Identification of Coronary Heart Disease through Iris using Gray Level Co-occurrence Matrix and Support Vector Machine Classification

Vincentius Abdi Gunawan1*
Department of Informatics Engineering
Universitas Palangka Raya
Palangka Raya, Indonesia

Leonardus Sandy Ade Putra2, Fitri Imansyah3, Eka Kusumawardhani4
Department of Electrical Engineering
Universitas Tanjungpura
Pontianak, Indonesia

Abstract—Now-a-days, coronary heart disease is one of the deadliest diseases in the world. An unfavorable lifestyle, lack of physical activity, and consuming tobacco are the causes of coronary heart disease aside from genetic inheritance. Sometimes the patient does not know whether he has abnormalities in heart function or not. Therefore, this study proposes a system that can detect heart abnormalities through the iris, known as the Iridology method. The system is designed automatically in the iris detection to the classification results. Feature extraction using five characteristics is applied to the Gray Level Co-occurrence Matrix (GLCM) method. The classification process uses the Support Vector Machine (SVM) with linear kernel variation, Polynomial, and Gaussian to obtain the best accuracy in the system. From the system simulation results, the use of the Gaussian kernel can be relied on in the classification of iris conditions with an accuracy rate of 91%, then the Polynomial kernel accuracy reaches 89%, and the linear kernel accuracy reaches 87%. This study has succeeded in detecting heart conditions through the iris by dividing the iris into normal iris and abnormal iris.

Keywords—Iris; iridology; coronary heart; circle hough transform; gray level co-occurrence matrix; support vector machine

I. INTRODUCTION

Coronary heart disease is the number one cause of death worldwide. According to data from the World Health Organization (WHO), there are 17 million people in the world who die from coronary heart disease. In Indonesia, coronary heart disease is the highest cause of death after stroke, with a mortality rate of 12.9% in 2014 [1]. Every year there are 1.9 million people die of coronary heart disease due to consuming tobacco [2]. An unhealthy lifestyle and lack of physical activity are the leading causes of coronary heart disease [3]. The death rate is higher among the older age population [4]. Consuming foods high in carbohydrates or fat and obesity are factors that cause constriction of blood vessels in the heart.

Examination to determine coronary heart disease will be advised by checking with an Electrocardiogram (EKG), which uses electricity to determine heart rhythm. The use of echocardiography is also sometimes done to see the part, pump function, and valve function of the heart. Taking some actions to check the current heart condition costs quite a bit so that a person will be reluctant to examine his heart condition.

Early prevention can be done to reduce the risk of coronary heart disease, namely by consuming enough fruits and vegetables every day [5], exercising or doing physical activities regularly in daily life, and do not consume tobacco [1].

Nowadays, technology to detect heart conditions from an early stage has been carried out by scientists in the medical field and computing technology. One of them is knowing the heart condition through the eye's iris, which is directly connected to the brain [6]. The brain is a human organ that receives 15% of blood flow from the heart and accounts for 20% of oxygen consumption in the body, making it susceptible to vascularization in the human brain [7].

The iris is one of the unique organs in the human body. Iris is usually used as electronic security or biometric identification system [8,9]. Nowadays, the research on the iris in the medical field is increasingly widespread and is being seriously studied by experts. Iris can provide information about the condition of human organs, known as the Iridology method.

