Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Use Of Audio Transfer Learning To Analyse Heart Sounds For Detecting Heart Diseases

Sathyanarayanan Swaminathan*, Srikanta Murthy K., Chandrashekar Gudada, Satish Kumar Mallappa

Department of Mathematical and Computational Sciences
Sri Sathya Sai University for Human Excellence, Gulbarga, Karnataka - 585313, India
**: sathyanarayanan.s@sssuhe.ac.in/sathyanarayanan.brn@gmail.com*

Cardiovascular disease (CVDs) is a major cause of mortality worldwide. Timely identification of CVD is essential to reduce the fatalities associated with this condition. The integration of AI into cardiovascular care can make it more accessible and reduce the mortality rates. In this study, an innovative artificial intelligence (AI) technique using a pretrained audio transfer learning layer was developed to detect valvular heart disease. This study incorporated the YAMNet pretrained audio layer into a specialised neural network called the Heart Sounds Analysis Network (HSANet). The methodologies demonstrated strong performance on a dataset containing 957 heart sound recordings, encompassing one normal case and four distinct valvular heart diseases. The binary classification task, which distinguished normal cases from abnormal cases, achieved an accuracy of 90.86%. This approach yielded an accuracy of 99.71% for the five-class classification task. The model yielded favourable outcomes even when confronted with imbalanced data and operated efficiently in a modest system. The AI-powered diagnostic approach holds significant promise for the early detection of heart diseases owing to its accuracy and noninvasive characteristics, rendering it suitable for various healthcare environments.

Keywords: Heart sounds, auscultation, cardiovascular diseases, phonocardiogram, transfer learning.

1. Introduction

Researchers have been making efforts to integrate AI into healthcare to enhance its availability and make it more effective and affordable. Several machine learning (ML) techniques are being used in various healthcare domains, such as medical imaging, drug discovery, clinical decision support systems, surgical robotics, precision medicine, and population health management, to improve patient care and reduce pain points in the healthcare system[1].

Cardiovascular disease (CVD) is a major cause of death worldwide. CVDs affect around 500 million individuals worldwide, resulting in 20.5 million deaths in 2021, as per the report of the World Heart Foundation in 2023. Early identification and treatment can prevent approximately 80% of strokes and heart attacks [2]. Congenital heart disease (CHD) affects approximately one in hundred children worldwide [3]. Approximately 10% of children diagnosed with CHD in less-developed countries have access to appropriate healthcare.

Significant progress has been made in cardiovascular medicine over the last five decades. Advancements in techniques have been made and the availability of tools has increased which has improved cardiovascular health. However, these resources are mainly found in affluent nations [4]. This disparity must be addressed through the integration of Artificial Intelligence (AI) into healthcare systems to ensure that cardiac care and overall healthcare services are accessible to underserved populations.

Clinicians have relied on auscultation, that is, listening to sounds emanating from the heart to detect heart diseases. However, this method requires training and is thus subjective. To overcome these limitations, researchers have turned to the application of phonocardiograms (PCG), which provide a visual representation of heart sounds during normal operations [5]. Researchers began using digital signal processing (DSP) to analyse PCGs to detect heart disease more than two decades ago. Researchers have applied deep learning (DL) and ML techniques to implement a heart sound classification algorithm to detect heart diseases by analysing the recordings of heart sounds during the last two decades.

ML techniques primarily depend on manual feature extraction, which is affected by noise, data variability, and limited generalisation. The ability to accurately differentiate between specific heart diseases remains a challenge. DL methods were implemented to reduce preprocessing and eliminate the feature extraction phases.

Heart sounds

The heart is a muscular organ that pumps blood throughout the human body through a complex set of heart valves. The closing of the atrioventricular (AV) and semilunar valves produces the sounds “lub” and “dub”, called S1 and S2, respectively [6]. In addition, two additional sounds, S3 and S4, may occur. The third sound, S3, is characterised by a low pitch and is usually observed in children and young individuals. Its occurrence in adults may indicate cardiac issues. S4 is a low-pitched sound produced during the presystolic phase resulting from atrial contraction immediately before ventricular contraction. S4 serves as an indicator of cardiovascular disease (CVD) [7]. An illustration of an abnormal heart sound through a PCG is shown in Figure 1.

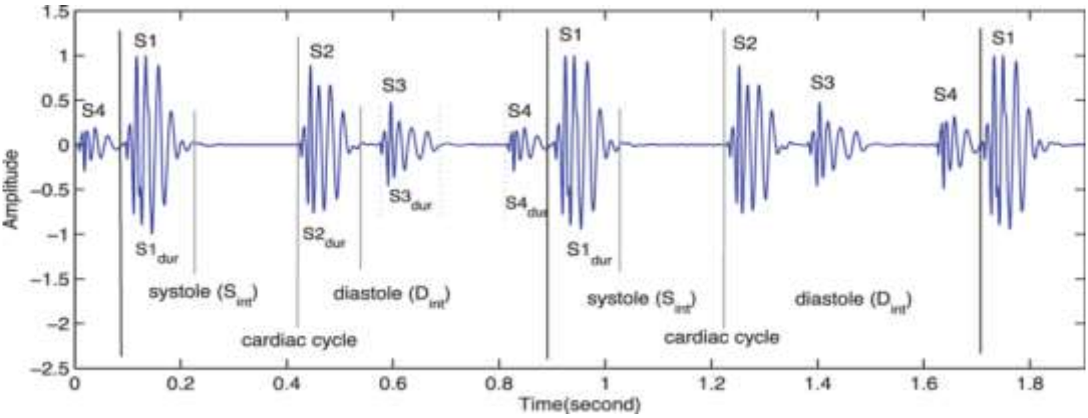

Figure 1: PCG signal depicting fundamental heart sounds S1 and S2 and additional heart sounds S3 and S4. Ref.: Bassiouni et al. [8]

Heart murmurs are abnormal sounds produced by turbulence in blood flow throughout the cardiac cycle. These murmurs can be categorised as either innocent or pathological, with innocent murmurs stemming from physiological factors and pathological murmurs indicating severe heart disease. Pathological murmurs typically result from structural abnormalities within the heart and serve as an indication of a critical cardiac condition [9].

