COVID-19 Chest X-Ray Detection Performance through Variations of Wavelets Basis Function

I Gusti Ayu Agung Diatri Indradewi[1], Ni Wayan Sumartini Saraswati[2], Ni Wayan Wardani[3]
[1]Universitas Pendidikan Ganesha,Indonesia
[2]STMIK STIKOM Bali, Indonesia

Article Info	ABSTRACT
	Our previous work regarding the X-Ray detection of COVID-19 using Haar wavelet feature extraction and the Support Vector Machines (SVM) classification machine has shown that the combination of the two methods can detect COVID-19 well but then the question arises whether the Haar wavelet is the best wavelet method. So that in this study we conducted experiments on several wavelet methods such as biorthogonal, coiflet, Daubechies, haar, and symlets for chest X-Ray feature extraction with the same dataset. The results of the feature extraction are then classified using SVM and measure the quality of the classification model with parameters of accuracy, error rate, recall, specification, and precision. The results showed that the Daubechies wavelet gave the best performance for all classification quality parameters. The Daubechies wavelet transformation gave 95.47% accuracy, 4.53% error rate, 98.75% recall, 92.19% specificity, and 93.45% precision.

Corresponding Author:

Ni Wayan Sumartini Saraswati,
Department of Informatics Engineering,
STMIK STIKOM Indonesia,
Email: sumartini.saraswati@stiki-indonesia.ac.id

1.   INTRODUCTION

No one ever expected that in early 2020 the world began to be hit by a pandemic caused by the corona virus or better known as COVID-19. The pandemic caused researchers in various fields of science to compete to research COVID-19 [1]. This virus as a whole attacks the human respiratory system, so that the lungs are the organs most affected by COVID-19.

Digital image processing is one of the fields of science that can be implemented in the health sector, especially in terms of detecting diseases based on input data in the form of digital images. X-ray image, especially chest X-ray image, is one type of image that is commonly used as input data in research related to respiratory disease detection. Yee's research supports the use of chest X-ray images and machine learning as input data for the pneumonia diagnosis system [2]. The diagnostic system in this study was built using a convolutional neural network (CNN) based on feature extraction. The features that have been extracted are then used to train three classification algorithm models using the Kaggle dataset.

Early features of serious lung disease based on classified chest X-ray images were investigated in [3]. The features of the chest x-ray image were extracted by gray level co-occurrence matrix (GLCM) and a probabilistic neural network were used as a classifier. The probabilistic neural network was used as a classifier into normal and abnormal classes. Detection of the presence of pneumonia clouds on chest X-Ray images is done using image processing techniques [4]. Indigenous algorithms have been developed to crop and extract lung regions from images. The Otsu thresholding method is used to detect pneumonia clouds by separating the healthy lung from the pneumonia-infected lung. Chest X-Ray image data was used to detect the presence of pneumothorax disease in this study [5]. In the first stage, the local binary pattern (LBP) is used to extract features from the lung image. The support vector machine (SVM) is then used for the classification of the pneumothorax. In the second stage, the detection method is based on multiscale intensity texture segmentation by removing the background and noise on the chest X-Ray image to segment the abnormal lung areas. Another study that uses chest X-Ray images is research on feature extraction on chest X-Ray images to analyze tuberculosis (TB) [6]. Texture and shape-based features are extracted using the concept of image processing. The extracted features were then analyzed using principal component analysis (PCA) and kernel principal component analysis (kPCA) techniques. DarkNet model used on [7] as a classifier for chest X-Ray images obtained in real-time. The model is implemented using 17 convolutional layers and uses different filtering on each layer. The results of these studies can be used by radiologists to validate the results of initial examinations and can also be used via the cloud to immediately examine patients. Based on some previous studies that used chest X-Ray images to detect respiratory diseases, in this study, chest X-Ray images were also used as input data.

Disease detection based on input data in the form of images requires a feature extraction process to obtain special features that can be used to perform classifications related to the diagnosis of the disease under study. One method that can be used to perform feature extraction is wavelet transform. Wavelets are a widely used choice for data analysis and have been applied in various fields from biology to engineering. The discrete wavelet transform (DWT)-based feature extraction has a high discriminating power in defining the target idea [8]. Discrete wavelet transform (DWT) was used on [9] to extract features from MRI images of the brain. DWT is an important tool used to decompose images into different resolution levels so that significant features can be extracted from the input image. The coefficients generated from the DWT are used as feature vectors which are then used as input vectors for the classification algorithm.

