CNN-Driven Detection of Abnormalities in PCG Signals Using Gammatonegram Analysis

Sireesha Moturi, Sumanth Tata, Somnath Katragadda, Venkata Pavan Kumar Laghumavarapu,
Brahmaiah Lingala, Dodda Venkata Reddy
Department of Computer Science and Engineering
Narasaraopeta Engineering College, Narasaraopet-522601, Palnadu, Andhra Pradesh, India
sireeshamoturi@gmail.com, tatasumanth2003@gmail.com, katragaddasomnath22@gmail.com,
pavanklaghumavarapu@gmail.com, lingalabrahmaiah399@gmail.com, doddavenkatareddy@gmail.com

2024 First International Conference for Women in Computing (InCoWoCo) | 979-8-3315-1894-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/INCOWOCO64194.2024.10863151

***Abstract*—Phonocardiogram (PCG) signals carry essential data about heart health, offering significant potential for early disease detection. This study presents a method to identify heart abnormalities using Gammatonegram images derived from PCG signals, which are analyzed with a Convolutional Neural Network (CNN). The Gammatonegram, a time-frequency representation mimicking human auditory processing, captures both temporal and spectral features of PCG signals. The images are preprocessed through resizing and data augmentation, the images are input into the CNN model for classification. The model distinguishes between healthy and unhealthy heart sounds, and its performance is evaluated using several optimizers—SGD, Adagrad, Adadelta, RMSprop, and Adam. Adam yielded the best performance, achieving a test accuracy of 100%. This method provides a robust, non-invasive solution for heart disease diagnosis, leveraging deep learning to improve accuracy in detecting heart abnormalities from PCG signals.**

***Index Terms*—PCG signals, Gammatonegram images , Proposed model CNN, Classification, Optimizers.**

I. Introduction

The heart is one of the major organs highly involved in sustaining life by pumping blood throughout the body. It is responsible for ensuring oxygen and other essential nutrients reach tissues while at the same time facilitating the removal of carbon dioxide and other waste products. Heart problems, also referred to as cardiovascular disorders or heart diseases, are many conditions that may involve changes in the structure, function, or electrical conduction of the heart. Such conditions can lead to serious health concerns; therefore, early diagnosis and treatment are of the essence for any good outcomes, preventing major complications and improving results among patients. These abnormalities can arise due to various factors, including genetic predisposition, lifestyle choices, infections, or underlying medical conditions. Common heart problems include arrhythmias (irregular heartbeats), coronary artery disease (narrowing or blockage of the heart's arteries), heart valve issues, cardiomyopathies (diseases of the heart muscle), and congenital heart defects (structure issues present at birth) [1]. These conditions can lead to symptoms such as chest pain, shortness of breath, palpitations, fatigue, and in severe cases, heart failure or sudden cardiac arrest.

Heart diseases are still one of the top causes of death throughout the world, and there is an urgent need to develop trustful and non-invasive methods to diagnose them [2]. Early detection and diagnosis of heart conditions are essential for good care and treatment. Advances in medical technology, such as electrocardiograms (ECGs), echocardiograms, and cardiac imaging techniques, have significantly improved the ability to detect heart abnormalities [3].

Deep learning has transformed many fields, including medical diagnostics. This is because CNNs are effective for image classification tasks as they have the ability to learn features automatically from data. The CNN can examine images, including X-rays, MRI scans, and in this case, Gammatonegram representations of PCG signals, during a medical diagnosis. Phonocardiogram, or PCG, signals, which record the sounds of the heart, serve as a useful modality toward the early detection of heart conditions. The analysis of these signals can provide critical insights into cardiac health, enabling timely intervention and treatment [4].

Traditional methods of abnormality detection in PCG signals generally include two major steps: feature extraction and classification [5]. The Gammatonegram, as in this work, is used here as one of the major techniques for feature extraction from PCG signals. This Gammatonegram represents the time-frequency features of a signal that will most closely match how the human ear actually responds to sound. By converting PCG signals into Gammatonegram images, we can capture both temporal and spectral information that is crucial for distinguishing between healthy and abnormal heart sounds. Compared to other methods similar to the Mel-frequency cepstral coefficients (MFCCs) or Short-Time Fourier Transform (STFT) [6], Gammatonegram provides a more physiologically relevant analysis, especially for biomedical signals like PCG. By converting PCG signals into Gammatonegram images, the intricate patterns of heart sounds are captured effectively, allowing the CNN model to learn from these rich representations [7]. This approach enhances the model's ability to detect abnormalities with higher accuracy, leveraging the detailed auditory features embedded within the Gammatonegram images.

