Journal of Information Systems and Informatics
Vol. 4, No. 2, June 2022 e-ISSN: 2656-4882 p-ISSN: 2656-5935
http://journal-isi.org/index.php/isi Published By DRPM-UBD

Classification of Fetal Heart Based on Images Augmentation Using Convolutional Neural Network Method

Muhammad Taufik Roseno[1], Hadi Syaputra*[1,2]

[1]Doctoral Program in Engineering Science , Universitas Sriwijaya, Palembang, Indonesia
[2]Teknik Informatika, Universitas Bina Darma, Palembang, Indonesia
Email: [1]mtroseno@gmail.com, [2]hadisyaputra@binadarma.ac.id

Abstract

Standard fetal heart echocardiography view consists of several specific views that can be prolific to optimize the visualization of various structures and anomalies including three vessel and trachea view, right ventricular outflow tract view, four chamber view, left ventricular outflow tract, and right ventricular outflow tract. With the use of current technological developments specifically deep learning, it can classify images from the visualization of the echocardiography point of view obtained. One of the deep learning models that has the best performance in image recognition and classification is the Convolutional Neural Network. Research consists of several stages, namely data collection, data pre-processing, data augmentation, data sharing designing the Convolutional Neural Network model architecture, training, testing, and results. 5 types of echocardiography videos were used based on the echocardiography point of view, resulting in 3,995 images consisting of 3,196 training data and 799 test data. The implementation of convolutional neural networks for the classification of fetal echocardiography images based on point of view obtained good results. The Convolutional Neural Network used consists of 2 convolution layers, 2 layers, 1 flatten layer, 2 dense layers, and 2 Dropout layers. The accuracy rate obtained from the CNN model with a learning rate value of 0.01 and the number of epochs of 50 gets an accuracy value of 98%.

Keywords: Convolutional Neural Networks, Deep Learning, Echocardiography, classification.

1. INTRODUCTION

Ultrasound (USG) is one of the efforts to examine the womb or Ante Natal Care (ANC) in pregnant women which aims to determine the condition of the fetus in the body. Examination of the womb using ultrasound can determine the presence or absence of pregnancy, the life or absence of the fetus, the location of the placenta, and the gestational age [1]. The use of ultrasound in the world of health is also known as obstetric sonography, carried out in pregnancy which

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

aims to determine the anatomy of the fetus. Anatomical examination of the fetus is part of the standard obstetric ultrasound examination performed to identify the main internal organs of the fetus [2]. Ultrasound can be an option for diagnosing other things in the body and pregnancy checks because it does not cover the risks that are severe such as the risks that are applied from diagnostic equipment that uses radioactive substances [3].

Special ultrasound to examine the fetal heart is called echocardiography (ECG), ECG is an examination method that uses high-frequency sound waves to capture a picture of the structure of the heart organ. ECG aims to check for abnormalities in the structure of the heart, blood vessels, blood flow, as well as the ability of the heart muscle to pump blood, this imaging method can be used to detect heart disease find the right treatment, and evaluate the treatment given. The ECG has several points of view, including three vessel and trachea view (3VT), right ventricular outflow tract view (3VV), four chamber view (4C), left ventricular outflow tract (LVOT), and right ventricular out tract flow (RVOT) , tends to have a greater chance of detecting congenital heart disease. In general, heart defects can be seen based on size, spatial position, ductus arteriosus, aortic arch, superior vana kava, trachea, as well as holes (defects) that will be very helpful in determining cardiac disability [4].

Based on previous research, where it tends to refer to not achieving good results. This is due to several factors such as the small size of the object, low image quality, and the diversity of the shape of the object in everyone. There are several factors that can cause a low-quality image, one of which is the continuous and unstructured change in shape [5]. This research uses fetal heart ECG data which aims to recognize the pattern of the fetal heart perspective by a computer. So, to recognize the perspective of the fetal heart on the side of computer technology or application programs, there are many methods used, one of which requires understanding deep learning.

