

A New Feature Extraction Method for Classifying Heart Wall from Left Ventricle Cavity

Riyanto Sigit^{a,1}, Achmad Basuki^b, Anwar^{a,2}

^a Department of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya, 60111, Indonesia
E-mail: ¹riyanto@pens.ac.id; ²anwar@pasca.student.pens.ac.id

^b Department of Creative Multimedia Technology, Politeknik Elektronika Negeri Surabaya, 60111, Indonesia
E-mail: basuki@pens.ac.id

Abstract— Echocardiography is a method of examination with high-frequency sound waves to obtain images of heart organs. Examination of heart health conditions with echocardiography as an imaging method, serves to detect the potential for heart disease, thus that the right treatment from the evaluation results can be decided. Examination of the source of heart disease with echocardiography was performed using several views, namely the long axis, short axis, two-chamber, and four-chamber. However, the assessment of cardiac function is still carried out conventionally. Thus it is necessary to build a system that can assess cardiac function. This study proposes a feature extraction method for the classification of heart disease based on the left ventricular motion on the short-axis. In this method, feature extraction uses 24 good features for the process of tracking the movement of the left ventricle with optical flow. Each good feature produces four features, namely direction (negative direction and positive direction) and distance (negative distance and positive distance) from the results of left ventricular tracking and produces 96 attributes for the whole process. The features that have been obtained are then processed using several classification algorithms with validation techniques that are, k-folds, and leave one out. The result is a classification algorithm with a gradient boosting classifier method that has the best accuracy. Gradient boosting classifier produces accuracy values with validation techniques for k-folds 90.98%, and leave one out 93.23%. This shows that the gradient boosting classifier can be relied upon for the classification of heart disease using the proposed feature extraction method. In this study, we developed a new feature extraction method from the results of tracking the heart wall using optical flow. This algorithm can produce feature values from the tracking results that can be used to build a knowledge system for the classification of heart health conditions.

Keywords— ultrasound images; left ventricle; optical flow; feature extraction; gradient boosting classifier.

I. INTRODUCTION

The development of echocardiography technology continues to grow today and provides a major contribution to the diagnosis of heart disease directly without having to do surgery to provide an assessment. Echocardiogram refers to a test that uses sound waves to visualize the structure of the heart in the body. The aim is to find the source of the disease and help diagnose problems with organs in the body. During the examination process, ultrasound is emitted and will reflect the visualization of a moving heart. Therefore, it allows doctors to see the anatomy of the heart with a variety of different angles for the examination process. This matter likewise recommended by the American Heart Association (AHA) to examine the heart at various angles of shooting, namely the long axis, short axis, two-chamber, and four-chamber in the left ventricle to determine the condition of the heart [1]. By measuring the movement of systole toward

diastole in the left ventricle, it can provide parameter values for diagnosing heart conditions. The assessment obtained can be a linear measurement (one dimension), area (two dimensions), or volume (three dimensions). Fundamentally, the assessment of function in the left ventricle is usually by estimating the ejection fraction [2]. Ejection fraction is the percentage of blood released during left ventricular movement in the final systole process with the total volume of final diastole. Assessment can also be seen visually, where the left ventricular function is assessed based on the size of the cavity during the process of systole to diastole. If the heart's function is disrupted, then less blood will be removed, and movement of the left ventricular wall of the heart will decline.

Assessment of heart health conditions usually uses the Simpson method to measure volume in the left ventricle. This method is generally used in the apical two and four-chambers to measure the function of the left ventricle [3].

However, this method still involves manually tracing the endocardium of the left ventricle in the end-systole and end-diastole, thus requiring a high level of concentration. Other limitations on ultrasound images are of poor quality. Therefore it is not possible to trace the contours of the heart cavity wall. This is due to an error in the process of sending data on the ultrasound device, thus that the ultrasound image produces speckle noise. Therefore the doctor conducts an initial examination on a short-axis view because it is considered to be easy and fast to make an assessment based on the results of visualization of heart movements. Based on observations made at Jemursari Hospital in Surabaya, the assessment of movement on the short axis view is still carried out qualitatively, so a system capable of conducting an assessment to identify diseases with image processing techniques is needed.

The application of various ultrasound image processing techniques has been explored in the literature, with the aim of automatically tracing the contours of the heart's left ventricle. Reference [4] traced the ultrasound image of the left ventricular part of the combination of the non-parametric combination of texture descriptors method and obtained good segmentation results. Reference [5] uses a deep convolutional neural network method for the segmentation process on the left ventricle. This method performs a training process that requires a large dataset. The resulting network model is capable of achieving an accuracy similar to automatic segmentation. Reference [6], [7] developed a method for segmenting the left ventricle automatically. The proposed method obtains the smallest value in the 108-millisecond computational process and a cavity error of 8.18% compared to the snake method of 19.94% and watershed of 15.97%. This method is also used in the process of left ventricular segmentation by proposing the development of a high-boost filter method to improve the results of heart segmentation from two and four-chambers, with an accuracy of 89.409% [8].

Several previous studies have carried out the application of the tracking process, which aims to visualize the movement of the left ventricle. Reference [9] performed the process of segmenting the left ventricle on the short axis automatically using the active shape model method and described the visualization of tracking using semi-automatic optical flow with the initialization of points on the cavity. In reference [10], a semi-automatic left ventricular contour formation process was conducted; the process proceeded to the tracking process using optical flow to calculate the volume of left ventricular function. The semi-automatic approach in tracing the process for the classification of heart movements in the short axis viewpoint is seen in [11]. The tracking process uses an optical flow method that produces direction and distance displacement features. The features that have been obtained are trained and validated, which have an accuracy rate of 81.82%. However, the dataset used is 11; therefore, it cannot represent the population of the whole data. A semi-automatic approach to the classification of heart conditions was also carried out in the previous research [12]. However, the proposed method is felt to be less than optimal in classifying heart conditions, and the dataset used is 21, thus unable to represent the population of the whole data.

