heart Sound Classification Using MFCC Features and Convolutional Recurrent Neural Networks

M. Deng et al. proposed an innovative approach at (2020) for extracting Mel-Frequency Cepstral Coefficients (MFCC) without segmenting heart sound signals. Their method enhances standard MFCC features by incorporating first- and second-order differential parameters, capturing the dynamic characteristics of consecutive heartbeats. The classification model combines a 2D CNN for feature extraction with a Long Short-Term Memory (LSTM) network to model long-term dependencies. This hybrid Convolutional Recurrent Neural Network (CRNN) approach enables more accurate classification, leveraging CNNs for spatial feature extraction and LSTMs for temporal pattern recognition. The method was extensively tested on the PhysioNet 2016 Challenge Database, achieving a classification accuracy of 98% for distinguishing pathological from non-pathological heart sounds, surpassing many existing approaches.

- *Research Gap*

Although the proposed method shows high accuracy, its evaluation was limited to the PhysioNet 2016 dataset, which has specific characteristics and quality constraints. The study does not explore the model's generalizability across real-world datasets with varying recording conditions, background noise, and acoustic environments. Additionally, the impact of environmental noise and artifacts on heart sound quality remains unexamined. CRNN models are computationally intensive, posing challenges for real- time implementation in low-resource clinical settings. Their "black box" nature also raises concerns regarding interpretability, making it difficult for medical professionals to trust the decision-making process. While the method performs well in controlled settings, its practical adoption in real-world clinical environments requires further validation.

C. CNN-Based Heart Sound Classification Using Time-Frequency Features for Pathology Detection

A study at (2018) by Baris Bozkurt et al. focused on improving automatic pediatric heart disease detection using digital PCG signals. The authors aimed to enhance CNN-based classification systems by refining segmentation techniques and time-frequency feature extraction. The study compared various feature types, including MFCC, Mel-Spectrograms, and sub-band envelopes, finding that sub-band envelopes provided superior classification performance. Different segmentation strategies and CNN architecture were also evaluated, with period-synchronous windowing yielding the best results. The research highlights sub-band envelopes as a promising feature set for improving heart sound classification accuracy. The study further contributes by making its code available for replication and validation.

- *Research Gap*

Most heart sound classification systems rely on standard features like MFCC and Mel-Spectrograms, which may not fully capture the complexities of pathological heart sounds, particularly in pediatric congenital heart disease (CHD). This study highlights the potential of sub-band envelopes as a more effective feature but calls for further research into optimized CNN architectures and alternative feature extraction methods. Additionally, current data augmentation techniques for heart sound analysis remain limited, affecting model robustness. The study identifies a need for improved segmentation strategies, enhanced feature extraction, and optimized CNN models to enhance early detection of CHD. Further exploration is also required to determine the effectiveness of different feature representations in real- world clinical applications.

V. CONCLUSION

This review explored the application of convolutional neural networks (CNNs) for the automatic diagnosis of cardiovascular diseases through heart sound analysis. Given their ability to extract meaningful features directly from raw phonocardiogram (PCG) signals, CNN-based methods have demonstrated high accuracy in detecting conditions such as heart murmurs, valve diseases, and congenital defects. Compared to traditional approaches, CNNs offer superior performance, particularly when combined with appropriate preprocessing techniques and diverse datasets. Despite these advancements, several challenges remain. Model generalization across different datasets, data quality issues, and integration with other diagnostic tools require further attention. The need for larger, more diverse datasets, effective strategies for addressing class imbalance, and improved model interpretability are critical areas for future research. Additionally, integrating multimodal data sources, developing real- time implementations, and optimizing CNNs for resource-constrained environments will be essential for broader clinical adoption.

Future studies should also explore personalized healthcare applications, cross-disease models, and adaptive CNN architectures that can accommodate variations in heart sound characteristics. Addressing these challenges will enhance the reliability and applicability of CNN-based heart sound analysis, ultimately improving early detection and diagnosis of cardiovascular diseases and leading to better patient outcomes.

References

[1] Liu, C., & Murray, A. (2017). Applications of complexity analysis in clinical heart failure. In Complexity and nonlinearity in cardiovascular signals (pp. 1–25). Springer.