Iridology is based on the analysis of one of the most complex tissue structures contained in the iris. This method can determine the condition of organs and systems in the body from the marks on the iris. Iridology cannot diagnose a disease but instead helps to identify existing or any potential problems in a particular organ [10]. According to Iridologists, the condition of the overall organ is reflected on the surface of the iris [11]. In healthcare practitioners, the iris is used to determine the systemic health, innate strengths, or weaknesses of an individual's personality [12]. Image processing and data mining processing techniques have been used as disease diagnosis tools in biomedical applications. There are parameters in improving the quality of the iris diagnosis technique, especially by using iris images by focusing on the PRISMA-ScR guidelines [13]. Based on this statement, the research obtained 89.63% classification accuracy of diabetic patients with iris diagnosis [14]. Heydari M. and Teimouri M. have also conducted a test on type 2 diabetic patients and obtained the highest accuracy rate of 97.44% using Artificial
Neural Networks in Iran. Young-Wool Lim and Young-Bae Park [11] have also conducted a study to examine the relationship between the iris and the characteristics differences of each individual. The results show that iris parameters can be used more definitively related to characteristics than functional changes. Research conducted by M. Gopalan and Gopal S. Pollai stated that iridology could be used to determine the health of humans organs contained in the character of the iris [12]. Recognition of the iris identifies biometrics and can diagnose the presence of cholesterol in the blood [15]. Iris is used in providing information of human's physical condition to determine vehicle driver fatigue by detecting the behavior of the iris that has been carried out by [16], with an accuracy of 80%. Research by [17] has diagnosed the heart with the iris using Principal Component Analysis (PCA) and has achieved 80% success using the SVM classification. The use of the GLCM method was also carried out by [18] by using four characteristics of feature extraction with an accuracy rate of 85.6%.

Previous studies have used the iris as an early diagnosis in detecting abnormal conditions in human organs. The use of different methods leads to different results in identification accuracy. The GLCM method using six characteristics of feature extraction has not been carried out until now. So this is a new scope for in-depth research to be carried out in the identification process. Separating the iris from the rest of the eye is carried out automatically using CHT by utilizing circle edge detection, an important part of the identification system designed in this paper.

In this study, iris samples were collected from healthy individuals and patients with coronary heart disease to explore new directions and methodologies in diagnosing coronary heart disease with computer imaging and machine learning techniques. The determination of the iris section has been carried out automatically using the Hough transformation method, which is considered more efficient in reading multiple iris samples [19]. Gray Level Co-occurrence Matrix as a texture extraction method was chosen in this study because of its ability to perform feature extraction in grayscale images properly [20]. The selection of GLCM as a feature extraction method is based on previous research that has not tested the condition of the heart through the iris using the six characteristics of GLCM. The GLCM method is considered to be better at reading images with grayscale values to obtain accurate results. Determination of the iris that reflects normal and abnormal heart conditions is carried out using the Support Vector Machine as a classification method. The SVM method can separate data into classes with linear solutions in high dimensions [21]. So that the final results of this study can help provide information on normal heart conditions and heart conditions with abnormal conditions through the iridology method. The system design flow can be shown in Fig. 1.

![Fig. 1. Shows the Flow of the Designed System; a) the flow of Training on the Recognition System; b) Test Flow on the Recognition System.](image-url)
II. RESEARCH METHOD

A. Iridology

Iridology is known as a diagnostic method using the human iris. In the medical world, the iris can interpret the condition of the human organs [22]. Iridology divides the iris into 60 parts, and each part represents the condition or function of its organs. The right iris will reflect the condition of the right organs of the body, and vice versa, where the left iris will reflect the condition of the left organs [23].

Dr. Bernard Jensen has created a chart that describes each part of the 60 sectors into an image that is mirrored in a circular image like a clock and is divided into sectors according to the part of the iris that reflects the organ. The chart has been described according to the division in the left eye and right eye [10], [24].

![Fig. 2. Left Iris and Right Iris Iridology Chart](image)

We can see in Fig. 2, the position of the heart is on the left side of the iris. The location of the heart is reflected through the left iris, which is shown in the iris zone 02.00 – 03.15.

B. Pre-Processing

The image captured through a digital camera has an RGB (Red, Green, Blue) format, so it is necessary to do a pre-processing process. In the pre-processing the iris image will be converted from an RGB image to a Grayscale image. Grayscale images have a simpler color value with a color intensity of 0 – 255 pixel thus shortening the computation time. The next step is to cut the unused part outside the iris, leaving a part that is close to the iris. Then, The eye image will be resized so that the entire image has the same pixel size when it enters the extraction process [25].