Based on previous research discussing contour tracking and tracking on the left ventricle for medical purposes, most of those focused on image processing and tracking without extracting features that can be used to build knowledge systems with large amounts of data. This paper proposes a feature extraction method from the results of left ventricular tracking used to establish a classification system of heart health conditions. This study uses 133 data and has had classes in each data. The amount of data for the abnormal class is 65, and the normal class is 68. The feature extraction results are analyzed to determine the characteristics of the data produced and find an appropriate classification algorithm using several validation techniques, namely, k-folds, and leave one out. It is expected that the system is able to classify heart health conditions with a high degree of accuracy.

II. MATERIAL AND METHOD

This study focused on feature extraction from the results of tracking left ventricular movements and the classification of features produced using several classification algorithms. In this section, we described the procedure for building a cardiac classification system that involves pre-processing, segmentation, tracking, classification, and validation. Figure. 1 shows the system diagram developed to implement the objectives of this study.

A. Pre-processing

An ultrasound image required pre-processing aimed at clarifying the difference between the left ventricle and the heart wall. The first stage used a median filter with a 27×27 kernel. The use of large kernels was intended to reduce speckle noise. However, it still maintained the edges of the image and increased the margins in the left ventricle, even if the resulting image was blurred [8]. After the speckle noise was reduced, the brightness of the heart wall in the image needed to be further increased. Our research improved the brightness of the heart wall using the high-boost filter method. The working principle of a high boost filter was to increase the value at high frequencies and maintain a low frequency at image values [13]. The high-boost filter can be

expressed as $G_{(x,y)}$, where A is the gain and $F_{(x,y)}$ is the original image.

$$G_{(x,y)} = A.F_{(x,y)} - \text{lowpass}(F_{(x,y)}) \quad (1)$$

$$G_{(x,y)} = (A - 1).F_{(x,y)} + [F_{(x,y)} - \text{lowpass}(F_{(x,y)})] \quad (2)$$

$$G_{(x,y)} = (A - 1).F_{(x,y)} + \text{highpass}(F_{(x,y)}) \quad (3)$$

The use of morphological operation method is carried out to reduce the noise from the results of high-boost filters in the image. The combination of opening capable handle noise. Opening morphology can be stated as:

$$D \circ F = (D \oplus (-F)) \oplus F \quad (4)$$

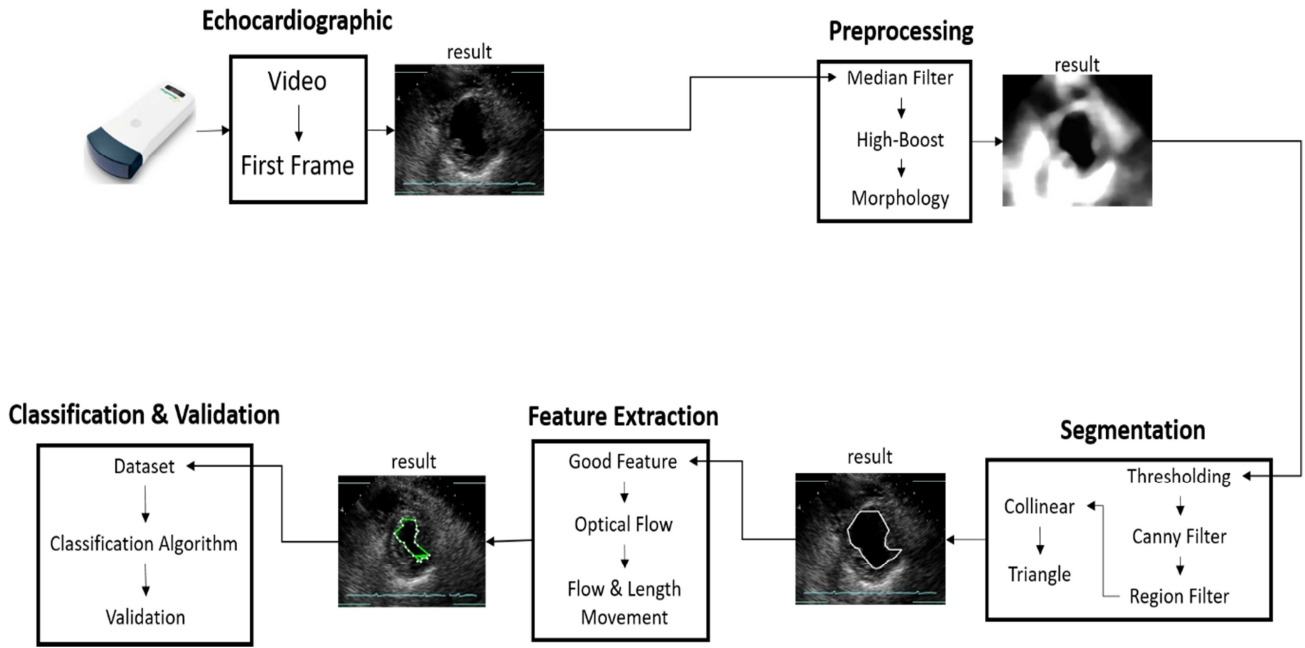


Fig. 1. System diagram

Where D is the original image and F is the structuring element. This method was able to eliminate the noise though also made the left ventricle disappear in some parts. Therefore, morphology closing was employed to restore the left ventricle, and it could be expressed as:

$$D \bullet F = (D \oplus F) \oplus (-F) \quad (5)$$

B. Segmentation

Segmentation is the process of separating the left ventricle from the heart wall. In this study, pre-processing images were segmented using the global thresholding method with a predetermined threshold value based on the characteristics of the results of the pre-processing image. Furthermore, the thresholding results are processed to detect the outline of the left ventricle using the canny edge detection method. Detection of edge lines using canny still needs to be improved to eliminate small noise around the left ventricle. The filter region method was able used to eliminate small contours by counting the number of points each contour and providing thresholds with values [14]. However, the results of the region filter have not been able to remove the large contours on the left ventricular side. This problem can be overcome by the collinear method. The working principle of collinear by finding the centroid of all contours with an Equation:

$$C = \left\{ \frac{\sum_{k=1}^n Xk}{n}, \frac{\sum_{k=1}^n Yk}{n} \right\} \quad (6)$$

A collinear equation was carried out from the center of the boundary to the centroid of each contour by finding the

slope and intercept [7]. The collinear equation used is shown in Equation (7 – 9).

$$y = wx + b \quad (7)$$

Where w and b can be expressed as:

$$w = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \quad (8)$$

$$b = \bar{y} - w \bar{x} \quad (9)$$

The next step was to use the triangle equation to search for the smallest angle in the contours of the open heart cavity [6]. This method is able to close the open contour by connecting between two separate contour lines. The method of triangle equations is illustrated in Figure 2, where angles A, B, and C are interconnected, and produce a distance value between two angles with values a , b , c .