Deep Learning is one of the classes of machine learning algorithms that has multiple screens consisting of nonlinear processing units [6], one of the deep learning models that has the best performance in image recognition and classification is the Convolutional Neural Network (CNN) [7]. CNN is able to carry out an independent learning process for object recognition, object abstraction and image calcification [8]. Several studies on image processing using CNN method got good accuracy results, namely research conducted by Isna Wulandari classification of digital images of herbs and spices with 2 convolutional layers, with a training data accuracy value of 0.9875 and a loss value of 0.0769, a testing data accuracy value of 0.85 and a loss value of 0.4773 [9] . Nur Fadlia's research on vehicle type classification using the convolutional Neural Network (CNN) method with a total of 120 images consisting of 3

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

classes, namely car, motorcycle, and bicycle images. With an accuracy result of 94.4% at the training stage and 73.3% at the testing stage [10]. Budi Nugroho's research was entitled The Work of the CNN Method for Pneumonia Classification with variations in input image size, by producing an accuracy value of 8.81% and an F1 analysis score of 0.0729 [11]. Based on the description from the background that has been presented, this study discusses how to build a fetal heart ECG classification model based on ECG video on the fetal heart from the point of view of 3VT, 3VV, 4C, LVOT, and RVOT by implementing the CNN method and analyze the performance of the model that has been built. Based on the identification of the problem, it is necessary to have a problem limit so that the scope of the study becomes clear, then the limitations in this study are that the model built can only classify 5 types of ECG viewpoint images on the fetal heart, namely: 3VT, 3VV, 4C, LVOT, and RVOT.

2. METHODS

This study aims to classify fetal heart images based on fetal heart ECG results with a point of view of 3VT, 3VV, 4C, LVOT, and RVOT, using CNN. The research stage can be seen in figure 1.

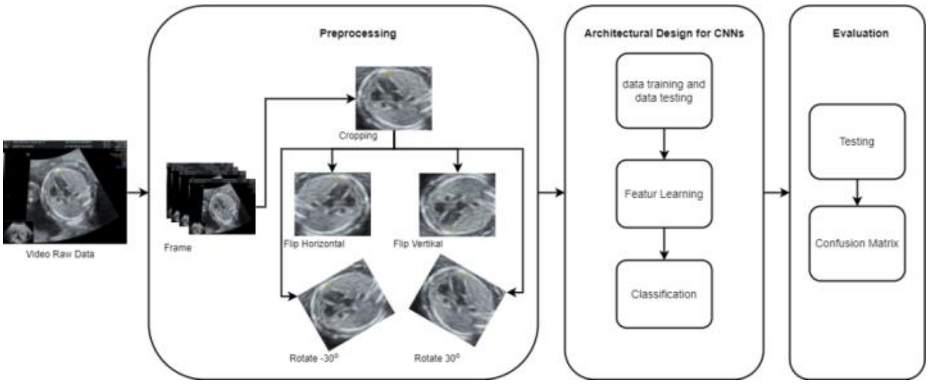


Figure 1. Research Scheme

Based on figure 1, raw data videos are pre-processed by converting videos into imagery, these images display the point of view of the fetal heart 3VT, 3VV, 4C, LVOT, and RVOT, then the imagery that has been obtained is carried out a crop process to obtain a focused image according to the type of viewpoint of the fetal heart. After that, the crop image is carried out the image augmentation process by multiplying the image source that has been cropped into a horizontal flip image, a vertical flip image, a -30^{o} rotation image and

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

a30 o 30º r o image. with the preprocessing process the next stage is the sharing of image data, namely data for training and testing, the data that has been divided will be continued with the design of the CNN model. The CNN model that has been designed and built is then used for the process of testing and measuring accuracy.

2.1. Pre-processing Data

Before the classification process using CNN is carried out, the image is first pre-processed. The 3VT, 3VV, 4C, LVOT, and RVOT angle-of-view images that have been collected have different pixel sizes during the process of converting video data into imagery. Therefore, the stage of data pre-processing that is carried out is to change the pixel size of the original image so that each image has the same size which is 512 x 512 pixels.