2. Background

2.1 Transfer Learning

The use of the knowledge obtained from learning one task to learn a related task is called transfer learning. In ML, transfer learning is used to train a model using a model pretrained on a related task. Transfer learning is particularly useful in domains in which data availability is limited. Medical tasks can be handled efficiently through transfer learning [10]. A block diagram of the transfer learning model is shown in Figure 2.

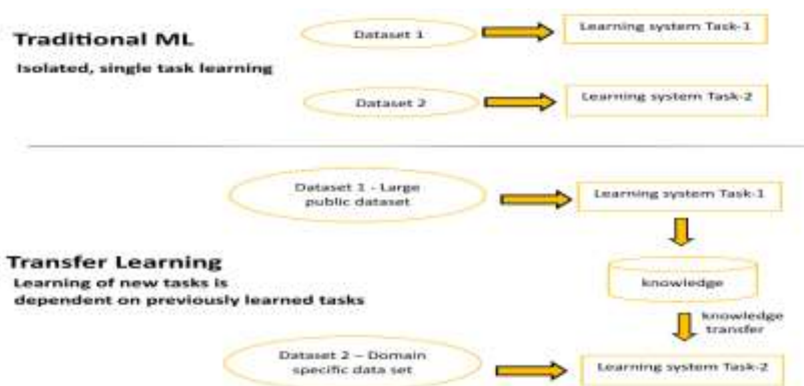

Figure 2: Difference between traditional ML and transfer learning [11]

2.2 YAMNet

YAMNet is an abbreviation for Yet Another Mobile Network, an already-trained CNN created by Google Research from Google AudioSet Ontology, specifically for classifying audio. YAMNet can predict various audio events from a selection of 521 classes, such as laughter, barking, human sounds, or sirens, making it a valuable tool for transfer learning in the audio classification field [12]. By using its pre-existing knowledge, YAMNet can reduce the training time and resource requirements, as well as boost the efficiency and accuracy of new tasks, even when working with limited datasets. The computational efficiency and adaptability of YAMNet make it possible to use it in real time with limited computational power and be embedded in devices in a variety of areas, including monitoring of the environment and analysis of multimedia content.

The network includes layers that have been trained to extract significant features from audio data and capture the key aspects of audio content. These features represent learned information about audio in a general sense rather than just the specific categories on which it was trained. YAMNet's design involves feature extraction and a fresh classification layer that is constructed on top of the extracted features. This methodology enables streamlined training on smaller datasets specific to specific classification tasks.

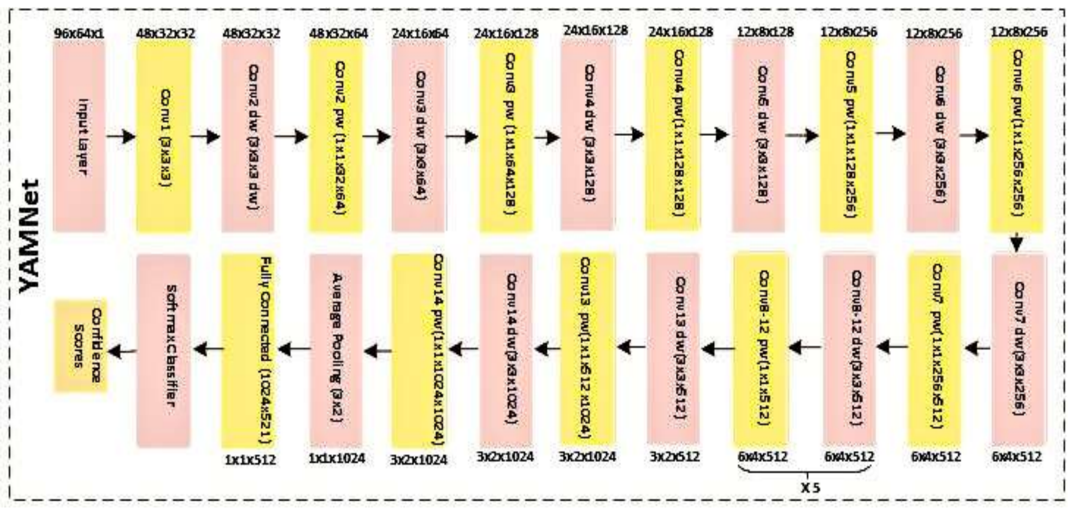

Figure 3: YAMNet architecture

YAMNet was constructed based on the architecture of MobileNet [13], a framework devised to accommodate low-latency AI models in mobile applications, where computational resources are typically limited. In MobileNet structures, depth-wise separable convolution replaces standard convolution. This change led to a decrease in the computational costs by up to nine times that of the standard convolution. YAMNet consists of an input layer, 27 convolutional layers, global average pooling layer, fully connected layer, and output layers. Standard convolution and depth-wise separable convolutions are stacked sequentially up to the pooling layer. The convolutional layers use the ReLU function and batch normalisation processes. Finally, the output layer delivers sound class prediction using a softmax function [14]. Figure 3 depicts the architecture of YAMNet [12].

The number of parameters in AlexNet, which uses standard convolution, totals 61 million across eight layers, compared with YAMNet which contains approximately 3.7 million parameters spread over 30 layers. This is because of the impact of depth-wise convolutions, and fewer learnable weights in such architectures reduce the risk of overfitting [14].