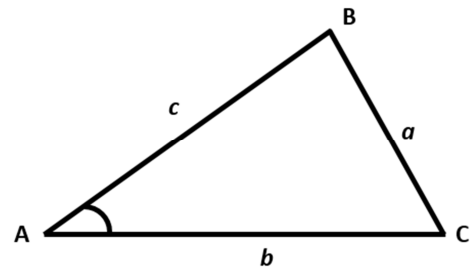


Fig. 2. Triangle Equation

Point A is the center of the contour boundary of the heart cavity. Points B and C respectively represent the point of the contour-open line. The value of a is the distance between points A and B that can calculate the distance between points

AB. Likewise, the values of b and c are the distance between the points AC and AB. If the values a , b , and c are obtained, then each angle can be calculated using Equations (10-11).

$$a^2 = b^2 + c^2 - 2bc(\cos A) \quad (10)$$

$$A = a \cos\left(\frac{b^2 + c^2 - a^2}{2bc}\right) \quad (11)$$

C. Feature Extraction

Based on previous studies, the use of *good features* to track the left ventricular obtained good results [9], [10]. This study uses a good feature to track the movement of diastole to systole in the left ventricle. Tracking on the left ventricle with good features, then forwarded to the optical flow for tracking the heart cavity in the next frame. A good feature search was performed automatically. Good features were obtained from the left ventricular contour of image segmentation results. However, not all contour points were used as good features for the tracking process, only a few sample points representing the shape of the contour were selected. Figure 3 is an illustration of the search for good features from contour lines.

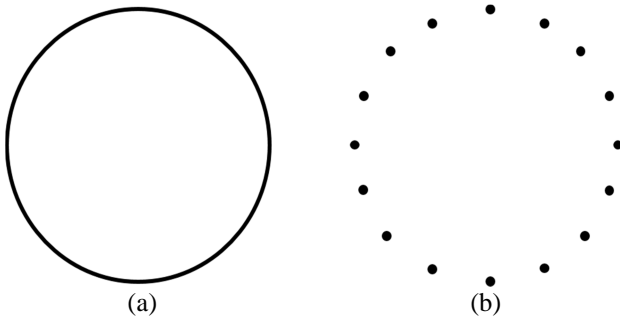


Fig. 3. Good feature result (a) contour, (b) good feature

The sample point that represents the contour shape is a good feature, which is then used for the tracking process in the next frame image. Movement in space was able to be described as an area of motion and was realized through gradients of different grayscale distributions. The motion field in space was transformed into an image, which is known as an optical flow field that shows techniques for characterizing image motion [15]. In this study, *good features* on the left ventricular wall were used for the tracking process with the optical flow using the Lucas-Kanade method. Reference [16], using good features obtained from crossing lines on the contours is used to track left ventricles. The proposed method shows the results of calculations with a sensitivity of 90% and an accuracy of 87.451%.

Lucas-Kanade is a method used to predict the movement of an object based on the intensity of light in the image. The application to the image can provide a percentage describing the extent to which an image pixel moves to the next frame. This method assumes that the displacement object between the two closest frames is small and approximately constant in the environment of that point. Therefore, the optical flow equation is assumed to apply to all pixels in an image

centered on p . That is, local image flow vector (speed) (V_x, V_y) must meet:

$$I_x(q_1)V_x + I_y(q_1)V_y = -I_t(q_1)$$

$$I_x(q_2)V_x + I_y(q_2)V_y = -I_t(q_2)$$

...

$$I_x(q_n)V_x + I_y(q_n)V_y = -I_t(q_n) \quad (12)$$

Where q_1, q_2, \dots, q_n , is the pixels in the image, and $I_x(q_i), I_y(q_i), I_t(q_i)$ is a partial derivative of the image in relation to the positions x, y and time t , evaluated at the point q_i and at run time. This equation can be described in the form of a matrix $Au=b$, where:

$$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \cdot & \cdot \\ \cdot & \cdot \\ I_x(q_n) & I_y(q_n) \end{bmatrix} v = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \cdot \\ \cdot \\ -I_t(q_n) \end{bmatrix} \quad (13)$$

The Lucas-Kanade method provides a solution with the principle of least squares; that is, by solving the 2×2 system

$$A^T A v = A^T b \quad (14)$$

As A^T is the transpose of matrix A , it can calculate the equation (15). The central matrix in the equation is the inverse matrix. The amount starts from $i = 1$ to n . $A^T A$ matrix is often called the tensor structure of the image at point p .

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum I_x(q_i)^2 & \sum I_x(q_i)I_y(q_i) \\ \sum I_y(q_i)I_x(q_i) & \sum I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum I_x(q_i)I_t(q_i) \\ -\sum I_y(q_i)I_t(q_i) \end{bmatrix} \quad (15)$$

Abnormalities in cardiac function can be assessed qualitatively based on visualization of the movement of diastole to systole in the heart cavity. Based on observations, the normal heart has large and fast movements, whereas abnormal heart conditions have small and slow movements. Heart movement can be used as a parameter to assess the condition of heart health. Left ventricular tracking with optical flow, was able to produce direction and distance features from the good features specified in the first frame. To obtain the value of the direction of movement features of the heart wall, the researchers propose a simple algorithm, as shown in Figure 4 and 5.

Figures 4 and 5 show three points, namely points A, B, and C. Point A is the origin, which will be a reference to determine the direction of displacement at point B. Point C is the center point on the contour of the cavity connected to point A, so the tracking results the direction will not affect

the tilt. Figure 4 explains the displacement of point A, moving forward towards point B.

