CLASSIFICATION OF ARRHYTHMIA DISEASES BY THE CONVOLUTIONAL NEURAL NETWORK METHOD BASED ON ECG IMAGES

Agustian Arditya Pratama¹*, Agus Maman Abadi²

¹,²Department of Mathematics Education, FMIPA, Yogyakarta State University
Karangmalang Campus, Colombo Street No.1, Yogyakarta, 55281, Indonesia

Corresponding author’s e-mail: *agustianardityap@gmail.com

ABSTRACT

Arrhythmia is a heart disorder that refers to an abnormal heartbeat rhythm. Arrhythmia detection uses an electrocardiogram (ECG) to describe the heart’s electrical activity. This research aimed to know the performance of the Convolutional Neural Network method in classifying arrhythmia diseases based on ECG signal images. Several stages were used to classify arrhythmias: the pre-processing data stage, CNN model formation stage, model compiling, training, model testing, and evaluation. The CNN model architecture that is formed involves 7 Convolution Layers, 7 Pooling Layers, 2 Dropout Layers, 2 Dense Layers, and 1 Flatten Layer, as well as ReLu and Softmax activation functions. The input variable in the classification process with CNN is an ECG image. The output variable is the classification of ECG signals into 17 classes, including normal sinus and pacemaker rhythms. The processed data are 1000 images; the division scenario is 750 training data and 250 testing data. The result of arrhythmia’s classification based on ECG image testing data using the CNN model shows the levels of Accuracy, Precision, Recall, and F1-score levels are 81%, 80%, 71%, and 73%, respectively. With the F1-score value as a measurement reference, the CNN model performs well in classifying ECG images.

How to cite this article:


Copyright © 2023 Author(s)
Journal homepage: https://ojs3.unpatti.ac.id/index.php/barekeng/
Journal e-mail: barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id

Research Article • Open Access
1. INTRODUCTION

Arrhythmia is a heart disorder that refers to an abnormal heart rate or rhythm, such as a heart rate that is too fast (tachycardia), too slow (bradycardia), and irregular [1]. There are various arrhythmias; one type is the most common, namely atrial fibrillation [2]. The condition of this type of arrhythmia is that the heart beats irregularly and quickly, increasing sufferers' risk of stroke and heart failure. Rohmantri and Surantha, in their journal, suggested that arrhythmia is a life-threatening heart disorder [3]. One way to identify and diagnose this disorder as early as possible is to use an electrocardiogram (ECG) [4].

An electrocardiogram checks the heart's condition and assesses the effectiveness of cardiac treatment by measuring and recording the heart's electrical activity [5]. The process of recording and detecting the heart's electrical impulses uses a device called an electrocardiograph. On the other hand, through Rohmantri and Surantha's journal, Klein said that an ECG is a diagram produced by an electrocardiograph sensor that records electrical impulses in the heart [6]. The results of electrocardiograph sensors are in the form of graphs or signals that can be used to diagnose arrhythmias in the heart. In addition to ECG signals, a phonocardiogram (PCG) can be used to diagnose heart disease [7], [8].

An ECG signal consists of a single heartbeat wave containing P waves, QRS complexes, and T waves with peak/primary waves P, Q, R, S, and T, respectively [9], [10]. However, this ECG process takes a long time, and the lack of experts to handle these cases makes it difficult to do the classification process manually. Therefore, to overcome this problem, a classification method was used.

ECG signal classification methods such as Support Vector Machine (SVM) [11], [12], Neural Network (NN) [13], and Deep Learning Convolutional Neural Network (CNN) methods are widely proposed in arrhythmia classification research [14]. Each method has advantages and disadvantages in its application. Wibawa et al., through their journal, stated that the SVM method could solve the over-fitting problem even though the optimal parameters are difficult to choose [15]. In the NN method, various noise data are highly tolerated, but the processing is very long and complex to interpret. Finally, the CNN method automatically learns filters to extract specific and relevant features from the input data, although the model training process can take a long time. It follows Bajaj and Kumar's statement that CNN has the advantage of combining image extraction and classification [16]. The feature extraction part has the function of automatically extracting images from ECG signals. In contrast, the classification part has the process of accurately classifying the signals using image extraction and is expected to produce a reasonably good accuracy [17].

