





Classification of Heart Sounds Using Grey Level Co-occurrence Matrix and Logistic Regression

Istiqomah^{*}, Rizqy Y.D. Wardhana, Achmad Rizal[†]

School of Electrical Engineering, Telkom University, Bandung 40257, Indonesia

Corresponding Author Email: istiqomah@telkomuniversity.ac.id

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.110720>

ABSTRACT

Received: 23 February 2024

Revised: 7 May 2024

Accepted: 15 May 2024

Available online: 31 July 2024

Keywords:

heart sound, GLCM, short-time Fourier transform, logistic regression

The heart is a person's fundamental organ. Heart sounds can help support healthcare workers by aiding in the early diagnosis of irregular heart rhythms. This study developed a system to categorize heart sounds by employing logistic regression as the classifier and the grey-level co-occurrence matrix as the classifier. For this reason, the GLCM technique was assessed in this work for feature extraction in the heart sound categorization. Moreover, the diagnostic heart sound analysis and classification procedure can be greatly improved by visualizing heart sounds using the Grey Level Co-occurrence Matrix (GLCM). The three data classifications for heart sounds are artifact, murmurs, and normal. Moreover, the heart sound is converted into the time-frequency domain using the short-time Fourier transform (STFT). The gray-level co-occurrence matrix approach is a useful tool for extracting the energy distribution in STFT. Dissimilarity, correlation, homogeneity, contrast, energy, and angular second moment (ASM) are the characteristics of the GLCM extraction. With dissimilarity offering the most feature extraction, logistic regression yields an 82% classification accuracy. The AUC value of 0.7 for the murmur class indicated that the feature and classification model had reduced sensitivity, but it performed well for the normal and artifact classes. This is because there are too few datasets for the murmur class. More abnormal class datasets are hoped to be contributed in the future in order to improve the classifier model.

1. INTRODUCTION

Every human being's heart is their central organ. The purpose of the heart is to circulate blood throughout the body. The heart is situated between the fourth and sixth ribs on the left side of the chest. Because the heart beats instinctively and autonomously, it is an organ whose function is not controlled by the brain [1]. The diagnosis results are highly influenced by the subjectivity of the physician, as they need expertise and sensitivity to diagnose normal or abnormal heart sounds, often known as pathological murmurs. An essential diagnostic technique for cardiac disease is auscultation, a low-cost screening technique. Issues with valvular function and heart disease diagnosis are related to heart valves. The range of frequencies for normal heart sounds is 20 Hz to 500 Hz, whereas the range for aberrant heart sounds is up to 1000 Hz. A murmur is the result of an improper valve opening or stenosis, which makes blood flow through a restricted aperture and causes blood to backflow [2]. Digital signal processing techniques have been used in numerous researches to diagnose cardiac problems because heart sound assessment is still highly subjective [3].

Early detection of abnormal heart sounds is needed and can also be used to support healthcare professionals [4]. Several studies have been carried out to develop heart sound

classification applications. Hamidi et al. classified cardiac sound signals using fractal dimension and curve fitting [5]. Using the closest neighbor as a classifier, the approach was evaluated on three datasets and yielded the maximum accuracy of 92%. The fractal approach was also employed by Juniati et al. [6] and Komalasari et al. [7] with a promising accuracy of up to 100% for regular and murmurs. Fauzi et al., meantime, categorizes cardiac sounds using statistical traits and empirical mode decomposition (EMD). When using KNN as a classifier [8], the highest accuracy is 98.2%. Other researchers have also processed cardiac sounds using the time-frequency domain. Wang et al. extracted features using wavelet-time entropy [2]. Excellent classification findings have been obtained from some of these experiments. However, another viewpoint is required to visualize and analyze heart sound texture while conducting research. Similar to the subsequent study, a heart sound spectrogram acquired from STFT was processed using the Grey Level Difference Matrix (GLDM) [9]. 73% accuracy was the best accuracy achieved when using SVM as a classifier. Nonetheless, the findings of this study are still regarded as inadequate and have room for improvement.