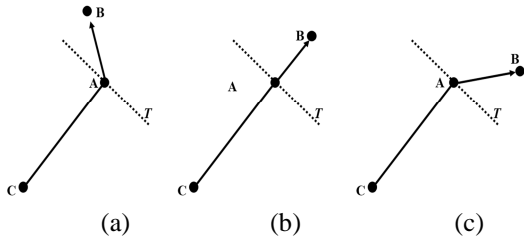


Fig. 4. Direction feature negative value

The resulting feature value is a positive direction because the point is inside the T-line threshold ($0^\circ - 180^\circ$). Whereas Figure 5 explains the displacement of point A that moves out towards point B. The resulting feature value is a negative direction because the point is outside the T-line threshold ($181^\circ - 360^\circ$).

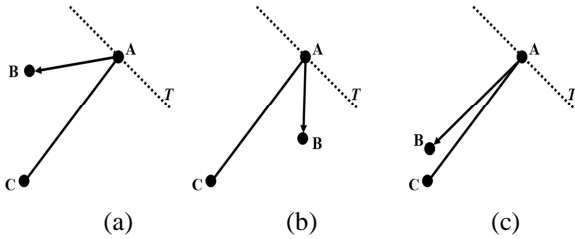


Fig. 5. Direction feature positive value

Left ventricular wall motion abnormalities were commonly observed in a variety of medical conditions; reference [17] describes the asymmetrical movement of the left ventricular wall, giving an indication of cardiac abnormalities. From the observations, the healthy heart has a displacement of systole to diastole moving forward expeditiously, while abnormal heart conditions have slow movement. Assessment of cardiac function can use the optical flow method to track the movement of diastole toward systole. Thus, the thickening of the heart wall indicates abnormalities in heart function. Optical flow tracking will produce direction and distance displacement features from the left ventricle using *good features*. The features obtained use Equation (16) (17) for the direction and distance features.

$$\cos A = \frac{b^2 + c^2 - a^2}{2bc} \quad (16)$$

$$d_{(i,j)} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (17)$$

The features that have been obtained were then normalized to scale the data. Normalization using the min-max method is one of the most common ways to normalize data. This method transforms the minimum value of the feature to 0, the maximum value gets transformed into a 1, and every other value obtains transformed into a value between 0 - 1. The normalization min-max can be expressed by Equation 18. Where data N is rescaled such that any

specific N will now be $0 \leq N \leq 1$, and is done through the formula below:

$$N = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (18)$$

D. Validation

In machine learning applications, the features in the data set have the most significant effect on the results of machine learning models, therefore determining the system's accuracy in predicting. The testing data set is a separate portion of the same data set from which the training set is derived. The main purpose of using the testing data set was to examine the generalization ability of a trained model [18]. The classification algorithm assessment used several validation techniques to obtain the error rate of the trained model. Validation is done by comparing three validation models, namely, k-folds, and leave one out. The performance results from validation can be calculated using a confusion matrix.

TABLE I
CONFUSION MATRIX

Class		Actual Class	
		Yes	No
Predict Class	Yes	True Positive (TP)	False Positive (FP)
	No	False Negative (FN)	True Negative (TN)

The explanation of each part of the Table I can be explained as follows:

- 1) True Positive: the positive value predicted correctly, which means that the value of the actual class is yes, and the value of the predicted class is also yes.
- 2) True Negative: the predicted negative value is true, which means that the value of the actual class is no, and the value of the predicted class is also no.
- 3) False Positive: true value is no, and class prediction is yes.
- 4) False Negative: actual class value is yes but predicted class no.

Confusion matrix can be used for performance evaluation that represents predictions and the actual conditions of the data generated by the classification algorithm. Performance resulting from the confusion matrix is as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (19)$$

$$Precision = \frac{TP}{TP + FP} \quad (20)$$

$$Recall = \frac{TP}{TP + FN} \quad (21)$$

$$F1-Score = \frac{2.(Recall.Precision)}{Recall + Precision} \quad (22)$$

III. RESULT AND DISCUSSION

This research focuses on the feature extraction process to build a classification system for heart disease in the short-axis view of the left ventricular cavity, and several processes are needed to obtain the features of the left ventricular movement. The resulting feature is the direction and distance of displacement from the tracking process using Lucas-Kanade optical flow. This study examines the value of features obtained from the results of tracking the left ventricle. The features that have been obtained are then tested using the classification algorithm.

Testing of the results of feature extraction using classification algorithms available in the library Scikit-Learn who uses the Python programming language. Testing involves several classification algorithms such as KNN with $K = 3$, Naive Bayes, Logistic Regression, Decision Tree, SVM (Support Vector Machine) with linear kernels, and conduct an assessment with two validation techniques namely, k-folds, and leave one out.

A. Pre-processing and segmentation result

The results of the pre-processing and image segmentation process in each step are shown in Figure 6 and Figure 7. The initial stage is done by improving the image using a 27×27 median kernel filter, which aims to reduce the speckle noise in ultrasound images. The use of a large kernel is intended to reduce speckle noise and to retain the image edge portion for improving the line edge cavity. However, the produced image was blurry. To clarify the difference between the left ventricle and wall heart, image enhancements such as high-boost filters and morphology operations need to be performed. By increasing the ultrasound image in the pre-processing stage, the heart cavity will become darker, and the heart wall will be brighter [8].

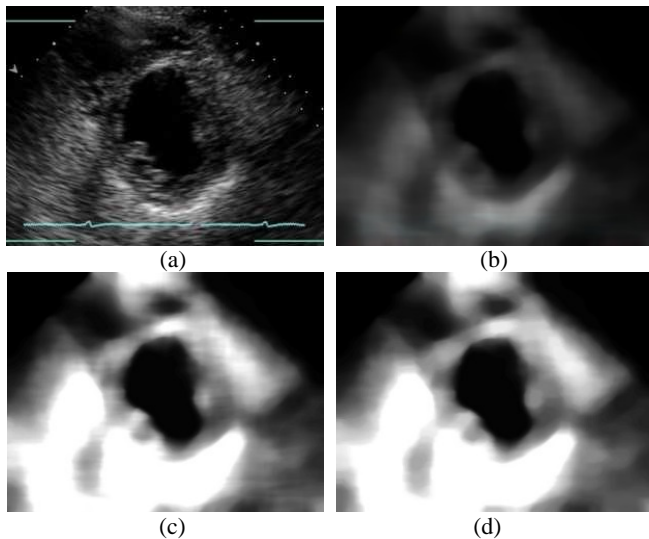


Fig. 6. Image pre-processing result. (a) original image; (b) median; (c) high-boost filter; (d) morphology operation.