In previous research [18], Fansyuri implemented the 1-Dimensional CNN method to classify heart disease classes based on ECG signal data into five scenarios. The study used the primary dataset, namely the PTB database, and datasets for adding courses, namely BIDMC Congestive Heart Failures, BIT-MIH Normal Sinus Arrhythmia, and The China Physiological Signal Challenge 2018. From the fifth trial scenario, the best results were obtained with accuracy, sensitivity, specificity, precision, F1-score, and error of 100.00%, 99.98%, 100.00%, 100.00%, and 0.00.0%, respectively. In other research with the CNN method, the classification of data based on images was carried out using the CNN-2D method [19], [20].

Based on the description of the problem above, this study aimed to determine the performance of the Convolutional Neural Network method in classifying arrhythmia disease based on ECG signal images. The selection of the CNN method is due to the input variable in the form of ECG signal images. Based on previous research, CNN can combine image extraction and classification with reasonably good accuracy results. This research was developed using Python for model building and GUI applications. The difference between this research and previous research lies in the type of CNN model used, namely 2D CNN, and the types of arrhythmia diseases classified, namely 17 classes (including normal sinus rhythm and pacemaker rhythm) based on ECG images from the MIT-BIH Arrhythmia Database.

2. RESEARCH METHODS

This research will discuss the performance of the CNN application as a method for arrhythmia classification. The classification data is an ECG signal dataset from the MIT-BIH Arrhythmia database obtained through the Mendeley Data website [21]. The processed data was in the form of ECG images, a total of 1000 image fragments, and consisted of 17 classes, including Normal Sinus Rhythm (NSR), Pacemaker Rhythm (PR), and 15 types of arrhythmic heart disorders. The steps are taken to classify arrhythmic diseases, namely:
a. Performing data pre-processing, images (.JPG) converted from numerical data (.MAT) using Matlab were allocated to the training data by 75% and testing data by 25% and then uploaded to Google Drive so that it can be connected to Google Colab. The images placed into the testing data were randomly selected by looking at the number of ECG signal images in each patient within the same class. Both datasets will go through several other pre-processing stages: gray-scale, resize, cropping, augmentation, and image data labeling.

b. Establishing a Convolutional Neural Network architecture to classify disease classes based on images. Each CNN input layer has a 3-dimensional neuron arrangement of width, height, and depth. The amount of output obtained is bound by the results of the titration of the previous layer and the number of filters used. In general, CNN layers are divided into two types, namely:

b.1. Feature Extraction Layer
The layer receives input from the image directly and processes it with a convolutional layer, max pooling layer, and ReLu (Rectifier Linear Unit) activation function. The equation for max-pooling is defined as \( f(x) = \max(0, x) \).

b.2. Classification Layer
It is a type of layer composed of neurons that are fully connected to other layers. The transformation output is the class accuracy for classification with the Sigmoid or Softmax activation function.

c. Compiling the model with a learning rate of 0.001.

d. Conducting model training to determine the accuracy of the CNN architecture that has been created. The number of epochs (iterations) used in training is 80, and the batch size value specified is 20 images. During the training process, data will be randomly taken from as many as 20 images from all dataset samples for each epoch until all epochs meet the sample limit.

e. Testing the model by classifying arrhythmia disease images on testing data using the CNN method.

f. Measuring the performance of a classification model by referring to several parameters, namely the accuracy rate, precision value, recall value, and f1-score value. To calculate these measurement parameters, a contingency table called Confusion Matrix is required. The contingency table is represented as a 2 x 2 matrix, as shown in Table 1.