![Fig. 3. Image Pre-Processing Process](image)

Fig. 3 shows the pre-processing chart. The resized image will be processed into the CHT method to obtain the iris and eliminate the pupil area.

C. Circle Hough Transform

Images that have gone through the RGB to Grayscale conversion process will be separated between the iris and other objects that are not used, especially the pupil. The CHT method is used to determine the iris part automatically without human assistance in determining the coordinates of the iris. CHT can detect circles in the iris image and know between the outer iris and the outer pupil.

\[(x - a)^2 + (y - b)^2 = r^2\]  

Equation 1 describes the center circle \((a, b)\) and has a radius of \(r\). With \((x, y)\) is a pixel at the edge of the circle, then it can be represented in the form of a circle as [26]:

\[
\begin{align*}
    x &= a + r \cos \theta \\
    y &= b + r \sin \theta 
\end{align*}
\]  

Equation 2

Fig. 4 describes the CHT method that is implemented into the iris image. The iris image at point \(a\) has RGB format, which needs to be converted into a greyscale format. The histogram in the grayscale image shows high values in several white areas of objects that occur from light reflections and black colors that have values from the middle pupil of the iris.

![Fig. 4. CHT Process in Detecting Circles on the Iris](image)
Then the image at point b is an image display in determining the point of the circumference of the pupil and iris ring. If a part of the noise is not circular in this process, it will be removed. The image is then given a thickening process on the line to determine the circle in the image. Furthermore, the part that has been detected as a circle will be combined with the initial RGB image and generate an image like point c.

The CHT algorithm is used to separate the iris from the pupil and sclera. This automatic determination is carried out using edge detection in the form of a circle with a diameter value to determine the inner circle and outer circle of the iris. Edge detection is carried out to find objects with a diameter of less than 3 mm for the object removal process. Then the system will detect a circle with a diameter of more than 3 mm with a shape close to a perfect circle that is selected as the inner iris circle. In determining the outline of the outer iris circle, the system looks for the diameter of the circle measuring 12 mm. Determination of the diameter of the circle is adjusted to the process of taking pictures using constant light so that the images in each data taken to have the same size.

D. Normalization

The texture in the image has coordinates that represent the dimensions of the iris image, such as pupil dilation [27]. The iris image segmentation method aims to normalize the image in a different form but still has the exact resolution [28]. The iris can be modeled by two non-concentric circles and different textures within an iris circle. The center of the pupil can be used as a reference point for the circle on the iris [8]. A radial line runs through the area in the iris, known as radial resolution. Since the pupil is elastic to the iris, it is necessary to rescaling the reference point. The scaling equation is calculated based on the angle around the iris circle, with the equation:

$$r' = \sqrt{a\beta^2 - a - r_i^2} + \sqrt{a\beta}$$  \hspace{1cm} (3)

where,

$$a = a_x^2 + a_y^2$$  \hspace{1cm} (4)

$$\beta = \cos(\pi - \arctan(\frac{a_2}{a_1} - \theta))$$  \hspace{1cm} (5)

$r'$ is the distance between the pupil and the iris, while $\theta$ is the edge angle based on the radius of the iris. $a_x$ and $a_y$ is the displacement from the center of the pupil to the displacement of the center of the iris [29].

The Iris image that has a circular shape needs to be normalized based on the angle. The circular iris will be formed into a 2D array with horizontal dimensions at angles and vertical dimensions at radial. The circular iris will be formed into a 2D array with horizontal dimensions at angles and vertical dimensions at radial so that it will produce a rectangular or polar image shape as shown in Fig. 5 [30, 31].

E. Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE is a method to improve image quality by limiting the histogram value [32]. In this study, the CLAHE method is used to increase image intensity so that there will be a lot of detail that can be improved on the image. The results of using CLAHE are shown in Fig. 6.