The domain of this research falls within the fields of Biomedical Signal Processing and Medical Diagnostics, with a focus on Cardiovascular Health. To put it precisely, deep learning techniques, particularly those on Convolutional Neural Networks, which have been applied in the analysis and classification of Phonocardiogram signals, the audio recordings of heart sounds. The project also incorporates aspects of Time-Frequency Analysis, leveraging Gammatonegram representations to enhance the detection of abnormalities in heart signals. This interdisciplinary domain blends elements of healthcare, signal processing, and artificial intelligence [8].

The technology applied for this is Deep Learning, specifically a variation called the Convolutional Neural Network. CNNs are useful in such tasks as classifying images and as such are highly suitable for Gammatonegram pictures looking at PCG signals [9]. The Gammatonegram, a time-frequency representation of audio signals, is computed using signal processing techniques, providing a rich feature set that repeat human auditory perception. The application is developed with the Python programming language, which is quite robust in handling machine learning and deep learning. Important libraries include TensorFlow or PyTorch for the building and training of the CNN model. These libraries offer the tools necessary to design the neural network, handle the data, and improve the model's performance. NumPy and Pandas are used to handle and prepare data efficiently. Matplotlib or Seaborn helps in showing data and results, and OpenCV is also there, with GPU acceleration making the calculations faster. Jupyter Notebooks offer the place to work on this complete and modern way of analyzing heart sounds.

At the end of the research, this application offers a powerful tool for transforming cardiovascular healthcare delivery, particularly in outside environments [10]. By helping to find problems early, making tests more available, and working with telemedicine and portable devices, it can greatly improve public health, especially in areas with few healthcare services. This progress could save lives, lower healthcare costs, and make life better for many people around the world.

II. Literature Study

In recent years, heart sound analysis has focused on addressing imbalanced signals in cardiovascular diagnostics. A significant study by ASHINIKUMAR SINGH et al. demonstrated the effectiveness of ensemble learning with CNNs and Gammatonegram images for predicting heart sounds. The Turkish Journal has published their research, which attained an accuracy rate of 99.51% on the PhysioNet 2016 dataset, that proves advanced deep learning techniques can contribute to the betterment of heart sound classification and, as a result, cardiovascular disease diagnosis [11].

A study by Taneja, Arora, and Verma enhances heart sound signal classification by using the PhysioNet CinC 2016 dataset. The novel approach introduced involves using Gammatonegram—biologically-inspired images that mimic auditory processing in the cochlea—to enhance event detection in cardiac audio signals. By extracting texture-related features from these Gammatonegrams, including Linear Ternary Pattern (LTP) and Local Phase Quantization (LPQ), the study achieved superior classification performance. The accuracy it achieved was 94.00%. Precision and F1 scores also reached 91.77 and 93.61, respectively. This was better as compared to other image representations [12]. Another study discusses the benefits of transfer learning through pre-trained convolutional neural networks for automatic classification of limited data. It uses standard time-frequency representations as input features, such as spectrogram, log-Mel spectrogram, and scalogram, for fine-tuning lightweight models pre-trained on audio and images for the classification of PCG. Four varieties of heart sound data are classified using the transfer learning method based on YAMNet. The proposed approach obtains an overall accuracy of 99.83%, sensitivity of 99.59%, and specificity of 99.90%. At the same time, it classified the PhysioNet/CinC Challenge 2016 data into two classes with an accuracy of 92.23% [13].

From another paper research a non-invasive method for detecting left ventricular diastolic dysfunction (LVDD) using a PCG-based transfer learning CatBoost model. Features were extracted by four spectrogram representations namely, STFT, MFCCs, S-transform and gammatonegram combined with pre-trained CNNs such as Xception, ResNet50, VGG16 and InceptionResNetv2. Features were fused and classified using CatBoost, outperforming other machine learning classifiers. The best model, with LDA feature fusion, achieved an AUC of 91.1%, 88.2% accuracy, 82.1% sensitivity, 92.7% specificity, and an F1-score of 0.892. This approach could help in non-invasive LVDD detection [14].

III. Proposed work

A. Gammatonegram

The Gammatonegram is in fact a graphical representation of an audio signal, just the name says, similar to the process in which the human auditory system copes with sound. It is based on the Gammatone filterbank, which is a series of bandpass filters modeled after the frequency response of the human ear, particularly the cochlea. Each filter in the bank is tuned to a different frequency, allowing the Gammatonegram to capture the distribution of energy across different frequencies over time, like how the cochlea analyzes sound.

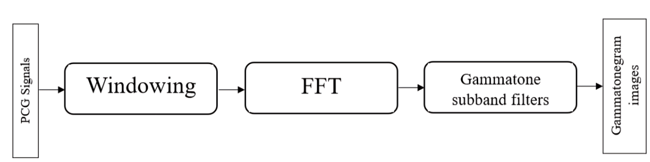

Fig. 1: PCG Signals to Gammatonegram images.