<table>
<thead>
<tr>
<th>Predicted Class</th>
<th>True Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Positives</td>
</tr>
<tr>
<td>Negative</td>
<td>Negatives</td>
</tr>
</tbody>
</table>

Table 1. Confusion matrix

Based on Table 1, the four possible condition values are as follows.

- True Positive (TP), where the model predicts data in the Positive class and the actual data is in a Positive class.
- False Positive (FP), where the model predicts data in the Positive class, but the actual data is in the Negative class.
- True Negative (TN), where the model predicts data in the Negative class and the actual data is in the Negative class.
- False Negative (FN), where the model predicts data in the Negative class, but the actual data is in the Positive class.

The accuracy, precision, recall, and f1-score levels can be calculated using the following confusion table formulas.

\[ \text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \]

\[ \text{Precision} = \frac{TP}{TP + FP} \]

\[ \text{Recall} = \frac{TP}{TP + FN} \]
\[
d) \quad \frac{1}{F_1} = \frac{1}{2} \left( \frac{1}{\text{precision}} + \frac{1}{\text{recall}} \right)
\]

The following is a diagram of the research steps in classifying arrhythmia diseases in the heart.

![Diagram of research steps](image)

**Figure 1. Diagram of research steps**

### 3. RESULTS AND DISCUSSION

#### 3.1. Results

Image data processing was done with the Python programming language and the CNN method. The first stage after data collection is the process of converting an image to make it easier to understand so that it can increase the accuracy of the analysis results. This process is called image data pre-processing. One of the results of converting numerical data (.MAT) into image data (.JPG) is shown in **Figure 2**.

![Visualization of ECG signal data](image)

**Figure 2. Visualization of ECG signal data**

Because the output from Matlab has a size that is too large, the image resizes process was carried out to facilitate the processing of the image dataset. All ECG image data (**Figure 2**), initially 1058 x 530 pixels in size, were reduced to 1058 x 265 pixels. Furthermore, the color image (RGB) from the resize process was converted into a gray-scale to simplify the image model. Gray-scale images can be filtered to realize a more optimal image, as shown in **Figure 3**.

![Gray-scale image with filters](image)

**Figure 3. Gray-scale image with filters**

Before the data labeling process, cropping of the gray-scale image from 1058 x 265 pixels to 265 x 265 pixels was performed. It aimed to identify the characteristics of each ECG image according to each class. Image cropping was done without equalizing the wavelength of each ECG signal image fragment, and the center point is the midpoint of the image (**Figure 3**).
Data augmentation was a stage in image data processing where the data was altered or modified so that the model would detect that the altered image was different. Although the model considers it an additional image, humans can tell that the modified image is the same as in Figure 5. The augmentation process aimed to improve the accuracy of the trained CNN model. It is because the model will get additional data to make the model better at concluding.

The last data pre-processing stage is labeling the training and testing data images. Labeling used decimal numbers 0 (zero) to 16 (sixteen) according to the number of classes in the dataset and in the form of an array. The order of labeling given to each class is shown in Table 2.

<table>
<thead>
<tr>
<th>Label</th>
<th>Class</th>
<th>Label</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>AFIB</td>
<td>6</td>
<td>NSR</td>
</tr>
<tr>
<td>1</td>
<td>AFL</td>
<td>7</td>
<td>PR</td>
</tr>
<tr>
<td>2</td>
<td>APB</td>
<td>8</td>
<td>PVC</td>
</tr>
<tr>
<td>3</td>
<td>FVNB</td>
<td>9</td>
<td>RBBBB</td>
</tr>
<tr>
<td>4</td>
<td>IVR</td>
<td>10</td>
<td>SDHB</td>
</tr>
<tr>
<td>5</td>
<td>LBBBB</td>
<td>11</td>
<td>SVTA</td>
</tr>
<tr>
<td>12</td>
<td>VBY</td>
<td>13</td>
<td>VFL</td>
</tr>
<tr>
<td>14</td>
<td>VT</td>
<td>15</td>
<td>VTY</td>
</tr>
<tr>
<td>16</td>
<td>WPW</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The formation of CNN network architecture can affect the accuracy of the model. The architecture was used in the model training process to form the CNN model, as shown in Table 3. The ECG signal image processed by the model is 265 x 265 pixels (Figure 4) gray-scale with JPG format, where each ECG signal fragment from one patient can be in two or more arrhythmia classes.
Table 3. Convolutional neural network model architecture