[2] Baghel, N., Dutta, M. K., & Burget, R. (2020). Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network. Computer Methods and Programs in Biomedicine, 197, 105750. https://doi.org/10.1016/j.cmpb.2020.105750

[3] Chen, W., Sun, Q., Chen, X., Xie, G., Wu, H., & Xu, C. (2021). Deep learning methods for heart sounds classification: A systematic review. Entropy, 23(6), 667. https://doi.org/10.3390/e23060667

[4] Springer, D. B., Tarassenko, L., & Clifford, G. D. (2016). Logistic regression-HSMM-based heart sound segmentation. IEEE Transactions on Biomedical Engineering, 63(4), 822–832. https://doi.org/10.1109/TBME.2015.2475278

[5] Dwivedi, A. K., Imtiaz, S. A., & Rodriguez-Villegas, E. (2019). Algorithms for automatic analysis and classification of heart sounds—A systematic review. IEEE Access, 7, 8316–8345. https://doi.org/10.1109/ACCESS.2018.2889437

[6] Li, S., Li, F., Tang, S., & Xiong, W. (2020). A review of computer-aided heart sound detection techniques. BioMed Research International, 2020, Article 5846191. https://doi.org/10.1155/2020/5846191

[7] Thalmayer, A., Zeising, S., Fischer, G., & Kirchner, J. (2020). A robust and real-time capable envelope-based algorithm for heart sound classification: Validation under different physiological conditions. Sensors, 20(3), 972. https://doi.org/10.3390/s20030972

[8] Kapen, P. T., Youssoufa, M., Kouam, S. U. K., Foutse, M., Tchamda, A. R., & Tchuen, G. (2020). Phonocardiogram: A robust algorithm for generating synthetic signals and comparison with real- life ones. Biomedical Signal Processing and Control, 60, 101983. https://doi.org/10.1016/j.bspc.2020.101983

[9] Giordano, N., & Knaflitz, M. (2019). A novel method for measuring the timing of heart sound components through digital phonocardiography. Sensors, 19(8), 1868. https://doi.org/10.3390/s19081868

[10] Wei, W., Zhan, G., Wang, X., Zhang, P., & Yan, Y. (2019). A novel method for automatic heart murmur diagnosis using phonocardiogram. In Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced

[11] Manufacturing (AIAM) (pp. 1–6). Dublin, Ireland. https://doi.org/10.1145/3358331.3358345

[12] Malarvili, M., Kamarulafizam, I., Hussain, S., & Helmi, D. (2003). Heart sound segmentation algorithm based on instantaneous energy of electrocardiogram. Computers in Cardiology, 30, 327– 330.

[13] Oliveira, J. H., Renna, F., Mantadelis, T., & Coimbra, M. T. (2018). Adaptive sojourn time HSMM for heart sound segmentation. IEEE Journal of Biomedical and Health Informatics, 23(2), 642– 649. https://doi.org/10.1109/JBHI.2018.2835535

[14] Kamson, A. P., Sharma, L., & Dandapat, S. (2019). Multi-centroid diastolic duration distribution based HSMM for heart sound segmentation. Biomedical Signal Processing and Control, 48, 265–272. https://doi.org/10.1016/j.bspc.2018.12.019

[15] Renna, F., Oliveira, J. H., & Coimbra, M. T. (2019). Deep convolutional neural networks for heart sound segmentation. IEEE Journal of Biomedical and Health Informatics, 23(6), 2435–2445. https://doi.org/10.1109/JBHI.2019.2894222

[16] Liu, C., Springer, D., & Clifford, G. D. (2017). Performance of an open-source heart sound segmentation algorithm on eight independent databases. Physiological Measurement, 38(8), 1730– 1745. https://doi.org/10.1088/1361-6579/aa9044

[17] Chen, T. E., Yang, S. I., Ho, L. T., Tsai, K. H., Chen, Y. H., Chang, Y. F., & Wu, C. C. (2017). S1 and S2 heart sound recognition using deep neural networks. IEEE Transactions on Biomedical Engineering, 64(2), 372–380. https://doi.org/10.1109/TBME.2016.2559800

