Diagnosis of Heart Diseases Using Heart Sound Signals with the Developed Interpolation, CNN, and Relief Based Model

Muhammed Yildirim

Department of Computer Engineering, Malatya Turgut Ozal University, Malatya 44210, Turkey

Corresponding Author Email: muhammed.yildirim@ozal.edu.tr

ABSTRACT

The majority of deaths today are due to heart diseases. Early diagnosis of heart diseases will lead to early initiation of the treatment process. Therefore, computer-aided systems are of great importance. In this study, heart sounds were used for the early diagnosis and treatment of heart diseases. Diagnosing heart sounds provides important information about heart diseases. Therefore, a hybrid model was developed in the study. In the developed model, first of all, spectrograms were obtained from audio signals with the Mel-spectrogram method. Then, the interpolation method was used to train the developed model more accurately and with more data. Unlike other data augmentation methods, the interpolation method produces new data. The feature maps of the data were obtained using the Darknet53 architecture. In order for the developed model to work faster and more effectively, the feature map obtained using the Darknet53 architecture has been optimized using the Relief feature selection method. Finally, the obtained feature map was classified in different classifiers. While the accuracy value of the developed model in the first dataset was 99.63%, the accuracy rate in the second dataset was 97.19%. These values show that the developed model can be used to classify heart sounds.

1. INTRODUCTION

When the data of the World Health Organization are examined, it is seen that the majority of deaths today are caused by heart diseases [1]. Therefore, it is clear that cardiovascular diseases are a serious threat to humans. There are advanced methods such as electrocardiography and computed tomography to detect heart diseases. However, these methods are both costly and not available in most centers. In addition, the implementation of these methods is time-consuming [2].

It is possible to obtain important information about heart diseases by examining heart sounds. Listening to heart sounds by experts is an alternative method for detecting the deformation of heart sounds. However, the analysis of heart sounds by auscultation of the patient is mostly dependent on the expertise and skills of the practitioner [3]. For this reason, recording and analysis of heart sounds via computer is a preferred method [4].

In the diagnosis of cardiovascular disorders, automatic classification of heart sounds is critical [5]. There has been a growing focus on creating deep learning algorithms for heart sound categorization since the arrival of medical big data and artificial intelligence technology. Despite major advancements in this area, there are still limitations due to a lack of data, inadequate training, and effective models. Heartbeat sound classification is still a challenging problem [6]. For this reason, the classification of heart sounds has been made in the study. Classification of sound signals, thanks to the artificial intelligence-based model used, will facilitate the work of specialists in the diagnosis stage, and this model is of great importance for the early diagnosis of the disease.

1.1 Related works

Heart sound comprises a number of key variables that can help with the early detection of heart disease. In recent years, many methods for the classification of heart sounds have been proposed for the detection of heart disease.

Raza et al. [7] classified heartbeat sounds in three categories as Normal, Murmur, and Extra systole in their study. In this context, the researchers aimed to remove the noise from the heartbeat audio signal by applying a band filter to the heart sounds. After this step, they fixed the size of the sampling rate of each audio signal and then applied down-sampling techniques to gain more distinctiveness and reduce the size of the frame rate. They also classified heart sounds using a Recurrent Neural Network (RNN) model consisting of LSTM, Dropout, Dense, and Softmax layers.

Noman et al. used 1D-CNN and 2D-CNN architectures to classify heart sounds in their study. 2D-CNN architectures are given features obtained by the MFCC method. In addition, the features obtained using CNN architectures in the study were classified using SVM and Markov models. In the study, the researchers used the PhysioNet challenge 2016 dataset. The accuracy value obtained in the proposed model was 89.22% [8].

In their study, Xiao et al. [9] proposed a three-part system based on deep learning methods for cardiovascular disease prediction. The proposed model consists of three parts. These stages are; Preprocessing is majority voting for 1-D waveform heart sound classification, attention mechanism, and final estimation of heart sound recordings using deep convolutional neural network.

Demir et al. [10] first created the spectrograms of the sound

signals in the method they proposed for the classification of heart sounds. Then, feature maps of the spectrogram images were extracted using deep learning methods. Finally, these feature maps are classified. The proposed model has been tested on two different datasets. The researchers stated that the model they proposed could be used in the classification process of heart sounds.

Cheng et al. [11] proposed a new heart sound classification model for the classification of heart sounds. This model, which they call LHSNN, consists of three parts. First of all, using the audio signals in the dataset, spectrogram images were obtained from these audio signals. Then, spectrogram images were analyzed with the proposed neural network models. Finally, they stated that the model they proposed works through mobile terminals.

Er [12] used Local Binary Pattern (LBP) and Local Ternary Pattern (LTP) methods to classify heart sounds in his study. Then, the feature maps obtained using these two feature extraction methods were combined. The relief method is preferred for the optimization of feature maps. It was stated that this proposed model was tested on two different datasets.

Boulares et al. [13] used CNN architectures, supervised learning, and unsupervised learning methods in their studies. These proposed methods were tested on two datasets. In the study, the spectrogram process was applied in the preprocessing step. Obtained images were classified using supervised and unsupervised learning methods.