GLDM seeks to extract the heart sound spectrogram's texture [9]. Other techniques, such as GLCM, can be applied in addition to GLDM. This is the reason that the GLCM approach was evaluated in this work for feature extraction in the

classification of heart sounds. Furthermore, visualizing heart sounds using GLCM can significantly enhance the diagnostic heart sound analysis and categorization process. GLCM is used for feature extraction in image processing, which is then used for classification with machine learning. The results of previous research are good, especially for biomedical image processing, with an accuracy level above 80% [10-13]. Previous research was also carried out for sound classification with GLCM, with accuracy results of up to 90% [14]. This could be the basis for a similar development of heart sounds to classify abnormal classification. GLCM uses textural features from heart sound spectrograms to discover patterns linked to various heart conditions [15]. To convert the signal from one dimension to two dimensions, heart sounds are converted into the time-frequency domain. After that, normalization is done to make the STFT value equal to the image's pixel value. The transformed image from the heart sound spectrogram is subjected to GLCM, and its properties are computed. It is anticipated that the accuracy that results would be better than the GLDM approach in earlier studies. This study makes use of a number of models created to identify which GLCM feature produces the most accurate model for categorizing heart sounds. All GLCM features such as Dissimilarity, Correlation, Homogeneity, Contrast, ASM and Energy with directions 0°, 45°, 90° and 135°.

2. MATERIAL AND METHODS

The research methodology is presented in Figure 1. STFT is used to convert heart sounds into spectrograms. This spectrogram is then transformed into an 8-bit picture with a 0-255 pixel value range. Then, with directions of 0°, 45°, 90°, and 135° and a one-pixel distance, GLCM is utilized to extract features such as Dissimilarity, Correlation, Homogeneity, Contrast, ASM, and Energy. The GLCM process produces four feature extractions from a single GLCM technique, producing six GLCM methods with 24 feature extractions from all the feature extraction processes. Each GLCM method for all directions, for example, dissimilarity with 0°, 45°, 90° and 135° (four features), is trained with several machine learning methods using GridSearchCV to get the best parameter per model with third cross-validation [16]. The process's primary goal is to choose the best GLCM feature with a validation process machine learning model. KKN, SVM, AdaBoost, Random Forest, Stochastic Gradient Descent, Decision Tree, and Logistic Regression are a few machine-learning techniques used in this research [17]. The final result from all the training processes is that Logistic regression is the best machine learning model.

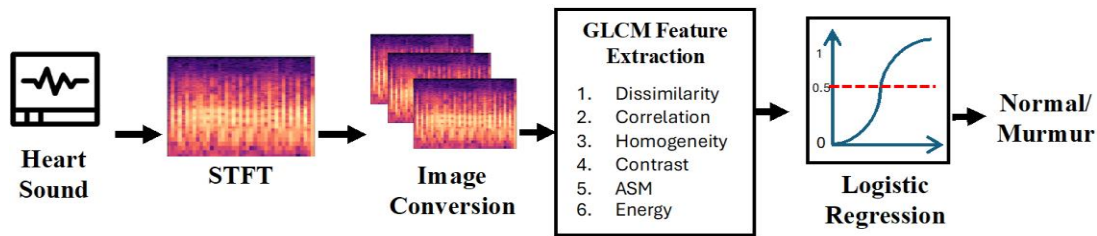


Figure 1. Diagram block of proposed method

2.1 Dataset

The present study employed a dataset to investigate potential anomalies in cardiac auscultation. The dataset utilized in this study was acquired from Kaggle with the purpose of studying the application of machine learning techniques. Digital stethoscopes are utilized to collect recordings of cardiac sounds. The data was obtained from two distinct sources: (A) The general population, who provided their data through the iStethoscope Pro iPhone application, and (B) A hospital study that utilized the Digiscope digital stethoscope [18, 19]. Every recording is appended with labels designating the S1, S2, systole, and diastole phases of the cardiac cycle. The dataset used has been labelled, and there are three sound classes: 1. Normal Heart Sounds 2. Murmur Heart Sounds 3. Artifact. An artifact exhibits functionality when the auditory signal differs from the characteristic heart sounds or heart murmurs. The duration of the dataset ranges from 1 second to 30 seconds. In the dataset pertaining to normal cardiac sounds, there were 320 instances, accounting for 70.32% of the total. Additionally, there were 95 occurrences of murmurs, representing 20.87% of the dataset, while artifacts constituted 40 instances, making up 8.79% of the dataset. Metrics other than accuracy are essential to grasp performance when assessing models for unbalanced datasets. Confusion metrics, accuracy, recall, F1 score ROC curve, in particular, offer a more complex picture of a model's efficacy [20]. The dataset typically contains recordings from multiple subjects

covering a range of ages and health conditions. Both pediatric and adult heart sounds are included to account for variations in heart sound characteristics across different demographics [19].

