

Classification of Melanoma Skin Cancer using Convolutional Neural Network

Rina Refianti¹, Achmad Benny Mutiara², Rachmadinna Poetri Priyandini³
Faculty of Computer Science and Information Technology, Gunadarma University
Jl. Margonda Raya No. 100, Depok 16424, Jawa Barat

Abstract—Melanoma cancer is a type of skin cancer and is the most dangerous one because it causes the most of skin cancer deaths. Melanoma comes from melanocyte cells, melanin-producing cells, so that melanomas are generally brown or black coloured. Melanomas are mostly caused by exposure to ultraviolet radiation that damages the DNA of skin cells. The diagnoses of melanoma cancer are often performed manually by using visuals of the skilled doctors, analyzing the result of dermoscopy examination and match it with medical sciences. Manual detection weakness is highly influenced by human subjectivity that makes it inconsistent in certain conditions. Therefore, a computer assisted technology is needed to help classifying the results of dermoscopy examination and to deduce the results more accurately with a relatively faster time. The making of this application starts with problem analysis, design, implementation, and testing. This application uses deep learning technology with Convolutional Neural Network method and LeNet-5 architecture for classifying image data. The experiment using 44 images data from the training results with a different number of training and epoch resulted the highest percentage of success at 93% in training and 100% in testing, which the number of training data used of 176 images and 100 epochs. This application was created using Python programming language and Keras library as Tensorflow back-end.

Keywords—Convolutional neural network; deep learning; image classification; LeNet-5; melanoma skin cancer; python

I. INTRODUCTION

The skin is a vital organ that covers the entire outside of the body, forming a protective barrier against pathogens and injuries from the environment. But because it is located on the outer part, the skin is prone to disease. One of these diseases is known as skin cancer. Skin cancer is an abnormality in skin cells caused by mutations in cell DNA. One of the most dangerous types of skin cancer is melanoma cancer. Melanoma is a skin malignancy derived from melanocyte cells, the skin pigment cells that produces melanin. Because these cells are still able to form melanin, melanoma is mostly brown or black colored [1].

Common symptoms of melanoma are the appearance of new moles or changes in existing moles. Changes to the mole can occur due to exposure to ultraviolet light that damages the DNA of skin cells and genes that control cell growth and division resulting in the formation of malignant cells.

One of the first steps to diagnosing melanoma is to do a physical examination using dermoscopy. With this dermoscopy examination, it can assess the size, color, and

texture of moles that being suspected as melanoma. To determine a person with melanoma, a dermatologist conducts research from the results of dermoscopy examinations obtained and matched them with medical science to produce conclusions, but the detection weaknesses are strongly influenced by human subjectivity that makes it inconsistent in certain conditions. Research with image-based can be maximized by utilizing information technology products, such as deep learning.

Deep learning has become a hot topic discussed in the machine learning world because of its significant capability in modeling various complex data such as images and sound. Convolutional Neural Network (CNN) is one of deep learning's methods that has the most significant result in image recognition because it tries to imitate the same way of recognizing images in visual cortex as humans so that they are able to process the same information [2,3].

The aim of this research is to build a system that can classify melanoma cancer through the images from the dermoscopy examination with Deep Learning training using the CNN method.

In the rest of paper, we show the theoretical background of CNN and the related work in Section II. In Section III the research methodology is presented. The experiments and results related to data of melanoma skin cancer are also shown in Section IV. The last section is conclusion and future work of our research.

II. THEORY AND RELATED WORK

A. Melanoma Cancer

Melanoma comes from melanocyte cells, melanin-producing cells that are usually present in the skin. Because most melanoma cells still produce melanin, melanoma is often brown or black. Fig. 1 shows the form of melanoma skin cancer.

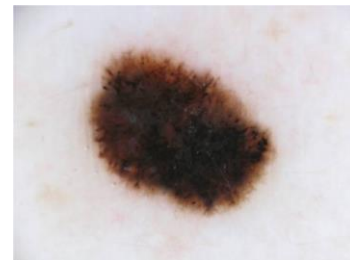


Fig. 1. Dermoscopy Image of Melanoma Cancer.

Melanoma can appear on normal skin, or can appear as a mole or other area of the skin that undergoes changes. Some moles that arise at birth can develop into melanoma. In addition, melanoma can also occur in the eyes, ears, gingival of the upper jaw, tongue, and lips. Melanoma cancer is often characterized by the appearance of new moles or when there is a change in shape from an old mole. Normal moles usually have one color, round or oval, and are less than 6 millimeters in diameter [1], while melanoma has these characteristics:

- 1) Has more than one color
- 2) Has an irregular shape
- 3) Its diameter is greater than 6 mm
- 4) It feels itchy and can bleed

To distinguish normal moles from melanoma, it can be examined for its form with the ABCDE list, as follows:

- 1) Asymmetrical: melanoma has an irregular shape and cannot be divided in half.
- 2) Border: melanoma has an uneven and rough edge, unlike normal moles.
- 3) Color: melanoma is usually a mixture of two or three colors.
- 4) Diameter: melanoma is usually larger than 6 millimeters in diameter, and is different from ordinary moles.
- 5) Enlargement or evolution: moles that change shape and size after a while will usually become melanoma.

B. Deep Learning

Deep learning is a machine learning technique that utilizes many layers of nonlinear information processing to perform feature extraction, pattern recognition, and classification [2]. Deep Learning utilizes artificial neural networks to implement problems with large datasets. Deep Learning techniques provide a very strong architecture for Supervised Learning. By adding more layers, the learning model can better represent labelled image data. In deep learning, a computer learns to classify directly from images, text, or sound. Just as a computer is trained to use large numbers of data sets and then change the pixel value of an image to an internal representation or vector feature where classifiers can detect or classify patterns in the input [5].