F. Region of Interest A

The process of diagnosing the heart through the iris requires certain parts so that it does not require the whole to be processed. According to Iridology, the iris that reflects the heart is on the left iris, as shown in the sector 02.00 – 03.15 according to Fig. 1.

Region of Interest (ROI) is a process where we only need a specific part of the image to be processed. These sections will be cropped, leaving only the heart's reflection in the iris of the left, as shown in Fig. 7.

---

---
G. Gray Level Co-occurrence Matrix (GLCM)

GLCM is a texture analysis technique on images with a gray level. GLCM has a relationship between 2 neighboring pixels, which is determined by the intensity of gray, a certain distance, and angle. The equation for GLCM is shown below [33]:

\[ G(\Delta x, \Delta y)(a, b) = \sum_{i=1}^{P} \sum_{j=1}^{Q} 1\{I(i, j) = a\} \text{ and } 1\{I(+\Delta x, j + \Delta y) = b\} \]  

(6)

\(I(i, j)\) is the gray value of the column (i) and row (j) pixels, \((a, b)\) is a gray value that occurs at the same time as the calculation of \(G(\Delta x, \Delta y)(a, b)\). Then, \(1\{I(+\Delta x, j + \Delta y) = b\}\) is the indicator of \(\Delta x\) as a direction from \(x\) and \(\Delta y\) as a direction of \(y\) which is determined by the distance of the \(x\) and \(y\). \(P\) and \(Q\) shows the rows and columns of the corresponding image. There are 4 angles used in GLCM, 0°, 45°, 90° and 135°, described in Fig. 8.

The illustration in Fig. 9 is the use of GLCM for image pixels. Point a shows the use of distances that have a value of \(d = 1\) with 4 different directions where \(d\) is the distance between pixels. Point b shows the GLCM Usage calculation on an image with \(\theta = 0°\) and the distance between pixels is 1.

The image that has been calculated for certain distances and angles will then be transposed to the values obtained, and then the GLCM matrix values are added to the transpose results. The results will be normalized using the following equation:

\[ GLCM_{Norm} = \frac{GLCM_{Value}}{\sum_{i}^{N} GLCM_{Value}} \]  

(7)

where,

\(GLCM_{Value} = \) value of each pixel

The normalization results can be used to determine the texture characteristics of the image by obtaining information, such as Contrast, Dissimilarity, Homogeneity, Angular Second Moment (ASM), Energy and Correlation. The function of the contrast characteristic is the calculation of the difference in intensity between adjacent pixels in the entire image. Dissimilarity is the process of measuring the difference in a texture with a significant value if it is random, and vice versa will have a small value if it is uniform. Homogeneity functions to show the homogeneity of intensity variations in the image. ASM is a uniformity measurement process that produces a high value if the pixel values are similar to each other and vice versa will have a low value if the pixel values are different. Energy is used to measure the concentration of intensity pairs in the matrix, and Correlation is a linearity measurement of several pixel pairs. The equation to obtain each characteristic is shown in Table I [34]:
TABLE I. CHARACTERISTIC EQUATIONS ON GLCM

<table>
<thead>
<tr>
<th>Texture Characteristics of GLCM</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>( \sum_{a,b=0}^{level-1} P_{a,b} (a - b)^2 )</td>
</tr>
<tr>
<td>Dissimilarity</td>
<td>( \sum_{a,b=0}^{level-1} P_{a,b}</td>
</tr>
<tr>
<td>Homogeneity</td>
<td>( \sum_{a,b=0}^{level-1} \frac{P_{a,b}}{1 + (a - b)^2} )</td>
</tr>
<tr>
<td>ASM</td>
<td>( \sum_{a,b=0}^{level-1} P_{a,b}^2 )</td>
</tr>
<tr>
<td>Energy</td>
<td>( \sqrt{\sum_{a,b=0}^{level-1} P_{a,b}^2} )</td>
</tr>
<tr>
<td>Correlation</td>
<td>( \sum_{a,b=0}^{level-1} P_{a,b} \left( \frac{(a - \mu_a)(b - \mu_b)}{\sqrt{\sigma_a^2(\sigma_b^2)}} \right) )</td>
</tr>
</tbody>
</table>

where,

\( a, b \) = Pixel coordinates on the matrix

\( level \) = The range of grayscale value between 0-255 (level = 256)