2.2 Short time Fourier transform dataset

STFT is a widely used technique in signal analysis that enables the examination of signal frequencies and the segmentation of signals into distinct time intervals. The Fast Fourier Transform (FFT) is utilized to convert the segmented signal into the frequency domain. STFT is a signal processing technique that allows for the visualization of an input signal in both the time and frequency domains. This is achieved by applying a window function to the signal, which enables the analysis of specific time intervals within the signal. The mathematical representation of STFT calculation is given by Eq. (1) [21]:

$$XSTFT[m, n] = \sum_{k=0}^{L-1} x[k]g[k-m]e^{-j2\pi nk/L} \quad (1)$$

The signal $x[k]$ is denoted as the input, whereas $g[k]$ represents the L-point window function. The Short-Time Fourier Transform (STFT) is a mathematical operation that applies the Fourier transform to a given signal $x[k]$, utilizing a window function $g[k]$. This research uses the Window function are, the Kaiser-Bessel window, with the number of FFT at 512.

STFT spectrogram converted to an 8-bit picture with a 0-255 pixel value range.

2.3 Feature extraction using GLCM

The GLCM technique is a texture analysis approach that computes the co-occurrence of two pixels separated by a specific distance [22]. The spatial distribution of pixels in an image can be computed using GLCM. The GLCM approach generates a co-occurrence matrix, from which several attributes are computed. A pair of pixels' coordinates are separated from the orientation angle θ by a distance d . Angles are expressed in degrees, and distances are expressed in pixels. Angle forms in angle orientation include 0° , 45° , 90° , and 135° , with a pixel-to-pixel spacing of one in this research. In GLCM, the extraction of characteristics or features is used in this research. The features used are as follows:

$$Energy = \sum_i \sum_j P(i, j)^2 \quad (2)$$

$$Homogeneity = \sum_i \sum_j \frac{P(i, j)}{1 + |i - j|} \quad (3)$$

$$Correlation = \sum_i \sum_j \frac{(i - \mu_i)(j - \mu_j)P(i, j)}{\sigma_i \sigma_j} \quad (4)$$

$$Contrast = \sum_i \sum_j (i - j)^2 P(i, j) \quad (5)$$

$$ASM = \sum_i \sum_j (P(i, j))^2 \quad (6)$$

$$Dissimilarity = \sum_i \sum_j |i - j| \cdot P(i, j) \quad (7)$$

2.4 Logistic regression

In this research, several machine learning methods were used and tested, such as KKN, SVM, AdaBoost, Random Forest, Stochastic Gradient Descent, Decision Tree, and Logistic Regression. The GridSearchCV function is also used to help find the best parameters for each model. The results obtained from pre-research found that simple models, such as logistic regression, were the best models of all those tried. The representation of results described is the best model of logistic regression.

One type of regression technique used for binary classification is called logistic regression. The class output is defined by this approach by utilising the probability of instances that measure with the sigmoid function. The output is predicted using Eq. (8).

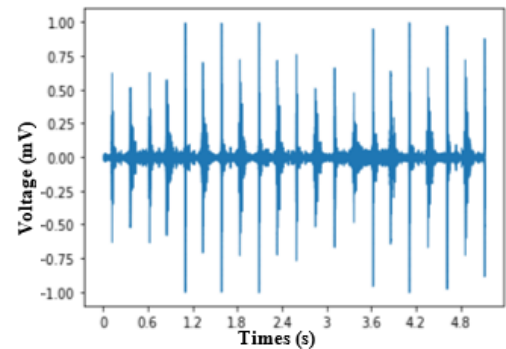
$$\hat{y} = \begin{cases} 0 & \text{if } \hat{p} < 0.5 \\ 1 & \text{if } \hat{p} \geq 0.5 \end{cases} \quad (8)$$

Because the problem in this research is multiclass, the Softmax function is used to decide the classification results from the output value of the function for each class by calculating the probability of each class is shown in Eqs. (9) and (10). The highest probability is the prediction result from the logistic regression model. From the description above, the logistic regression model is simple and is even the basic algorithm of the classification model [23].