C. CNN

Convolutional Neural Network (CNN) is one of deep learning's algorithms that is claimed to be the best model for solving problems in object recognition. CNN is the development of Multilayer Perceptron (MLP) which is designed to process two-dimensional data. CNN is included in the type of Deep Neural Network because of the high network depth and many applied to image data. In the case of image classification in research on virtual cortex on cat's visual sense, MLP is less suitable for use because it does not store spatial information from image data and considers each pixel to be an independent feature that results in unfavourable results.

CNN was first developed by Kunihiko Fukushima under the name NeoCognitron. This concept was later developed by Yann LeCun for numerical recognition and handwriting. In 2012, Alex Krizhevsky successfully won the 2012 ImageNet

Large Scale Visual Recognition Challenge competition with his CNN application. This is the moment of proof that the Deep Learning method with CNN method has proven to be successful in overcoming other Machine Learning methods such as SVM in the case of object classification in images [3].

In general, the layer type on CNN is divided into two, namely:

6) *Feature extraction layer*: Located at the beginning of the architecture is composed of several layers and each layer is composed of neurons connected to the local area (local region) of the previous layer. The first type layer is the convolutional layer and the second layer is the pooling layer. Each layer applies the activation function with its intermittent position between the first type and the second type. This layer accepts image input directly and processes it until it produces an output in the form of a vector to be processed in the next layer.

7) *The classification layer*: Composed of several layers and each layer is composed of fully connected neurons with other layers. This layer accepts input from the output feature image extraction layer in the form of a vector then transformed like Multi Neural Networks with the addition of several hidden layers. The output is class accuracy for classification.

CNN is thus a method for transforming the original image layer per layer from the image pixel value into the class scoring value for classification, where each layer has a hyper parameter and some do not have parameters (weight and bias on neurons).

On CNN there are four types of layers used, namely:

8) *Convolutional layer*: The Convolution Layer performs convolution operations at the output of the previous layer. Convolution operations are operations on two functions of real value arguments. This operation uses image input to produce the output function as a Feature Map. These inputs and outputs are two real-value arguments. Convolution operations in general can be written with the formula below:

$$s(t) = (x * \omega) \quad (1)$$

The equation $s(t)$ gives results in the form of a Feature Map as a single output with the first argument used is the input expressed as x and the second argument used is the kernel or filter which is stated as ω . Because the input used is an image that has two dimensions, it can be expressed as t as a pixel and replace it with the arguments i and j . Therefore, convolution operations with more than one dimension input can be written as follows:

$$s(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i - m, j - n) K(m, n) \quad (2)$$

The above equation is the basic calculation for convolution operations where the pixels of the image are expressed as i and j . The calculation is commutative and appears when K as a kernel can be reversed relative to I as input. Convolution operation can be seen as matrix multiplication between image input and kernel where the results can be calculated with dot products. In addition, the output volume of each layer can be

adjusted using hyperparameters. Hyperparameter is used to calculate how many activation neurons in one output are stated in the equation below:

$$\frac{(W-F+2P)}{S+1} \tag{3}$$

From the equation above, the spatial size of the output volume can be calculated by the hyperparameter used is the volume size (W), filter (F), Stride applied (S), and the number of zero padding used (P). Stride is the value used to shift filters through image input and Zero Padding is the value to place zeros around the image border. In image processing, convolution means applying a kernel (yellow box) to the image in all possible offsets as shown in Fig. 3.

The green box as a whole is the image that will be convoluted. The kernel moves from the upper left corner to the lower right. So that the convolution of the image can be seen in the picture on the right. The purpose of convolution on image data is to extract features from the input image.

9) *Pooling layer*: Pooling Layer is a layer that uses functions with Feature Map as input and processes it with various statistical operations based on the nearest pixel value. Pooling layer on the CNN model is usually inserted regularly after several convolution layers. The Pooling layer in the CNN model architecture that is inserted between the convolution layers can progressively reduce the size of the output volume in the Feature Map, so as to reduce the number of parameters used and calculations on the network, and to control Overfitting. In most CNNs, the pooling method used is max pooling. Max pooling divides the output of the convolution layer into several small grids and then takes the maximum value from each grid to compile a reduced image matrix as shown in Fig. 4.

Grids that are red, green, yellow and blue are the grid groups whose maximum values will be selected. So that the results of the process can be seen in the grid collection on the right. The process ensures that the features obtained will be the same even though the image object experiences translation (shift). Using the CNN pooling layer aims to reduce the size of the image so that it can be easily replaced with a convolution layer with the same stride as the corresponding pooling layer. This form of pooling will reduce the Feature Map up to 75% of its original size.

10) *Fully connected layer*: Fully Connected Layer is a layer in which all activation neurons from the previous layer are connected all with neurons in the next layer and aim to transform data dimensions so that data can be classified linearly. Every neuron in the convolution layer needs to be transformed into one-dimensional data before it can be inserted into a fully connected layer. This causes data to lose spatial information and is not reversible so that the fully connected layer can only be implemented at the end of the network. The difference between the fully connected layer and the ordinary convolution layer is that in the convolution layer, the neurons are only connected to a certain area of the input, while the fully connected layer has neurons that are

completely connected. However, the two layers still operate dot products, so the function is not so different.

11) *Activation function*: In this paper the activation functions used are ReLu (Rectified Linear Units) and Softmax Classifier. ReLu activation increases the non-linear nature of decision making functions and all networks without affecting the receptive fields of Convolutional Layer. ReLu is also widely used because it can train neural networks faster. Softmax activation for this layer is another form of Logistic Regression algorithm that can be used to classify more than two classes. The usual classification used by the Logistic Regression algorithm is the classification of binary classes. Softmax provides more intuitive results and has a better probabilistic interpretation than other classification algorithms. Softmax makes it possible to calculate the probability for all labels. From the existing label, a real value vector is taken and converts it to a vector with a value between zero and one which, if all are added, will be worth one.