The reliability of the wavelet transform in feature extraction is shown in [10]. The study compared transformation-based image texture analysis to evaluate the quality of bananas. Based on this research, it is known that the wavelet transform shows the most reliable results for all reference parameters, which is then followed by the Tamura and Gabor transformations. Other research related to wavelet transform is [11]. In this study, multiple sclerosis was detected using biorthogonal wavelet transform, RBF kernel principal component analysis, and logistic regression. Based on the statistical analysis of the detection results, it was found that the combination of these methods far outperformed the five previous studies which were designated as state of the art.

In previous research [12], the author has used the Haar wavelet to generate feature vectors which are then used to detect COVID-19 on chest X-Ray images. This study shows that the classification model accuracy is 93.91%, the error rate is 6.09%, the recall is 98.75%, the specificity is 89.06%, and the precision is 91.26%. The vertical feature coefficient is the best feature coefficient in the classification model that is built because it produces the best values for the accuracy and recall variables. This feature coefficient also performs well on precision measurements. However, the research that the author has done is limited to using Haar wavelets for the feature extraction stage, the question arises whether Haar wavelets are the best method of chest X-Ray feature extraction for COVID-19 detection so that in this study feature extraction was carried out using other types of wavelets to determine the type of wavelet capable of producing the best performance classification based on accuracy, error rate, recall, specificity, and precision. The types of wavelet transform used in this study consist of biorthogonal, coiflet, Daubechies, haar, and symlets. The result of feature extraction is in the form of a feature vector which becomes the input data for the SVM classifier.

In the case of disease detection through images, the special features generated at the feature extraction stage are used as input data at the classification stage. One of the classification methods that are proven to have good classification abilities is the support

vector machine (SVM). Based on research comparing SVM and random forest (RF) classification methods in remote sensing image classification, one conclusion is obtained which states that the average classification accuracy value of SVM outperforms RF when the data used contains many features [13]. Another research is about defect detection in logistic packaging boxes [14] using SVM as a classifier. The simulation showed that the approach employed can detect two types of defects that commonly occur in logistic packaging boxes with higher accuracy values and less computation time. Both of these are under industry requirements regarding the classification and detection of defects in machine vision detection systems. Research conducted by [15], using the SVM method to classify the port wine stain (PWS). Classification is carried out on the quantitative features of normal skin images and skin images containing PWS tissue. Input image data was recorded from optical coherence tomography (OCT). The results obtained from this study indicate that, when the classification is carried out with a complete feature set, the accuracy value is 92.7%, the sensitivity value is 92.3%, and the specificity value is 93.8%. Research on the classification of cape gooseberry fruit based on maturity level used four machine learning techniques consisting of artificial neural networks, SVM, decision trees, and K-nearest neighbor done in [16]. The study compared the results of the classification carried out by the four machine learning techniques. Based on the results obtained, it is known that the classification model based on the L * a * b * color space and the SVM classification method shows the highest f-measure value regardless of the color space used.

COVID-19 detection was carried out in this study using data from chest X-ray scans as input. A wavelet transform is used to extract features from the input data. Biorthogonal, coiflet, daubechies, haar, and symlets are the wavelet transformations employed in this study. The feature vector obtained as a consequence of feature extraction is used as input data for the SVM classification engine. The accuracy, error rate, recall, specificity, and precision of the final categorization are all evaluated. In comparison to earlier works [11], the contribution of this study is to give a comparison of multiple wavelet basis functions that can provide the greatest classification performance in cases of COVID-19 detection using chest X-ray images as input data. The study's findings are provided in a systematic format that includes a research method, results and analysis, and a conclusion.

2. RESEARCH METHOD

2.1. Dataset

Chest X-Ray images in the dataset are categorized into normal chest X-Ray images and chest X-Ray images diagnosed with COVID-19, respectively represented by Figure 1. and Figure 2. Normal chest X-Ray image obtained from Ganesha General Hospital located in Gianyar, Bali. Chest X-Ray image of diagnosed COVID-19 downloaded from [17]. Visually, the two types of images are distinguished based on the presence or absence of a fog area covering the lung area. If there is a fog area in the image, then this image is a chest X-Ray image diagnosed with COVID-19. To ensure the validity of all data used, an expert (doctor) has been involved to verify and validate the data. The total amount of data used in this study was 164 data consisting of 82 normal chest X-Ray images and 82 chest X-Ray images diagnosed with COVID-19.