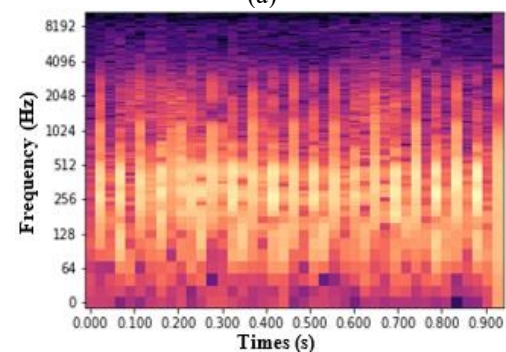
$$S_k(x) = X^T \theta^{(k)} \quad (9)$$

$$\hat{P}_k = \sigma(s(x))_k = \frac{\exp(S_k(x))}{\sum_{j=1}^k \exp(S_j(x))} \quad (10)$$

3. RESULTS AND DISCUSSION

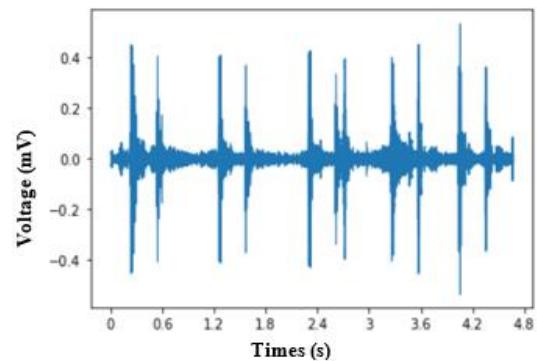


(a)

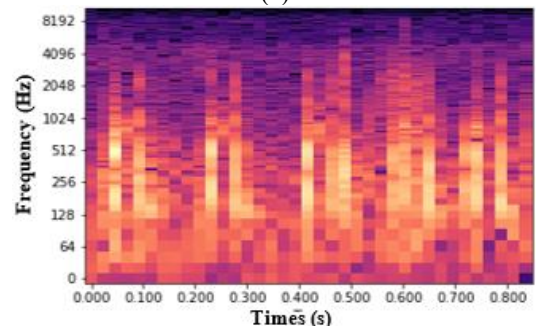


(b)

Figure 2. (a) Normal heart sound; (b) Spectrogram of normal heart sound



(a)



(b)

Figure 3. (a) Murmur heart sound; (b) Spectrogram of murmur heart sound

Figures 2 and 3 present visual representations of both normal cardiac sound signals and murmurs, together with the corresponding spectrums that come from these acoustic phenomena. A spectrogram can be visualized as a two-dimensional image that exhibits a texture indicative of the energy distribution of a signal over both the time and frequency domains. Prior to performing feature extraction using GLCM, a conversion process is undertaken to transform the values obtained from the spectrogram into integer values ranging from 0 to 255, as has been previously demonstrated in prior research [9].

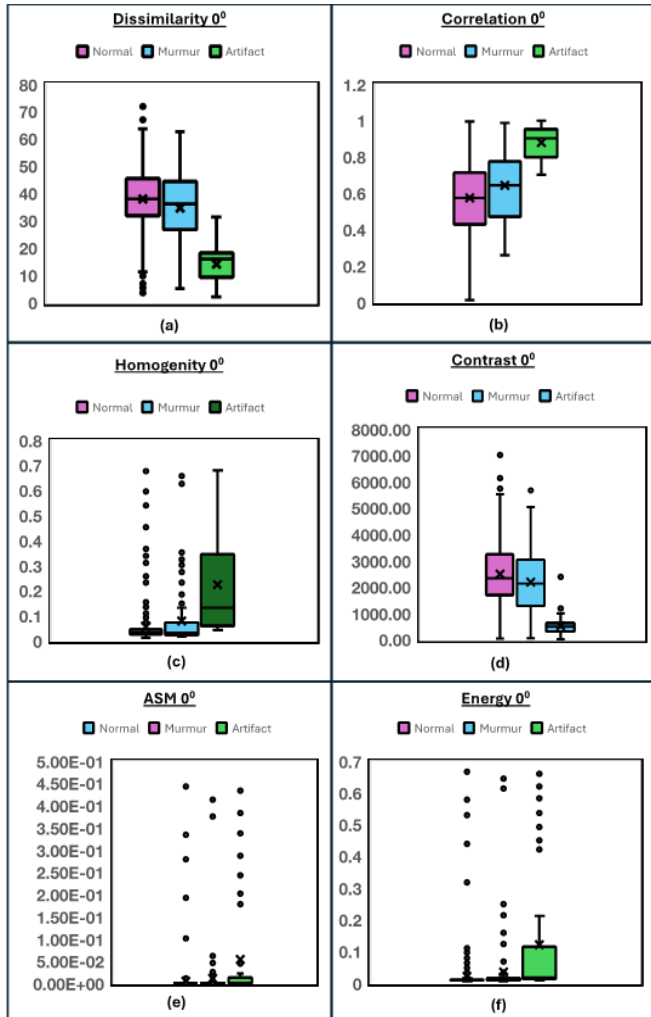


Figure 4. Distribution of several GLCM in the 0° direction (a) Dissimilarity; (b) Correlation; (c) Homogeneity; (d) Contrast; (e) ASM; (f) Energy