The next step is the segmentation of the left ventricle; it is performed after the difference between the left ventricle, and the heart wall is seen. Stages that need to be performed is to convert images from the results of pre-processing into binary images with global thresholding. The result is a binary image

that has values 0 and 1 as the boundary between two different regions in an image. Therefore the canny edge is implemented to obtain the left ventricular contour. However, the results of the canny method need to be improved since there is noise around the contour. Therefore, the filter region and collinear are used to remove noise that is not part of the left ventricular contour.

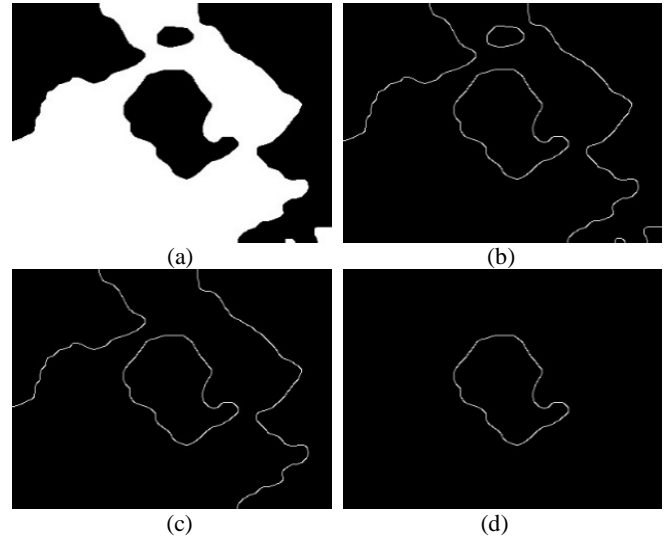


Fig. 7. Image segmentation result. (a) thresholding; (b) canny; (c) region filter; (d) collinear.

However, there are problems where the results of segmentation on the left ventricle have open cavity contour lines, as shown in Figure 8. Thus, the triangle equation method is adopted to improve contour lines. This method serves to calculate the length of the line through two sides of the line and the consinus of the angle facing the calculated line. In addition, the triangle equation is also used to calculate angles through the three lines that make up the angle. In this study, the triangle equation method is used to connect two separate points.

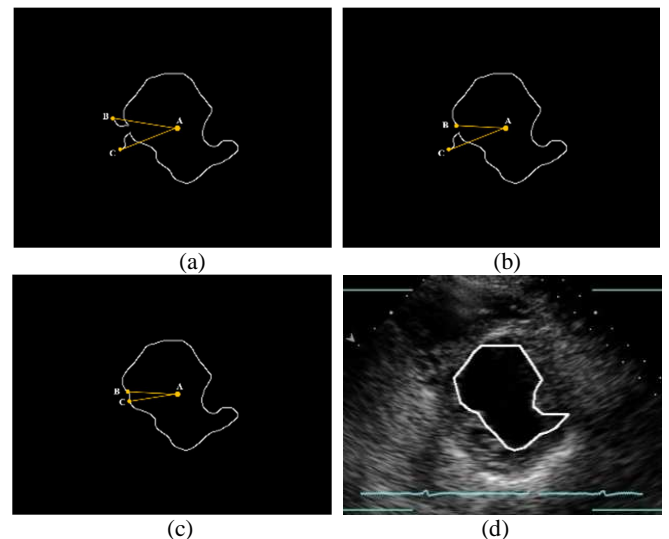


Fig. 8. Triangle equation result

The proposed segmentation method is able to segment automatically. The results of image segmentation with the

proposed method need to be tested to determine the level of success in segmenting the left ventricle on the short axis view. Testing is carried out by comparing several methods of image segmentation with the proposed method. Figure 9 shows the results of image segmentation using manual, snake, watershed, and triangle methods (proposed method).

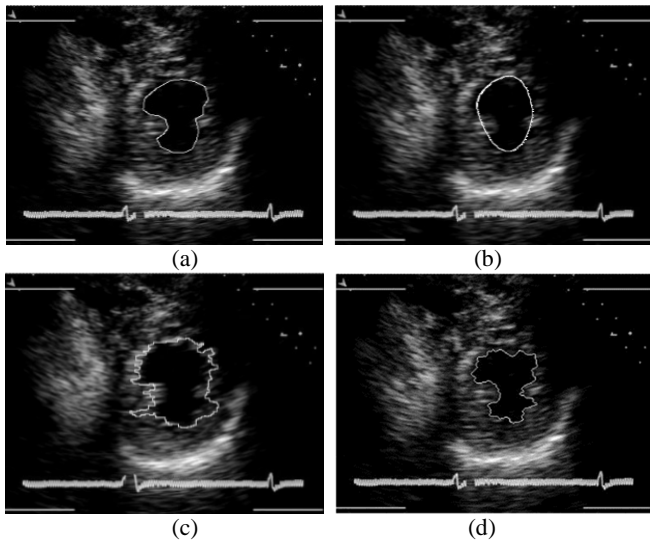


Fig. 9. Image segmentation result. (a) manually; (b) snake; (c) watershed; (d) triangle.

The segmentation assessment is based on the area of the heart cavity by the method of snake, watershed and triangle equations compared to the results of segmentation manually as in Figure 9 (a). Based on the results of tests conducted, the triangle equation segmentation method obtained a value of 91.82%, snake 80.06% and watershed 84.03%. These results indicate the triangle equation method has a higher success value than the snake and watershed methods.