2.2. Splitting Data

The angle-of-view image data of 3VT, 3VV, 4C, LVOT, and RVOT are further divided into two types, namely training data and test data. Training data is used to conduct learning, and test data is used to conduct testing after training. The training data was 3,196 images, the data and testing data were 799 images.

2.3. CNN Model Architecture Design

According to the CNN model in this study, the input image used was 512 x 512 x 3. With 512 x 512 is the length and width of the image and 3 are the color components owned, namely RGB (Red, Green, Blue). The CNN model to be developed consists of 2 convolution layers and 2 pooling layers. The inserted image will be convoluted in the first step with a layer filter size of 3 x 3, where this convolution process will reduce the rows and columns in the image. The next stage is to perform a max pooling operation. The operation performed is matrix multiplication between the results of the previous convolution measuring 3×3 with a 2 x 2 pooling filter. In the second stage, convolution operations with a layer filter size of 2×2 and a pooling filter of 2×2 are carried out. After the convolution and pooling process has been completed, the flatten and fully connected process is carried out. This process aims to convert the feature map resulting from the pooling layer into a vector form. figure 2 is a design of the CNN model used in this study.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

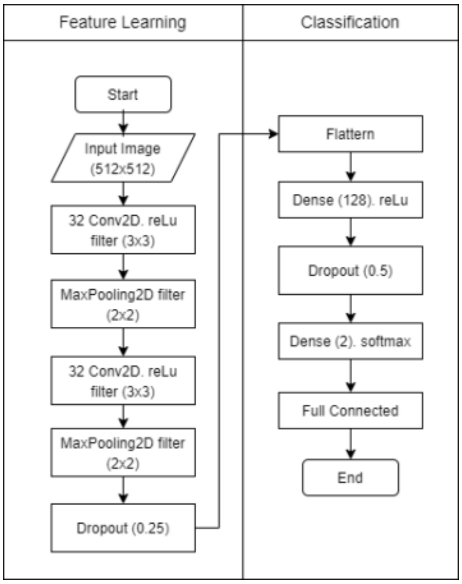


Figure 2. CNN Model Architecture Design

2.4. Training

In the training process, there are several parameter values that are initialized, including the number of epochs, and learning rate. The number of epochs and learning rates that provide optimum accuracy is not yet known. Therefore, in this study, training was carried out using several numbers of epochs and several learning rates values so that the number of epochs and learning rates were obtained that provide optimum accuracy. According to the CNN model in this study, the input image used was 512 x 512 x 3. With 512 x 512 is the length and width of the image and 3 are the color components owned, namely RGB (Red, Green, Blue). For its optimization using Stochastic Gradient Descent and the momentum value is 0.9

2.5. Testing

After conducting a training process on the CNN model, a testing process was carried out to test the model to classify the image according to its class. The testing process will be carried out using test data of 799 images.

2.6. Classification Result

The determination of whether or not the performance of a classification model is good can be seen from its performance measurement parameters, namely the level of accuracy, sensitivity, and precision [12]. To calculate these factors requires a matrix commonly called a confusion matrix [14]. Some of the values in confusion matrix are: True Positive (TP), True Negative (TN), False Positive

(FP), and False Negative (FN). The entire probability of the actual event is positive (P) and the whole possibility of the actual event is negative (N). The value can be used to calculate accuracy with Equation (1).

$$\textbf{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}} \tag{1}$$

Meanwhile, to calculate the level of precision of predictions, the equation (2) can be used.

$$\textbf{Precission} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

Sensitivity can be calculated using equation (3)

$$\textbf{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

3. RESULTS AND DISCUSSION

Figure 3. Sample of raw data or 3VT, 3VV, 4C, LVOT, and RVOT (from the left to the right). The illustrated of the fetal heart is taken 3VT, 3VV, 4C, LVOT, and RVOT

3.1. Result of Pre-processing Data

Previously, the data was in the form of ECG videos that were processed first to produce images. The image obtained displays the entire display captured from the ECG video, figure 4 is the imagery display of the video result that is converted into imagery.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Figure 4. The result of the video's change to imagery.