[18] Liu, Q., Wu, X., & Ma, X. (2018). An automatic segmentation method for heart sounds. Biomedical Engineering Online, 17(1), 22. https://doi.org/10.1186/s12938-018-0458-8

[19] Deng, S.-W., & Han, J.-Q. (2016). Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps. Future Generation Computer Systems, 60, 13–21. https://doi.org/10.1016/j.future.2015.11.024

[20] Abduh, Z., Nehary, E. A., Wahed, M. A., & Kadah, Y. M. (2019). Classification of heart sounds using fractional Fourier transform based Mel-frequency spectral coefficients and stacked autoencoder deep neural network. Journal of Medical Imaging and Health Informatics, 9(1), 1–8. https://doi.org/10.1166/jmihi.2019.2570

[21] Nogueira, D. M., Ferreira, C. A., Gomes, E. F., & Jorge, A. M. (2019). Classifying heart sounds using images of motifs, MFCC and temporal features. Journal of Medical Systems, 43(6), 168. https://doi.org/10.1007/s10916-019-1297-2

[22] Soeta, Y., & Bito, Y. (2015). Detection of features of prosthetic cardiac valve sound by spectrogram analysis. Applied Acoustics, 89,28–33. https://doi.org/10.1016/j.apacoust.2014.08.019

[23] Chakir, F., Jilbab, A., Nacir, C., & Hammouch, A. (2018). Phonocardiogram signals processing approach for PASCAL Classifying Heart Sounds Challenge. Signal, Image and Video Processing, 12(6), 1149–1155. https://doi.org/10.1007/s11760-018-1258-0

[24] Bohadana, A., Izbicki, G., & Kraman, S. S. (2014). Fundamentals of lung auscultation. New England Journal of Medicine, 370(8), 744–751. https://doi.org/10.1056/NEJMra1302901

[25] Xavier, G. N., Duarte, A. C., Melo-Silva, C. A., dos Santos, C. E., & Amado, V. M. (2014). Accuracy of pulmonary auscultation to detect abnormal respiratory mechanics: A cross-sectional diagnostic study. Medical Hypotheses, 83(6), 733–734. https://doi.org/10.1016/j.mehy.2014.09.029

[26] Oliveira, A., & Marques, A. (2014). Respiratory sounds in healthy people: A systematic review. Respiratory Medicine, 108(4), 550–570. https://doi.org/10.1016/j.rmed.2014.01.004

[27] Abella, M., Formolo, J., & Penney, D. (1992). Comparison of the acoustic properties of six popular stethoscopes. The Journal of the Acoustical Society of America, 91(4), 2224–2228. https://doi.org/10.1121/1.403655

[28] Murphy, R., Vyshedskiy, A., Power-Charnitsky, V. A., Bana, D. S., Marinelli, P. M., & Wong-Tse, A. (2004). Automated lung sound analysis in patients with pneumonia. Respiratory Care, 49(12), 1490–1497.

[29] Einthoven, W., & Geluc, M. A. J. (1894). Die Registrierung der Herztöne. Pflügers Archiv für die Gesamte Physiologie des Menschen und der Tiere, 57, 617–622.

[30] Frank, O. (1904). Die Unmittelbare Registrierung Der Herztöne. Münchener Medizinische Wochenschrift, 51, 593–595.

[31] Obaidat, M. S. (1993). Phonocardiogram signal analysis: Techniques and performance comparison. Journal of Medical Engineering & Technology, 17(6), 221–227. https://doi.org/10.3109/03091909309006329

[32] Meziani, F., Debbal, S. M., & Atbi, A. (2013). Analysis of the pathological severity degree of aortic stenosis (AS) and mitral stenosis (MS) using the discrete wavelet transform (DWT). Journal of Medical Engineering & Technology, 37(1), 61–74. https://doi.org/10.3109/03091902.2012.733058

[33] Sepehri, A. A., Kocharian, A., Janani, A., & Gharehbaghi, A. (2016). An intelligent phonocardiography for automated screening of pediatric heart diseases. Journal of Medical Systems, 40(1), 16. https://doi.org/10.1007/s10916-015-0359-3