Tariq et al. [14] proposed a feature-based fusion network to classify heart and lung sounds. In the study, three different feature extraction methods were used to extract the features of audio signals. Spectrogram, MFCC, and Chromagram feature extraction methods are the methods used in this study. Researchers combined the features obtained from these three feature extraction methods and classified them over the Softmax layer.

Kui et al. [15] introduced a method for feature extraction and classification of heart sound signals. In the first step of the study, the noises in the heart sound signals were removed. In the method, the time-dependent hidden Markov method, log Mel-frequency spectral coefficients (MFSC), CNN, and majority voting algorithm were also used. In this study, two-class and multi-classification models were created.

1.2 Contribution and novelty

The aim of the study was to diagnose heart diseases, which are a serious problem today.

In the study, it has been shown that the interpolation method is successful in datasets where the number of data is low.

Mel-spectrograms of the audio data were obtained and feature selection was made with the Darknet53 architecture, which is one of the pre-trained architectures accepted in the literature.

Relief feature selection method was used to reduce the size of the obtained feature maps. In this way, unnecessary features are eliminated. Therefore, it was possible to run the model faster.

The feature map, which was optimized in the last stage, was classified with the SVM classifier accepted in the literature.

The proposed model has been tested using the two datasets. While the accuracy value of the proposed model in the first dataset was 99.63%, the accuracy value in the second dataset was 97.19%. The obtained accuracy values show that the proposed model is successful in classifying heart sounds.

1.3 Organization of paper

The rest of the work is organized in the following manner. The second part constitutes the materials and methods part of the study. In this section, the dataset used and the technologies used in the proposed approach are presented. Section 3 contains the experimental results of the proposed method. Section 4 is the discussion section where the results of the study are discussed. Finally, Section 5 is the conclusion section.

2. MATERIAL AND METHODS

In order to understand the study, it is important to explain the dataset used and to understand the techniques used in the method. This section presents the techniques used in the implementation of the proposed model. In the last section, the proposed model is given. Figure 1 depicts the proposed structure.
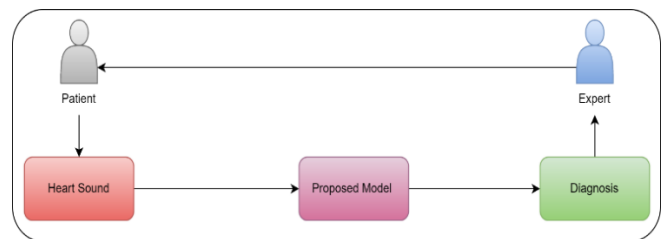

Figure 1. Block diagram of proposed model

2.1 Dataset

The dataset from the Kaggle was used for the classification of the heart sounds [16]. There are studies using the relevant dataset in the literature [7, 10]. Data in the dataset were collected from two different sources, from the general public via the stethoscope Pro iPhone app and from a clinical trial in hospitals using the digital stethoscope DigiScope [16, 17]. The relevant dataset is divided into two sources, Set_a and Set_b.

Set_a: Contains tags and metadata for heartbeats collected from the public via an iPhone app. Heart sounds are classified into 4 classes: Artifact, murmur, normal, and extrasystole. Figure 2 shows some sample sounds of the dataset.
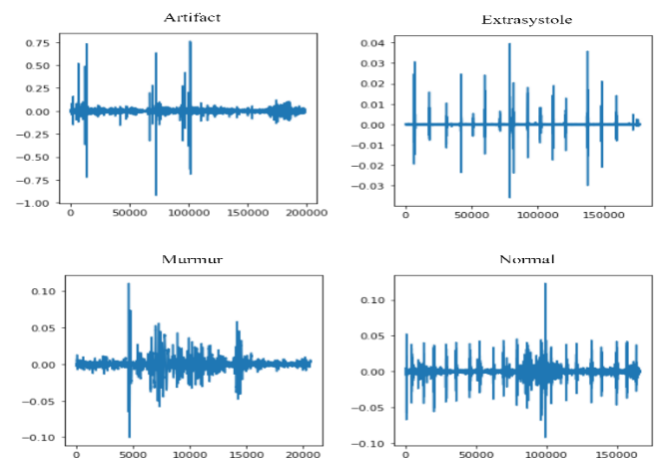

Figure 2. Examples from the Set_a dataset

Set_b: Set_b dataset contains tags and metadata for heartbeats collected from a clinical trial using a digital stethoscope in hospitals. The dataset categorizes heart sounds into three classes. These classes are; Normal, Murmur and Extrasystole. Figure 3 illustrates some of the sounds in the dataset.
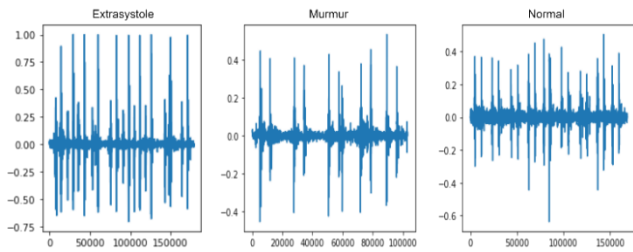

Figure 3. Examples from the Set_b dataset

2.2 Spectrogram, interpolation, and feature extraction

The spectrogram is a visual heat map in which the horizontal axis indicates the signal's time and the vertical axis represents the signal's frequency. In other words, the Spectrogram is a visual representation of an audio signal that changes with time [18, 19]. In the study, the Mel-Spectrogram method was used to obtain spectrograms of audio signals. The parameters used are given in Table 1.