After the process of obtaining an image, the next preprocessing is the process of image augmentation (figure 5), which is the process of doubling a data by translating, transforming, adding / reducing noise, rotation, magnification, or flipping of the image dataset [13] . The main purpose of data augmentation itself is to find out the characteristics of an optimal image which will be used as an analysis, through the process of taking information on an object, as well as for the process of compressing or reducing data for the purpose of storing data.

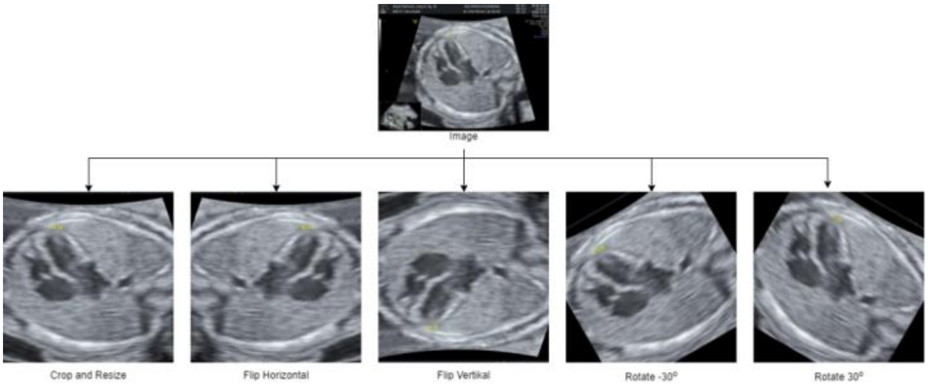


Figure 5. Image Augmentation.

3.2. Result of Comparing Data

After the pre-processing of image data is carried out, it is further divided into two types of data with different amounts. The training data is 3,196 images and the test data is 799 images which can be seen in Table 1.

Table 1. Amount of Training and Testing Data

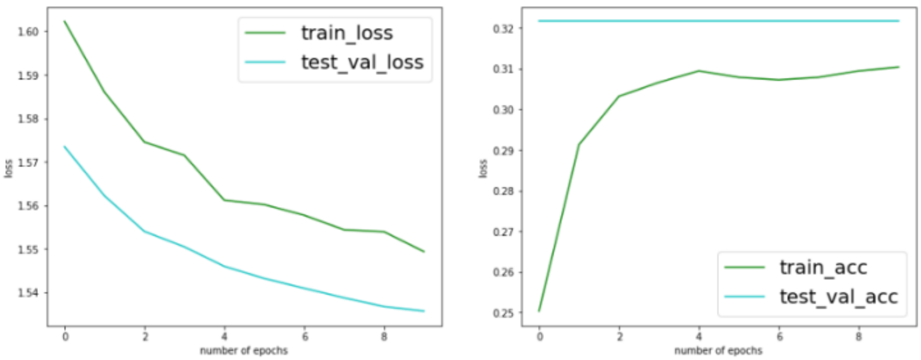
Dataset	Training Data	Testing Data	Total
Citra 3VT	1.000	250	1.250
Citra 3VV	636	159	795
Citra 4C	512	128	640

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Citra VOT	584	146	730
Citra VOT	464	116	580
Total	3.196	799	3.995

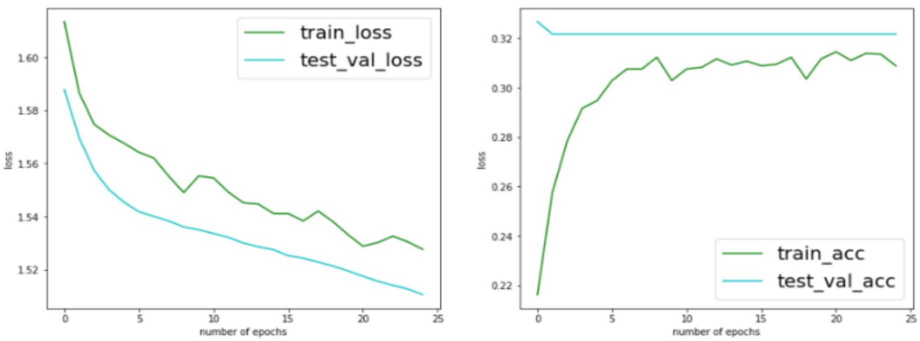
3.3. Training Results with Multiple Epochs