As stated earlier, this study utilized six features that were computed at four different angles: 0°, 45°, 90°, and 135°. In Figure 4, a boxplot illustrates the computed values distribution of data for each class for each feature in the 0° orientation. There is a tendency for all distribution data per each class trait to overlap with one another, as shown in each graph for each feature. This pattern of distribution data per class suggests that this could lead to a comparatively reduced level of accuracy.

Table 1 displays the results of classification tests for each feature in each direction. The contrast in the 90° direction produced the highest accuracy when utilizing a single characteristic, at 75%, while the 90° direction dissimilarity gave the second-highest accuracy, at 74%. This outcome is better than earlier studies that extracted features using GLDM

with only two classes of data, yielding an accuracy of 73% for features in a single direction [9]. The pretty good accuracy obtained with only one feature demonstrates the significant potential for increased accuracy by including other parameters.

Table 1. Training results for all features

Feature	Accuracy (%)			
	0°	45°	90°	135°
Dissimilarity	72	70	74	70
Correlation	70	70	70	70
Homogeneity	71	71	70	71
Contrast	71	70	75	70
ASM	70	70	70	70
Energy	70	70	70	70

Next, testing was done with all features (24 features) and all degree of each type of feature extraction (every 4 features), with training data comprising 90% and test data comprising 10% of the 455 total features. It is evident that even with an increase in the number of attributes, the rise is not statistically significant. In addition to using all the features, other classification methods, such as SVM, KNN, and Logistic regression, are also compared. Determining the optimal parameters for all classifier is the next step towards improving accuracy. This is why GridSearchCV, a method for selecting the best parameters from a range of parameters applied to a subset of data, is used with third cross validation [24]. It was found that logistic regression obtained greater accuracy than SVM and KNN. The Dissimilarity and Contrast features have the best accuracy from the previous model test in Table 1, and this combined test also show in Table 2, that the highest accuracy is in the same feature. Dissimilarity had the highest score, 82.64%, among the other categories.

Table 2. All degree GLCM feature extraction accuracy machine learning

Feature	Accuracy		
	SVM	Logistic Regression	KNN
Dissimilarity	72.97%	82.64%	77.14%
Correlation	72.75%	70.33%	72.97%
Homogeneity	70.33%	70.33%	74.51%
Contrast	70.33%	76.49%	76.92%
ASM	69.45%	70.33%	69.23%
Energy	70.33%	70.33%	70.11%
All feature	70.33%	78.03%	76.92%

Dissimilarity is the best feature that can be extracted from GLCM. Therefore, we reserve this feature's use for testing many other performance metrics. The performance results from the all machine learning model using the dissimilarity feature are shown in Figure 5 and Table 3, which in this example include the confusion matrix, precision, recall, and F1 score.

Because each class's data size is not balanced, measurements using these parameters are required. These findings show that although the accuracy of the model with feature dissimilarity is generally not outstanding, the precision, recall, and F1 Score values for the normal class are still excellent when compared to other classes, is shown in Figure 5. However, only the logistic regression method could classify the murmur class well, as seen in Figure 5(a) and Table 3, with a recall value reaching 100%. Meanwhile, for SVM, the model cannot classify this class well with a recall value of even 0%,

as can be seen in Figure 5(b) and Table 3. The artifact class is also not classified well for each model, even though the boxplot in Figure 4 clearly illustrates this class's pattern.