Mel-spectrogram features: YAMNet incorporates a Mel-spectrogram feature map. Humans perceive sound nonlinearly in terms of frequency content. Humans are better at perceiving pitch at lower frequency ranges, as opposed to higher ones, where tones might sound similar. The Mel scale, derived from the word "Melody", and the Mel Spectrogram are intended to mimic the logarithmic nature of human hearing. The calculation of the Mel scale is constructed on this perception of pitch by humans, allowing for the conversion of frequencies f into Mel scale as given in Equation (1).

$$\mathrm{Mel}(f) = 2595 \log(1 + f / 100) \qquad - (1)$$

3. Related Work

Sathyanarayanan et al. [15] surveyed the use of ML techniques to detect heart diseases. Dey et al. [15] used discrete wavelet transform (DWT) on spectrograms to classify heart sounds as normal or abnormal. Khan et al. [16] proposed an automatic residual neural network model [16][16][16][16][16][16][16][16][16][16][16][16][16][16][16] using power spectrograms of PCG audio samples as inputs for the diagnosis of multiple crucial heart disorders. Baghel et al. [17] proposed a CNN architecture for multiclass

[17][17][17][17][17][17][17][17][17]classification of cardiac diseases. Shabbir et al. [18] classified heart murmurs using CNNs trained on different signal representations such as spectrograms and mel-frequency cepstral coefficients (MFCC) of PCGs. Deep CNNs were combined by Carter et al. [19] to extract temporal signatures in heart recordings, enabling multilabel classification and severity determination and incorporating explainable AI algorithms. Abubakar et al. [20] proposed a hybrid CNN model for diagnosing heart conditions by analysing heart sound signals to classify heart sounds into three classes: murmurs, extrasystole and normal. Feng Li et al. [21] used improved MFCCs and ResNet to get an accuracy of 94.43%. Sathyanarayan et al. used both CNN and transfer learning [22] and the use of audio features extracted from the heart sounds [23] and got excellent results.

Koike et al. [24] built a pretrained audio neural network (PANN), a transfer learning network built by pretraining on a large amount of audio data, compared to other transfer learning networks that are built on image data, to classify the PhysioNet CinC heart sounds dataset [25] and reported an accuracy of 89.7%. Unsegmented PCGs were used by Langley and Murray to classify heart sounds using classification trees with combined wavelet entropy and spectral entropy features.

A combination of textural features extracted from spectrogram and chromagram representations of heart sound signals was used by Taneja et al. [26] to obtain an accuracy of

94.87%. The frequency domain, time domain, and statistical features of heart sounds were extracted by Milani et al. [27] and input to LDA and ANN to achieve an accuracy of up to 93.33% on the PhysioNet dataset. Tao Li et al. [28] implemented a lightweight two-dimensional CNN model with frequency domain features after segmentation as input and reported an accuracy of up to 86%.

These studies highlight the depth and breadth of research in the field of heart sound analysis using AI. These represent ongoing efforts to refine AI techniques for better accuracy, efficiency, and applicability, contributing significantly to advancements in CVD diagnostics and treatment.

Several methods have been implemented by researchers to detect heart diseases using AI to analyse heart sounds. However, these methods have not yielded a satisfactory level of accuracy, require more computational resources, and have low confidence to implement them in an actual system. This study attempts to bridge this research gap by building a training model that incorporates audio transfer learning.

4. Dataset Details

The dataset utilised in this study comprised 1000 audio samples gathered and processed by Yaseen et al., who collected heart sounds from books, CDs, and websites. The dataset consisted of 200 normal heart sounds and 800 abnormal heart sounds, comprising 200 heart sounds from each of the four valvular diseases. The audio data were filtered and converted into mono-channel. Each sample in the dataset contained one channel of 16 bits. The audio samples were recorded at a rate of 8000 Hz and encoded at 128 kbps. The duration of the audio recordings ranged from 1 to 3 s, with the majority lasting 2 s. To ensure consistency in the training data, only two seconds of each audio clip were considered, excluding those shorter than two seconds.

Table 1: Details of the Yaseen dataset

Class	Type	Number of samples
Normal (N)	Normal	200
Aortic stenosis (AS)	Abnormal	200
Mitral regurgitation (MR)		184
Mitral stenosis (MS)		186
Mitral valve prolapse (MVP)		187
Total		957

To preprocess the audio data, audio samples shorter than two seconds were eliminated, and only the first two seconds of any audio sample longer than two seconds were taken into consideration. This was performed to prevent any errors arising from the nonuniformity of the training data. To reduce the variations in the frequency range of the various sound types, the audio samples were normalised. The performance was evaluated after performing experiments with all combinations.

5. Methodology

Algorithm for building a model for heart sound signals classification model called HSANet by building a customised neural network incorporating audio transfer learning and CNN

Input: Heart sounds dataset

Output: Heart sound signals classification model

Steps:

1) Acquire the heart sounds dataset and preprocess the audio data
2) Generate spectrograms from audio dataset.
3) Give input to YAMNet audio transfer learning layer for processing
4) Training and performance evaluation of the model based on various ML metrics

Figure 4: Steps in the classification system using a customised neural network named HSANet using audio transfer learning

The algorithm used in this methodology is shown in Figure 4. A pretrained audio transfer learning layer was incorporated to reduce the training time. The experiment involved two combinations of the datasets. In the first combination, the data were divided into two classes: normal and abnormal, that is, 200 normal heart sounds and 757 abnormal heart sounds. In the second combination, audio samples were categorised into five classes: normal and four distinct types of abnormal samples representing mitral stenosis (MS), mitral valve prolapse (MVP), aortic stenosis (AS), and mitral regurgitation (MR). Each abnormal category contained up to 200 audio samples.

The audio dataset was converted to spectrograms after preprocessing and input to the YAMNet layer for training. The performance was evaluated based on various ML metrics. The model architecture consists of an input layer, YAMNet layer, and output layer.

6. Results and Discussion

The customised neural network with an audio transfer learning layer provided validation accuracies of 99.71% and 90.86% for five- and two-class combinations, respectively. The best results were obtained with the Adam optimiser, with a learning rate of 0.0001, epochs set to 95, and mini-batch size set to 128. The model details are listed in Table 2. The training and validation accuracies and training loss for the 2-class classification are shown in Figure 3. Figure 4 shows the same for the 5-class classification.