\( P_{a,b} \) = value of coordinat pixel \( a, b \) on matrix GLCM

H. Support Vector Machine

Support Vector machine is one of the supervised machine learning that is commonly used in classification and regression analysis. SVM was widely introduced by Vapnik in the late 90s [35]. SVM can separate data by forming a hyperplane, maximizing margins and dividing data into different classes. The kernel function in SVM can separate data by looking for a hyperplane as a separator between data and finding the best hyperplane to put data in each class [36].

SVM as a classification method has two essential aspects: the first aspect is finding a hyperplane that can optimally separate two classes. The second aspect is the transformation of linearly inseparable classifications into separable ones [35].

Vector learning input and class \((x_i, y_i)\), where \( i = 1, 2, \ldots \) with \( x_i \in \mathbb{R}^n \) and \( y \in \{1, -1\} \). In solving problems using linear separation, hyperplane can define boundaries between classes -1 and +1 with the calculation representation as follows. [37]:

\[
\begin{align*}
\min_{\omega, b, \xi} & \quad \frac{1}{2} \omega^T \omega + C \sum_{i=1}^{n} \xi_i \\
\text{subject to} & \quad y_i (\omega^T \phi(x_i) + b) \geq 1 - \xi_i; \quad \xi_i \geq 0
\end{align*}
\]

(8)

Solving for classes that cannot be solved linearly, then vector \( x_i \) mapped into a higher dimension using the function \( \phi \) so that they can be separated linearly. Furthermore, SVM can define a hyperplane that separates linearly with a maximum margin of higher dimensional space. The SVM error parameter is known as \( C > 0 \). \( K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \) known as the Kernel function, shown in Table II [38].

TABLE II. SVM KERNEL EQUATION

<table>
<thead>
<tr>
<th>Kernel Name</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>( K(x_i, x_j) = x_i^T x_j )</td>
</tr>
<tr>
<td>Polynomial</td>
<td>( K(x_i, x_j) = (\gamma x_i^T + x_j + r)^d, \gamma &gt; 0 )</td>
</tr>
<tr>
<td>Gaussian</td>
<td>( K(x_i, x_j) = \exp \left( -\frac{</td>
</tr>
</tbody>
</table>
III. RESULTS AND ANALYSIS

In this study, system training was carried out using 75 normal iris data and 75 abnormal iris data. Normal iris data is the iris of people who have no history of heart disease; on the contrary for abnormal iris data is the iris of people who have heart disease.

Fig. 11 shows the training data using linear, polynomial, and gaussian kernel variations. Iris data in training can be separated according to normal (red) and abnormal (blue) classes. The results of linear kernel training separate the data into each class with an even distribution of data. The difference in the polynomial kernel training where the data has been separated and more centralized. The results of the training using the Gaussian Kernel resulted in a tighter grouping than using the two previous kernels. The results of the training are able to separate between classes according to existing characteristics, which can help in the classification of test data and affect the level of recognition accuracy. The further apart the hyperplane in the SVM that separates the classes, the higher the accuracy.

Tests were carried out on 50 normal iris data and 50 abnormal iris data. The data that has gone through the region of interest process is the iris data which reflects the heart. Then this section will be extracted using the GLCM method. The characteristics used in GLCM are Contrast, Dissimilarity, Homogeneity, Energy, and Correlation. The use of angle variations in the test is carried out to obtain optimal results. For example, Fig. 12 results from GLCM feature extraction at an angle with a distance of 1.