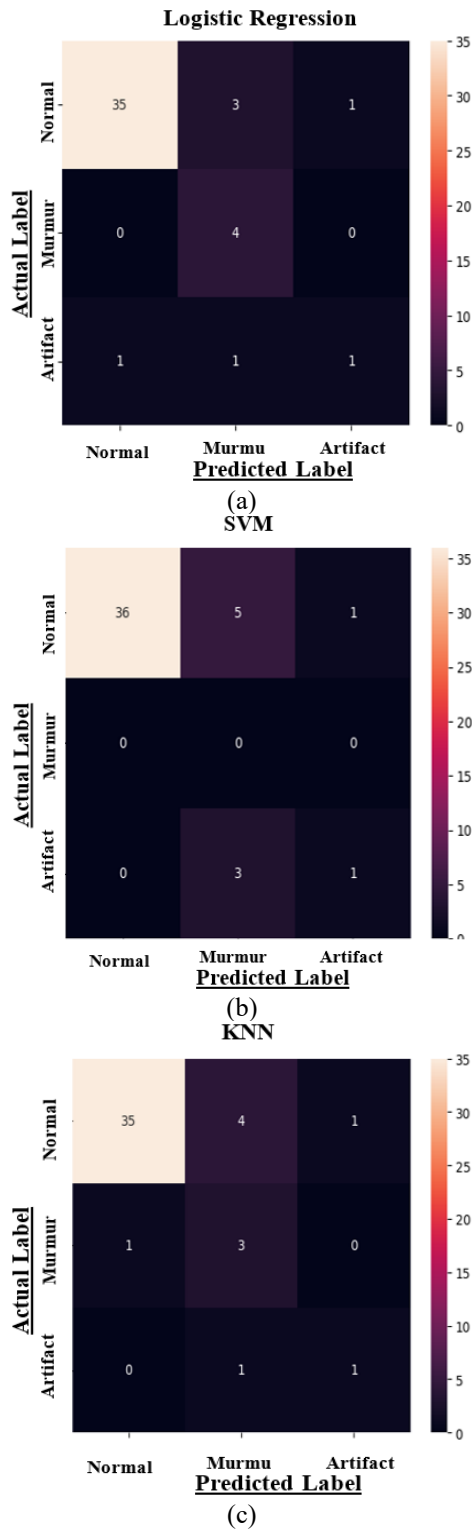


Figure 5. Confusion matrix dissimilarity feature from machine learning method: (a) SVM; (b) Logistic regression; (c) KNN

In Figure 6, other measurement results are carried out using the ROC curve from one class to another (one versus rest classifier) from logistic regression as best machine learning model. From the curve, the area under the ROC curve (AUC) value for the normal and artifact classes is considered very good because it is close to 1, but had lower sensitivity for the

murmur class, with an AUC value of 0.7. The average value of the AUC for each class is still 0.93. These results show that the resulting model is still able to classify each class well, even though the accuracy obtained is only 82%.

Table 3. Performance Softmax regression with dissimilarity feature training results for all classes

Classes	Precision		
	SVM	Logistic Regression	KNN
Normal	100%	97%	97%
Murmur	0%	50%	38%
Artifact	50%	50%	50%

Classes	Recall		
	SVM	Logistic Regression	KNN
Normal	0.86	90%	88%
Murmur	0	100%	75%
Artifact	0.25	33%	50%

Classes	F1-score		
	SVM	Logistic Regression	KNN
Normal	92%	93%	92%
Murmur	0%	67%	50%
Artifact	33%	40%	50%

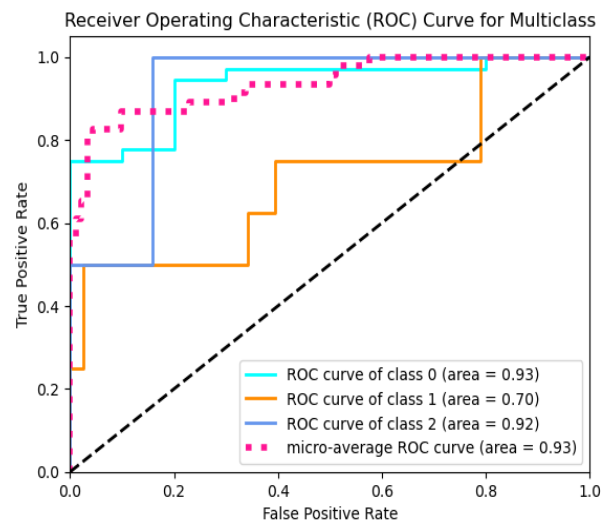


Figure 6. ROC for one versus rest for each class

The results of the research above are better than those of research on lung sound classification using GLDM, which had the highest accuracy of 73% [9]. This method uses the distribution of signal energy in the time-frequency domain, which is analyzed using the texture analysis method. The spectrogram analysis technique differs from that of Shanthakumari and Priya [25]. Shanthakumari uses morphological features in spectrograms, such as energy, entropy, variance, and waveform length. The weakness of texture analysis in heart sound analysis is the low image contrast because the heart sound energy is relatively low. Apart from that, the frequency of heart sounds is also low. The image produced from a spectrogram is also greatly influenced by spectrogram parameters, such as N-FFT, window width, and sampling frequency [24]. The resulting combination of parameters and characteristics can increase accuracy in subsequent research. Apart from that, the feature subset selection method can be carried out before exploring more advanced classification methods such as deep learning.