Table 2: Details of the customised neural network called HSANet

Pretrained audio network	**No. of layers**	**Solver**	**Initial learn rate**	**Mini batch size**	**Max epochs**	**Validation accuracy for 2 Class**	**Validation accuracy for 5 Class**
HSANet	86	Adam	0.0001	128	95	**90.86%**	**99.71%**

The proposed model provided higher accuracy for the 5-class classification compared for the binary classification. This is counterintuitive, and can be attributed to several factors. Each class in the 5-class classification had approximately 200 samples per class; hence, the distribution was balanced. This is in contrast to the binary classifier, which deals with class imbalance. The balanced dataset in the 5-class classification may be responsible for the improvement in the ability of the model to learn differentiating features in heart sound recordings. The YAMNet pretrained layer might be good at learning slight differences in the heart sounds that differentiate between the four classes. These differences may be less important in differentiating between abnormal and normal heart sounds. However, additional details can help distinguish between specific valvular diseases. Another reason is the possibility of overfitting the binary classification task. However, the 5-class classification model is forced to learn generalisable features to differentiate between a wider range of heart sounds. In addition, the features extracted by the YAMNet layer may be more interpretable or better aligned with the characteristics of valvular diseases, making it suitable for multiclass classifications.

The accuracy of 90.86% for the binary classification task is better than that reported by Li et al. [28]. Liu et al. used a 2D-CNN on the PhysioNet CinC dataset. The results are also better than those reported by Soto-Murillo et al. [29], who used logistic regression with audio features; Demir et al. [30], who used a 2D-CNN with spectrograms; and Raza et al. [31], who used a time series and LSTM. The lower accuracy may be due to an imbalance in data and minute differences between normal and abnormal categories in several cases.

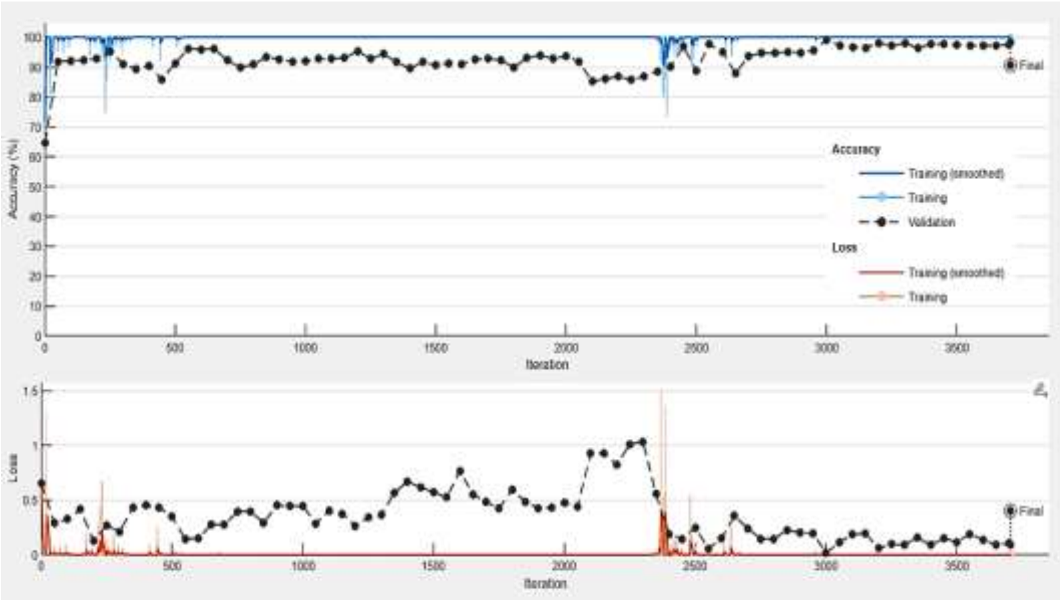

Figure 3: Training and validation accuracy for 2-class classification

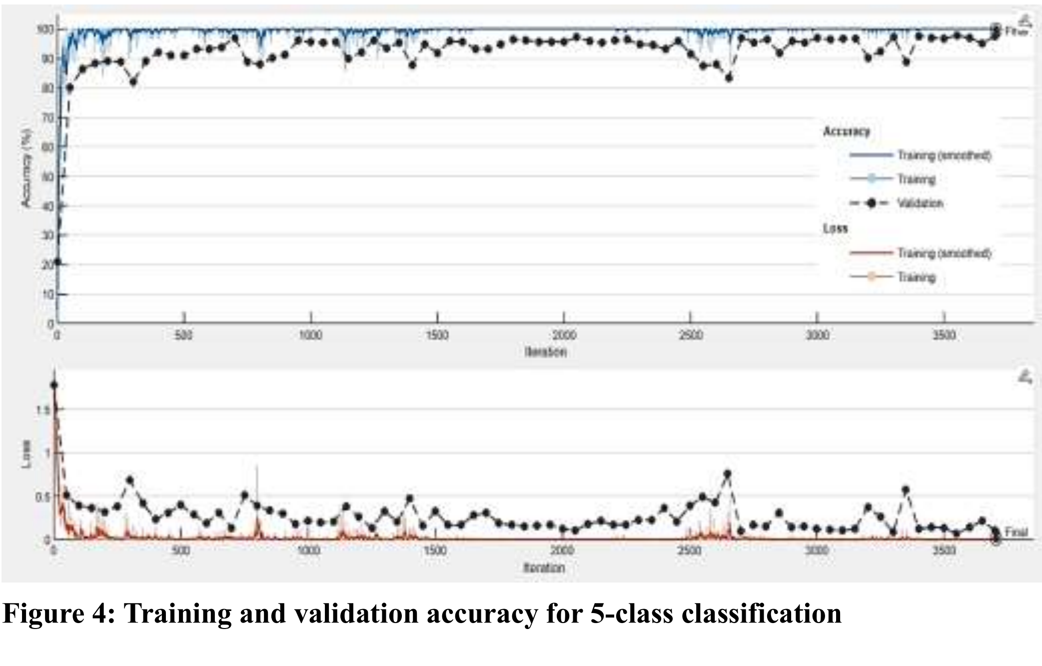

Figure 4: Training and validation accuracy for 5-class classification

The audio transfer learning techniques applied here achieved superior results for 5-class classification compared to several others including Chui et al. [32] who used WaveNet architecture, Upretee et al. [33] who used audio features and Baghel et al. [17] who used time series and CNN. This demonstrates the capability of the proposed model to classify valvular diseases with a high accuracy.