In the context of heart sound analysis, a Gammatonegram can be applied to phonocardiogram (PCG) signals to highlight specific features that may indicate abnormalities in heart function. Fig.1 shows the transformation from a sound signal to a Gammatonegram image. The process begins with a sound signal. Then, the signal is divided into small parts called windows or frames to make it easier to analyze. The next step is to use the Fast Fourier Transform (FFT) on these small parts. The FFT changes the signal from time-based to frequency-based, which helps to look at the different frequency parts of the signal separately. Once the frequency-domain representation is obtained, it is passed through Gammatone subband filters. These filters are designed to simulate the human auditory system's response to various frequencies, particularly by emphasizing frequencies that are more relevant to human hearing. The output from the Gammatone filters is then transformed into Gammatone images, which visually represent the frequency information in a way that reflects the perceptual characteristics of the sound as it would be interpreted by the human ear. This is shown in the figure below, with the healthy and unhealthy PCG signals samples filtered into images using gammatone subband filters.

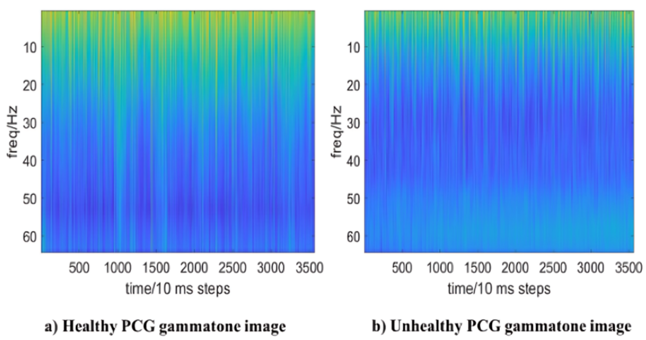

Fig. 2: Healthy and Unhealthy PCG signals samples filtered into images using gammatone subband filters.

B. Dataset

The dataset is crucial to the successful training and validation of the CNN model for detecting abnormalities in Phonocardiogram (PCG) signals. It consists of a total of 3,541 Gammatonegram images derived from PCG recordings, categorized into two main classes: 'healthy' and 'unhealthy'. These images represent the time-frequency analysis of heart sounds, capturing both temporal and spectral information that is essential for distinguishing normal heart function from potential abnormalities. For dataset refer [14]

Dataset Breakdown:

- **Training Data:** Out of the total, 3,240 images are utilized for training the CNN model. This big part of the dataset makes sure that a lot of diverse examples are seen by the model so that it learned in detail about both healthy and unhealthy heart sounds. Out of the total, 3240 images 2575 images are healthy and 665 images are unhealthy.
- **Validation Data:** The remaining 301 images form a validation set that will be very important for model performance evaluation. This subset is essential for ensuring that the model generalizes effectively to new, unseen data. The validation data has 301 images. Out of these 150 images are healthy and 151 images are unhealthy.

C. Preprocessing

Preprocessing is a vital step in machine learning and deep learning workflows, particularly when dealing with raw data. In the context of your project involving Phonocardiogram (PCG) signals and Gammatonegram images, preprocessing ensures that the data is clean, structured, and ready for input into a model, improving its performance and accuracy. The preprocessing techniques involved are normalization, resizing, data augmentation.

Normalization:

Normalization is a technique which is used to scale data into a standard range; often, a range of 0 to 1 or -1 to 1 is considered standard. Normalized data can help the model converge faster during training, as the optimization algorithms (like gradient descent) perform better when the input features are on a similar scale.

Resizing:

Resizing is the process of the altering the dimensions of the images to match the input size required by the neural network. CNNs typically require fixed-size inputs, so all images must be resized to ensure uniformity. Thus, the input images are resized to 150×150 and fed to the CNN model.

Augmentation:

Data augmentation is used to increase the size of a training dataset by altering the original images in some different ways. In this way, the model becomes powerful enough, and its performance increases when few examples are available in the dataset. Various classic transformations would involve random rotations, flipping, shifts, zooms, and cropping. Random horizontal flip is used on 3541 images because the dataset has few images. This technique is used to increase images in multiples of 2.

D. Proposed Workflow

The suggested model uses a Convolutional Neural Network (CNN) design to find problems in Phonocardiogram (PCG) signals by analyzing them with Gammatonegram. The process starts with the input PCG signals, which are changed using Gammatonegram computation to create time-frequency representations needed for capturing important features. Before feeding these representations into the CNN, several preprocessing techniques are applied. These include normalization to scale the data within a specific range, resizing the Gammatonegram images to a consistent dimension, and data augmentation to enhance the model's robustness by artificially expanding the training dataset through techniques like rotation, scaling, and flipping. Additionally, class weights are adjusted to address class imbalance, ensuring the model pays equal attention to underrepresented categories.