Figure 1. Normal Chest X-Ray

Figure 2. COVID-19 Chest X-Ray

2.2. Proposed Method

The stages used in this study are shown in Figure. 3. In the initial stage, preprocessing is carried out on all image data used. The purpose of preprocessing is to obtain image data with uniform characteristics so that it will facilitate the process at a later stage. The image preprocessing then goes through the feature extraction stage using wavelet transform. This study examines the classification results of the feature vectors produced by five types of wavelets consisting of haar (the type of wavelet used in previous studies) [12]), biorthogonal, coiflet, daubechies, and symlets. Feature extraction using wavelet transform aims to obtain feature vectors from normal chest X-Ray images and chest X-Ray images diagnosed with COVID-19. The resulting feature vectors are then divided using the k-fold cross-validation method to balance the composition of training data and test data [18]. This feature vector will be used at the classification stage such as in research [19], so that the SVM method can distinguish the two types of images.
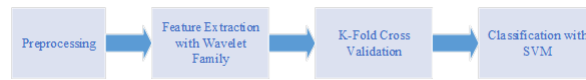
Figure 3. Stages of Proposed Method

2.3. Preprocessing

The preprocessing stage was conducted in this study to equalized the characteristics of the image data used. Image data with the position of the lung object that is not ideal, through the first cropping process. Furthermore, all image data, both normal chest X-Ray and chest X-Ray images diagnosed with COVID-19 are resized to 160×160 pixels. The initial format of the image data is in the form of a 24-bit RGB image, so the next process is to convert a 24-bit RGB image into an 8-bit grayscale image using the weighted summation of the R, G, and B components [20] as stated in (1):

$$f(x) = 0.2989 * R + 0.5870 * G + 0.1140 * B \tag{1}$$

2.4. Wavelet Transformation

Wavelet transformation is a method used to present data, functions, or operators into different frequency components. The frequency components are then analyzed based on the resolution with the appropriate scale [12]. When digital images are processed at multiple resolutions, discrete wavelet transform (DWT) was used [21]. DWT can extract local features by separating the feature vector components based on time and scale. Wavelet analysis involves two elements consisting of approximation and detail components [22]. In general, DWT is an image decomposition at the image subband frequency. DWT implementation is done by passing high frequency and low frequency signals to the image as shown in Figure. 4 [23].
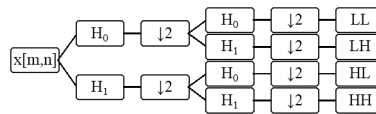

Figure 4. Illustration of 1-level 2D Wavelet Transformation (Source:[24])

Table 1 compares the parameter characteristics of the wavelet basis functions used in this study, namely haar, daubechies, biorthogonal, coiflets, and symlets. The five wavelet basis functions are used to generate feature vectors through the feature extraction process.

Table 1. Comparison of Parameter Characteristics of Wavelet Basis Functions (Source:[25])

Parameter Characteristics	Wavelet Basis Function				
	Haar	Daubechies	Biorthogonal	Coiflets	Symlets
Orthogonality	Yes	Yes	No	Yes	Yes
Biorthogonality	Yes	Yes	Yes	Yes	Yes
Compact Support	Yes	Yes	Yes	Yes	Yes
CWT	Yes	Yes	Yes	Yes	Yes
DWT	Yes	Yes	Yes	Yes	Yes
Support Width	1	2N-1	Reconstitution:2Nr+1 Decomposition:2Nd+1	6N-1	2N-1
Filter Length	2	2N	Max(2Nr,2Nd) +2	6N	2N
Symmetry	Symmetry	Approximate Symmetry	Dissymmetry Dissymmetry	Approximate Symmetry	Approximate Symmetry
Vanishing Moment of Wavelet Function	1	N	Nr-1	2N	N
Vanishing Moment of Scaling Function	-	-	-	2N-1	-

The size of the chest X-Ray image that has passed the preprocessing stage is 160×160 pixels. The preprocessed image is then processed with the basic functions of wavelet haar, daubechies (db4), biorthogonal (bior3.5), coiflets (coif3), and symlets (sym4) using the MATLAB application to generate feature vectors. Figure 5 is an example of the results of wavelet decomposition of normal chest X-Ray images using wavelet haar level 1.The image size after the level 1 wavelet decomposition process changes to $\frac{1}{2}$ times the original size, namely 80×80 pixels.
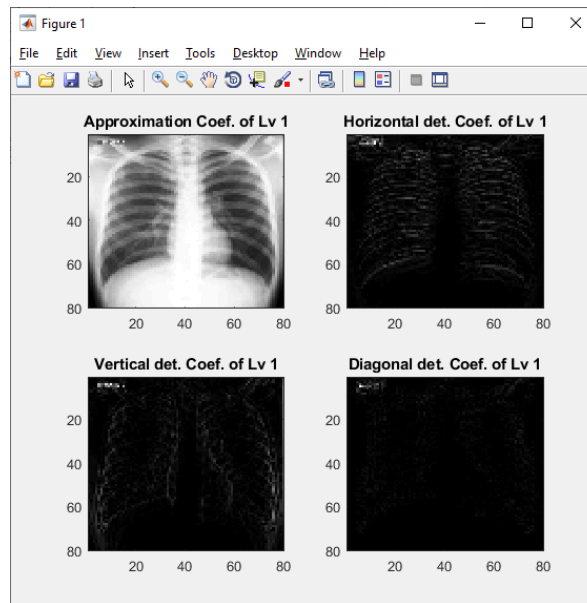