Table 1. Mel-Spectrogram parameters

FrequencyRange	62.5,8e3
Window,hann	2048, periodic
OverlapLength	1024
FFT length	4096
NumBands	64

Since the number of data in the dataset is very small, the interpolation method was used to obtain new images from the obtained spectrograms. Thanks to this method, new spectrograms are produced from existing spectrograms [20, 21]. The working logic of the interpolation method is different from the data augmentation methods. In data augmentation methods, data augmentation is performed by pre-processing the existing data (scaling, rotation, etc.) [22]. In this respect, the Interpolation method works differently from the data augmentation methods. Because data is not augmented in the interpolation method, new data is produced. While in data augmentation methods, image multiplexing is done using one image, new images are obtained by using two images in the interpolation method.

The linear interpolation method is a curve-fitting method that provides new data by using known points of the data using linear polynomials [23, 24]. In this study, the interpolation method was used to produce new images. There are studies on the subject in the literature [20, 21]. When this method is applied to images, it first produces new images by using 2 RGB images in X1 and X2 format. Assuming that the value of each pixel in the X1 images is P1(R1, G1, B1) and the value of each pixel in the X2 images is P2(R2, G2, B2), the new pixel obtained from these values is Pnew (Rnew, Gnew, Bnew) can be expressed as. The new values obtained are calculated by Eqns. (1)-(3).

$$R_{new} = (1 - t) \times R_1 + t \times R_2 \text{ for } t \in [0,1] \qquad (1)$$

$$G_{new} = (1 - t) \times G_1 + t \times G_2 \text{ for } t \in [0,1] \qquad (2)$$

$$B_{new} = (1 - t) \times B_1 + t \times B_2 \text{ for } t \in [0,1] \qquad (3)$$

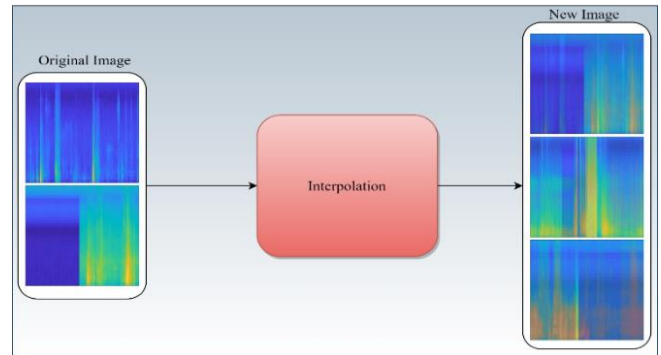Figure 4 shows 3 new images produced by the interpolation technique from the original 2 images.

Figure 4. Three new images were created by interpolation method from the original two images

When Figure 3 is examined, it is clear that the images produced by the interpolation method are different from the original images. While no new images are produced in other data augmentation methods, new images are produced in the interpolation method.

After obtaining new images with the interpolation method, feature extraction was done using Darknet53 architecture. Darknet53 architecture is a pre-trained architecture that has been accepted in the literature [25]. The size of the feature map obtained with the Darknet53 architecture is the number of images x 1000. The relief feature selection method was used to eliminate unnecessary features in the obtained feature map [26]. The size of the feature map obtained after the relief feature selection method was the number of images x 450. Finally, different classifiers accepted in the literature were used to classify the obtained feature map. Among the used classifiers, the highest performance ratio was achieved with the SVM classifier. The SVM classifier is one of the most preferred supervised learning techniques in classification problems.

In addition, the performances of different classifiers in the classification of heart sounds were also examined. The feature maps obtained using the MFCC method are k-nearest Neighbor (KNN) [27], Support Vector Machine (SVM) [28], Naive Bayes (NB) [29], Logistic Regression (LR) [30], Random Forest (RF) [31], Gradient Boosting Classifier (GBC) [32], XGBoost [33], Light Gradient Boosting Machine (LGBM) [34], and classified in CatBoost [35] model.

2.3 Proposed model

The creation and processing of audio data are very important. Sound processing can be used for different purposes in different areas. Our proposed model, it is aimed to classify heart sounds with a high accuracy value. When studies using the same dataset are examined, it is seen that our model has a high accuracy value. Figure 5 shows the proposed model.

When Figure 5 is examined, it is seen that the model we propose is a hybrid model. Our model first obtains spectrograms of the audio data. Obtained spectrogram samples are given in Figure 6.