The CNN architecture is composed of three parallel processing streams. Each stream is initiated with a Convolutional layer that extracts out the spatial features. Then, there is Batch Normalization to make learning more stable, with subsequent use of the Rectified Linear Unit (RELU) as an activation function to introduce non-linearity. Such outputs from these branches are combined into a central block consisting of several layers

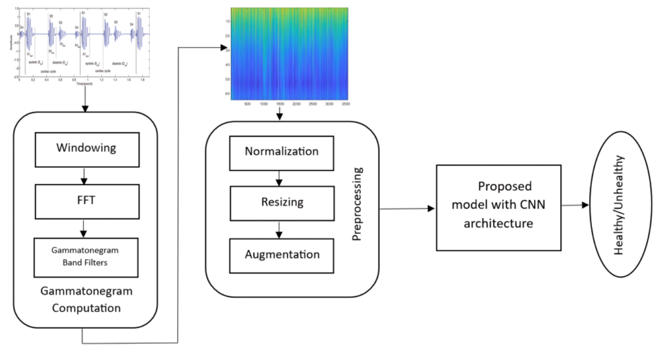

Fig. 3: Flowchart of CNN-Driven Detection of Abnormalities in PCG Signals using Gammatonegram Analysis.

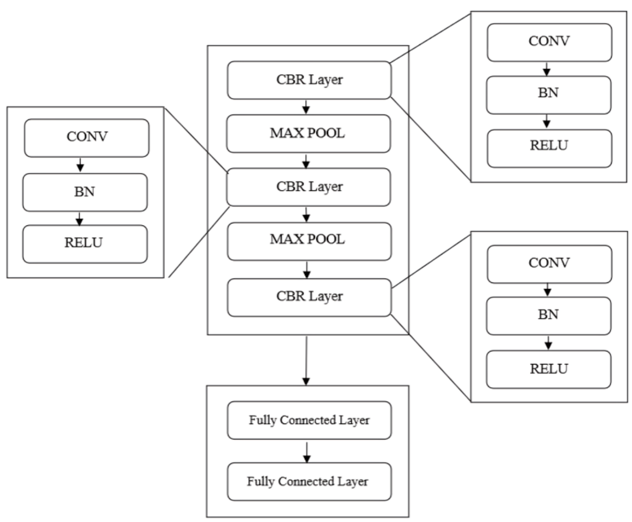

Fig. 4: Proposed model with CNN architecture.

of Convolution-BatchNorm-ReLU (CBR) combined with Max Pooling layers mixed in. Smaller feature maps are obtained by using the Max Pooling layers, which reduce the size of the feature maps without losing important patterns. The CBR layers are repeated in order to enhance feature extraction and allow the model to better capture small details in the signals.

After passing through these convolutional layers, the features are flattened and passed as input to two fully connected layers. These layers serve as the model's decision-making components, transforming the extracted features into classification outputs. The final layer outputs a decision indicating whether the input PCG signal corresponds to a healthy or unhealthy heart condition. To optimize the performance of the model, several optimizers are used to adjust the model weights by minimizing the loss function during training.

E. Model

The convolutional neural network is one of the deep learning methods where the feature choice has changed with model learning. A neural network usually consists of three layers they are input layer, hidden layers, and output layer.

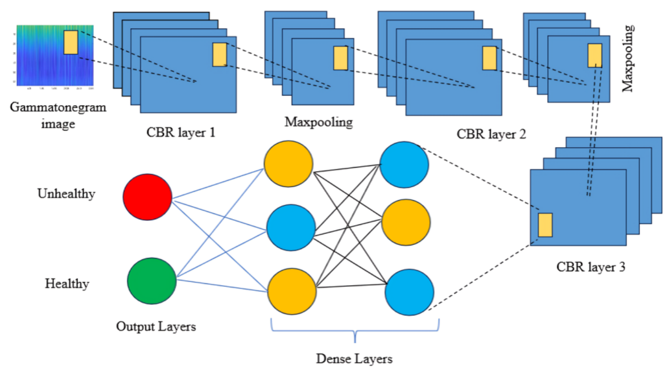

Fig. 5: Proposed CNN model.

As shown in Fig. 5. The proposal has been subjected to various layers of Conv2D, with the first layer having 12 filters, followed by 20 and 32 in the later stages. Each convolutional layer is followed by batch normalization and ReLU activation for normalization and introducing non-linearity for stabilized learning and modeling complex relations. Max-pooling with a filter size of 2x2 was used after the first two convolutional layers, reducing the feature maps from 150x150 to 75x75 and subsequently to 37x37. This retains only the most important elements of the feature maps. Deeper convolutional layers, represented by CBR layers 2 and 3, capture even more complex and high-level patterns. Additional filters were added to detect new patterns while reducing the image size in. It then flattens into a single line of 43,888 features and goes through fully connected layers (dense), which are 256, 128, and 2 in size. Finally, the sounds are classified as "Healthy" or "Unhealthy." This gradually increased filter size and decreased detail help the CNN systematically gather important features from the input Gammatonegram images, thereby coming out with precise heart sound classification.