Figure 5. Result of Level 1 Haar Wavelet Decomposition in Normal Chest X-Ray Image

Meanwhile, Figure 6 is an example of the results of wavelet decomposition of chest X-Ray images infected with COVID-19 using wavelet haar level 1.
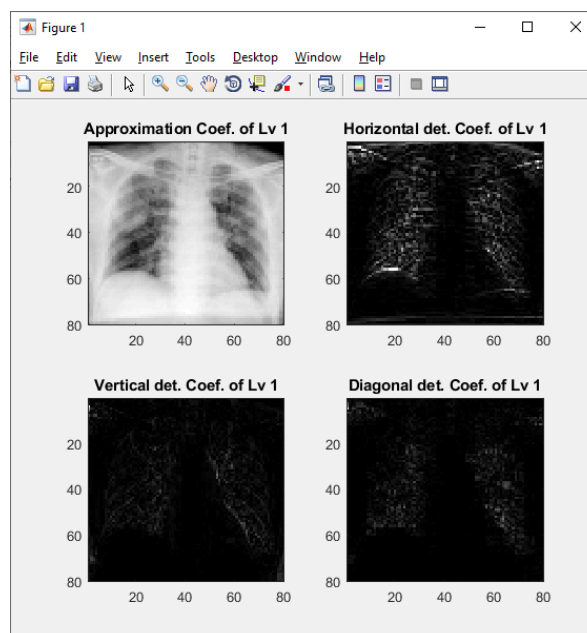

Figure 6. Result of Level 1 Haar Wavelet Decomposition in Covid Chest X-Ray Image

The four resulting images are then converted into vector shapes which are used as feature vectors. The resulting feature vector consists of the approximate coefficient, vertical detail coefficient, horizontal detail coefficient, and diagonal detail coefficient. All feature vectors from each of these coefficients are then used in the classification to determine the chest X-Ray image which is categorized as normal or diagnosed with COVID-19.

2.5. K-Fold Cross-Validation

One form of validation that uses k values to divide training data and test data is k-fold cross-validation. The main objective of this method is to keep the amount of training data and test data balanced such that overfitting does not occur. This method can guarantee that the classification algorithm is applied to each data in the dataset and that all data in the dataset is tested as training and testing data [18]. K-fold cross-validation works by dividing the initial sample size into k subsamples. One subsample is utilized for testing, while the other k-1 subsample is used for training. Cross-validation is carried out k times, with each subsample checked once [25]. Based on the first test, the accuracy value can be found using (2).

$$accuracy = \frac{t}{n} \tag{2}$$

where t is the number of samples that are true when tested, while n is the total number of samples in the dataset [18].

10-fold cross-validation is a cross-validation method that is very commonly used [25]. Therefore, in this study, the k value used for cross-validation is 10. The classification experiment was carried out using a feature vector consisting of the approximate coefficient, the vertical detail coefficient, the horizontal detail coefficient, and the diagonal detail coefficient. In each group of feature vectors, the training data and test data were divided using 10-fold cross-validation. The feature vector data used as test data is divided into 16 first datasets, 16 second datasets, and so on up to 16 tenth datasets. Characteristic vector data that is not used as test data is used as training data that forms a classification model.

2.6. Support Vector Machine

Support vector machine (SVM) is a classification machine based on linear or nonlinear classification functions that can distinguish training and test datasets based on two or more classes [26, 27]. SVM transforms data into a form that produces the widest plane separating the two data classes [28–30] as shown in Figure 7.
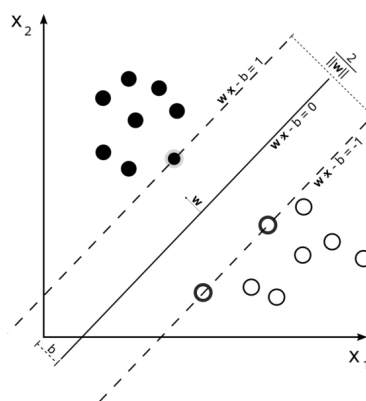

Figure 7. Data Classification with SVM

The SVM classification model used in this study is the same as previous research, which is built using the MATLAB application with the kernel type used as a linear kernel [12]. The process consists of two stages. The first stage is the formation of a classification model with training data input which will form a support vector and hyperplane. The second stage is the testing phase where the classification results are tested based on the test data and the previously formed models.