After obtaining the spectrogram images, new data were

produced using the interpolation method. In the interpolation method, new spectrograms are obtained by using 2 original spectrograms. At this stage, rotate, scale, etc. Data augmentation is not done. The data produced after the interpolation step are given as input to the Darknet53 architecture. The features obtained from the Conv53 layer of the Darknet53 architecture are optimized with the Relief method. Elimination of unnecessary features by using the Relief method is of great importance in making the model work faster and reducing the cost. Finally, the obtained feature maps were classified in the SVM classifier.
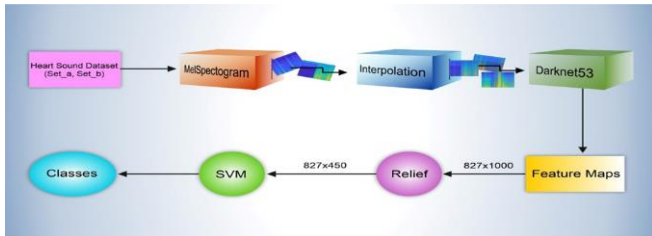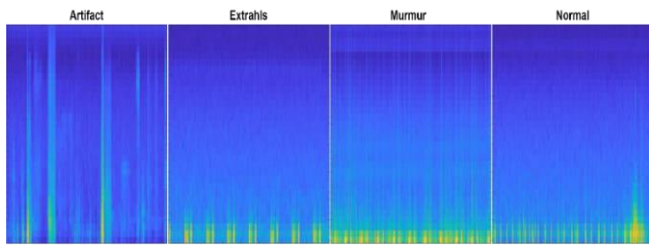

Figure 5. Proposed model

Figure 6. Spectrogram examples

3. EXPERIMENTAL RESULTS

Experimental results were obtained in Matlab and Python environments. First of all, the features obtained using the MFCC method were classified using 9 different classical machine learning classifiers. Different evaluation metrics were used by using a confusion matrix to measure the performance of the classifier and proposed model used in the study [36]. An example of a confusion matrix is given in Figure 7.
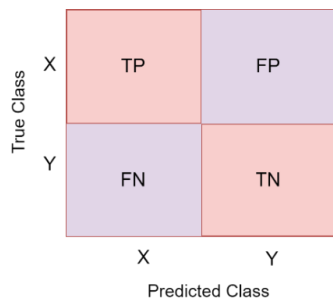

Figure 7. Confusion matrix

TP: The number of images that actually belong to the X class, estimated as X,

TN: The number of images actually belonging to the X class, estimated as Y,

FN: Number of images that actually belong to class Y, estimated as X,

FP: It actually belongs to the Y class and indicates the number of images estimated as Y.

The metrics given in Eqns. (4)-(12) were used to test the performance of the models used in the study. These evaluation criteria are the most commonly used metrics in classification problems [37, 38].

$$\text{Accuracy (ACC)= (TP+TN) / Total} \qquad (4)$$

$$\text{Specificity (SPE)= TN / (FP+TN)} \qquad (5)$$

$$\text{Sensitivity (SENS)= TP / (TP+FN)} \qquad (6)$$

$$\text{Precision (PRE)= TP / (TP+FP)} \qquad (7)$$

$$\text{Negative Predicted Value (NPV)= TN/(TN+FN)} \qquad (8)$$

$$\text{False Positive Rate (FPR)= FP / (FP+TN)} \qquad (9)$$

$$\text{F1 Score (F1) = 2TP / (2TP+FP+FN)} \qquad (10)$$

$$\text{False Discovery Rate (FDR)= FP / (FP+TP)} \qquad (11)$$

$$\text{False Negative Rate (FNR)= FN / (FN+TP)} \qquad (12)$$

3.1 Classification of the properties obtained by the MFCC method in different classifiers

In this section, the feature maps obtained using the MFCC method are classified into 9 different machine learning classifiers. In the study, there are 2 datasets, Set_a and Set_b. Set_a dataset consists of 4 classes and Set_b dataset consists of 3 classes. In order to classify the audio signals in these two datasets, firstly, feature maps of these audio signals were obtained. For the process, the MFCC method accepted in the literature was used. While creating the feature map with the MFCC method, 40 features were taken from each audio signal. 80% of the data in the obtained feature map is used for training the models, and the remaining 20% is reserved for testing the models. The accuracy values of 9 different classifiers obtained using the Set_a dataset are given in Table 2.

Table 2. Accuracy values obtained in Set_a dataset

Set_a Dataset (%Accuracy)			
KNN	72	GBC	72
SVM	68	XGB	68
NB	68	LGBM	60
LR	**84**	CatBoost	68
RF		68	