F. Optimizers

Optimizers, in general, play an important role in training neural networks, including CNNs. They adapt the model's weights in a manner that minimizes the loss function of the model, therefore making it perform better. Optimizers modify the model's parameters-weights-after the gradients are computed during backpropagation. The idea is to converge on an optimal set of parameters that minimize the loss function, hence yielding the optimal performance of the model. How this set of changes is accomplished may vary with the optimizer, which then reflects on the speed and manner in which the model will learn. The choice of optimizer can greatly influence convergence speed, escaping local minimum, and overall performance of a model. The optimizers used in the proposed model are SGD, Adagrad, Adadelta, RMSProp, and Adam. More information about these optimizers is discussed below.

*1) **Stochastic Gradient Descent (SGD)**:* Stochastic Gradient Descent changes the model weights based on the gradient of the loss concerning each parameter. Contrary to regular Gradient Descent, which uses the whole dataset to calculate the gradient, in SGD, updates of weights are made for every training example or small groups of them. This makes it quicker and, moreover, far better when big datasets are taken into consideration.

*2) **Adagrad (Adaptive Gradient Algorithm)**:* Adagrad updates the learning rate independently for every parameter. It makes the learning rate smaller by the magnitude of past gradients. As a result, parameters with large gradients get smaller updates while parameters with small gradients get larger updates. This can be useful when dealing with sparse data.

*3) **Adadelta**:* Adadelta can be seen as a variant of Adagrad that tries to solve the problem of decreasing learning rate. Rather than accumulating all formerly squared gradients, Adadelta restricts the accumulation of gradients to a fixed size. This allows the learning rate to remain adaptive over time.

*4) **RMSProp (Root Mean Squared Propagation)**:* RMSProp does the same thing as Adagrad but it uses an exponentially decaying average over the squared gradients, instead of accumulating all the squared gradients. This keeps the learning rate manageable and prevents it from shrinking too much.

*5) **Adam (Adaptive Moment Estimation)**:* Adam is one of the most popular optimizers, combining the advantages of both RMSProp and Momentum (another optimization technique that accelerates SGD). It computes adaptive learning rates for each parameter, and it includes momentum by keeping a running average of both the gradients and their second moments.

The optimizer parameters are

Number of Epochs: The algorithm is executed multiple times on the training dataset. In this case, we train the model for 25 epochs.

Batch Size: Batch size is among the most important hyperparameters in the training of neural networks. It defines the number of examples that are used in one model iteration or one forward and backward pass of the model. Batch Size gives the number of samples to update the model parameters. In this model, the batch size used is 32.

Learning rate: Learning rate is the scaling factor that is needed for updating the model weights. The proposed model uses different optimizers with different learning rates. It is a hyperparameter, which has a greater influence on both the improvement of training and activity of model which is based on the assessment metrics.

Loss Function: A loss function in general will measure how well your prediction model is doing with respect to predict the expected output. The loss used within this example model is cross entropy loss.

Class Weights: Class weights are used to handle imbalanced datasets where certain classes have significantly more samples than others. Applying class weights ensures that the model pays equal attention to all classes, preventing bias towards the majority class. The dataset used for this study consists of 3541 Gammatonegram images, separated into training and testing sets with a ratio of approximately 9:1.So class weights play an important role to help the model pay equal attention to all classes and preventing bias towards the majority class.

The below table shows the learning rates of optimizers.

TABLE I: Learning rates of optimizers

Optimizers	Learning rates
SGD	0.001
Adagrad	0.01
Adadelta	0.01
RMSProp	0.001
Adam	0.0001

IV. RESULT AND DISCUSSION

The Gammatonegram images derived from the raw PCG signals were processed through a CNN model, with the accuracy evaluated using different optimizers. The model was trained for 25 epochs, using a batch size of 32, and the training and testing performance was compared across various optimizers. The Adam optimizer showed exceptional results, achieving a test accuracy of 100% by the 25th epoch, with a learning rate of 0.0001.The below figures show the training and testing accuracies of each of the optimizers.

Fig 6 represents the graphs for the SGD optimizer, which reached the maximum training accuracy of 95.03% with a test accuracy of 89.70%. Fig. 7 represents the graph for the Adagrad optimizer, which reached the maximum training accuracy of 97.56% with a test accuracy of 96.01%. Fig. 8 Results with Adadelta optimizer shows that training accuracy is up to 96.76% and test accuracy reaches 92.69%. Fig. 9 Maximum of 91.98% on training accuracy and maximum of 82.39% on test accuracy achieved by RMSProp.Fig. 10 The Adam optimizer's performance is represented, which reached the highest training accuracy of 100.00% and best test accuracy of 100%.