Based on all of the tests mentioned above, the GLCM dissimilarity feature generates a machine-learning model that is reasonably good. The current features can effectively

identify all classes, including normal and artifacts, but they have poor sensitivity for class murmurs. The similarity in feature data distribution patterns between the normal and murmur classes accounts for this performance. Nevertheless, because there aren't enough datasets, the murmur class isn't very sensitive. This needs to be enhanced since high sensitivity to abnormal classes is necessary for real-world heart sound classification systems to detect them accurately [26]. The research is limited by the balanced dataset for each class, which prevents the classification model's accuracy from being deemed excessively high even when other metrics like precision, recall, AUC, and F1 are deemed to be rather powerful. Further datasets will be required in the future, particularly for the murmur class. A strategy can be carried out by collaborating with healthcare centers to collect datasets, especially for murmur classes. Another way is to use Augmentation Techniques to add murmur class datasets artificially using GAN or other methods [27].

4. CONCLUSION

This study applies GLCM to the heart sound signal spectrogram and suggests a feature extraction strategy for heart sound categorization. The present study employs the following GLCM characteristics: dissimilarity, correlation, homogeneity, contrast, ASM, and energy. Tests have shown that, despite the cleaning procedure being completed, every feature in GLCM overlaps for normal and murmur class. As a result, the accuracy of the results could be improved. The best features from the GLCM were found to be dissimilar and contrasted at degree 90, with the highest accuracy of 75%, when tested with one degree of one feature. Based on these findings, the dissimilarity feature yielded the highest accuracy of 82.64% when a machine learning model was built utilizing data aggregated at four degrees from each feature. Because the dataset is not balanced, measuring model performance with other parameters is used, such as precision, recall, AUC, and F1. It was found that the model only had high sensitivity to the normal and artifact classes but was still less sensitive to the murmur class, which is the most important class that must be classified in heart sound applications. This performance is due to the similarity of feature data distribution patterns for the normal and murmur classes and the lack of datasets for the murmur class, so this class has low sensitivity. For future research, it is hoped that more datasets will be added for the murmur class so that it can improve the resulting model and increase sensitivity to the murmur class. The application can display STFT and GLCM visualizations of the texture of heart sounds as additional information for healthcare professionals to analyze.

REFERENCES

[1] Kristomo, D., Hidayat, R., Soesanti, I., Kusjani, A. (2016). Heart sound feature extraction and classification using autoregressive power spectral density (AR-PSD) and statistics features. *AIP Conference Proceedings*, 1755(1): 090007. <https://doi.org/10.1063/1.4958525>

[2] Wang, Y., Li, W., Zhou, J., Li, X., Pu, Y. (2014). Identification of the normal and abnormal heart sounds using wavelet-time entropy features based on OMS-WPD. *Future Generation Computer Systems*, 37: 488-

495. <https://doi.org/10.1016/j.future.2014.02.009>

[3] Dwivedi, A.K., Imtiaz, S.A., Rodriguez-Villegas, E. (2018). Algorithms for automatic analysis and classification of heart sounds—A systematic review. *IEEE Access*, 7: 8316-8345. <https://doi.org/10.1109/ACCESS.2018.2889437>

[4] Chen, D., Xuan, W., Gu, Y., et al. (2022). Automatic classification of normal–Abnormal heart sounds using convolution neural network and long-short term memory. *Electronics*, 11(8): 1246. <https://doi.org/10.3390/electronics11081246>

[5] Hamidi, M., Ghassemian, H., Imani, M. (2018). Classification of heart sound signal using curve fitting and fractal dimension. *Biomedical Signal Processing and Control*, 39: 351-359. <https://doi.org/10.1016/j.bspc.2017.08.002>

[6] Juniati, D., Khotimah, C., Wardani, D.E.K., Budayasa, K. (2018). Fractal dimension to classify the heart sound recordings with KNN and fuzzy c-mean clustering methods. *Journal of Physics: Conference Series*, 953: 012202. <https://doi.org/10.1088/1742-6596/953/1/012202>

[7] Komalasari, R., Rizal, A., Suratman, F.Y. (2020). Classification of normal and murmur hearts sound using the fractal method. *International Journal*, 9(5): 8178-8183. <https://doi.org/10.30534/ijatse/2020/181952020>