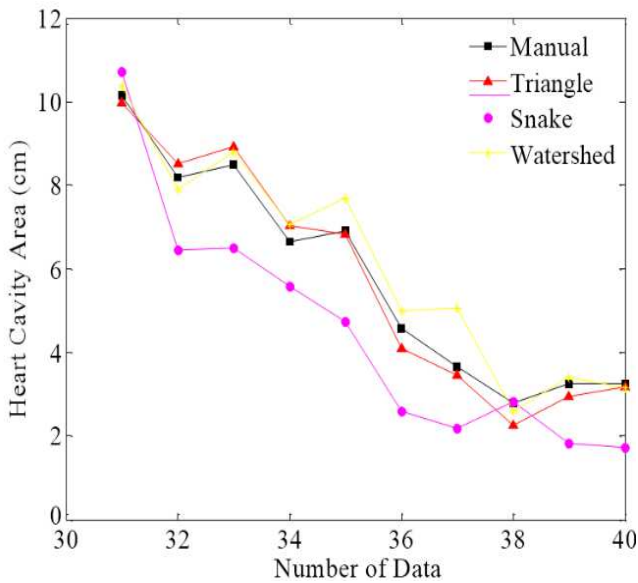


Fig. 10. Segmentation result. (a) manually; (b) snake; (c) watershed; (d) triangle.

B. Tracking result

As a whole, the process of finding good features for the tracking process utilizes left ventricular edge detection. A good feature definition uses several sample points that represent the contour shape of the cavity. There are 24 good features that are used to track left ventricular movements. Figure 11 is an illustration of the definition of good features from contour lines using several sample points.

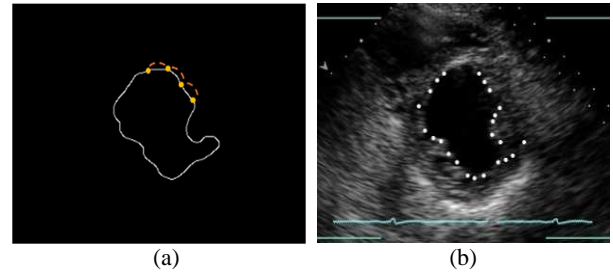


Fig. 11. Good feature result. (a) definition good feature; (b) good feature left ventricle.

Good features that have been obtained in the initial frame then proceed to use the optical flow method to track the movement of the left ventricular heart wall in the next frame. Figure 12 shows the results of the tracking process of diastole to systole, where good features are able to track well on the moving left ventricle. Tracking on good features with the optical flow will produce a feature extraction value that can be used to classify heart health conditions.

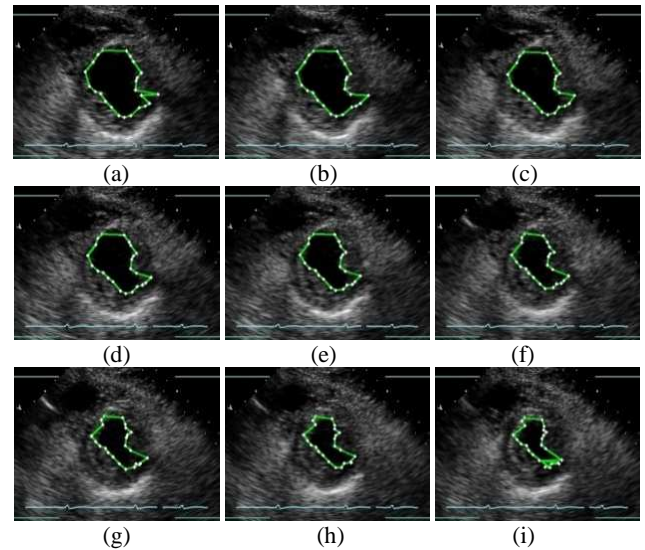


Fig. 12. Tracking optical flow. (a) frame 1; (b) frame 2; (c) frame 3; (d) frame 4; (e) frame 5; (f) frame 6; (g) frame 7; (h) frame 8; (i) frame 9.

The results of tracking need to be tested to determine the level of success of the tracking process of the left ventricle. Tests carried out by forming contours manually using the initialization point will then be compared with the results of the proposed method. Figure 13 shows the results of a comparison of manually formed contours (red lines) and contour formation from search results by the proposed method (green lines).

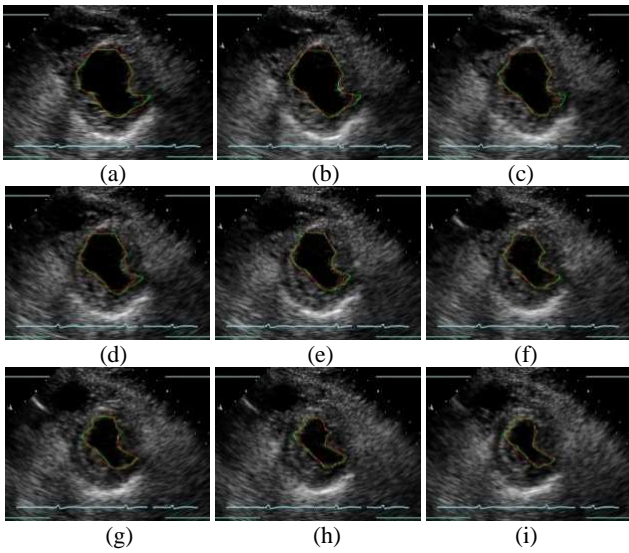


Fig. 13. Tracking optical flow. (a) frame 1; (b) frame 2; (c) frame 3; (d) frame 4; (e) frame 5; (f) frame 6; (g) frame 7; (h) frame 8; (i) frame 9.

Calculation of the results of the tracking is done by means, if there is a green contour line contained in the red line of the contour area then it is true positive (TP). If there is a green contour line outside the red line area then it is false positive (FP). If there is a green contour line that is not in the red contour line then it is false negative (FN). This study uses 10 data to conduct the test, based on Figure 14 the method of tracking on the left ventricle that is proposed to obtain calculation results with a sensitivity level of 92.33% and an accuracy of 87.73%.

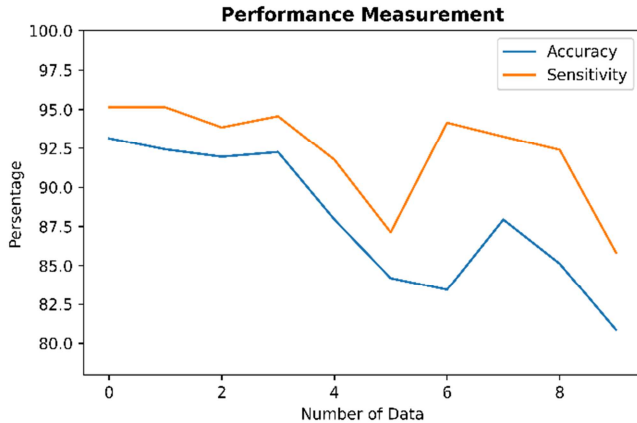


Fig. 14. Performance measurement.