The values given in Table 1 were obtained using the data in the Set_a dataset. For the test process, 20% of the data in this dataset was used. When the accuracy values obtained in 9 different classifiers are examined; It is seen that the highest accuracy value is reached in the Logistic Regression model. The accuracy value obtained in this model was 84%. Logistic Regression was followed by KNN and GBC models with an accuracy of 72%, respectively. The accuracy value obtained in SVM, NB, RF, XGB, and CatBoost models is 68%. The lowest accuracy value was obtained in the LGBM model with 60%. Confusion matrices of 3 architectures with the highest accuracy value among 9 different models used in the study are given in Figure 8.
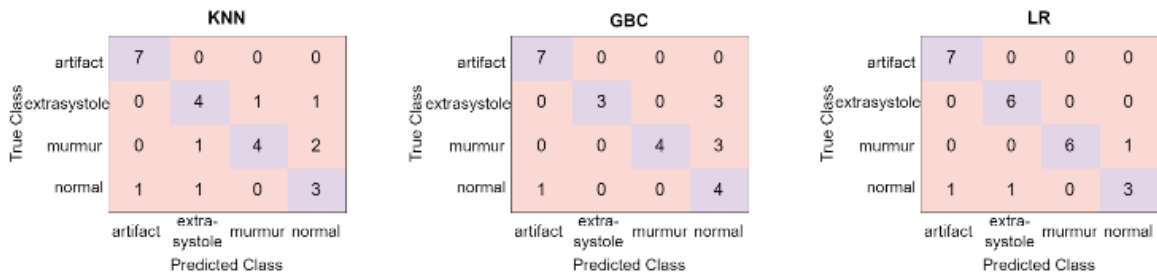
Figure 8. Confusion matrices of the highest models obtained using the Set_a Dataset
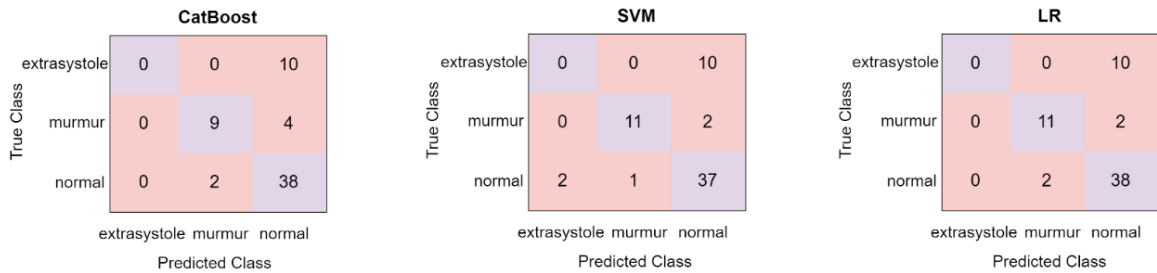

Figure 9. Confusion matrices of the highest models obtained using the Set_b dataset

When Figure 8 is examined, it is seen that the most successful classifier is Logistic Regression (LR). LR classified 21 of 25 test data as correct and 4 of them incorrectly.

The second dataset used in this study is the Set_b dataset. This dataset consists of 3 classes. For the classification of heart sounds in the Set_b dataset, the properties of the sound signals were obtained by using the MFCC method. While feature maps are obtained from this dataset, it is the number of features obtained from each audio signal. 80% of the audio signals in the Set_b dataset was used in the training of the models, and 20% in the testing of the models. The accuracy values obtained in 9 different models are in Table 3.

Table 3. Accuracy values obtained in Set_b dataset

Set_b Dataset (%Accuracy)			
KNN	71.42	GBC	61.90
SVM	76.19	XGB	66.66
NB	65.07	LGBM	63.49
LR	77.7	CatBoost	74.60
RF		73.01	

The values given in Table 3 were obtained using the data in the Set_b dataset. For the test process, 20% of the data in this dataset was used. When the accuracy values obtained in 9 different classifiers are examined; It is seen that the highest accuracy value was obtained in the Logistic Regression model with 77.7%. Logistic Regression was followed by SVM with 76.19% accuracy, Catboost with 74.60%, RF with 73.01%, KNN with 71.42%, XGB with 66.66%, NB with 65.07%, LGBM with 63.49% and GBC with 61.90%, respectively. The most unsuccessful classifier in this dataset was GBC. Confusion matrices of 3 architectures with the highest accuracy among 9 different models used in the study are given in Figure 9.

When Figure 9 is examined, it is seen that the most successful classifier is Logistic Regression (LR). LR 63 classified 49 of the test data correctly and 14 of them incorrectly. Among the classifiers used in the study, LR obtained the highest accuracy value in both Set_a and Set_b datasets.

3.2 Experimental results obtained in the proposed model

Classification of heart sounds provides early diagnosis of diseases and accelerates the treatment process. Two different datasets were used in the study. While the accuracy value obtained in the model we have proposed is 99.63% in the Set_a dataset, it is 97.19% in the Set_b dataset. In our proposed model, a hybrid model was created using the spectrogram, interpolation method, Darknet53 architecture, Relief feature extraction method, and SVM classifier. Confusion matrices obtained in two datasets using the proposed model are given in Figure 10.

When Figure 10 is examined, the proposed model classified 823 of the 826 images obtained using the Set_a dataset correctly and misclassified 3 of them. One of these 3 misclassified images belongs to the artifact class, while the other 2 images belong to the normal class. The proposed model incorrectly predicted 1 image belonging to the artifact class as a murmur. Similarly, he incorrectly predicted 2 images that actually belonged to the normal class as murmurs. Similarly, the proposed model correctly classified 588 of the 605 images obtained using the Set_b dataset, while misclassifying 17 of them. While the proposed model actually predicted an image belonging to the extrasystole class as normal, it actually predicted 3 of the 4 images belonging to the murmur class as extrasystole and 1 as normal. The maximum number of images that the model misclassified belongs to the normal class. The model predicted 7 of the 12 images belonging to the normal class as extrasystole and 5 of them as murmur. While the accuracy value of the proposed model in the Set_a dataset was 99.63%, the accuracy value in the Set_b dataset was 97.19%.