Taking five types of texture characteristics on iris data is needed to obtain more detailed information in training and testing. Normal and abnormal iris data have different characteristic values. The data can be classified into the appropriate class classification using SVM with three kernels: Linear, Polynomial, and Gaussian.

Simulation testing for variations angle in GLCM 0°, 45°, 90° and 135° with the SVM kernel variation is shown in Fig. 13. By using angle 0° the highest level of accuracy is obtained by using the Gaussian kernel which reaches 94%, linear kernel reaches 90% and Polynomial kernel 88%. The use of 45° obtained the highest accuracy rate with 88% Gaussian kernel, 86% polynomial kernel and 82% linear kernel. Then, for angle 90° the highest level of accuracy is achieved by Gaussian and Linear kernels which reach 92% and polynomial kernels which reach 88%. Meanwhile, with angle 135° the highest level of accuracy is achieved by 92% polynomial kernel, 90% Gaussian kernel and linear kernel which reaches 84%. From these data, it can be seen that the angle 0° and 90° has an average classification accuracy level of 90.67% at a distance of 1.

![Fig. 11. Results of Training using Kernel Variations; a) Linear kernel; b) Polynomial kernel; c) Gaussian kernel.](image-url)
GLCM Angle Efficiency Against SVM Kernel

From Table III, it can be concluded that the average for each angle variation using different SVM kernels has better performance in the classification process. By testing 100 normal and abnormal iris data, a high degree of accuracy was obtained. The use of linear SVM obtained an accuracy rate of 87%, then the use of polynomials obtained 89%. While the use of the kernel with the highest accuracy reaching 91%, was obtained using Gaussian.

Simulation tests with the Gaussian kernel proved to be superior in iris image classification. The Gaussian kernel

TABLE III. THE AVERAGE LEVEL OF ANGLE ACCURACY ON THE VARIATION OF THE SVM KERNEL

<table>
<thead>
<tr>
<th>SVM Classifiers</th>
<th>Training Time (s)</th>
<th>Accuracy Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linier</td>
<td>381.73</td>
<td>87%</td>
</tr>
<tr>
<td>Polynomial</td>
<td>422.28</td>
<td>89%</td>
</tr>
<tr>
<td>Gaussian</td>
<td>583.09</td>
<td>91%</td>
</tr>
</tbody>
</table>

Fig. 12. GLCM Characteristic Value using Angle 0° and d = 1 on Normal and Abnormal Iris Data.

Fig. 13. The Results of the Test with Variations of the GLCM Angle on the Variation of the SVM Kernel.

www.ijacsa.thesai.org
considers the probability of the density function of the standard deviation, squared, and variance. Gaussian can add data space into a higher dimensional vector to determine the intersection of hyperplanes more accurately. With flexible limiting performance, the Gaussian kernel can obtain a higher accuracy level than Linear and Polynomial kernels.

It is proven that normal and abnormal iris conditions can be identified using GLCM feature extraction and a classification process using the Gaussian method on the iris image. Identification is made to assist in providing an early diagnosis of heart conditions through the iris using the iridology method.

IV. CONCLUSION

This study has proposed a new method to determine the condition of the heart through the iris using the SVM classification with variations of the linear kernel, polynomial kernel, and Gaussian kernel. The use of GLCM characteristics as feature extraction has an essential role in the classification process. The main contribution in this study is not only limited to determining heart health conditions through the iris but also contributes to the automatic processing of the iris with CHT. The proposed system in determining the heart condition automatically is to optimize the classification by using angle 0° and 90° on GLCM with SVM classification on the Gaussian kernel to obtain a high level of accuracy. In ongoing research, the iris database can be added to improve the classification to make it more accurate. Different extraction methods can be used to get the results of image extraction with a smaller size so that it can increase the system's speed in iris identification.

ACKNOWLEDGMENT

This work was supported by the Kardiari Hospital in Semarang and the Department of Engineering at the Universitas Palangkaraya and the Department of Engineering at the Universitas Tanjungpura, Indonesia.

REFERENCES