In the training stage, several numbers of epochs are carried out , namely 10, 25, and 50. The results of the training with epoch 10 and the learning rate value of 0.0001 obtained a loss value of 160.22%, an accuracy value of 31.04%, a loss validation value of 157.35%, and an accuracy validation value of 32.17% with a time of 42 seconds. A graph of training results with epoch 10 can be seen in figure 6.



(a) (b)
Figure 6. Training Results Graph with Epoch 10

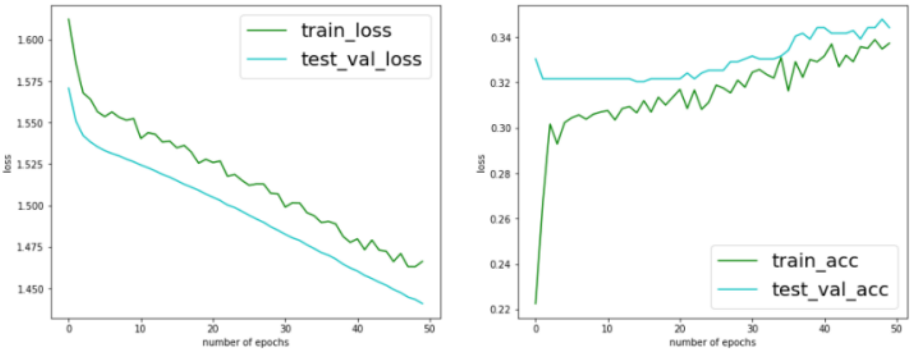
Figure 6(a) is a loss model chart and figure 6(b) is an accuracy model graph. Based on figure 6, the accuracy of the training result model with epoch 10 has not been optimal because the resulting graph is not convergent.



(a) (b)
Figure 7. Training Results Graph with Epoch 25

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Furthermore, training was carried out with an epoch of 25 and a learning rate of 0.0001. Based on the results of the training, the loss value was obtained, which was 161.34%, the accuracy value was 31. 45%, the loss validation value is 158.78% and the accuracy validation value is 32.67% with a time of 1 minute 35 seconds. Figure 7 shows a graph of training results with epoch 25. Based on Figure 7, the accuracy of the training result model with epoch 25 has not been optimal because the resulting graph is not convergent.



(a) (b)

Figure 8. Training Results Graph with Epoch 50

A graph of training results with an epoch of 50 and a learning rate value of 0.0001 can be seen in figure 8. The loss value of the training results is 157.07%, the accuracy value is 33.89%, the loss validation value is 157.07%, and the accuracy validation value is 34.79% with a time of 3 minutes 13 seconds. Based on figure 8, the accuracy of the training result model with an epoch of 50 is also not optimal because the resulting graph is not convergent but there is starting to be an improvement in terms of accuracy.

Table 2. Epoch Comparison

	10	25	50
Loss	**148.50%**	161.34%	161.24%
Accuracy	33.54%	31.45%	**33.89%**
Loss Validation	**146.52%**	158.78%	157.07%
Accuracy Validation	33.04%	32.45%	**34.79%**