C. Feature extraction result

This study uses ultrasound video data on short-axis views. Video data was obtained at Darmo Hospital Surabaya from 2 February to 1 October 2019. The amount of data collected is 133 data and has a class in each data.

TABLE II
NUMBER OF DATA

No	Class Data	Total
1	Normal	68
2	Abnormal	65
Total Data		133

Data that have been collected are then processed to obtain the feature direction and distance of movement from the results of tracking systole to diastole. Processes involving *good features* on the left ventricle are continued with optical flow for tracking. Each *good feature* produces four values, namely direction (negative direction and positive direction) and distance (negative distance and positive distance). This study employed 24 *good features* for tracking, thus resulting in 96 features for each dataset. The acquired features need to be processed first by data normalization. Normalization is a technique that is often applied as part of data preparation for machine learning. The purpose of normalization is to change the numeric column values in a dataset into a general scale, without distorting the differences in the range of values.

TABLE III
FEATURE EXTRACTION

No	Good Feature	Movement	Feature	Value
1	1	Direction +	F1	0.599
2		Direction -	F2	-0.599
3		Distance +	F3	0.061
4		Distance -	F4	-0.631
..
93	24	Direction +	F93	0.556
94		Direction -	F94	-0.556
95		Distance +	F95	0.980
96		Distance -	F96	-0.387

The features that have been obtained are then tested using the classification algorithm. Testing of the results of feature extraction using classification algorithms available in the library Scikit-Learn who uses the Python programming language. Testing involves several classification algorithms such as Support Vector Machine, Decision Tree, Nearest Neighbor, Naïve Bayes, Random Forest, Gradient Boosting, Neural Network, and conduct an assessment with two validation techniques namely k-folds, and leave one out.

TABLE IV
CLASSIFICATION ALGORITHM

Classification Algorithm	Validation (Accuracy %)	
	Leave one out Validation	K-Folds Cross Validation
Support Vector Machine	84.96%	86.42%
Decision Tree	91.72%	89.39%
K-Nearest Neighbor	81.20%	82.69%
Naïve Bayes	84.21%	83.40%
Random Forest	89.47%	89.39%
Gradient Boosting	93.23%	90.98%
Neural Network	84.21%	82.63%

Table IV shows that the proposed feature extraction method can produce the highest accuracy values from several classification algorithms. Based on the results of experiments that have been conducted, the classification algorithm using the gradient boosting classifier method has

the highest level of accuracy from the results of testing several validation models. This classification algorithm produces accuracy values with validation techniques, k-fold 90,98%, and leave one out 93,23%. This study also calculated the value of precision, recall, and f1-score from several classification algorithms. From the results of several validations, the gradient boosting classification algorithm has the highest results compared to other classification algorithms. Thus, gradient boosting classifiers are suitable for use in the classification of heart disease using the feature extraction method proposed. Figures 15 and 16 show the results obtained from two validation models as follows:

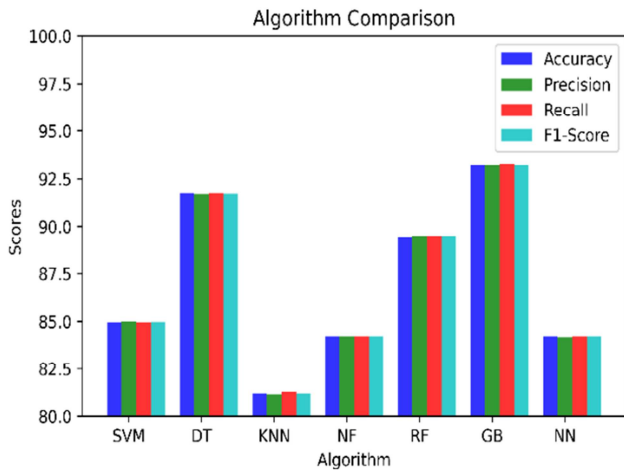


Fig. 15. Leave one out validation.

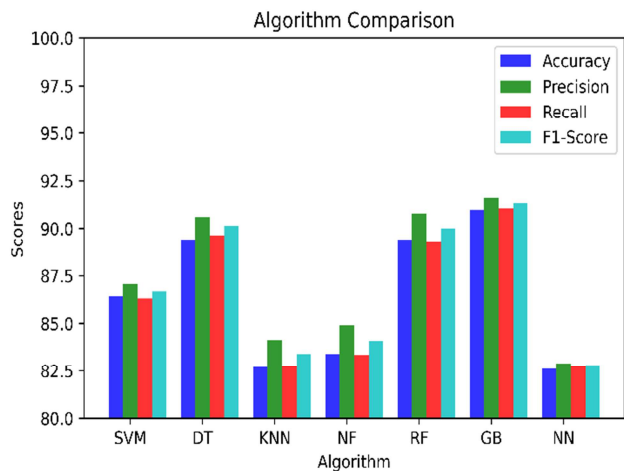


Fig. 16. K-Folds validation.

IV. CONCLUSIONS

This research focuses on the feature extraction process to establish a classification system of heart disease in the short axis left ventricular cavities, and several processes are needed to obtain the features of the left ventricular movement. The next step segment of the left ventricle aims to get the contour of the cavity. The next step is to segment the image in the left ventricle to obtain contour in the cavity. The segmentation using the proposed method is the triangle equation obtains a success value of 91.82% higher than snake 80.06% and watershed 84.03%. A good feature search is carried out automatically. Good features are obtained from

the left ventricular contour of image segmentation results. However, not all contour points are used as good features for the tracking process, only a few sample points that represent the shape of the contour. Good features that have been obtained in the initial frame processing then continued using the optical flow method to track the movement of the left ventricular heart wall.