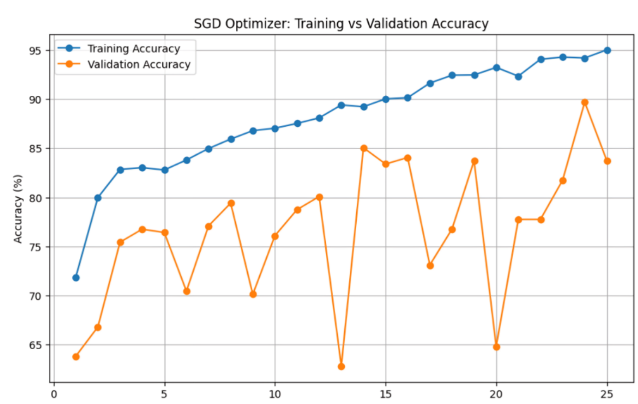

Fig. 6: Differentiation of training and testing accuracy with SGD optimizer

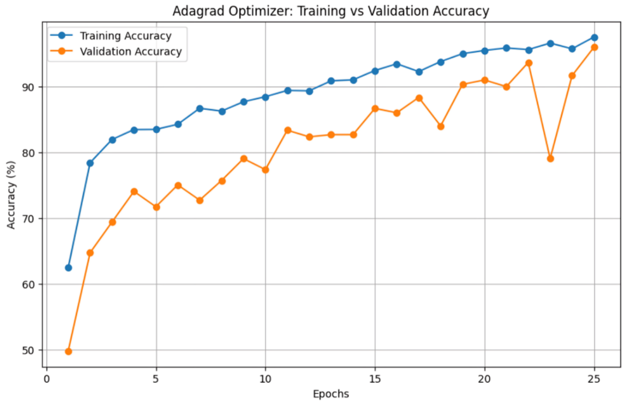

Fig. 7: Differentiation of training and testing accuracy with Adagrad optimizer

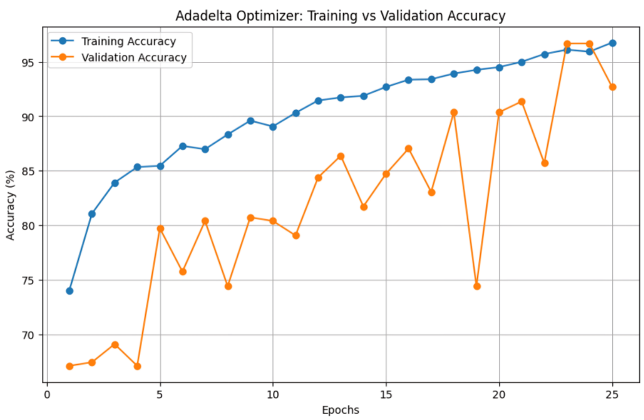

Fig. 8: Differentiation of training and testing accuracy with Adadelta optimizer

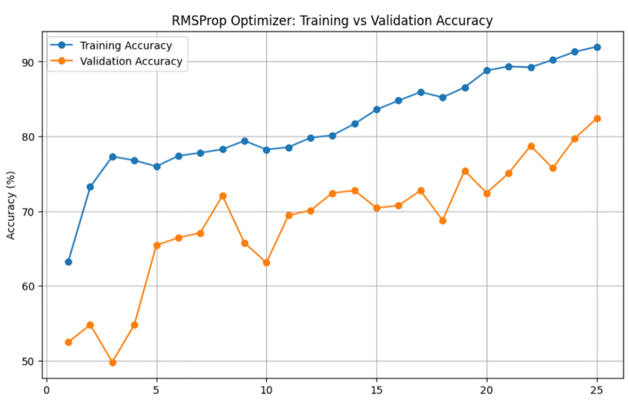

Fig. 9: Differentiation of training and testing accuracy with RMSProp optimizer

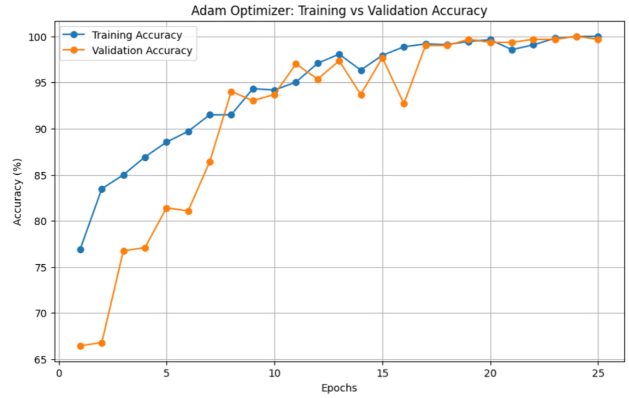

Fig. 10: Differentiation of training and testing accuracy with Adam optimizer

The comparison of training and testing accuracy for all optimizers is represented in below figure.

