Computer Engineering and Applications Vol. 8, No. 3, October 2019

Skin Lesion Classification Based on Convolutional Neural Networks

Renny Amalia Pratiwi[1*], Siti Nurmaini[2], Dian Palupi Rini[3], Firdaus[2]

[1]*Master of Informatic Engineering, Faculty of Computer Science, Universitas Sriwijaya*
[2]*Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya, Indonesia*
[3]*Faculty of Computer Science, Universitas Sriwijaya, Indonesia*
**amalia.renny@ymail.com*

ABSTRACT

Melanoma causes the majority of skin cancer deaths. The population level of melanoma has increased over the past 30 years. It kills around 9.320 people in the US every year. Melanoma can often be found early, when it is most likely to be cured. Medical diagnoses using digital imaging with machine learning methods have become popular because of their ability to recognize patterns in digital images. Image diagnosis accuracy allows disease cured at an early stage. This paper proposes a simulation that can be used for early detection of skin cancer that can help dermatologists to distinguish melanomas from other pigmented lesions on the skin. Some researchers have developed a system using machine learning algorithms used to classify skin lesions from dermoscopy images of human skin. In this study, we proposed Convolutional Neural Network (CNN) to our model. CNN is very efficient for image processing because feature extractors can be optimized, applied to each feature image position. The results of skin lesion classification of benign melanocytic nevi and malignant melanoma based on CNN models produces good classification with accuracy is 87.64%, sensitivity is 96.40%, specificity is 81.66%, precision is 78.21% and F1 score is 86.36%.

Keywords: Melanoma, Dermoscopy Image, Skin Lesion Classification, Convolutional Neural Network.

1. INTRODUCTION

One of the most dangerous types of skin cancer that can cause death is melanoma. Melanoma is cancer that occurs in melanocytes, pigment cells in the skin that produce skin color or melanin [1]. Melanoma usually appears on normal skin and starts in a certain type of skin cell from a mole. Melanoma is dangerous because it can spread and move deep into the skin where it can move to the lymph nodes and throughout the body. While it is not one of the most common cancers in Indonesia, melanoma kills around 9.320 people in the US every year. Of that number, 5.990 were male and 3.330 were female [2]. The patient can almost always be cured if the melanoma is known and treated early. Melanoma becomes difficult to treat when the cancer spread to other parts of the body.

Medical diagnoses using digital imaging with machine learning methods have become popular because of their ability to recognize patterns in digital images. Therefore, information is needed from the medical image itself. Image diagnosis accuracy allows disease cured at an early stage. Medical diagnoses using digital

imaging for melanoma classification can help dermatologists to distinguish melanomas from other pigmented lesions on the skin.

Dermoscopy image is used for classification benign and malignant skin lesion of human skin. In previous work, some machine learning algorithms used to classify malignant and benign lesions from dermoscopy images of human skin. Teck Yan Tan et al. diagnose skin cancer using the Particle Swarm Optimization (PSO) algorithm based on dermoscopic images for feature optimization. The proposed PSO model reaches an average global minimum of more than 30 times for almost all testing functions [3]. PSO is effectively used as an optimizer in various functions. It has a great relationship with social life and is associated with evolutionary computation [4]. Cicero et al approach uses a special dataset of skin diseases with augmentation techniques in their data, then they apply transfer learning from convolutional neural networks with the ResNet architecture and get training accuracy of 100% but testing accuracy only 61.7 % [5]. A. Pathiranage proposed a deep highway CNN with batch normalization and the system produced training accuracy of 98% and validation accuracy of 64.57% [6]. Romero et al. classified dermoscopic images containing malignant and benign skin lesions. In their study, VGGNet architecture was used and produced a sensitivity value of 78.66% [7].

We proposed a deep learning algorithm to get better performance in accuracy, sensitivity, and specificity in melanoma classification based on previous research that produced an average training accuracy of 98%, validation accuracy only 64.57% and sensitivity value of 78.66%. As we know to process data that has a grid pattern, such as images requires a CNN deep learning model. CNN is very efficient for image processing because feature extractors can be optimized, applied to each feature image position [8].

This paper is organized as follows. Section 2, describes briefly Convolutional Neural Networks. Section 3, presents data preprocessing and method. Section 4, describes the proposed CNN model. Section 5, explains the performance evaluation using the ROC curve and its confusion matrix in binary classification. The last, section 6 shows the conclusions of this work.

2. CONVOLUTIONAL NEURAL NETWORK CLASSIFIER

Convolutional Neural Network (CNN) is a typical deep learning architectures and variations of the Multilayer Perceptron (MLP) that inspired by human neural networks [9]. Actually, CNN consists of nodes that are connected to each other by weight, as neural network structures. CNN constructs of convolutional layers, pooling layers, and fully connected layers [9]. The output of each layer, there is a nonlinear activation function. As we know in the classification of the image, CNN input is data that consists of three 2D arrays that have height, width and depth. An illustration of the CNN architecture in general is shown in Figure 1.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