2.7. Evaluation Measures for Classification Model

To calculate the level of accuracy of the classifier, the accuracy measurement is used on [12, 31, 32], and [33] is a confusion matrix. The confusion matrix element consists of True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN). TP and TN stated that the classifier was capable of correct recognition. This means that positive tuples are recognized as such, whereas negative tuples are recognized as such. FP and FN, on the other hand, indicated that the classifier machine made an error while identifying tuples, with negative tuples being identified as positive and positive tuples being identified as negative. In this study, normal chest X-Ray images represent positive tuples, while chest X-Ray images diagnosed with COVID-19 represent negative tuples. The evaluation measure of the classification model used in this study is shown in Table 2.

Table 2. Classification Model Measurement Tools

Measures	Formulas
Accuracy	$\frac{TP+TN}{P+N}$
Error rate	$\frac{FP+FN}{P+N}$
Recall	$\frac{TP}{P}$
Specivicity	$\frac{TN}{N}$
Precision	$\frac{TP}{TP+FP}$

Where:
P = number of positive test data/normal chest X-ray
N = number of negative test data/COVID-19 chest X-ray.

3. RESULT AND ANALYSIS

The classification results with the SVM classifier show good performance based on the accuracy, error rate, recall, specificity, and precision values. Table 3 shows the results of the performance of the SVM classification machine in the case of COVID-19 detection using input data in the form of chest X-Ray images.

Table 3. Measurement Result of Classification Model

Wavelet Basis Function	Measurement Tools				
	Accuracy	Error Rate	Recall	Specificity	Precision
Haar	93.91%	6.09%	98.75%	89.06%	91.26%
Daubechies (db4)	95.47%	4.53%	98.75%	92.19%	93.45%
Biorthogonal (bior3.5)	94.84%	5.16%	98.13%	91.56%	92.97%
Coiflets (coif3)	95.00%	5.00%	98.13%	91.88%	93.20%
Symlets (sym4)	94.84%	5.16%	98.75%	90.94%	92.55%
Average	94.81%	5.19%	98.50%	91.13%	92.69%

The accuracy, error rate, recall, specificity, and precision values of each wavelet basis function shown in Table 3 are the average values of each measurement tool in each classification experiment carried out using the 10-fold cross-validation method. Based on the performance results in Table 3, it can be seen that the wavelet basis function with SVM can classify chest X-Ray images diagnosed with COVID-19 well. This can be seen from the average value of the overall accuracy of the wavelet basis function is 94.81%. This accuracy is higher than [12]. The recognition rate for normal chest X-Ray images for all datasets was very good, represented by the average recall value, which was 98.50%. The recognition rate for chest X-Ray images diagnosed with COVID-19 for the entire dataset shows good results, although the values are lower than the recognition rate for normal chest X-Ray images. This can be seen from the specificity value of 91.13%.

3.1. Accuracy

The quality of the classification model used can be seen from the accuracy value obtained. Based on Figure 8 it can be seen that the highest accuracy value is produced by wavelet daubechies with 95.47%, while the lowest accuracy value is produced by haar wavelets with 93.91%. These results indicate that the classification model using the feature vector generated by the daubechies wavelet can classify chest X-Ray images diagnosed with COVID-19 and normal chest X-Ray images better when compared to other types of wavelet basis function used in this study.
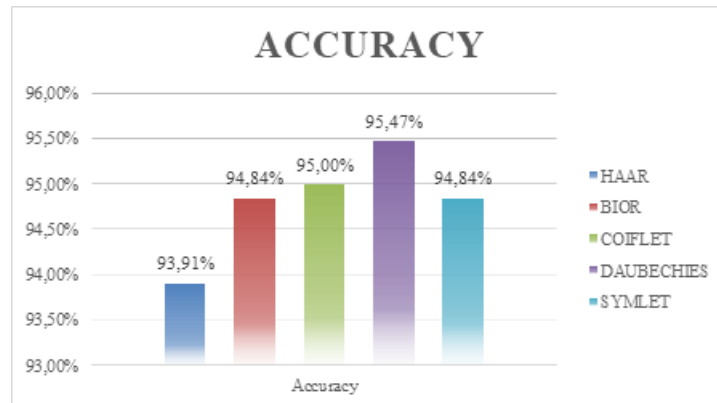

Figure 8. Comparison of The Accuracy Value of Each Wavelet Basis Function

3.2. Error Rate

When the accuracy value obtained is high, the error rate will be low. The relationship between the accuracy value and the error rate is inversely proportional, this can be seen in Figure 9 If the accuracy value, wavelet haar produces the lowest percentage, then at the error rate value, wavelet haar produces the highest percentage, which is 6.09%. Daubechies wavelet produces the lowest error rate, which is 4.53%. This shows that the feature vector produced by using haar wavelets has the highest number of misclassification, both errors in classifying chest X-Ray images diagnosed with COVID-19 and errors in classifying normal chest X-Ray images. In contrast, the feature vector generated using daubechies wavelet has the lowest number of misclassification.