[34] Gharehbaghi, A., Ekman, I., Ask, P., Nylander, E., & Janerot-Sjoberg, B. (2015). Assessment of aortic valve stenosis severity using intelligent phonocardiography. International Journal of Cardiology, 198, 58–60. https://doi.org/10.1016/j.ijcard.2015.06.126

[35] Wang, S., Liu, M., Fang, F., Liang, Y., Li, Y., & Huang, Q. (2016). Prognostic value of acoustic cardiography in patients with chronic heart failure. International Journal of Cardiology, 219, 121–126. https://doi.org/10.1016/j.ijcard.2016.06.004

[36] Leng, S., Tan, R. S., Chai, K. T., Wang, C., Ghista, D., & Zhong, L. (2015). The electronic stethoscope. BiomedicalEngineering

Online, 14(1), 66. https://doi.org/10.1186/s12938-015-0056-y

[37] Lakhe, A., Sodhi, I., Warrier, J., & Sinha, V. (2016). Development of digital stethoscope for telemedicine. Journal of Medical Engineering & Technology, 40(1), 20–24. https://doi.org/10.3109/03091902.2015.1116633

[38] Rabinowitz, P., Taiwo, O., Sircar, K., Aliyu, O., & Slade, M. (2006). Physician hearing loss. American Journal of Otolaryngology, 27(1), 18–23. https://doi.org/10.1016/j.amjoto.2005.05.014

[39] Mangione, S., & Nieman, L. Z. (1997). Cardiac auscultatory skills of internal medicine and family practice trainees: A comparison of diagnostic proficiency. JAMA, 278(9), 717–722. https://doi.org/10.1001/jama.1997.03550090041030

[40] Wang, P., Fan, E., & Wang, P. (2021). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. Pattern Recognition Letters, 141, 61–67. https://doi.org/10.1016/j.patrec.2020.12.003

[41] Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., & Shaalan, K. (2019). Speech recognition using deep neural networks: A systematic review. IEEE Access, 7, 19143–19165. https://doi.org/10.1109/ACCESS.2019.2896880

[42] Otter, D. W., Medina, J. R., & Kalita, J. K. (2020). A survey of the usages of deep learning for natural language processing. IEEE Transactions on Neural Networks and Learning Systems, 32(2), 604–624.

https://doi.org/10.1109/TNNLS.2020.2979670

[43] Oh, S. L., Hagiwara, Y., Raghavendra, U., Yuvaraj, R., Arunkumar, N., Murugappan, M., & Acharya, U. R. (2020). A deep learning approach for Parkinson's disease diagnosis from EEG signals. Neural Computing and Applications, 32, 10927–10933. https://doi.org/10.1007/s00521-018-3689-5

[44] Jamshidi, M., Lalbakhsh, A., Talla, J., Peroutka, Z., Hadjilooei, F., Lalbakhsh, P., & Mohyuddin, W. (2020). Artificial intelligence and COVID-19: Deep learning approaches for diagnosis and treatment. IEEE Access, 8, 109581–109595. https://doi.org/10.1109/ACCESS.2020.3001973

[45] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

https://doi.org/10.1038/nature14539

[46] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

http://www.deeplearningbook.org

[47] Bruna, J., & Mallat, S. (2013). Invariant scattering convolution networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1872–1886. https://doi.org/10.1109/TPAMI.2013.86

[48] Mallat, S. (2012). Group invariant scattering. Communications on Pure and Applied Mathematics, 65(10), 1331–1398. https://doi.org/10.1002/cpa.21413

[49] Wiatowski, T., & Bolcskei, H. (2015). A mathematical theory of deep convolutional neural networks for feature extraction. arXiv:1512.06293.

[50] Huang, F. J., & LeCun, Y. (2006). Large-scale learning with SVM and convolutional nets for generic object categorization. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Vol. 1, pp. 284–291). IEEE.

[51] Humayun, A. I., Ghaffarzadegan, S., Feng, Z., & Hasan, T. (2018). Learning front-end filter-bank parameters using convolutional neural networks for abnormal heart sound detection. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1408–1411). IEEE.