D. LeNet-5

LeNet-5 is a multi-layer network based on CNN, introduced by Yann LeCun. LeNet-5 is the development of the LeNet-1 and LeNet-4 where LeNet-5 has a greater number of free parameters or layers than its predecessor (Fig. 2).

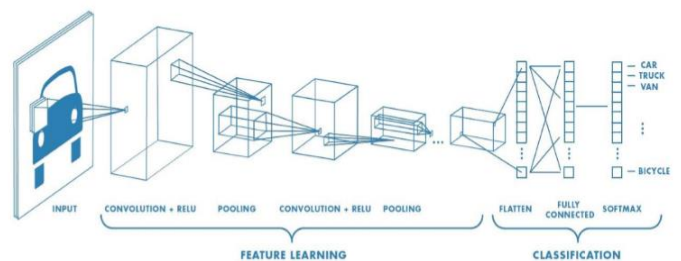


Fig. 2. Example of CNN with Multiple Layers [4].

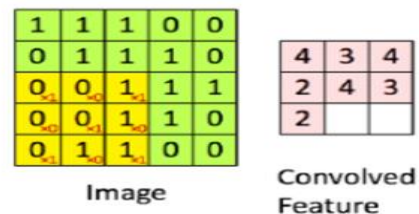


Fig. 3. Convolutional Operation.

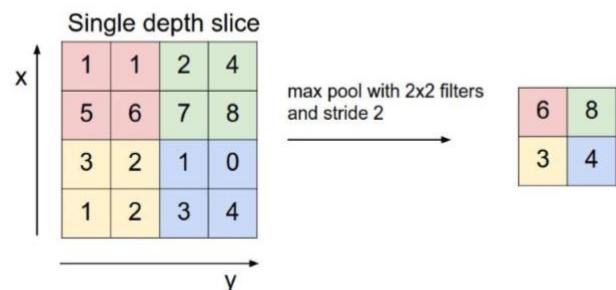


Fig. 4. Max Pooling Operation [3].

LeNet-5 consists of 7 layers where the input layer is not calculated. The LeNet-5 input layer is a 32x32 pixel image. The convolution layer in Fig. 5 is marked with the Cx symbol, the subsampling layer is marked with the Sx symbol, the fully connected layer is marked with the Fx symbol, and the last is the output layer which is the fully connected layer for class classification [6].

In the first layer there is a convolutional layer that studies 20 convolution filters with each 5x5 size and uses ReLu activation. Then the second layer is a pooling layer, using 2x2 size of MaxPooling. The third layer is a convolutional layer which studies 50 convolution filters with each 5x5 size and uses ReLu activation. The filter size is getting bigger in each layer, which is useful to deepen the architectural network studied by the system. Then the system proceed with the fourth layer, which is the pooling layer using 2x2 size of MaxPooling. In the fifth layer, the results of the previous layer process will be flattened into a vector. This layer is called a fully connected layer, where there are 120 nodes that are connected to each other. After that, the process continued with the sixth layer in the form of a fully connected layer with 84 connected nodes. Finally, at the seventh layer is a fully connected layer with softmax activation which connects 2 nodes as the end result of the class to be classified [6].

E. Keras

Keras is a high-level neural network library written in python and able to run on TensorFlow, CNTK, or Theano. This library provides features that are used with a focus on facilitating deeper development of deep learning.

F. Tensorflow

Tensorflow is an open-source software library, developed by the Google Brain team in order to support smart computing to support the search and learning of their products. Computing stated using Tensorflow can be executed with a variety of systems, ranging from mobile devices such as cellphones and tablets to hundreds of large-scale distributed systems of machines and thousands of computing devices such as GPU Cards. The system is flexible and can be used to express a variety of algorithms, including training and inference algorithms for deep neural network models, and has been used to conduct research and to spread machine learning systems to production in more than a dozen fields of computer science and other fields, including voice recognition, computer vision, robotics, information retrieval, natural language processing, geographical information extraction, and others [7].

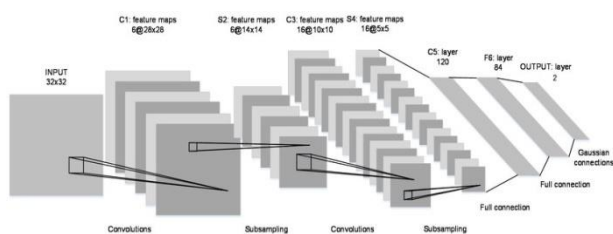


Fig. 5. LeNet-5 Architecture [6].

G. Related Work

The research has related with the works of: first, Andre Esteva et.al. 2017 [8], i.e. “level classification of skin cancer with deep neural networks; and second, T.J. Binker et al. 2018 [9], i.e. “Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review”. Both researches are about general skin cancer, but our research is more specific for melanoma skin cancer.

III. RESEARCH METHODOLOGY

A. Data Collection

The dataset is obtained from ISIC (International Skin Imaging Collaboration) website, contains 220 images of dermoscopy examination. These 220 images consist of 110 melanoma cancer images and 110 non-melanoma cancer images.

B. Data Acquisition

The aim of data acquisition is to determine which objects will be used as research objects. The object of research is in the form of two-dimensional images in JPG format which contain images of melanoma cancer and non-melanoma, as in Fig. 6.

C. Pre-Processing

The dataset contains of different image resolution which require high cost of computation. It is necessary to rescale all the images to 32 x 32 pixels for this deep learning network.

D. Data Augmentation

Data augmentation is used to multiply the variation of images from the dataset by rotating the image, increasing or decreasing the image’s length and width, zooming in the image, and also flipping the image horizontally. The example of this data augmentation can be seen in Fig. 7.