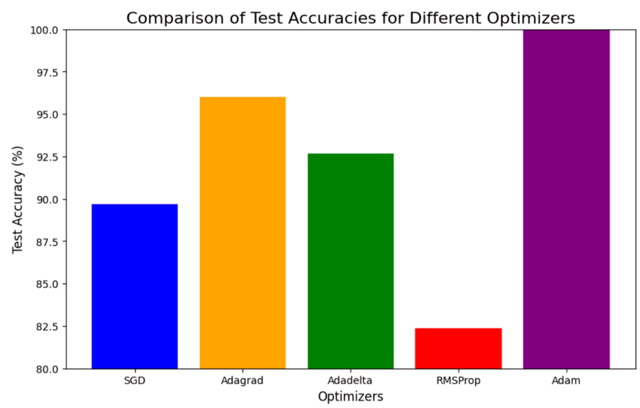

Fig. 11: Comparison of test accuracies across various optimizers

The above Fig. 11 compares the maximum test accuracies of all optimizers. The Adam optimizer achieved the highest training accuracy of 100%. The Adagrad and Adadelta optimizers also performed well, with test accuracies of 96.01% and 92.69%, respectively. In contrast, the SGD and RMSProp optimizers exhibited lower performance, with test accuracies of 89.70% and 82.39%. These results underscore the superior effectiveness of the Adam optimizer for the task of PCG signal classification using CNNs. Overall, the Adam optimizer proved to be the most efficient, yielding the highest accuracy and faster convergence, while other optimizers like Adagrad and Adadelta delivered moderate performance.

From the below table II , we can observe that various advanced techniques have been employed to significantly improve heart sound classification. The study by ASHINIKUMAR SINGH et al. applied ensemble learning with CNN models on the PhysioNet 2016 dataset, obtained accuracy of 99.51% [11]. Another study focused on heart sound classification using Gammatonegram images and texture-based

TABLE II: Accuracy comparsion for different models

S.No	Study & author	Models/Techniques	Accuracy
1	ASHINIKUMAR SINGH et al. Ensemble Learning for Accurate Prediction	CNN, Ensemble Learning	99.51%
2	Heart sound classification method using gammatonegram	CNN, Texture Feature Extraction	94%
3	Transfer Learning Research	YAMNet-based Transfer Learning (TL)	92.23% (PhysioNet), 99.83 (HVD dataset)
4	LVDD Detection using CatBoost	Pre-trained CNNs (VGG16, Xception, ResNet50), CatBoost with LDA	91.1%, 88.2%, 82.1%, 92.7%
5	Classification of Phonocardiogram Signals Using the Wavelet Scattering Transform and Equilibrium Optimization Approach.	K-nearest neighbour (KNN)	99.5%
6	Gammatonegram image classification	Proposed CNN model	100%

feature extraction methods like Linear Ternary Pattern (LTP) and Local Phase Quantization (LPQ), also on the PhysioNet CinC 2016 dataset. This approach reached an accuracy of 94% [12]. A study fine-tuned YAMNet-based pre-trained models using time-frequency representations, including spectrograms and scalograms, on the PhysioNet/CinC Challenge 2016 dataset. This achieved accuracy of 92.23% on the PhysioNet dataset and improved accuracy 99.83% on the HVD dataset [13]. A study on detecting left ventricular diastolic dysfunction (LVDD) used pre-trained CNNs, such as VGG16, Xception, and ResNet50, combined with CatBoost classifiers. This approach resulted 88.2% accuracy [14]. Cross Wavelet Transform with the AlexNet model on the PhysioNet dataset resulted in an accuracy of 98% [15]. Finally, a CNN model using Gammatonegram images was applied to the PhysioNet Challenge 2016 dataset, where it achieved a perfect accuracy of 100%.

V. Conclusion

In conclusion, the proposed method using Gammatonegram-based analysis with CNN offers a simple yet effective approach for detecting abnormalities in PCG signals. By transforming PCG signals into Gammatonegram images, and applying preprocessing techniques like resizing and random horizontal flip augmentation, the model can accurately classify healthy and unhealthy heart sounds. Among the optimizers tested, the Adam optimizer achieved the highest accuracy of 100%, demonstrating its effectiveness. This study confirms that Gammatonegram representations are highly suitable for the classification of PCG signals, offering a reliable method for heart disease detection.