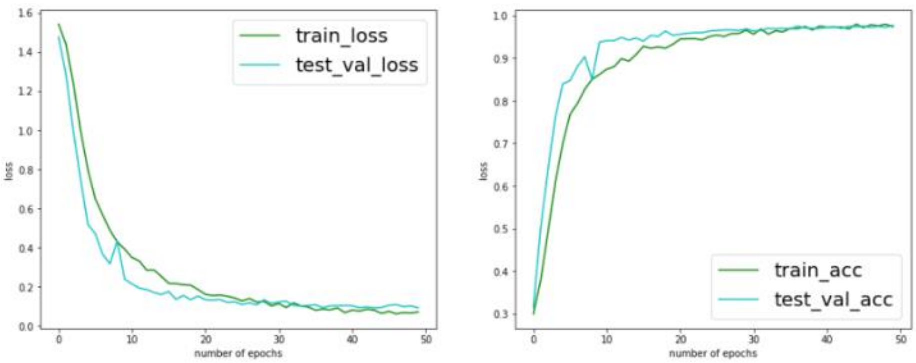
Table 2 is a comparison of loss values, accuracy, loss validation, and accuracy validation with several numbers of epochs, namely 10, 25, and 50. Based on Table 2, the smallest loss occurs at epoch 10 which is 148.50%, the highest accuracy occurs at epoch 50 which is 33.89%, the lowest loss validation occurs at epoch 10 which is 146.52%, and the highest accuracy validation occurs

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

at epoch 50 which is 34.79%. Therefore, it can be concluded that on epoch 50 it has an accuracy of more than epoch 10 and 25.

3.4. Training Results with Multiple Learning Rates

Judging from the results of training with some of the highest accuracy epochs occurring at epoch 50, then epoch 50 is used to carry out the training process again with different learning rates. Starting with epoch 50 with a learning rate value of 0.01, the loss value is 154.07%, the accuracy value is 98.00%, the loss validation value is 147.52%, and the accuracy validation value is 97.75% with a time of 3 minutes 18 seconds. Figure 9 shows a graph of training results with a learning rate of 0.01. The accuracy of the training model with a learning rate value of 0.01 is optimal because the resulting graph is convergent.

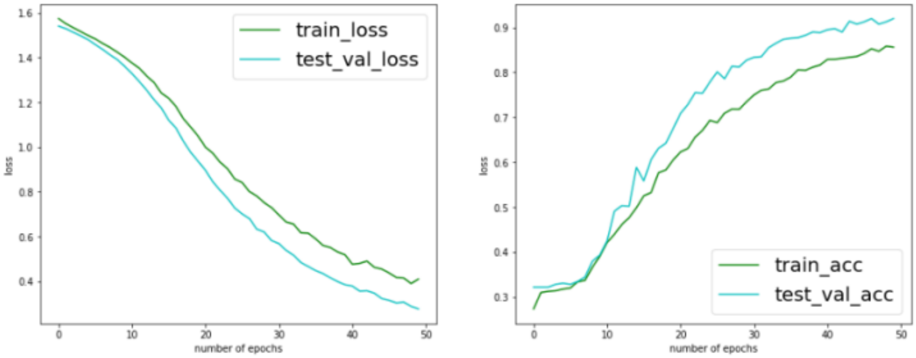


(a) (b)

Figure 9. Graph of Training Results with Learning Rate 0.01

The results of the training with a learning rate value of 0.001 and epoch 50 obtained a loss value of 157.47%, an accuracy value of 85.83%, a loss validation value of 154.09%, and an accuracy validation value of 91.99% with a time of 3 minutes 38 seconds. The accuracy of the training model with a learning rate value of 0.001 is not optimal because the resulting graph is not convergent as seen in figure 10.

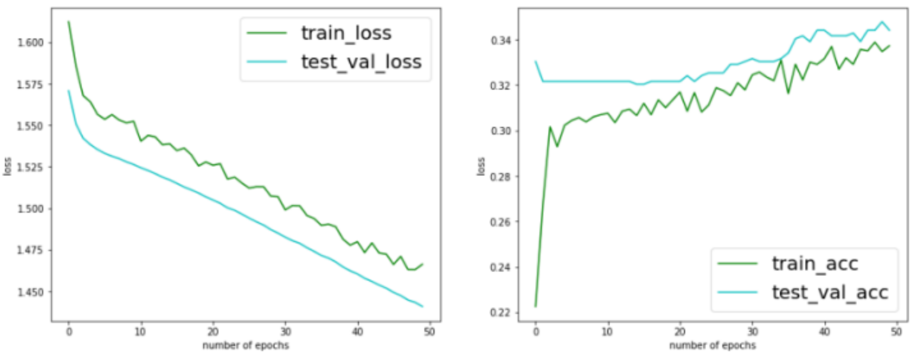
Journal of Information Systems and Informatics
Vol. 4, No. 2, June 2022
p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882