The results obtained by other researchers are compared with the results of the proposed model in Table 2.

Table 2: Comparison of results

Reference	Author(s)	Methodology	No. of classes	Dataset used	Accuracy (%)
[32]	Chui et al.	WaveNet architecture and 10-fold validation	5	Yaseen dataset	97
[33]	Upretee et al.	Spectral centroid frequency with kNN and SVM	5	Yaseen dataset	96.50
[17]	Baghel et al.	Time series and 1D-CNN	5	Yaseen dataset	96.2
Proposed method		**HSANet – a customised NN with audio transfer learning**	**5**	Yaseen dataset	**99.71**
[28]	T. Li, & Yin et al.	Frequency domain features and 2D-CNN	2	Yaseen dataset	86
[24]	Koike et al.	Pretrained audio neural network	2	PhysioNet 2016	89.7
[29]	Soto-Murillo et al.	Audio features and logistic regression	3	Pascal dataset	80.49
[30]	Demir et al.	Spectrograms and 2D-CNN	3	Pascal dataset	80
[34]	Langley and Murray	Classification tree with unsegmented PCGs and combined features	2	PhysioNet 2016	79
[31]	Raza et al.	Time series and LSTM	3	Pascal dataset	80.8
[28]	T. Li, & Yin et al.	Frequency domain features and 2D-CNN	2	PhysioNet 2016	86
Proposed method		**HSANet – a customised NN with audio transfer learning**	**2**	PhysioNet 2016	**90.86%**

7. Conclusion

In this study, an innovative method using a pretrained audio transfer learning layer was developed. The study integrated the YAMNet pretrained audio layer into a specialised neural network called the Heart Sounds Analysis Network (HSANet). The technique showed good performance on a dataset with 957 heart sound recordings, including one normal case and four valvular heart diseases: normal, mitral valve prolapse, mitral stenosis, aortic stenosis, and mitral regurgitation.

The binary classification task aimed at distinguishing normal cases from abnormal cases achieved an accuracy of 90.86%. The proposed method performed excellently in a five-class classification task and achieved an accuracy of 99.71%. Notably, the model produced positive results, even when dealing with imbalanced data.

The advantages of audio transfer learning include increased efficiency as pretrained models are reused which results in saving time and computational resources when training on new datasets. The results of the new model were excellent, despite the limited number of datasets. Overfitting is reduced and the generalisation capabilities of the model also increase. Resource optimisation is appropriate for real-time applications.

The proposed approach is novel. These findings indicate that the proposed method has immense potential as a valuable tool for the early detection and diagnosis of valvular diseases. It has the potential for real-time implementation by integration into a device such as an electronic stethoscope. The accuracy and non-invasiveness of the proposed method make it suitable for various settings such as primary healthcare clinics and rural communities.

It should be noted that audio transfer learning models are adaptable and can be used for different applications such as environmental monitoring and multimedia content analysis.

Limitations and Future Work: The availability of data in the medical domain has always been an issue for researchers. The model can be improved by training with more samples in each category. The model can also be enhanced by adding more types of diseases that it can detect and classify. Further customisation of the network can be attempted to increase accuracy and reduce the amount of computation required. The model can be modified to use other pretrained audio transfer learning networks to improve its performance. Explainable AI features can increase the user acceptance.

The integration of AI in cardiovascular care and healthcare in general is a way forward to make healthcare more accessible, detect diseases early and quicker, reduce costs, and make the treatment more effective.

Funding. This study was supported by a grant from the Sri Sathya Sai University for Human Excellence, Gulbarga, Karnataka, India.

Acknowledgement. We would like to thank Mrs. Shobha Sathyanarayanan for her assistance.

Conflict of Interest. The authors declare that they have no conflicts of interest.

8. References

[1] S. Sathyanarayanan and S. Chitnis, “A Survey of Machine Learning in Healthcare,” in Artificial Intelligence Applications for Health Care, I., M. K. Ahirwal, N. D. Londhe, and A. Kumar, Eds. Boca Raton: Taylor & Francis Ltd, 2022, pp. 1–22.