References

[1] Cheung, Ka Wang, M. Bonet, K. A. Frank, O. T. Oladapo, G. J. Hofmeyr, WHO Intrapartum Care Algorithms Working Group, Livia Ciabati et al. "Clinical algorithms for management of fetal heart rate abnormalities during labour." BJOG: An International Journal of Obstetrics & Gynaecology 131 (2024): 102-111.

[2] Tunnell, Natalie C., Sarah E. Corner, Andres D. Roque, Juliet L. Kroll, Thomas Ritz, and Alicia E. Meuret. "Biobehavioral approach to distinguishing panic symptoms from medical illness." Frontiers in Psychiatry 15 (2024): 1296569.

[3] Duca, Ştefania-Teodora, Ionuţ Tudorancea, Radu-Ştefan Miftode, Adriana Chetran, Alexandru-Dan Costache, D. Robert Pavăl, and Irina-Iuliana Costache-Enache. "Heartbeat Chronicles: Decoding the Interplay of Echocardiography and Heart Rate Variability in Chronic Heart Failure Patients–Unraveling the Mysteries with Traditional and Advanced 24-Hour Holter ECG Parameters." Romanian Journal of Cardiology.

[4] Lee, Joon-Hyop, EunKyung Ku, Yoo Seung Chung, Young Jae Kim, and Kwang Gi Kim. "Intraoperative detection of parathyroid glands using artificial intelligence: optimizing medical image training with data augmentation methods." Surgical Endoscopy 38, no. 10 (2024): 5732-5745.

[5] M.Sireesha, S. N. TirumalaRao, Srikanth Vemuru, Optimized Feature Extraction and Hybrid Classification Model for Heart Disease Prediction International Journal of Recent Technology and Engineering Vol - 7, No 6, Mar - 2019 ISSN - 2277-3878, Pages – 1754 – 1772.

[6] Lee, T. H., E. F. Shair, A. R. Abdullah, K. A. Rahman, and N. Nazmi. "Comparison of Short Fast Fourier Transform and Continuous Wavelet Transform in Study of Stride Interval."

[7] M.Sireesha, Srikanth Vemuru, S.N.Tirumala Rao "Classification Model for Prediction of Heart Disease Using Correlation Coefficient Technique" International Journal of Advanced Trends in Computer Science and Engineering, Vol. 9, No. 2, March - April 2020, Pages- 2116 – 2123.

[8] Chambi, Edwin M., Jefry Cuela, Milagros Zegarra, Erasmo Sulla, and Jorge Rendulich. "Benchmarking Time-Frequency Representations of Phonocardiogram Signals for Classification of Valvular Heart Diseases Using Deep Features and Machine Learning." Electronics 13, no. 15 (2024): 2912.

[9] Arslan, Muhammad, Muhammad Mubeen, Muhammad Bilal, and Saadullah Farooq Abbasi. "1D-CNN-IDS: 1D CNN-based Intrusion Detection System for IIoT." arXiv preprint arXiv:2409.08529 (2024).

[10] Ahmad, Ahsan, Ali Husnain, Ashish Shiwlani, Ahad Hussain, Mahnoor N. Gondal, and Ayesha Saeed. "Ethical and clinical implications of AI integration in cardiovascular healthcare." World Journal of Advanced Research and Reviews 23, no. 3 (2024): 2479-2501.

[11] ASHINIKUMAR SINGH, S. I. N. A. M., S. I. N. A. M. Ajitkumar SINGH, and A. H. E. I. B. A. M. DINAMANI SINGH. "Ensemble Learning for Accurate Prediction of Heart Sounds using Gammatonegram Images." Turkish Journal of Electrical Engineering and Computer Sciences 32, no. 4 (2024): 555-573.

[12] Taneja, Kriti, Vinay Arora, and Karun Verma. "Heart sound classification method using gammatonegram and SVM." Multimedia Tools and Applications (2024): 1-37.

[13] Maity, Arnab, Akanksha Pathak, and Goutam Saha. "Transfer learning-based heart valve disease classification from Phonocardiogram signal." Biomedical Signal Processing and Control 85 (2023): 104805.

[14] Zheng, Yineng, Xingming Guo, Yang Yang, Hui Wang, Kangla Liao, and Jian Qin. "Phonocardiogram transfer learning-based CatBoost model for diastolic dysfunction identification using multiple domain-specific deep feature fusion." Computers in Biology and Medicine 156 (2023): 106707.

[15] Ismael, Mustafa R., Haider J. Abd, and Raad Z. Homod. "Classification of Phonocardiogram Signals Using the Wavelet Scattering Transform and Equilibrium Optimization Approach." (2024).

[16] For dataset refer this link : https://www.kaggle.com/datasets/brundavanamsatyasai/gammatone-images-of-heart-sounds