Figure 9. Comparison of The Error Rate of Each Wavelet Basis Function

3.3. Recall

The recall value shows that the classification model used can recognize positive data sets as positive data, or in other words, the classification model can recognize normal chest X-Ray images. Figure 10 shows that haar and daubechies wavelets have the highest percentage of recall values when compared to other types of wavelet basis functions, namely 98.75%. The lowest recall value is produced by bior and coiflet wavelets, which is 98.13%. However, on average, the recall value of all types of wavelet basis functions

used in this study has a very high percentage, as shown in Table 3, which is 98.50%. The average percentage of recall values indicates that the five types of wavelet basis functions used in this study have excellent capabilities in terms of recognizing normal chest X-Ray images from positive data sets.
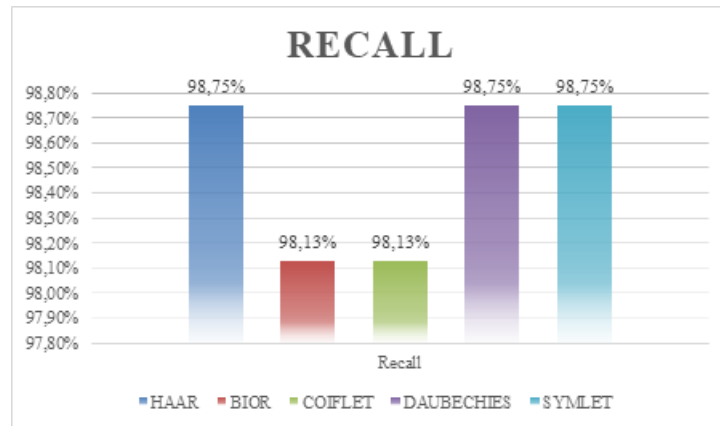

Figure 10. Comparison of The Recall of Each Wavelet Basis Function

3.4. Specificity

The ability of the classification model to recognize negative data sets as negative data is measured by the value of specificity. This value in this study illustrates that the classification model used can recognize a set of chest X-Ray images diagnosed with COVID-19. In Figure 11 it can be seen that the feature vector produced by daubechies wavelets has the highest percentage value of specificity, namely 92.19%, followed by coiflet, biorthogonal, and symlets wavelets. Haar wavelet has the lowest specificity value, which is 89.06%. When compared with the percentage of recall values, this shows that the feature vector produced by haar wavelets only has the equivalent of daubechies wavelet when recognizing a set of normal chest X-Ray images.
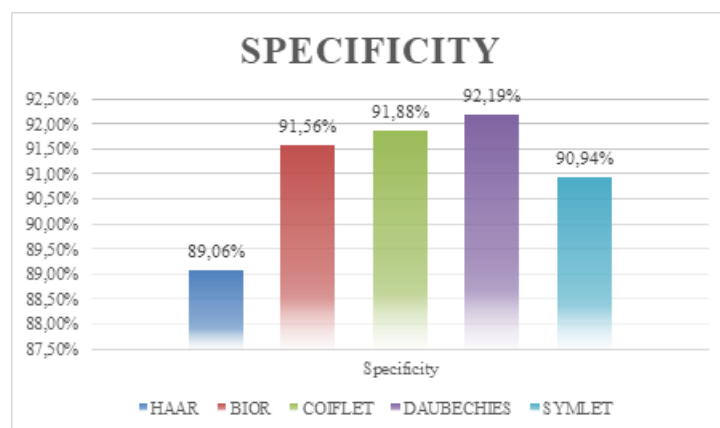

Figure 11. Comparison of The Specificity of Each Wavelet Basis Function

3.5. Precision

The precision value shows a measure of the certainty of the classification model in recognizing positive data sets that match the actual data labels. In this study, this means that how much the percentage of classification results on the set of normal chest X-Ray images are recognized according to the actual data label. Figure 12 shows that the percentage of the highest precision value is generated by the classification using the feature vector of the wavelet daubechies, which is 93.45%. The next percentage precision values are followed by coiflet, biorthogonal, and symlets wavelets. The feature vector produced by Haar wavelet has the lowest