[8] Fauzi, H., Rizal, A., Oktarianto, A., Said, Z. (2023). Classification of normal and abnormal heart sounds using empirical mode decomposition and first order statistic. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 5(2): 82-88. <https://doi.org/10.35882/jeeemi.v5i2.287>

[9] Rizal, A., Handzah, V.A.P., Kusuma, P. D. (2022). Heart sounds classification using short-time Fourier transform and gray level difference method. *Ingénierie des Systèmes d'Information*, 27(3): 369-376. <https://doi.org/10.18280/isi.270302>

[10] Chen, Q., Agu, E. (2015). Exploring statistical GLCM texture features for classifying food images. In 2015 International Conference on Healthcare Informatics, Dallas, USA, pp. 453-453. <https://doi.org/10.1109/ICHI.2015.71>

[11] Mall, P.K., Singh, P.K., Yadav, D. (2019). GLCM based feature extraction and medical X-RAY image classification using machine learning techniques. In 2019 IEEE Conference on Information and Communication Technology, Allahabad, India, pp. 1-6. <https://doi.org/10.1109/CICT48419.2019.9066263>

[12] Tan, J., Gao, Y., Liang, Z., et al. (2019). 3D-GLCM CNN: A 3-dimensional gray-level Co-occurrence matrix-based CNN model for polyp classification via CT colonography. *IEEE Transactions on Medical Imaging*, 39(6): 2013-2024. <https://doi.org/10.1109/TMI.2019.2963177>

[13] Shabu, S.J., Jayakumar, C. (2020). Brain tumor classification with MRI brain images using 2-level GLCM features and sparse representation based segmentation. In 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, pp. 793-799. <https://doi.org/10.1109/ICISS49785.2020.9315971>

[14] Kazenzeza, M., Amenyedzi, D.K., Vodacek, A., Hanyurwimfura, D., Ndashimye, E. (2023). Bird sound classification using GLCM features and LightGBM

- applied to farm monitoring. In 2023 11th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC), Chongqing, China, pp. 20-24. <https://doi.org/10.1109/ICWOC57905.2023.10200229>
- [15] Yuenyong, S., Nishihara, A., Kongprawechnon, W., Tungpimolrut, K. (2011). A framework for automatic heart sound analysis without segmentation. *Biomedical Engineering Online*, 10: 13. <https://doi.org/10.1186/1475-925X-10-13>
- [16] Anggoro, D.A., Afdallah, N.A. (2022). Grid search CV implementation in random forest algorithm to improve accuracy of breast cancer data. *International Journal on Advanced Science, Engineering and Information Technology*, 12(2): 515-520. <https://doi.org/10.18517/ijaseit.12.2.15487>.
- [17] Géron, A. (2022). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc.
- [18] Heartbeat Sounds. <https://www.kaggle.com/datasets/kinguistics/heartbeat-sounds>, accessed on Sep. 25, 2023.
- [19] Bentley, P., Nordehn, G., Coimbra, M., Mannor, S., Getz, R. (2024). Classifying heart sounds challenge. <https://istethoscope.peterjbentley.com/heartchallenge/index.html#downloads>.
- [20] Paula, B., Torgo, L., Ribeiro, R. (2015). A survey of predictive modelling under imbalanced distributions. *arXiv preprint arXiv, 1505(01658)*.
- [21] Samiee, K., Kovacs, P., Gabbouj, M. (2014). Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform. *IEEE Transactions on Biomedical Engineering*, 62(2): 541-552. <https://doi.org/10.1109/TBME.2014.2360101>
- [22] Haralick, R.M., Shanmugam, K., Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6): 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>
- [23] James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). *An Introduction to Statistical Learning*. New York: Springer.
- [24] Li, J.R., Hong, Y. (2020). Wheeze detection method based on spectrogram entropy analysis. *ACTA ACUSTICA*, 45: 131-136.
- [25] Shanthakumari, G., Priya, E. (2022). Spectrogram-based detection of crackles from lung sounds. In 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, pp. 1-6. <https://doi.org/10.1109/IC3IoT53935.2022.9768007>
- [26] Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Computing and Applications*, 29: 685-693. <https://doi.org/10.1007/s00521-016-2604-1>
- [27] Abayomi-Alli, O.O., Damaševičius, R., Qazi, A., Adedoyin-Olowe, M., Misra, S. (2022). Data augmentation and deep learning methods in sound classification: A systematic review. *Electronics*, 11(22): 3795. <https://doi.org/10.3390/electronics11223795>