(a) (b)

Figure 9. Graph of Training Results with Learning Rate 0.001

The results of the training with a learning rate value of 0.0001 and epoch 50 obtained a loss value of 161.24%, an accuracy value of 33.89%, a loss validation value of 157.07%, and an accuracy validation value of 34.79% with a time of 3 minutes 13 seconds. The accuracy of the training result model with a learning rate value of 0.0001 is not optimal because the resulting graph is not convergent as seen in figure 11.



(a) (b)

Figure 9. Graph of Training Results with Learning Rate 0.0001

Table 3 shows the comparison of loss values, accuracy, loss validation, and accuracy validation of several learning rates, namely 0.01, 0.001 and 0.0001 at epoch 50. Based on Table 3, the smallest loss occurs at a learning rate of 0. 01 is 154.07%, the highest accuracy occurs at the time of learning rate 0. 01 is 98.00%, the lowest loss validation occurs at the time of learning rate 0. 01 is 147.52%, and the highest accuracy validation occurs at a learning rate of 0. 01 i.e., 97.75%.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Therefore, the learning rate value is 0. 01 already provides optimum accuracy and convergent model graphics.

Table 3. Comparison of Learning Rate with Epoch 50

	0.01	0.001	0.0001
Loss	**154.07%**	157.47%	161.24%
Accuracy	**98.00%**	85.83%	33.89%
Loss Validation	**147.52%**	154.09%	157.07%
Accuracy Validation	**97.75%**	91.99%	34.79%

3.5. Test Results

The results of the testing process will be displayed in a confusion matrix in Table 4. Based on Table 4, the results obtained from testing the models that have been trained are quite good. This can be seen by the correct predictions obtained. Predictions of 3VT, 3VV, 4C, LVOT, and RVOT angle-of-view images are correctly classified into 3VT as many as 166 images with an error of 2, 3VV as many as 120 images with errors of 7,4C as many as 253 images with errors of 4, LVOT a total of 125 images with 3 errors, and RVOT as many as 117 images with 2 errors.

Table 4. Confusion Matrix

Matrix	Prediction Class				
	3VT	3VV	4C	LVOT	RVOT
3VT	166	0	0	0	2
3VV	1	120	3	0	3
4C	0	4	253	0	0
LVOT	0	0	1	125	2
RVOT	0	0	1	1	117

Table 5 shows the experimental results of the CNN model by displaying the precision, recall, and f1-score values per class of viewpoint of 3VT, 3VV, 4C, LVOT, and RVOT images.

Table 5. Accuracy of the model CNN

class	precision	recall	f1-score	support
3VT	0.99	0.99	0.99	168
3VV	0.97	0.94	0.96	127
4C	0.98	0.98	0.98	257
LVOT	0.99	0.98	0.98	128
RVOT	0.94	0.98	0.96	119

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Based on Table 5, the highest precision value in the 3VT class and LVOT class with a value of 0.99, the highest recall value in the 3VT class with a value of 0.99, and the highest f1-score value in the 3VT class with a value of 0.99. for the average accuracy value is 0.98 with a total of 799 testing data.

4. CONCLUSION

The implementation of CNN for the classification of fetal heart images based on the ECG video point of view obtained good results. The CNN used consists of 2 convolution layers, 2 2 x 2 pooling layers, 1 flatten layer, 2 dense layers, and 2 Dropout layers. The accuracy rate obtained from the CNN model with a learning rate value of 0.01 and the number of epochs of 50 received validation accuracy of 98% and loss validation of 2.6%. therefore, it can be concluded that deep learning implementation using CNN able to carry out the classification of fetal heart image based on the point of view well. The classification results on the testing data as many as 799 image data have an accuracy of 98%. The classification of fetal heart images based on the point of view in this study still uses the division of datasets from the same ECG video source. Furthermore, the testing process uses different ECG video sources so that it can be seen the capabilities of more CNN models that have been developed.