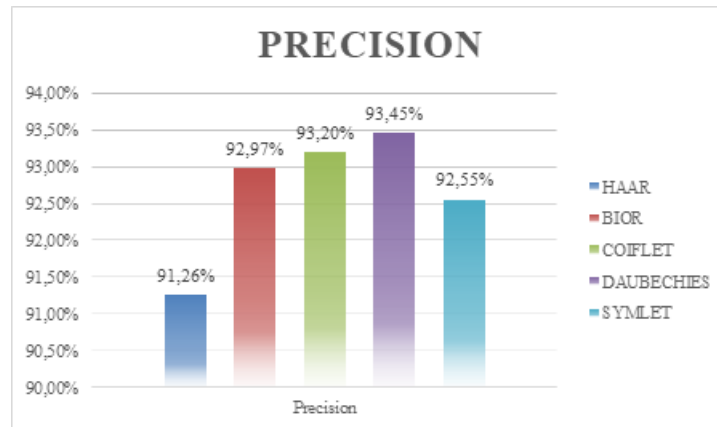precision value in the classification, which is 91.26%.

Figure 12. Comparison of The Precision of Each Wavelet Basis Function

4. CONCLUSION

Based on the findings of the research, it is possible to infer that the classification model for detecting normal chest X-Ray images and diagnosing COVID-19 with the best performance is achieved by utilizing the feature vector generated by the daubechies wavelet. This implies that the daubechies wavelet outperforms the haar wavelet applied in the [12]. This is represented by the percentages of the highest accuracy, recall, specificity, and precision values obtained when the feature vector from the wavelet daubechies is used. Also, the percentage error rate value of the daubechies wavelet is the lowest when compared to other wavelet basis functions used in this study. For further research, it is suggested that the chest X-ray picture be segmented using the segmentation approach, having wavelet decomposition performed just in the lung area.

REFERENCES

[1] I. N. Sweta, "Perancangan Sistem Penentuan Objek Wisata di Bali Masa Pandemi COVID-19 dengan Metode Weighted Product yang Dimodifikasi," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 2, pp. 367–378, 2021.

[2] S. L. K. Yee and W. J. K. Raymond, "Pneumonia Diagnosis Using Chest X-ray Images and Machine Learning," *ICBET 2020: Proceedings of the 2020 10th International Conference on Biomedical Engineering and Technology*, vol. 10, no. 1, pp. 101–105, 2020.

[3] A. Zotin, Y. Hamad, K. Simonov, and M. Kurako, "Lung Boundary Detection for Chest X-Ray Images Classification Based on GLCM and Probabilistic Neural Networks," *Procedia Computer Science*, vol. 159, no. 1, pp. 1439–1448, 2019.

[4] A. Sharma, D. Raju, and S. Ranjan, "Detection of Pneumonia Clouds in Chest X-Ray Using Image Processing Approach," *2017 Nirma University International Conference on Engineering (NUiCONE)*, vol. 20, no. 1, pp. 1–4, 2017.

[5] Y.-H. Chan, Y.-Z. Zeng, H.-C. Wu, M.-C. Wu, and H.-M. Sun, "Effective Pneumothorax Detection for Chest X-Ray Images Using Local Binary Pattern and Support Vector Machine," *Journal of Healthcare Engineering*, vol. 20, no. 1, pp. 98–104, 2018.

[6] R. H and A. T, "Feature Extraction of Chest X-Ray Images and Analysis Using PCA and kPCA," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 5, pp. 3392–3398, 2018.

[7] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated Detection of COVID-19 Cases Using Deep Neural Networks with X-Ray Images," *Computers in Biology and Medicine*, vol. 121, no. 4, pp. 103–109, 2020.

[8] J. Too, A. R. Abdullah, and N. M. Saad, "Classification of Hand Movements Based on Discrete Wavelet Transform and Enhanced Feature Extraction," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 6, pp. 83–89, 2019.

[9] T. G. Krishna, K. V. N. Sunitha, and S. Mishra, "Detection and Classification of Brain Tumor from MRI Medical Image Using Wavelet Transform and PSO Based LLRBFNN Algorithm," *International Journal of Computer Sciences and Engineering*, vol. 6, no. 1, pp. 18–23, 2018.

[10] N. Hashim, S. Adebayo, K. Abdan, and M. Hanafi, "Comparative Study of Transform-Based Image Texture Analysis for The Evaluation of Banana Quality Using An Optical Backscattering System," *Postharvest Biology and Technology*, vol. 135, no. 1, pp. 38–50, 2018.

[11] S.-H. Wang, T. M. Zhan, Y. Chen, Y. Zhang, M. Yang, H.-M. Lu, H.-N. Wang, B. Liu, and P. Phillips, "Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression," *IEEE Access*, vol. 4, no. 1, pp. 7567–7576, 2016.

[12] N. W. S. Saraswati, N. W. Wardani, and I. G. A. A. D. Indradewi, "Detection of Covid Chest X-Ray Using Wavelet and Support Vector Machines," *International Journal of Engineering and Emerging Technology*, vol. 5, no. 2, pp. 116–121, 2020.