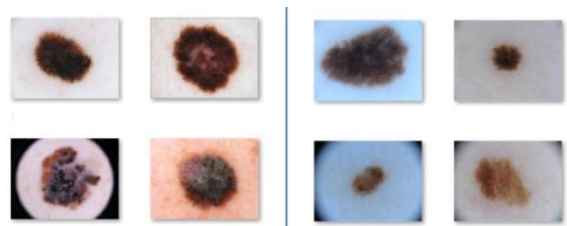


Fig. 6. Images of Melanoma and Non-Melanoma.

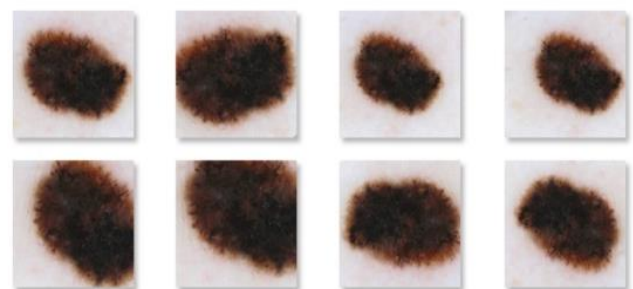


Fig. 7. Data Augmentation of Melanoma Cancer Image.

E. Training Data Process

Fig. 8 shows the process flow in conducting training on the dataset using CNN with LeNet-5 architecture.

The training process starts by reading the model name and number of epoch and batch size received from the user. Then the system reads the dataset with melanoma and non-melanoma categories. Then all the images from the dataset are resized into 32x32 pixel and the dataset augmentation will be generated. The system initializes LeNet-5 architecture and starts to train the network as much the number of epoch inputted by the user earlier. The training will produce a probability value for the two classification classes, where the class with the greatest probability value is the classification class predicted by the program. The training results are then stored in the form of a model file. After completing the training, the system will save the model and plot from the results of the training.

In this training there are parameters that are run constantly throughout the procedure, namely learning rate and batch size. The learning rate used is 0.001, where this parameter states the constants for learning speed from the network layer used. While the batch size parameter serves to determine the total amount of data used in one batch of training. In this paper, the batch size used is 32. Determination of batch size is considered from the memory capability of the device used to conduct the training process.

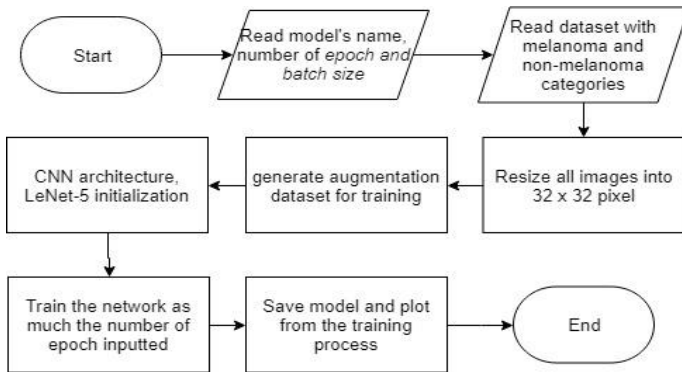


Fig. 8. Training Data Process Flow.

IV. EXPERIMENT AND RESULT

In this paper, the experiment was carried out by determining a different number of training data and epoch to get the best accuracy result. There were two section of training data, the first one used 154 images and the second one used 176 images. Each of the training data section was trained with 50 epochs and 100 epochs. Then, all the model resulted from the training were tested against 44 images of test data and calculated the percentage of precision, recall, and accuracy from the results of the testing using confusion matrix as in Table I.

$$\text{Precision} : \frac{TP}{TP+FP} \times 100\%$$

$$\text{Recall} : \frac{TP}{TP+FN} \times 100\%$$

$$\text{Accuracy} : \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

TABLE I. CONFUSION MATRIX

Matrix		Actual Class	
		Melanoma	Non-Melanoma
Prediction Class	Melanoma	TP (True Positive)	FP (False Positive)
	Non-Melanoma	FN (False Negative)	TN (True Negative)

A. Experiment using 154 Train Data and 50 Epochs

The training process of this experiment was carried out using 50 epochs on 154 train data which consist of 77 images of melanoma and 77 images of non-melanoma. The plot result from this training can be seen in Fig. 9.

In Fig. 9 can be seen that from epoch 0 to 49 shows that training accuracy has increased with the final result of 0.92 while training loss has decreased with the final result of 0.28. The model from this training then was tested against 44 test data. This testing result can be seen in Table II.

Based on Table II with 44 images being tested, there are 40 correct images and 4 incorrect images in classification. From the table, the results in confusion matrix are shown in Table III.

$$\text{Precision} : \frac{TP}{TP+FP} \times 100\%$$

$$\frac{21}{21+3} \times 100\% = \frac{21}{24} \times 100\% = 88\%$$

$$\text{Recall} : \frac{TP}{TP+FN} \times 100\%$$

$$\frac{21}{21+1} \times 100\% = \frac{21}{22} \times 100\% = 95\%$$

$$\text{Accuracy} : \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$\frac{21+19}{21+19+3+1} \times 100\% = \frac{40}{44} \times 100\% = 91\%$$

The calculation of the confusion matrix above results 88 % of precision, 95 % of recall, and 91 % of accuracy.

A. Experiment using 154 Train Data and 100 Epochs

The training process of this experiment was carried out using 100 epochs on 154 train data which consist of 77 images of melanoma and 77 images of non-melanoma. The plot result from this training can be seen in Fig. 5.

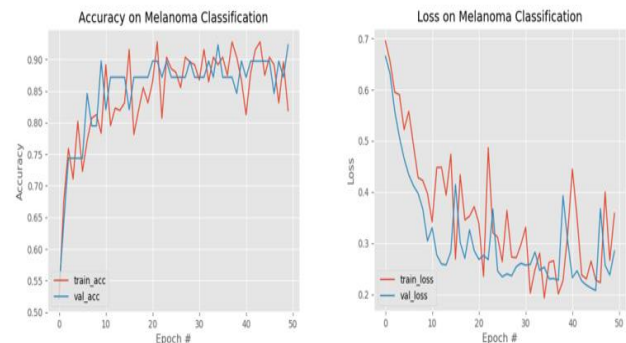


Fig. 9. Plot of Training Result with 154 Training Data and 50 Epochs.