<table>
<thead>
<tr>
<th>No.</th>
<th>Layer</th>
<th>Size</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Input</td>
<td>265 x 265 x 3</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>Conv2d_1</td>
<td>265 x 265 x 4</td>
<td>((4<em>4</em>3)+1)*4 filter = 196</td>
</tr>
<tr>
<td>2</td>
<td>MaxPool2D_1</td>
<td>133 x 133 x 4</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Conv2d_2</td>
<td>133 x 133 x 16</td>
<td>((4<em>4</em>4)+1)*16 filter = 1040</td>
</tr>
<tr>
<td>4</td>
<td>MaxPool2D_2</td>
<td>67 x 67 x 16</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Conv2d_3</td>
<td>67 x 67 x 32</td>
<td>((4<em>4</em>16)+1)*32 filter = 8224</td>
</tr>
<tr>
<td>6</td>
<td>MaxPool2D_3</td>
<td>34 x 34 x 32</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Conv2d_4</td>
<td>34 x 34 x 32</td>
<td>((4<em>4</em>32)+1)*32 filter = 16416</td>
</tr>
<tr>
<td>8</td>
<td>MaxPool2D_4</td>
<td>17 x 17 x 32</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Conv2d_5</td>
<td>17 x 17 x 64</td>
<td>((4<em>4</em>32)+1)*64 filter = 32832</td>
</tr>
<tr>
<td>10</td>
<td>MaxPool2D_5</td>
<td>9 x 9 x 64</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>Conv2d_6</td>
<td>9 x 9 x 128</td>
<td>((4<em>4</em>64)+1)*128 filter = 131200</td>
</tr>
<tr>
<td>12</td>
<td>MaxPool2D_6</td>
<td>5 x 5 x 128</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>Conv2d_7</td>
<td>5 x 5 x 512</td>
<td>((4<em>4</em>128)+1)*512 filter = 1049088</td>
</tr>
<tr>
<td>14</td>
<td>MaxPool2D_7</td>
<td>3 x 3 x 512</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>Flatten</td>
<td>4608</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>Dense</td>
<td>512</td>
<td>(4608*512)+512 = 2359808</td>
</tr>
<tr>
<td>17</td>
<td>Output</td>
<td>17</td>
<td>0</td>
</tr>
</tbody>
</table>

The results obtained after the model training process are the accuracy level on training data of 1.0000 and the loss value of 0.0014. Then, for accuracy on test data of 0.8120, the loss value is 1.9445. Figure 6 shows a graph of accuracy and loss of model training results on training and testing data in each iteration.

![Training and Validation Accuracy Graph](image1)

![Training and Validation Loss Graph](image2)

Next, the model was tested by classifying the ECG signal image in the testing data. The test was carried out by inputting image information into the model. Then the model provides prediction results that best match the image information obtained. The Python script used for the testing process is as follows.

```python
img = load_img(picture, target_size=(265,265))
img = img_to_array(img)
img = np.expand_dims(img, axis=0)
p = model.predict(img)
result = p[0]
```
The results of testing the CNN model against the testing data are presented visually through a confusion table (confusion matrix) to make it easy to read and evaluate. The confusion matrix is shown in Figure 7.