[13] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, and P. Ghamisi, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, no. 1, pp. 6308 – 6325, 2020.

[14] X. Yang, M. Han, H. Tang, Q. Li, and X. Luo, "Detecting Defects with Support Vector Machine in Logistics Packaging Boxes for Edge Computing," *IEEE Access*, vol. 8, no. 1, pp. 64 002 – 64 010, 2020.

[15] C. Gui, "Analysis of Imbalanced Data Set Problem: The Case of Churn Prediction for Telecommunication," *Artificial Intelligence Research*, vol. 6, no. 2, p. 93, 2017.

[16] W. Castro, J. Oblitas, D.-L.-T. Miguel, C. Cotrina, K. Bazan, and H. Avila-George, "Classification of Cape Gooseberry Fruit According to its Level of Ripeness Using Machine Learning Techniques and Different Color Spaces," *IEEE Access*, vol. 7, no. 1, pp. 27 389–27 400, 2019.

[17] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "COVID-19 Image Data Collection: Prospective Predictions are The Future," *arXiv*, pp. 1–38, 2020.

[18] Y. A. Sari, A. G. Hapsani, S. Adinugroho, L. Hakim, and S. Mutrofin, "Preprocessing of Skin Images and Feature Selection for Early Stage of Melanoma Detection Using Color Feature Extraction," *International Journal of Artificial Intelligence Research*, vol. 4, no. 2, p. 95, 2021.

[19] I. Candradewi, A. Harjoko, and B. Sumbodo, "Intelligent Traffic Monitoring Systems: Vehicle Type Classification Using Support Vector Machine," *International Journal of Artificial Intelligence Research*, vol. 5, no. 1, 2021.

[20] K. Padmavathi and K. Thangadurai, "Implementation of RGB and Grayscale Images in Plant Leaves Disease Detection - Comparative Study," *Indian Journal of Science and Technology*, vol. 9, no. 6, pp. 4–9, 2016.

[21] R. Gonzalez, R. Woods, and S. Eddins, *Digital Image Processing Using MATLAB*, 3rd ed. Gatesmark Publishing, 2020.

[22] H. Kutlu and E. Avc, "A Novel Method for Classifying Liver and Brain Tumors Using Convolutional Neural Networks, Discrete Wavelet Transform and Long Short-Term Memory Networks," *Sensors (Basel, Switzerland)*, vol. 19, no. 9, 2019.

[23] D. Putra, *Pengolahan Citra Digital*. Yogyakarta: Andi Offset, 2010.

[24] L. Cheng, D. Li, X. Li, and S. Yu, "The Optimal Wavelet Basis Function Selection in Feature Extraction of Motor Imagery Electroencephalogram Based on Wavelet Packet Transformation," *IEEE Access*, vol. 7, no. Cc, pp. 174 465–174 481, 2019.

[25] N. Li, F. He, W. Ma, R. Wang, and X. Zhang, "Wind Power Prediction of Kernel Extreme Learning Machine Based on Differential Evolution Algorithm and Cross Validation Algorithm," *IEEE Access*, vol. 8, pp. 68 874–68 882, 2020.

[26] P. R. Sihombing and I. F. Yuliati, "Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 2, pp. 417–426, 2021.

[27] F. D. Ananda and Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 2, pp. 407–416, 2021.

[28] A. Sridharan, A. S. Remya Ajai, and S. Gopalan, "A Novel Methodology for The Classification of Debris Scars Using Discrete Wavelet Transform and Support Vector Machine," *Procedia Computer Science*, vol. 171, no. 2019, pp. 609–616, 2020.

[29] B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline, "Machine Learning for Medical Imaging," *Radiographics*, vol. 37, no. 2, pp. 505–515, 2017.

[30] S. B. Rakhmetulayeva, K. S. Duisebekova, A. M. Mamyrbekov, D. K. Kozhamzharova, G. N. Astaubayeva, and K. Stamkulova, "Application of Classification Algorithm Based on SVM for Determining The Effectiveness of Treatment of Tuberculosis," *Procedia Computer Science*, vol. 130, pp. 231–238, 2018.

[31] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," *IEEE Access*, vol. 8, pp. 90 847–90 861, 2020.

[32] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The Impact of Class Imbalance in Classification Performance Metrics Based on The Binary Confusion Matrix," *Pattern Recognition*, vol. 91, pp. 216–231, 2019.

[33] N. A. Melo Riveros, B. A. Cardenas Espitia, and L. E. Aparicio Pico, "Comparison Between K-means and Self-Organizing Maps Algorithms Used for Diagnosis Spinal Column Patients," *Informatics in Medicine Unlocked*, vol. 16, no. July, p. 100206, 2019.