TABLE II. TESTING RESULT USING 154 TRAIN DATA AND 50 EPOCHS

No	Filename	Category	Result	Answer
1	ISIC_0000030	Melanoma	Melanoma	True
2	ISIC_0000157	Melanoma	Melanoma	True
3	ISIC_0000288	Melanoma	Melanoma	True
4	ISIC_0000295	Melanoma	Melanoma	True
5	ISIC_0000301	Melanoma	Melanoma	True
6	ISIC_0000303	Melanoma	Melanoma	True
7	ISIC_0010036	Melanoma	Non Melanoma	*False
8	ISIC_0010217	Melanoma	Melanoma	True
9	ISIC_0010652	Melanoma	Melanoma	True
10	ISIC_0011112	Melanoma	Melanoma	True
11	ISIC_0011492	Melanoma	Melanoma	True
12	ISIC_0011508	Melanoma	Melanoma	True
13	ISIC_0011515	Melanoma	Melanoma	True
14	ISIC_0011517	Melanoma	Melanoma	True
15	ISIC_0011518	Melanoma	Melanoma	True
16	ISIC_0011693	Melanoma	Melanoma	True
17	ISIC_0011772	Melanoma	Melanoma	True
18	ISIC_0011774	Melanoma	Melanoma	True
19	ISIC_0011873	Melanoma	Melanoma	True
20	ISIC_0011878	Melanoma	Melanoma	True
21	ISIC_0011879	Melanoma	Melanoma	True
22	ISIC_0011944	Melanoma	Melanoma	True
23	ISIC_0000008	Non Melanoma	Non Melanoma	True
24	ISIC_0000012	Non Melanoma	Non Melanoma	True
25	ISIC_0000015	Non Melanoma	Non Melanoma	True
26	ISIC_0000018	Non Melanoma	Non Melanoma	True
27	ISIC_0000021	Non Melanoma	Non Melanoma	True
28	ISIC_0000058	Non Melanoma	Non Melanoma	True
29	ISIC_0000059	Non Melanoma	Non Melanoma	True
30	ISIC_0000064	Non Melanoma	Non Melanoma	True
31	ISIC_0000067	Non Melanoma	Non Melanoma	True
32	ISIC_0000086	Non Melanoma	Non Melanoma	True
33	ISIC_0000089	Non Melanoma	Non Melanoma	True
34	ISIC_0000093	Non Melanoma	Non Melanoma	True
35	ISIC_0000096	Non Melanoma	Melanoma	*False
36	ISIC_0000108	Non Melanoma	Non Melanoma	True
37	ISIC_0000109	Non Melanoma	Melanoma	*False
38	ISIC_0000113	Non Melanoma	Non Melanoma	True
39	ISIC_0000116	Non Melanoma	Non Melanoma	True
40	ISIC_0000121	Non Melanoma	Non Melanoma	True
41	ISIC_0000124	Non Melanoma	Non Melanoma	True
42	ISIC_0000125	Non Melanoma	Melanoma	*False
43	ISIC_0000128	Non Melanoma	Non Melanoma	True
44	ISIC_0000134	Non Melanoma	Non Melanoma	True

TABLE III. CONFUSION MATRIX FROM THE TESTING RESULT USING 154 TRAIN DATA AND 50 EPOCHS

Matrix		Actual Class	
		Melanoma	Non-Melanoma
Prediction Class	Melanoma	TP = 21	FP = 3
	Non- Melanoma	FN = 1	TN = 19

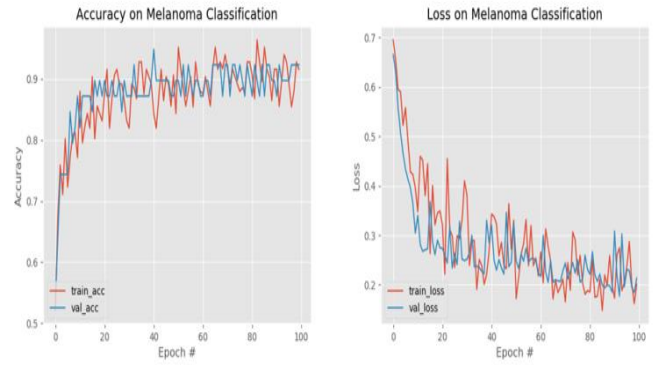


Fig. 10. Plot of Training Result with 154 Training Data and 100 Epochs.

In Fig. 10 can be seen that from epoch 0 to 99 shows that training accuracy has increased with the final result of 0.92 while training loss has decreased with the final result of 0.21. The model from this training then was tested against 44 test data. This testing result can be seen in Table IV.

Based on Table IV with 44 images being tested, there are 41 correct images and 3 incorrect images in classification. From the table, the results in confusion matrix are shown in Table V.

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\%$$

$$\frac{21}{21+2} \times 100\% = \frac{21}{23} \times 100\% = 91\%$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\%$$

$$\frac{21}{21+1} \times 100\% = \frac{21}{22} \times 100\% = 95\%$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$\frac{21+20}{21+20+2+1} \times 100\% = \frac{41}{44} \times 100\% = 93\%$$

The calculation of the confusion matrix above results 91 % of precision, 95% of recall, and 93% of accuracy.

B. Experiment using 176 Train Data and 50 Epochs

The training process of this experiment was carried out using 50 epochs on 176 train data which consist of 88 images of melanoma and 88 images of non-melanoma. The plot result from this training can be seen in Fig. 11.