The performance metrics of the proposed model are given in Table 4.

When the performance evaluation metrics given in Table 4 are observed, it is seen that the proposed model can be used in the classification of heart sounds.

The proposed model produced successful results on 2 different datasets consisting of heart sounds.
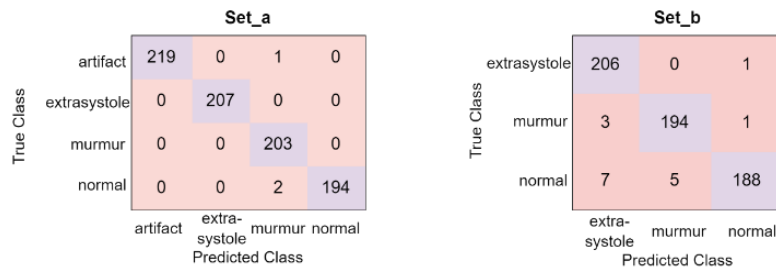
Figure 10. Confusion matrices obtained in Set_a and Set_b datasets

Table 4. Performance metrics of the proposed model on two different datasets

Set_a	Accuracy	Sensitivity	Specificity	Precision	FPR	FDR	FNR	F_1-score
1	99.54	100	100	99.54	0.16	0.45	0	99.77
2	100	100	100	100	0	0	0	100
3	100	98.54	99.51	100	0	0	1.45	100
4	98.97	100	100	98.97	0.31	1.02	0	99.48
Set_b								
1	99.51	95.37	97.48	99.51	0.25	0.48	4.62	97.39
2	97.97	97.48	98.77	97.97	0.98	2.02	2.51	97.73
3	94	98.94	99.50	94	2.89	6	1.05	96.41

4. DISCUSSION

Heart diseases are diseases with high mortality and they constitute an important part of hospital admissions today. Early diagnosis and appropriate treatment are extremely important. In the evaluation of heart diseases, methods such as auscultation, electrocardiography, echocardiography, computed tomography, and conventional angiography are used. Auscultation is an examination method in which the heart sound is listened to with devices called stethoscopes. This method constitutes one of the first and most important steps of cardiac examination [39]. In addition, this method is a cheap, easily accessible, and fast evaluation method compared to other methods. Normal heart sounds are heard during auscultation of a healthy heart. However, in some heart diseases, normal sounds change, become indistinct, disappear, or abnormal additional sounds are heard. Accurate classification of abnormal auscultation findings is important for early diagnosis and the determination of appropriate treatment. It also provides guidance to determine whether further investigations are needed. Murmur is an abnormal blowing sound [40] caused by some heart diseases. Extra systole is the abnormal additional sound heard due to the extra beating of the heart. Apart from these, sometimes artifact sounds can be heard due to environmental or patient-related reasons. In this study, it has been tried to distinguish different classes consisting of normal heart sounds and abnormal sounds such as artifact sounds, murmur, and extra systole.

The proposed model is compared with similar studies in the literature in Table 5.

When Table 5 is examined, it is seen that the proposed model has a high accuracy value. However, there are some limitations of the study. The proposed model has been tested on 2 datasets. The proposed model should be tested and developed using data from patients in different regions. Among our goals is to collect multi-centered data and design a model that can work online.

Table 5. Comparison of the proposed model with similar studies

Authors/Year	Dataset	Classes	Method	Accuracy
[7]/2019	Dataset-B	Normal, Murmur, Extra-systole	RNN	0.80
[8]/2019	PhysioNet/CinC 2016	Normal, Abnormal	2D CNN	0.89
[9]/2020	PhysioNet/CinC 2016	Normal, Abnormal	1D CNN	0.93
[10]/2019	Classifying Heart Sounds Challenge (CHSC) Dataset-A	Artifact, Extra sound, Murmur, Normal	2D CNN	0.80
[10]/2019	Classifying Heart Sounds Challenge (CHSC) Dataset-B	Normal, Murmur, Extrasystole	2D CNN	0.79
[11]/2019	PhysioNet/CinC 2016)	Normal, Abnormal	2D CNN	0.89
[12]/2021	PASCAL	Normal, Murmur, Extra, Artifact	1D CNN	0.91
[12]/2021	PhysioNet/CinC 2016	Normal, Abnormal	1D CNN	0.91
[13]/2021	PASCAL/PhysioNet	Normal, Abnormal	2D CNN	0.87/0.97
[14]/2022	ICHBI 2017	Normal, Murmur, Extrasystole	2D CNN	0.97
[15]/2021	Collected from various hospitals	Normal, Abnormal	2D CNN	0.93
[15]/2021	Collected from various hospitals	Normal, Ventricular septal defect, Atrial septal defect, Patent ductus arteriosus	2D CNN	0.86
2022	Classifying Heart Sounds Challenge (CHSC) Dataset-A	Artifact, Extrasystole, Murmur, Normal	Proposed Model	0.9963
2022	Classifying Heart Sounds Challenge (CHSC) Dataset-B	Normal, Murmur, Extrasystole	Proposed Model	0.9719

5. CONCLUSION

Artificial intelligence methods are frequently preferred in the biomedical field. In this paper, it was used for the classification of heart sounds. In this study, before the heart sounds were classified, the interpolation method was applied, and feature maps of the data were obtained using the Darknet53 architecture, which is one of the pre-trained models. Then, feature selection was made with the Relief method so that the proposed model could work faster and produce more accurate results. When the proposed model is compared with similar studies in the literature, it has been observed that the proposed model is successful. It shows that the proposed model can be used in the classification of heart sounds. Also the proposed model has been tested on two different datasets. While the accuracy value of the proposed model in the first dataset was 99.63%, the accuracy value in the second dataset was 97.19%.