REFERENCES

[1] L. T. Coilal, L. Anggraeni, and I. Gustina, "Gambaran Tingkat Pengetahuan Ibu Hamil Tentang Manfaat Ultrasonografi (Usg) Dalam Pemeriksaan Kehamilan," p. 4.
[2] E. L. Utari, "ANALISA DETEKSI TEPI JANIN DENGAN MENGGUNAKAN METODE PREWITT DAN CANNY," p. 10.
[3] S. Imardi and K. Ramli, "Pengembangan Dan Pengkayaan Fungsi Antarmuka Perangkat Lunak Untuk Visualisasi Dan Analisis Citra Ultrasonografi," p. 10.
[4] P. Veronese, G. Bogana, A. Cerutti, L. Yeo, R. Romero, and M. T. Gervasi, "A Prospective Study of the Use of Fetal Intelligent Navigation Echocardiography (FINE) to Obtain Standard Fetal Echocardiography Views," Fetal Diagn. Ther., vol. 41, no. 2, pp. 89–99, 2017, doi: 10.1159/000446982.
[5] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," IEEE Trans. Image Process., vol. 26, no. 7, pp. 3142–3155, Jul. 2017, doi: 10.1109/TIP.2017.2662206.
[6] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Netw., vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[7] A. Peryanto, A. Yudhana, and R. Umar, "Klasifikasi Citra Menggunakan Convolutional Neural Network dan K Fold Cross Validation," J. Appl. Inform. Comput., vol. 4, no. 1, pp. 45–51, May 2020, doi: 10.30871/jaic.v4i1.2017.
[8] F. F. Maulana and N. Rochmawati, "Klasifikasi Citra Buah Menggunakan Convolutional Neural Network," J. Inform. Comput. Sci. JINACS, vol. 1, no. 02, pp. 104–108, Jan. 2020, doi: 10.26740/jinacs.v1n02.p104-108.
[9] I. Wulandari, H. Yasin, and T. Widiharih, "Klasifikasi Citra Digital Bumbu Dan Rempah Dengan Algoritma Convolutional Neural Network (CNN)," J. Gaussian, vol. 9, no. 3, pp. 273–282, Aug. 2020, doi: 10.14710/j.gauss.v9i3.27416.
[10] N. Fadlia and R. Kosasih, "Klasifikasi Jenis Kendaraan Menggunakan Metode Convolutional Neural Network (CNN)," J. Ilm. Teknol. Dan Rekayasa, vol. 24, no. 3, pp. 207–215, 2019, doi: 10.35760/tr.2019.v24i3.2397.
[11] B. Nugroho and E. Y. Puspaningrum, "Kinerja Metode CNN untuk Klasifikasi Pneumonia dengan Variasi Ukuran Citra Input," J. Teknol. Inf. Dan Ilmu Komput., vol. 8, no. 3, p. 533, Jun. 2021, doi: 10.25126/jtiik.2021834515.
[12] Z. F. Abror, "Klasifikasi Citra Kebakaran Dan Non Kebakaran Menggunakan Convolutional Neural Network," J. Ilm. Teknol. Dan Rekayasa, vol. 24, no. 2, pp. 102–113, 2019, doi: 10.35760/tr.2019.v24i2.2389.
[13] K. H. Mahmud and S. A. Faraby, "Klasifikasi Citra Multi-Kelas Menggunakan Convolutional Neural Network," p. 10.
[14] Muzakir A, Ependi U. Model for Identification and Prediction of Leaf Patterns: Preliminary Study for Improvement. Scientific Journal of Informatics. 2021 Nov 30;8(2):244-50.