[2] World Heart Federation, “World Heart Report 2023: Confronting the World’s Number One Killer,” Geneva, 2023. [Online]. Available: https://world-heart-federation.org/wp-content/uploads/World-Heart-Report-2023.pdf.
[3] K. W. Giang et al., “Congenital heart disease: changes in recorded birth prevalence and cardiac interventions over the past half-century in Sweden,” European Journal of Preventive Cardiology, vol. 30, no. 2, pp. 169–176, 2023, doi: 10.1093/eurjpc/zwac227.
[4] Z. Su et al., “Global, regional, and national time trends in mortality for congenital heart disease, 1990–2019: An age-period-cohort analysis for the Global Burden of Disease 2019 study,” eClinicalMedicine, vol. 43, no. January, p. 101249, 2022, doi: 10.1016/j.eclinm.2021.101249.
[5] P. Lubaib, A. M. Kv, and A. R. V, “PHONOCARDIOGRAM BASED DIAGNOSTIC SYSTEM,” International Journal of Biomedical Engineering and Science, vol. 2, no. 3, pp. 1–10, 2015.
[6] C. Liu et al., “An open access database for the evaluation of heart sound algorithms,” Physiological Measurement, vol. 37, no. 12, pp. 2181–2213, 2016, doi: 10.1088/0967-3334/37/12/2181.
[7] Yaseen, G. Y. Son, and S. Kwon, “Classification of heart sound signal using multiple features,” Applied Sciences (Switzerland), vol. 8, no. 12, 2018, doi: 10.3390/app8122344.
[8] M. Bassiouni, W. Khalifa, E. S. A. El Dahshan, and A. B. M. Salam, “A study on PCG as a biometric approach,” 2015 IEEE 7th International Conference on Intelligent Computing and Information Systems, ICICIS 2015, no. December, pp. 161–166, 2016, doi: 10.1109/IntelCIS.2015.7397215.
[9] S. Swaminathan, S. Murthy, S. K. Mallappa, and C. Gudada, “Machine Learning Approach Using HOG and LBP Features of Spectrograms-based Heart Sounds Analysis for the Detection of Heart Diseases,” in 15th Soft Computing and Pattern Recognition (SoCPaR’23), 2023, pp. 1–11.
[10] P. Kora et al., “Transfer learning techniques for medical image analysis: A review,” Biocybernetics and Biomedical Engineering, vol. 42, no. 1, pp. 79–107, 2022, doi: 10.1016/j.bbe.2021.11.004.
[11] R. Ribani and M. Marengoni, “A Survey of Transfer Learning for Convolutional Neural Networks,” in Proceedings - 32nd Conference on Graphics, Patterns and Images Tutorials, SIBGRAPI-T 2019, 2019, no. February, pp. 47–57, doi: 10.1109/SIBGRAPI-T.2019.00010.
[12] R. Mahum, A. Irtaza, and A. Javed, “EDL-Det: A Robust TTS Synthesis Detector Using VGG19-Based YAMNet and Ensemble Learning Block,” IEEE Access, vol. 11, pp. 134701–134716, 2023, doi: 10.1109/ACCESS.2023.3332561.
[13] S. Parvaneh, Z. Ardalan, J. Song, K. Vyas, and C. Potes, “Heart Murmur Detection Using Ensemble of Deep Learning Classifiers for Phonocardiograms Recorded from Multiple Auscultation Locations,” Computing in Cardiology, vol. 2022-Septe, no. March 2023, 2022, doi: 10.22489/CinC.2022.241.
[14] E. Brusa, C. Delprete, and L. G. Di Maggio, “Deep transfer learning for machine diagnosis: From sound and music recognition to bearing fault detection,” Applied Sciences (Switzerland), vol. 11, no. 24, 2021, doi: 10.3390/app112411663.
[15] S. Sathyanarayanan, Srikanta Murthy K., and S. Chitnis, “A Comprehensive Survey of Analysis of Heart Sounds using Machine Learning Techniques to Detect Heart Diseases,” The Journal of Population Therapeutics and Clinical Pharmacology, vol. 30, no. May, pp. 375–384, 2023, doi: 10.47750/jptcp.2023.30.11.038.
[16] J. S. Khan, M. Kaushik, A. Chaurasia, M. K. Dutta, and R. Burget, “Computer Methods and Programs in Biomedicine Cardi-Net : A deep neural network for classification of cardiac disease using phonocardiogram signal,” Computer Methods and Programs in Biomedicine, vol.

219, p. 106727, 2022, doi: 10.1016/j.cmpb.2022.106727.
[17] N. Baghel, M. K. Dutta, and R. Burget, "Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network," Computer Methods and Programs in Biomedicine, vol. 197, 2020, doi: 10.1016/j.cmpb.2020.105750.
[18] M. Shabbir, X. Liu, M. Nasseri, and S. Helgeson, "Heart Murmur Classification in Phonocardiogram Representations Using Convolutional Neural Networks," Proceedings of the International Florida Artificial Intelligence Research Society Conference, FLAIRS, vol. 36, no. Ml, pp. 2020–2022, 2023, doi: 10.32473/flairs.36.133189.
[19] T. S. Carter, G. H. Yang, G. Loke, and W. Yan, "Deciphering simultaneous heart conditions with spectrogram and explainable-AI approach," Biomedical Signal Processing and Control, vol. 85, p. 104990, 2023, doi: https://doi.org/10.1016/j.bspc.2023.104990.
[20] M. M. Abubakar and T. Tuncer, "Heart Sounds Classification Using Hybrid CNN Architecture," in 5th International Students Science Congress Proceedings, 2021, pp. 1–4, doi: 10.52460/issc.2021.023.
[21] F. Li, Z. Zhang, L. Wang, and W. Liu, "Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning," Frontiers in Physiology, vol. 13, no. December, pp. 1–16, 2022, doi: 10.3389/fphys.2022.1084420.
[22] S. Sathyanarayanan and S, Murthy K, "Heart Sound Analysis Using SAINet Incorporating CNN and Transfer Learning for Detecting Heart Diseases," Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), vol. 15, no. 2 (June), pp. 152–169, 2024, doi: 10.58346/JOWUA.2024.I2.011.
[23] S. Sathyanarayanan, S. Murthy K, C. Gudada, S. Mallappa, and N. Ail, "Heart Sound Analysis with Machine Learning Using Audio Features for Detecting Heart Diseases," International Journal of Computer Information Systems and Industrial Management Applications, vol. 16, no. 2, pp. 131–147, 2024, [Online]. Available: https://cspub-ijcisim.org/index.php/ijcisim/article/view/628.
[24] T. Koike, S. Member, K. Qian, Q. Kong, S. Member, and M. D. Plumbley, "Audio for Audio is Better ? An Investigation on Transfer Learning Models for Heart Sound Classification," pp. 31–34, 2020.
[25] C. Liu, D. Springer, and G. Clifford, "Performance of an open-source heart sound segmentation algorithm on eight independent databases," Physiological Measurement, vol. 38, pp. 1730–1745, Aug. 2017, doi: 10.1088/1361-6579/aa6e9f.
[26] K. Taneja, V. Arora, and K. Verma, "Classifying the heart sound signals using textural-based features for an efficient decision support system," Expert Systems, vol. 40, no. 6, p. e13246, Feb. 2023, doi: 10.1111/exsy.13246.
[27] M. G. M. Milani, P. E. Abas, L. C. De Silva, and N. D. Nanayakkara, "Abnormal heart sound classification using phonocardiography signals," Smart Health, vol. 21, no. February, p. 100194, 2021, doi: 10.1016/j.smhl.2021.100194.
[28] T. Li, Y. Yin, K. Ma, S. Zhang, and M. Liu, "Lightweight end-to-end neural network model for automatic heart sound classification," Information (Switzerland), vol. 12, no. 2, pp. 1–11, 2021, doi: 10.3390/info12020054.
[29] M. A. Soto-Murillo et al., "Automatic evaluation of heart condition according to the sounds emitted and implementing six classification methods," Healthcare (Switzerland), vol. 9, no. 3, pp. 1–15, 2021, doi: 10.3390/healthcare9030317.
[30] F. Demir, A. Şengür, V. Bajaj, and K. Polat, "Towards the classification of heart sounds based on convolutional deep neural network," Health Information Science and Systems, vol. 7, no. 1, 2019, doi: 10.1007/s13755-019-0078-0.
[31] A. Raza, A. Mehmood, S. Ullah, M. Ahmad, G. S. Choi, and B. W. On, "Heartbeat sound signal