REFERENCES

[1] Cardiovascular diseases, available online: http's://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1, accessed on 28 Feb. 2021.

[2] Tariq, Z., Shah, S.K., Lee, Y. (2020). Automatic multimodal heart disease classification using phonocardiogram signal. In 2020 IEEE International Conference on Big Data (Big Data), pp. 3514-3521. https://doi.org/10.1109/BigData50022.2020.9378232

[3] Indrakumari, R., Shukla, P., Sehgal, A. (2021). Heart disease prediction using Tableau. In Exploratory Data Analytics for Healthcare, pp. 125-141.

[4] Baydoun, M., Safatly, L., Ghaziri, H., El Hajj, A. (2020). Analysis of heart sound anomalies using ensemble learning. Biomedical Signal Processing and Control, 62: 102019. https://doi.org/10.1016/j.bspc.2020.102019

[5] Yadav, A., Singh, A., Dutta, M.K., Travieso, C.M. (2020). Machine learning-based classification of cardiac diseases from PCG recorded heart sounds. Neural Computing and Applications, 32(24): 17843-17856. https://doi.org/10.1007/s00521-019-04547-5

[6] Shaikh Salleh, S.H., Noman, F.M., Chee-Ming, T., et al. (2021). Key techniques and challenges for processing of heart sound signals. In International Conference on Applied Intelligence and Informatics, pp. 136-149. https://doi.org/10.1007/978-3-030-82269-9_11

[7] Raza, A., Mehmood, A., Ullah, S., Ahmad, M., Choi, G.S., On, B.W. (2019). Heartbeat sound signal classification using deep learning. Sensors, 19(21): 4819. https://doi.org/10.3390/s19214819

[8] Noman, F., Ting, C. M., Salleh, S.H., Ombao, H. (2019). Short-segment heart sound classification using an ensemble of deep convolutional neural networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1318-1322. https://doi.org/10.1109/ICASSP.2019.8682668

[9] Xiao, B., Xu, Y., Bi, X., Zhang, J., Ma, X. (2020). Heart sounds classification using a novel 1-D convolutional neural network with extremely low parameter consumption. Neurocomputing, 392: 153-159. https://doi.org/10.1016/j.neucom.2018.09.101

[10] Demir, F., Şengür, A., Bajaj, V., Polat, K. (2019). Towards the classification of heart sounds based on convolutional deep neural network. Health Information Science and Systems, 7(1): 1-9. https://doi.org/10.1007/s13755-019-0078-0

[11] Cheng, X., Huang, J., Li, Y., Gui, G. (2019). Design and application of a laconic heart sound neural network. IEEE Access, 7: 124417-124425. https://doi.org/10.1109/ACCESS.2019.2934827

[12] Er, M.B. (2021). Heart sounds classification using convolutional neural network with 1D-local binary pattern and 1D-local ternary pattern features. Applied Acoustics, 180: 108152. https://doi.org/10.1016/j.apacoust.2021.108152

[13] Boulares, M., Alotaibi, R., AlMansour, A., Barnawi, A. (2021). Cardiovascular disease recognition based on heartbeat segmentation and selection process. International Journal of Environmental Research and Public Health, 18(20): 10952. https://doi.org/10.3390/ijerph182010952

[14] Tariq, Z., Shah, S.K., Lee, Y. (2022). Feature-based Fusion using CNN for Lung and Heart Sound Classification. Sensors, 22(4): 1521. https://doi.org/10.3390/s22041521

[15] Kui, H., Pan, J., Zong, R., Yang, H., Wang, W. (2021). Heart sound classification based on log Mel-frequency spectral coefficients features and convolutional neural networks. Biomedical Signal Processing and Control, 69: 102893. https://doi.org/10.1016/j.bspc.2021.102893

[16] Mannor, S., Bentley, P., Nordehn, G., Coimbra, M. (2011). The PASCAL classifying heart sounds challenge. Book The PASCAL Classifying Heart Sounds Challenge. http://www.peterjbentley.com/heartchallenge/index.html, accessed on 28 February 2022.