Measuring the performance of a classification model can be done through several methods, one of which is by finding the accuracy, precision, recall, and f1-score values. Based on Figure 7, the accuracy, precision, memory, and f1-score values on the testing data can be found using the confusion table formulas. However, the search for these measurement values can be done systematically using a Python program with the following script.

```python
y_pred = model.predict(X_test)
print(classification_report(y_test.argmax(axis=1),y_pred.argmax(axis=1),target_names=labels))
```

**Output:**

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
<th>f1-score</th>
<th>support</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFIB</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>34</td>
</tr>
<tr>
<td>AFL</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>5</td>
</tr>
<tr>
<td>APB</td>
<td>0.70</td>
<td>0.44</td>
<td>0.54</td>
<td>16</td>
</tr>
<tr>
<td>FVNB</td>
<td>1.00</td>
<td>0.33</td>
<td>0.50</td>
<td>3</td>
</tr>
<tr>
<td>IVR</td>
<td>1.00</td>
<td>0.67</td>
<td>0.80</td>
<td>3</td>
</tr>
<tr>
<td>LBBBB</td>
<td>0.93</td>
<td>0.96</td>
<td>0.94</td>
<td>26</td>
</tr>
<tr>
<td>NSR</td>
<td>0.84</td>
<td>0.90</td>
<td>0.87</td>
<td>70</td>
</tr>
<tr>
<td>PR</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>11</td>
</tr>
<tr>
<td>PVC</td>
<td>0.61</td>
<td>0.70</td>
<td>0.65</td>
<td>33</td>
</tr>
<tr>
<td>RBBBB</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>15</td>
</tr>
<tr>
<td>SDHB</td>
<td>0.75</td>
<td>1.00</td>
<td>0.86</td>
<td>3</td>
</tr>
<tr>
<td>SVTA</td>
<td>1.00</td>
<td>0.33</td>
<td>0.50</td>
<td>3</td>
</tr>
<tr>
<td>VBY</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>14</td>
</tr>
<tr>
<td>VFL</td>
<td>0.75</td>
<td>1.00</td>
<td>0.86</td>
<td>3</td>
</tr>
<tr>
<td>VT</td>
<td>1.00</td>
<td>0.67</td>
<td>0.80</td>
<td>3</td>
</tr>
<tr>
<td>VTY</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>3</td>
</tr>
<tr>
<td>WPW</td>
<td>0.67</td>
<td>0.80</td>
<td>0.73</td>
<td>5</td>
</tr>
</tbody>
</table>

**accuracy** 0.81  250
Based on the classification report output on the testing data, it can be seen that the processed dataset is more than 2 classes (multi-class) and unbalanced. The testing dataset is unstable because there is one class with data that is much different from the other classes, which is 70 images. The class is NSR (Normal Sinus Rhythm). Therefore, the model performance was measured based on the macro-F1 value because the average macro treats all classes equally, where all classes are equally important regardless of the number of images in each class.

Thus, the average value of the macro f1-score is 73%, with precision and recall values of 80% and 71%, respectively. The calculation of the average macro value through the classification report output of the testing data is as follows.

\[
\text{Precision}_{\text{macro}} = \frac{\sum \text{Class Precision Value}}{\sum \text{Total of classes}} \times 100\% = \frac{0.88 + 0.8 + 0.7 + \ldots + 0.67}{17} \times 100\% = 80\%
\]

\[
\text{Recall}_{\text{macro}} = \frac{\sum \text{Class Recall Value}}{\sum \text{Total of classes}} \times 100\% = \frac{0.88 + 0.8 + 0.44 + \ldots + 0.8}{17} \times 100\% = 71\%
\]

\[
\text{f1-score}_{\text{macro}} = \frac{\sum \text{Value f1-class score}}{\sum \text{Total of classes}} \times 100\% = \frac{0.88 + 0.8 + 0.54 + \ldots + 0.73}{17} \times 100\% = 73\%
\]

The macro precision value of 80% means that the CNN model’s ability to provide identification results from each testing data image as the same class is 80% correct. Meanwhile, the macro recall value of 71% means that each ECG image of arrhythmia disease classified with the CNN model (Table 3) has a relevance level of 71% with the classification results. Finally, the macro f1-score value of 73% means that the CNN model performs well on the testing data because the accuracy and relevance between the identification and classification results on the image on the ECG signal image of arrhythmia disease is 73%.