classification using deep learning," Sensors (Switzerland), vol. 19, no. 21, pp. 1–15, 2019, doi: 10.3390/s19214819.

[32] S. L. Oh et al., "Classification of heart sound signals using a novel deep WaveNet model," Computer Methods and Programs in Biomedicine, vol. 196, no. June, 2020, doi: 10.1016/j.cmpb.2020.105604.

[33] P. Upretee and M. E. Yüksel, "Accurate classification of heart sounds for disease diagnosis by a single time-varying spectral feature: Preliminary results," 2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science, EBBT 2019, pp. 1–4, 2019, doi: 10.1109/EBBT.2019.8741730.

[34] P. Langley and A. Murray, "Heart sound classification from unsegmented phonocardiograms," Physiological Measurement, vol. 38, no. 8, p. 1658, 2017, doi: 10.1088/1361-6579/aa724c.

[1] S. Sathyanarayanan and S. Chitnis, "A Survey of Machine Learning in Healthcare," in Artificial Intelligence Applications for Health Care, I., M. K. Ahirwal, N. D. Londhe, and A. Kumar, Eds. Boca Raton: Taylor & Francis Ltd, 2022, pp. 1–22.

[2] World Heart Federation, "World Heart Report 2023: Confronting the World's Number One Killer," Geneva, 2023. [Online]. Available: https://world-heart-federation.org/wp-content/uploads/World-Heart-Report-2023.pdf.

[3] K. W. Giang et al., "Congenital heart disease: changes in recorded birth prevalence and cardiac interventions over the past half-century in Sweden," European Journal of Preventive Cardiology, vol. 30, no. 2, pp. 169–176, 2023, doi: 10.1093/eurjpc/zwac227.

[4] Z. Su et al., "Global, regional, and national time trends in mortality for congenital heart disease, 1990–2019: An age-period-cohort analysis for the Global Burden of Disease 2019 study," eClinicalMedicine, vol. 43, no. January, p. 101249, 2022, doi: 10.1016/j.eclinm.2021.101249.

[5] P. Lubaib, A. M. Kv, and A. R. V, "PHONOCARDIOGRAM BASED DIAGNOSTIC SYSTEM," International Journal of Biomedical Engineering and Science, vol. 2, no. 3, pp. 1–10, 2015.

[6] C. Liu et al., "An open access database for the evaluation of heart sound algorithms," Physiological Measurement, vol. 37, no. 12, pp. 2181–2213, 2016, doi: 10.1088/0967-3334/37/12/2181.

[7] Yaseen, G. Y. Son, and S. Kwon, "Classification of heart sound signal using multiple features," Applied Sciences (Switzerland), vol. 8, no. 12, 2018, doi: 10.3390/app8122344.

[8] M. Bassiouni, W. Khalifa, E. S. A. El Dahshan, and A. B. M. Salam, "A study on PCG as a biometric approach," 2015 IEEE 7th International Conference on Intelligent Computing and Information Systems, ICICIS 2015, no. December, pp. 161–166, 2016, doi: 10.1109/IntelCIS.2015.7397215.

[9] S. Swaminathan, S. Murthy, S. K. Mallappa, and C. Gudada, "Machine Learning Approach Using HOG and LBP Features of Spectrograms-based Heart Sounds Analysis for the Detection of Heart Diseases," in 15th Soft Computing and Pattern Recognition (SoCPaR'23), 2023, pp. 1–11.

[10] P. Kora et al., "Transfer learning techniques for medical image analysis: A review," Biocybernetics and Biomedical Engineering, vol. 42, no. 1, pp. 79–107, 2022, doi: 10.1016/j.bbe.2021.11.004.

[11] R. Ribani and M. Marengoni, "A Survey of Transfer Learning for Convolutional Neural Networks," in Proceedings - 32nd Conference on Graphics, Patterns and Images Tutorials, SIBGRAPI-T 2019, 2019, no. February, pp. 47–57, doi: 10.1109/SIBGRAPI-T.2019.00010.

[12] R. Mahum, A. Irtaza, and A. Javed, "EDL-Det: A Robust TTS Synthesis Detector Using VGG19-Based YAMNet and Ensemble Learning Block," IEEE Access, vol. 11, pp. 134701–134716, 2023, doi: 10.1109/ACCESS.2023.3332561.

[13] S. Parvaneh, Z. Ardalan, J. Song, K. Vyas, and C. Potes, “Heart Murmur Detection Using Ensemble of Deep Learning Classifiers for Phonocardiograms Recorded from Multiple Auscultation Locations,” Computing in Cardiology, vol. 2022-Septe, no. March 2023, 2022, doi: 10.22489/CinC.2022.241.