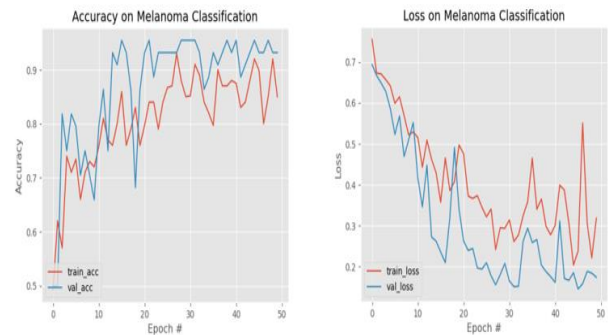


Fig. 11. Plot of Training Result with 176 Training Data and 50 Epochs.

In Fig. 6 can be seen that from epoch 0 to 49 shows that training accuracy has increased with the final result of 0.93 while training loss has decreased with the final result of 0.17. The model from this training then was tested against 44 test data. This testing result can be seen in Table IV.

Based on Table VI with 44 images being tested, there are 42 correct images and 2 incorrect images in classification. From the table, the results in confusion matrix are shown in Table VII.

TABLE IV. TESTING RESULT USING 154 TRAIN DATA AND 100 EPOCHS

No	Filename	Category	Result	Answer
1	ISIC_0000030	Melanoma	Melanoma	True
2	ISIC_0000157	Melanoma	Melanoma	True
3	ISIC_0000288	Melanoma	Melanoma	True
4	ISIC_0000295	Melanoma	Melanoma	True
5	ISIC_0000301	Melanoma	Melanoma	True
6	ISIC_0000303	Melanoma	Melanoma	True
7	ISIC_0010036	Melanoma	Non Melanoma	*False
8	ISIC_0010217	Melanoma	Melanoma	True
9	ISIC_0010652	Melanoma	Melanoma	True
10	ISIC_0011112	Melanoma	Melanoma	True
11	ISIC_0011492	Melanoma	Melanoma	True
12	ISIC_0011508	Melanoma	Melanoma	True
13	ISIC_0011515	Melanoma	Melanoma	True
14	ISIC_0011517	Melanoma	Melanoma	True
15	ISIC_0011518	Melanoma	Melanoma	True
16	ISIC_0011693	Melanoma	Melanoma	True
17	ISIC_0011772	Melanoma	Melanoma	True
18	ISIC_0011774	Melanoma	Melanoma	True
19	ISIC_0011873	Melanoma	Melanoma	True
20	ISIC_0011878	Melanoma	Melanoma	True
21	ISIC_0011879	Melanoma	Melanoma	True
22	ISIC_0011944	Melanoma	Melanoma	True
23	ISIC_0000008	Non Melanoma	Non Melanoma	True
24	ISIC_0000012	Non Melanoma	Non Melanoma	True
25	ISIC_0000015	Non Melanoma	Non Melanoma	True
26	ISIC_0000018	Non Melanoma	Non Melanoma	True
27	ISIC_0000021	Non Melanoma	Non Melanoma	True
28	ISIC_0000058	Non Melanoma	Non Melanoma	True
29	ISIC_0000059	Non Melanoma	Non Melanoma	True
30	ISIC_0000064	Non Melanoma	Non Melanoma	True
31	ISIC_0000067	Non Melanoma	Non Melanoma	True
32	ISIC_0000086	Non Melanoma	Non Melanoma	True
33	ISIC_0000089	Non Melanoma	Non Melanoma	True
34	ISIC_0000093	Non Melanoma	Non Melanoma	True
35	ISIC_0000096	Non Melanoma	Non Melanoma	True
36	ISIC_0000108	Non Melanoma	Non Melanoma	True
37	ISIC_0000109	Non Melanoma	Non Melanoma	True
38	ISIC_0000113	Non Melanoma	Non Melanoma	True
39	ISIC_0000116	Non Melanoma	Non Melanoma	True
40	ISIC_0000121	Non Melanoma	Non Melanoma	True
41	ISIC_0000124	Non Melanoma	Non Melanoma	True
42	ISIC_0000125	Non Melanoma	Non Melanoma	True
43	ISIC_0000128	Non Melanoma	Non Melanoma	True
44	ISIC_0000134	Non Melanoma	Non Melanoma	True

TABLE V. CONFUSION MATRIX FROM THE TESTING RESULT USING 154 TRAIN DATA AND 100 EPOCHS

Matrix		Actual Class	
		Melanoma	Non-Melanoma
Prediction Class	Melanoma	TP = 21	FP = 2
	Non- Melanoma	FN = 1	TN = 20