From the actual testing data, an accuracy value of 81% is obtained, which means that 81% of the images from the arrhythmia disease testing data can be classified according to their class accurately using the CNN model that has been created. The model’s accuracy value can be found using the confusion table formula point (a), based on Figure 7. The confusion table shows that the data that was successfully detected correctly amounted to 203 and detected incorrectly amounted to 47, so the model’s accuracy value is as follows.

\[
\text{Accuracy} = \frac{\sum \text{Data of correct detection}}{\sum \text{Total of test data}} \times 100\% = \frac{203}{250} \times 100\% = 81.2\%
\]

Looking at the results above, the accuracy obtained on the testing data is 81.2%.

3.2. Discussion

Based on the research that has been conducted, it is found that the CNN model used to classify arrhythmia diseases into 17 classes based on ECG images has good performance. The dataset obtained through Mendeley Data [21] was converted from numerical data (.MAT) to image (.JPG) using Matlab. Next, the image was put into the Pre-processing stage to alter the color image into a gray-scale image. In addition, image processing that was also carried out at this stage was resizing, cropping, augmentation, and labeling. The CNN model architecture used in ECG image detection and classification involved 7 Convolution Layer, 7 Pooling Layer, 2 Dropout Layer, 2 Dense Layer, and 1 Flatten Layer as ReLu and Softmax activation functions.

Performance measurement was based on the f1-score value of 73% with a precision value of 80%, a recall value of 71%, and an accuracy rate of 100% training and 81% testing. This result follows the results of Rohmantri and Surantha’s research. Rohmantri and Surantha [3], who used CNN-2D to classify heart
disorders into 7 classes, obtained an f1-score value of 98% with a precision value of 98%, a recall value of 98%, and an accuracy rate of 98%. In that study, the CNN model architecture used for classification involved some 2 Convolution Layer, 1 Pooling Layer, 2 Dropout Layers, 2 Dense Layer, and 1 Flatten Layer, as well as ReLu and Softmax activation functions.

Based on the results of model testing and model evaluation, prediction errors in ECG image testing data are likely due to determining the value of model training parameters that were still inappropriate. The parameters include the effect of image input size, training and testing data division scenarios, the impact of the number of epochs, kernel size, batch size, and learning rate. An arrhythmia disease classification application was created using the Python programming language and based on the Graphic User Interface (GUI) to implement the classification model quickly, as shown in Figure 8.

![Figure 8. Arrhythmia disease classification GUI application](image)

4. CONCLUSIONS

Based on the results of arrhythmia disease classification through ECG images (17 classes), the CNN method classification model was divided into several stages: Data Pre-processing, CNN Model, Model Compile, Training and Testing, and ends with an evaluation. The CNN model architecture used in the detection and classification of ECG images involves 7 Convolution Layer, 7 Pooling Layer, 2 Dropout Layer, 2 Dense Layer, and 1 Flatten Layer, as well as ReLu and Softmax activation functions. The detection and identification results based on ECG images obtained a precision value of 80%, a recall value of 71%, an f1-score discount of 73%, and an accuracy rate of 100% training and testing of 81%. With the f1-score value as a measurement reference, the CNN model performs well in classifying ECG images. The model in this study used an input image of 265 x 265 pixels, a learning rate of 0.001, a filter kernel of size 4 x 4, several epochs of 80, and a data division scenario for training data of 750 images and testing data of 250 prints.

REFERENCES

[2] S. H. Rampengan, Kardioversi Pada Fibrilasi Atrium [Cardioversion in Atrial Fibrillation], Jakarta: Faculty of Medicine, University of Indonesia, 2015.
Pratama, et. al.  

634  

CLASSIFICATION OF ARRHYTHMIA DISEASE BY CONVOLUTIONAL NEURAL...  