This method is able to obtain the results of calculations with a sensitivity level of 92.33% and an accuracy of 87.73% for the tracking process. Tracking on the left ventricle uses 24 good feature points that will be forwarded to optical flow for the tracking process in the next frame, thus resulting in 96 features for each dataset. The features that have been obtained are then tested using several classification algorithms with validation techniques, namely k-fold and leave one out. Based on the results of experiments that have been conducted, the classification algorithm using the gradient boosting classifier method has the highest level of accuracy from the results of testing several validation models. This classification algorithm produces accuracy values with validation techniques, k-fold 90.98%, and leaves one out 93.23%. Thus, the gradient boosting classifier method is suitable to be used in the classification of heart disease using the proposed feature extraction method. Future work, researchers are focused on the right amount of good features for tracking the left ventricle and hyper parameter optimization in machine learning for increased accuracy.

ACKNOWLEDGMENT

The authors would like to thank all the members of the SVG (Signal, Vision, and Graphics) laboratories and Politeknik Elektronika Negeri Surabaya (PENS) for their support and guidance in this research.

REFERENCES

- [1] American Heart Association, "What is Cardiovascular Disease," 2014. doi: 10.1016/j.cplett.2007.12.031.
- [2] T. A. Foley *et al.*, "Measuring left ventricular ejection fraction-techniques and potential pitfalls," *Eur. Cardiol.*, 2012, doi: 10.15420/ecr.2012.8.2.108.
- [3] M. Seckerdieck, P. Holler, P. Smets, and G. Wess, "Simpson's method of discs in Salukis and Whippets: Echocardiographic reference intervals for end-diastolic and end-systolic left ventricular volumes," *J. Vet. Cardiol.*, vol. 17, no. 4, pp. 271–281, 2015, doi: 10.1016/j.jvc.2015.08.002.
- [4] D. Attia and A. Benazza-Benyahia, "Left ventricle detection in echocardiography videos," *2018 4th Int. Conf. Adv. Technol. Signal Image Process. ATSIP 2018*, pp. 1–6, 2018, doi: 10.1109/ATSIP.2018.8364476.
- [5] E. Smistad, A. Ostvik, B. O. Haugen, and L. Lovstakken, "2D left ventricle segmentation using deep learning," *IEEE Int. Ultrason. Symp. IUS*, pp. 4–7, 2017, doi: 10.1109/ULTSYM.2017.8092573.
- [6] R. Sigit, A. R. Barakbah, and I. A. Sulistijono, "Improved segmentation of cardiac image using triangle and partial Monte Carlo," in *2016 International Conference on Knowledge Creation and Intelligent Computing, KCIC 2016*, 2017, doi: 10.1109/KCIC.2016.7883624.
- [7] R. Sigit, A. R. Barakbah, I. A. Sulistijono, and A. S. Aziz, "Automatic cardiac segmentation using triangle and optical flow," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 8, no. 2, pp. 315–326, 2017, doi: 10.11591/ijeecs.v8.i2.pp315-326.
- [8] A. Anwar, R. Sigit, A. Basuki, I. P. Adi, and S. Gunawan, "Automatic Segmentation of Heart Cavity in Echocardiography Images: Two & Four-Chamber View Using Iterative Process Method," *Proc. - Int. Electron. Symp. Knowl. Creat. Intell. Comput. IES-KCIC 2019*, pp. 177–182, 2019.

- [9] R. Sigit, C. A. Roji, T. Harsono, and S. Kuswadi, "Improved echocardiography segmentation using active shape model and optical flow," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 17, no. 2, p. 809, 2019, doi: 10.12928/telkomnika.v17i2.11821.
- [10] A. A. Pratiwi, R. Sigit, D. K. Basuki, and Y. H. Oktaviono, "Improved ejection fraction measurement on cardiac image using optical flow," *Proc. - Int. Electron. Symp. Knowl. Creat. Intell. Comput. IES-KCIC 2017*, vol. 2017-Janua, pp. 295–300, 2017, doi: 10.1109/KCIC.2017.8228602.
- [11] A. S. Aziz, R. Sigit, A. Basuki, and T. Hidayat, "Cardiac motions classification on sequential PSAX echocardiogram," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, no. 3, pp. 1289–1296, 2018, doi: 10.11591/ijeecs.v12.i3.pp1289-1296.
- [12] C. Nofindarwati, R. Sigit, and T. Harsono, "Detection of Heart Condition based on Echocardiography Image using Ultrasound," *IES 2019 - Int. Electron. Symp. Role Techno-Intelligence Creat. an Open Energy Syst. Towar. Energy Democr. Proc.*, pp. 522–526, 2019, doi: 10.1109/ELECSYM.2019.8901556.
- [13] I. P. A. S. Gunawan, R. Sigit, and A. I. Gunawan, "Vein Visualization System Using Camera and Projector Based on Distance Sensor," *2018 Int. Electron. Symp. Eng. Technol. Appl. IES-ETA 2018 - Proc.*, pp. 150–156, 2019, doi: 10.1109/ELECSYM.2018.8615501.
- [14] I. Putu Adi Surya Gunawan, R. Sigit, and A. I. Gunawan, "Veins projection performance based on ultrasonic distance sensor in various surface objects," *Indones. J. Electr. Eng. Comput. Sci.*, 2019, doi: 10.11591/ijeecs.v17.i3.pp1362-1370.
- [15] S. S. Sengar and S. Mukhopadhyay, "Moving object area detection using normalized self adaptive optical flow," *Optik (Stuttg.)*, vol. 127, no. 16, pp. 6258–6267, 2016, doi: 10.1016/j.ijleo.2016.03.061.
- [16] Anwar, R. Sigit, A. Basuki, and I. Putu Adi Surya Gunawan, "Implementation of optical flow: Good feature definition for tracking of heart cavity," *Indones. J. Electr. Eng. Comput. Sci.*, 2020, doi: 10.11591/ijeecs.v18.i2.pp1057-1065.
- [17] N. G. Pandian, D. J. Skorton, S. M. Collins, H. L. Falsetti, E. R. Burke, and R. E. Kerber, "Heterogeneity of left ventricular segmental wall thickening and excursion in 2-dimensional echocardiograms of normal human subjects," *Am. J. Cardiol.*, vol. 51, no. 10, pp. 1667–1673, 1983, doi: 10.1016/0002-9149(83)90207-2.
- [18] E. Alpaydin, *Introduction to Machine Learning 2nd*. 2010.