TABLE VI. TESTING RESULT USING 176 TRAIN DATA AND 50 EPOCHS

No	Filename	Category	Result	Answer
1	ISIC_0000030	Melanoma	Melanoma	True
2	ISIC_0000157	Melanoma	Melanoma	True
3	ISIC_0000288	Melanoma	Melanoma	True
4	ISIC_0000295	Melanoma	Melanoma	True
5	ISIC_0000301	Melanoma	Melanoma	True
6	ISIC_0000303	Melanoma	Melanoma	True
7	ISIC_0010036	Melanoma	Non Melanoma	*False
8	ISIC_0010217	Melanoma	Melanoma	True
9	ISIC_0010652	Melanoma	Melanoma	True
10	ISIC_0011112	Melanoma	Melanoma	True
11	ISIC_0011492	Melanoma	Melanoma	True
12	ISIC_0011508	Melanoma	Melanoma	True
13	ISIC_0011515	Melanoma	Melanoma	True
14	ISIC_0011517	Melanoma	Melanoma	True
15	ISIC_0011518	Melanoma	Melanoma	True
16	ISIC_0011693	Melanoma	Melanoma	True
17	ISIC_0011772	Melanoma	Non Melanoma	*False
18	ISIC_0011774	Melanoma	Melanoma	True
19	ISIC_0011873	Melanoma	Melanoma	True
20	ISIC_0011878	Melanoma	Melanoma	True
21	ISIC_0011879	Melanoma	Melanoma	True
22	ISIC_0011944	Melanoma	Melanoma	True
23	ISIC_0000008	Non Melanoma	Non Melanoma	True
24	ISIC_0000012	Non Melanoma	Non Melanoma	True
25	ISIC_0000015	Non Melanoma	Non Melanoma	True
26	ISIC_0000018	Non Melanoma	Non Melanoma	True
27	ISIC_0000021	Non Melanoma	Non Melanoma	True
28	ISIC_0000058	Non Melanoma	Non Melanoma	True
29	ISIC_0000059	Non Melanoma	Non Melanoma	True
30	ISIC_0000064	Non Melanoma	Non Melanoma	True
31	ISIC_0000067	Non Melanoma	Non Melanoma	True
32	ISIC_0000086	Non Melanoma	Non Melanoma	True
33	ISIC_0000089	Non Melanoma	Non Melanoma	True
34	ISIC_0000093	Non Melanoma	Non Melanoma	True
35	ISIC_0000096	Non Melanoma	Non Melanoma	True
36	ISIC_0000108	Non Melanoma	Non Melanoma	True
37	ISIC_0000109	Non Melanoma	Non Melanoma	True
38	ISIC_0000113	Non Melanoma	Non Melanoma	True
39	ISIC_0000116	Non Melanoma	Non Melanoma	True
40	ISIC_0000121	Non Melanoma	Non Melanoma	True
41	ISIC_0000124	Non Melanoma	Non Melanoma	True
42	ISIC_0000125	Non Melanoma	Non Melanoma	True
43	ISIC_0000128	Non Melanoma	Non Melanoma	True
44	ISIC_0000134	Non Melanoma	Non Melanoma	True

TABLE VII. CONFUSION MATRIX FROM THE RESULT OF EXPERIMENT USING 176 TRAIN DATA AND 50 EPOCHS

Matrix		Actual Class	
		Melanoma	Non-Melanoma
Prediction Class	Melanoma	TP = 20	FP = 0
	Non- Melanoma	FN = 2	TN = 22

$$\text{Precision: } \frac{TP}{TP+FP} \times 100\%$$

$$\frac{20}{20+0} \times 100\% = \frac{20}{20} \times 100\% = 100\%$$

$$\text{Recall : } \frac{TP}{TP+FN} \times 100\%$$

$$\frac{20}{20+2} \times 100\% = \frac{20}{22} \times 100\% = 91\%$$

$$\text{Accuracy : } \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$\frac{20+22}{20+22+0+2} \times 100\% = \frac{42}{44} \times 100\% = 95\%$$

The calculation of the confusion matrix above results 100% of precision, 91% of recall, and 95% of accuracy.

B. Experiment Using 176 Train Data and 100 Epochs

The training process of this experiment was carried out using 50 epochs on 154 train data which consist of 77 images of melanoma and 77 images of non-melanoma. The plot result from this training can be seen in Fig. 12.

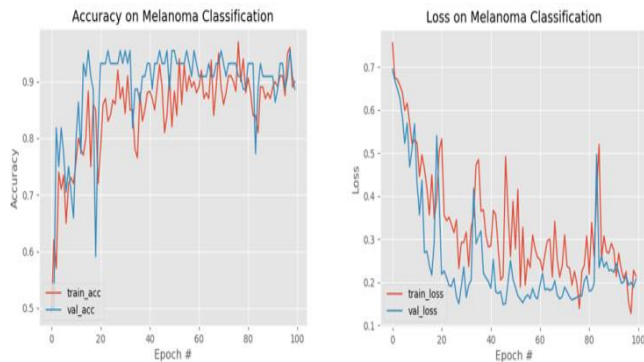


Fig. 12. Plot of Training Result with 176 Training Data and 100 Epochs.

In Fig. 12 can be seen that from epoch 0 to 99 shows that training accuracy has increased with the final result of 0.88 while training loss has decreased with the final result of 0.20. The model from this training then was tested against 44 test data. This testing result can be seen in Table VIII.

Based on Table VIII with 44 images being tested, there are 44 correct images and 0 incorrect images in classification. From the table, the results in confusion matrix are shown in Table IX.

$$\text{Precision: } \frac{TP}{TP+FP} \times 100\%$$

$$\frac{22}{22+0} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Recall : } \frac{TP}{TP+FN} \times 100\%$$

$$\frac{22}{22+0} \times 100\% = \frac{22}{22} \times 100\% = 100\%$$

$$\text{Accuracy : } \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$\frac{22+22}{22+22+0+0} \times 100\% = \frac{44}{44} \times 100\% = 100\%$$

The calculation of the confusion matrix above results 100 % of precision, 100% of recall, and 100% of accuracy.