[14] E. Brusa, C. Delprete, and L. G. Di Maggio, “Deep transfer learning for machine diagnosis: From sound and music recognition to bearing fault detection,” Applied Sciences (Switzerland), vol. 11, no. 24, 2021, doi: 10.3390/app112411663.

[15] S. Sathyanarayanan, Srikanta Murthy K., and S. Chitnis, “A Comprehensive Survey of Analysis of Heart Sounds using Machine Learning Techniques to Detect Heart Diseases,” The Journal of Population Therapeutics and Clinical Pharmacology, vol. 30, no. May, pp. 375–384, 2023, doi: 10.47750/jptcp.2023.30.11.038.

[16] J. S. Khan, M. Kaushik, A. Chaurasia, M. K. Dutta, and R. Burget, “Computer Methods and Programs in Biomedicine Cardi-Net : A deep neural network for classification of cardiac disease using phonocardiogram signal,” Computer Methods and Programs in Biomedicine, vol. 219, p. 106727, 2022, doi: 10.1016/j.cmpb.2022.106727.

[17] N. Baghel, M. K. Dutta, and R. Burget, “Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network,” Computer Methods and Programs in Biomedicine, vol. 197, 2020, doi: 10.1016/j.cmpb.2020.105750.

[18] M. Shabbir, X. Liu, M. Nasseri, and S. Helgeson, “Heart Murmur Classification in Phonocardiogram Representations Using Convolutional Neural Networks,” Proceedings of the International Florida Artificial Intelligence Research Society Conference, FLAIRS, vol. 36, no. Ml, pp. 2020–2022, 2023, doi: 10.32473/flairs.36.133189.

[19] T. S. Carter, G. H. Yang, G. Loke, and W. Yan, “Deciphering simultaneous heart conditions with spectrogram and explainable-AI approach,” Biomedical Signal Processing and Control, vol. 85, p. 104990, 2023, doi: https://doi.org/10.1016/j.bspc.2023.104990.

[20] M. M. Abubakar and T. Tuncer, “Heart Sounds Classification Using Hybrid CNN Architecture,” in 5th International Students Science Congress Proceedings, 2021, pp. 1–4, doi: 10.52460/issc.2021.023.

[21] F. Li, Z. Zhang, L. Wang, and W. Liu, “Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning,” Frontiers in Physiology, vol. 13, no. December, pp. 1–16, 2022, doi: 10.3389/fphys.2022.1084420.

[22] S. Sathyanarayanan and K. S. Murthy, “Heart Sound Analysis Using SAINet Incorporating CNN and Transfer Learning for Detecting Heart Diseases,” Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), vol. 15, no. 2 (June), pp. 152–169, 2024, doi: 10.58346/JOWUA.2024.I2.011.

[23] S. Sathyanarayanan, S. Murthy K, C. Gudada, S. Mallappa, and N. Ail, “Heart Sound Analysis with Machine Learning Using Audio Features for Detecting Heart Diseases,” International Journal of Computer Information Systems and Industrial Management Applications, vol. 16, no. 2, pp. 131–147, 2024, [Online]. Available: https://cspub-ijcisim.org/index.php/ijcisim/article/view/628.

[24] T. Koike, S. Member, K. Qian, Q. Kong, S. Member, and M. D. Plumbley, “Audio for Audio is Better ? An Investigation on Transfer Learning Models for Heart Sound Classification,” pp. 31–34, 2020.

[25] C. Liu, D. Springer, and G. Clifford, “Performance of an open-source heart sound segmentation algorithm on eight independent databases,” Physiological Measurement, vol. 38, pp. 1730–1745, Aug. 2017, doi: 10.1088/1361-6579/aa6e9f.

[26] K. Taneja, V. Arora, and K. Verma, “Classifying the heart sound signals using textural-based features for an efficient decision support system,” Expert Systems, vol. 40, no. 6, p. e13246,

Feb. 2023, doi: 10.1111/exsy.13246.
[27] M. G. M. Milani, P. E. Abas, L. C. De Silva, and N. D. Nanayakkara, “Abnormal heart sound classification using phonocardiography signals,” Smart Health, vol. 21, no. February, p. 100194, 2021, doi: 10.1016/j.smhl.2021.100194.
[28] T. Li, Y. Yin, K. Ma, S. Zhang, and M. Liu, “Lightweight end-to-end neural network model for automatic heart sound classification,” Information (Switzerland), vol. 12, no. 2, pp. 1–11, 2021, doi: 10.3390/info12020054.
[29] M. A. Soto-Murillo et al., “Automatic evaluation of heart condition according to the sounds emitted and implementing six classification methods,” Healthcare (Switzerland), vol. 9, no. 3, pp. 1–15, 2021, doi: 10.3390/healthcare9030317.
[30] F. Demir, A. Şengür, V. Bajaj, and K. Polat, “Towards the classification of heart sounds based on convolutional deep neural network,” Health Information Science and Systems, vol. 7, no. 1, 2019, doi: 10.1007/s13755-019-0078-0.
[31] A. Raza, A. Mehmood, S. Ullah, M. Ahmad, G. S. Choi, and B. W. On, “Heartbeat sound signal classification using deep learning,” Sensors (Switzerland), vol. 19, no. 21, pp. 1–15, 2019, doi: 10.3390/s19214819.
[32] S. L. Oh et al., “Classification of heart sound signals using a novel deep WaveNet model,” Computer Methods and Programs in Biomedicine, vol. 196, no. June, 2020, doi: 10.1016/j.cmpb.2020.105604.
[33] P. Upretee and M. E. Yüksel, “Accurate classification of heart sounds for disease diagnosis by a single time-varying spectral feature: Preliminary results,” 2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science, EBBT 2019, pp. 1–4, 2019, doi: 10.1109/EBBT.2019.8741730.
[34] P. Langley and A. Murray, “Heart sound classification from unsegmented phonocardiograms,” Physiological Measurement, vol. 38, no. 8, p. 1658, 2017, doi: 10.1088/1361-6579/aa724c.