[17] Chakir, F., Jilbab, A., Nacir, C., Hammouch, A. (2016). Phonocardiogram signals classification into normal heart sounds and heart murmur sounds. In 2016 11th International Conference on Intelligent Systems: Theories and Applications (SITA), pp. 1-4. https://doi.org/10.1109/SITA.2016.7772311

[18] Skoczylas, A., Stefaniak, P., Anufriiev, S., Jachnik, B. (2021). Belt conveyors rollers diagnostics based on acoustic signal collected using autonomous legged inspection robot. Applied Sciences, 11(5): 2299. https://doi.org/10.3390/app11052299

[19] Sun, Y., Kommers, D., Tan, T., et al. (2020). Automated discomfort detection for premature infants in NICU using time-frequency feature-images and CNNs. In Medical Imaging 2020: Computer-Aided Diagnosis, 11314: 113144B. https://doi.org/10.1117/12.2549250

[20] Keys, R. (1981). Cubic convolution interpolation for digital image processing. IEEE Transactions on Acoustics, Speech, and Signal Processing, 29(6): 1153-1160. https://doi.org/10.1109/TASSP.1981.1163711

[21] Chen, M.J., Huang, C.H., Lee, W.L. (2005). A fast edge-oriented algorithm for image interpolation. Image and Vision Computing, 23(9): 791-798. https://doi.org/10.1016/j.imavis.2005.05.005

[22] Lei, C., Hu, B., Wang, D., Zhang, S., Chen, Z. (2019). A preliminary study on data augmentation of deep learning for image classification. In Proceedings of the 11th Asia-Pacific Symposium on Internetware, pp. 1-6. https://doi.org/10.1145/3361242.3361259

[23] Krenk, S. (1975). On the use of the interpolation polynomial for solutions of singular integral equations.

Quarterly of Applied Mathematics, 32(4): 479-484. https://doi.org/10.1090/qam/474919

[24] Beatson, R.K., Light, W.A., Billings, S. (2001). Fast solution of the radial basis function interpolation equations: Domain decomposition methods. SIAM Journal on Scientific Computing, 22(5): 1717-1740. https://doi.org/10.1137/S1064827599361771

[25] Redmon, J., Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767. https://doi.org/10.48550/arXiv.1804.02767

[26] Dash, M., Liu, H., Yao, J. (1997). Dimensionality reduction of unsupervised data. In Proceedings Ninth IEEE International Conference on Tools with Artificial Intelligence, pp. 532-539. https://doi.org/10.1109/TAI.1997.632300

[27] Lowry, S.R., Isenhour, T.L., Justice, J.B., McLafferty, F.W., Dayringer, H.E., Venkataraghavan, R. (1977). Comparison of various K-nearest neighbor voting schemes with the self-training interpretive and retrieval system for identifying molecular substructures from mass spectral data. Analytical Chemistry, 49(12): 1720-1722. https://doi.org/10.1021/ac50020a022

[28] Joachims, T. (1999). Making large-scale support vector machine learning practical, Advances in Kernel Methods. Support Vector Learning.

[29] Lewis, D.D. (1998). Naive (Bayes) at forty: The independence assumption in information retrieval. In European Conference on Machine Learning, pp. 4-15. https://doi.org/10.1007/BFb0026666

[30] Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., Feinstein, A.R. (1996). A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology, 49(12): 1373-1379. https://doi.org/10.1016/S0895-4356(96)00236-3

[31] Castagna, J.P., Batzle, M.L., Eastwood, R.L. (1985). Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. Geophysics, 50(4): 571-581. https://doi.org/10.1190/1.1441933

[32] Natekin, A., Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in Neurorobotics, 7: 21. https://doi.org/10.3389/fnbot.2013.00021

[33] Chen, T., Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794. https://doi.org/10.1145/2939672.2939785

[34] Alzamzami, F., Hoda, M., El Saddik, A. (2020). Light gradient boosting machine for general sentiment classification on short texts: A comparative evaluation. IEEE Access, 8: 101840-101858. https://doi.org/10.1109/ACCESS.2020.2997330

[35] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. Advances in Neural Information Processing Systems, 31.

[36] Yildirim, M., Cinar, A. (2022). Classification with respect to colon adenocarcinoma and colon benign tissue of colon histopathological images with a new CNN model: MA_ColonNET. International Journal of Imaging Systems and Technology, 32(1): 155-162. https://doi.org/10.1002/ima.22623

[37] Saikumar, K., Rajesh, V., Babu, B.S. (2022). Heart disease detection based on feature fusion technique with augmented classification using deep learning technology. Traitement du Signal, 39(1): 31-42. https://doi.org/10.18280/ts.390104

[38] Cengil, E., Çınar, A., Yıldırım, M. (2022). A hybrid approach for efficient multi-classification of white blood cells based on transfer learning techniques and traditional machine learning methods. Concurrency and Computation: Practice and Experience, 34(6): e6756. https://doi.org/10.1002/cpe.6756

[39] Warriner, D., Michaels, J., Morris, P.D. (2019). Cardiac auscultation: Normal and abnormal. British Journal of Hospital Medicine, 80(2): C28-C31. https://doi.org/10.12968/hmed.2019.80.2.C28

[40] Klocko, D.J., Hanifin, C. (2019). Cardiac auscultation: Using physiologic maneuvers to further identify heart murmurs. Journal of the American Academy of PAs, 32(12): 21-25. https://doi.org/10.1097/01.JAA.0000604856.33701.ad