TABLE VIII. TESTING RESULT USING 176 TRAIN DATA AND 100 EPOCHS

No	Filename	Category	Result	Answer
1	ISIC_0000030	Melanoma	Melanoma	True
2	ISIC_0000157	Melanoma	Melanoma	True
3	ISIC_0000288	Melanoma	Melanoma	True
4	ISIC_0000295	Melanoma	Melanoma	True
5	ISIC_0000301	Melanoma	Melanoma	True
6	ISIC_0000303	Melanoma	Melanoma	True
7	ISIC_0010036	Melanoma	Melanoma	True
8	ISIC_0010217	Melanoma	Melanoma	True
9	ISIC_0010652	Melanoma	Melanoma	True
10	ISIC_0011112	Melanoma	Melanoma	True
11	ISIC_0011492	Melanoma	Melanoma	True
12	ISIC_0011508	Melanoma	Melanoma	True
13	ISIC_0011515	Melanoma	Melanoma	True
14	ISIC_0011517	Melanoma	Melanoma	True
15	ISIC_0011518	Melanoma	Melanoma	True
16	ISIC_0011693	Melanoma	Melanoma	True
17	ISIC_0011772	Melanoma	Melanoma	True
18	ISIC_0011774	Melanoma	Melanoma	True
19	ISIC_0011873	Melanoma	Melanoma	True
20	ISIC_0011878	Melanoma	Melanoma	True
21	ISIC_0011879	Melanoma	Melanoma	True
22	ISIC_0011944	Melanoma	Melanoma	True
23	ISIC_0000008	Non Melanoma	Non Melanoma	True
24	ISIC_0000012	Non Melanoma	Non Melanoma	True
25	ISIC_0000015	Non Melanoma	Non Melanoma	True
26	ISIC_0000018	Non Melanoma	Non Melanoma	True
27	ISIC_0000021	Non Melanoma	Non Melanoma	True
28	ISIC_0000058	Non Melanoma	Non Melanoma	True
29	ISIC_0000059	Non Melanoma	Non Melanoma	True
30	ISIC_0000064	Non Melanoma	Non Melanoma	True
31	ISIC_0000067	Non Melanoma	Non Melanoma	True
32	ISIC_0000086	Non Melanoma	Non Melanoma	True
33	ISIC_0000089	Non Melanoma	Non Melanoma	True
34	ISIC_0000093	Non Melanoma	Non Melanoma	True
35	ISIC_0000096	Non Melanoma	Non Melanoma	True
36	ISIC_0000108	Non Melanoma	Non Melanoma	True
37	ISIC_0000109	Non Melanoma	Non Melanoma	True
38	ISIC_0000113	Non Melanoma	Non Melanoma	True
39	ISIC_0000116	Non Melanoma	Non Melanoma	True
40	ISIC_0000121	Non Melanoma	Non Melanoma	True
41	ISIC_0000124	Non Melanoma	Non Melanoma	True
42	ISIC_0000125	Non Melanoma	Non Melanoma	True
43	ISIC_0000128	Non Melanoma	Non Melanoma	True
44	ISIC_0000134	Non Melanoma	Non Melanoma	True

TABLE IX. CONFUSION MATRIX FROM THE RESULT OF EXPERIMENT USING 176 TRAIN DATA AND 100 EPOCHS

Matrix		Actual Class	
		Melanoma	Non-Melanoma
Prediction Class	Melanoma	TP = 22	FP = 0
	Non- Melanoma	FN = 0	TN = 22

V. CONCLUSION AND FUTURE WORK

Classification of melanoma cancer images is carried out in 2 stages; the first stage is training the dataset to produce a model. The second stage is the process of classification which the system takes the image data, then initializes the model from the results of the training and makes predictions using the model, then the system takes the prediction results along with their probabilities and display the prediction results along with the image.

The experiment was conducted on 44 images of test data using a different number of training data (images) and epochs in the training process. The experiment obtained the highest accuracy of 93% in training result. Meanwhile, in testing result obtained 91% of accuracy for using 154 images and 50 epochs in training, then 93% accuracy for using 154 images and 100 epochs in training. The training conducted on 176 images and 50 epochs resulted in a 95% accuracy of testing result, while for the training using 176 images and 100 epochs resulted in 100% accuracy of testing result.

The experiment results prove that the amount of training data and epochs used for training affects the level of accuracy in classifying melanoma cancer images. The more data that is trained, the better the test results will be produced. Whereas 100 epochs is the optimal epoch to produce the best accuracy, this is supported by several other parameters of the LeNet-5 architecture, such as learning rate, number of layers, and the size of input pixels used during training.

Unfortunately, it is difficult to compare different classification methods because some approaches use non-public datasets for training and/or testing, thereby making

reproducibility difficult. Future publications should use publicly available benchmarks and fully disclose methods used for training to allow comparability.

REFERENCES

- [1] World Health Organization. 2018. *Skin Cancers*. Available: <http://www.who.int/uv/faq/skincancer/en/index1.html>.
- [2] L. Deng and Yu D., "Deep Learning: Methods and Applications", *Foundations and Trends in Signal Processing*, Vol. 7, No. 3 – 4, 2013
- [3] Y. LeCun, B. Boser, J.S. Denker, R.E. Howard, W. Hubbard, L.D. Jackel and D. Henderson, "Handwritten digit recognition with a back-propagation network", in *Advances in neural information processing systems 2*, pp. 396-404, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1990
- [4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016
- [5] Y. LeCun, Y., Y. Bengio, dan G. Hinton, 2015. "Deep Learning", *Nature*, Vol 521, pp. 436-444, 2015, doi:10.1038/nature14539
- [6] Y. LeCun, L.D. Jackel, Leon Bottou, A. Brunot, C. Cortes, J. S. Denker, H. Drucker, J. Guyon, U. A. Muller, E. Sackinger, P. Simard and V. Vapnik, " Learning algorithms for classification: a comparison on Handwritten digit recognition", 1995, <http://yann.lecun.com/exdb/pubs/lecun-95b.djvu>
- [7] J. Wallen, [Online] Cheat sheet: TensorFlow, an open source software library for machine learning. Available: <https://www.techrepublic.com/article/tensorflow-googles-open-source-software-library-for-machine-learning-the-smart-persons-guide/>, 2017
- [8] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks", *Nature*, Vol. 542, pp 115-118, 2017
- [9] T.J. Brinker, A. Hekler, J.S. Utikal, N.Grabe, D. Schadendorf, J. Klode, C. Berking, T. Steeb, A.H. Enk, and C.von Kalle, "Skin Cancer Classification Using Convolutional Neural Networks: Systematic Review", *Journal of Medical Internet Research*, Vol. 20, No.10, pp. 11936- 11946, 2018