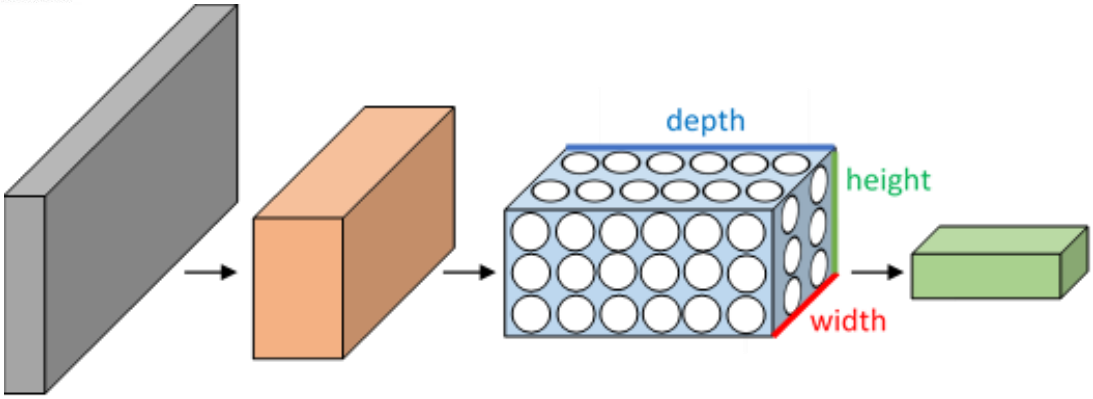


FIGURE 1. An Illustration of Convolutional Neural Network [10]

The basic component of CNN consists of a convolutional layer that is characterized by an input feature map F, filter kernel W, and bias b. Thus, the feature map of layer L is given as Equation (1) [11]:

$$F_L^k = \emptyset\,(W_L^k * F_{L-1}^k + b_L^k) \quad (1)$$

After the convolution stage, the image will enter the pooling layer. Pooling is a layer that reduces the dimensions of a feature map or better known as downsampling. One type of polling is Max Pooling, which in this layer takes the largest value for each filter change. The max-pooling function is shown in Equation (2) [12] :

$$f_m(v) = max_i\, x_i \quad (2)$$

We applied a nonlinier activation operation between the convolution and the pooling layer such as the rectification linear unit (ReLU) layer. The image size (feature map size) is reduced and more complex features are extracted after several convolutional and pooling layers [13]. The ReLu function in Equation (3) applies the values of matrix elements in the convolution layer. The ReLu layer will have the same value as the convolution layer except if the element is negative, in this case, the element in the ReLu layer will be 0 [11].

$$ReLu\,(x) = max\,(0, x) \quad (3)$$

Moreover, we use dropout layer for our network. The method used to overcome the problem of overfitting in deep learning is dropout. With this method, several

neuron units during training will be randomly removed. Then, each neuron units were retrained to provide certain probability p that some neuron activation disuse of the existing network. Dropout use approximately doubled the number of iterations required to converge [11], [13].

3. MATERIAL AND METHODS

The proposed CNN model in this study is VGGNet architecture. The CNN classification process uses transfer learning techniques by applying the VGG19 architecture which shows good results in the task of classification of the Image Large Scale Visual Recognition Challenge (ILSVRC). CNN classifiers are used to identify images of benign and malignant skin lesions. The proposed classification of skin lesions based on convolutional neural networks is divided into three main stages, namely pre-processing data, convolutional neural network classifiers, and evaluations. The block diagram system is displayed in Figure 2.

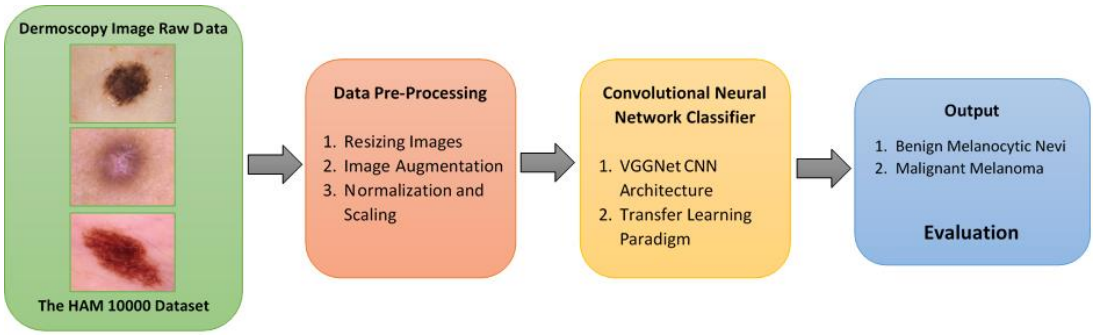
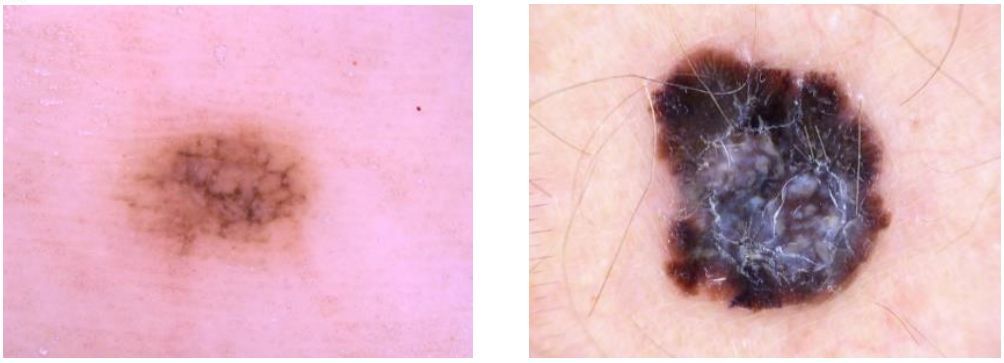


FIGURE 2. Process of Skin Lesion Classification based on CNN

3.1. DATA PREPARATION

Dermoscopy image is used for classification benign and malignant skin lesion of human skin. The dataset used as network input is The HAM 10000 dataset available through the 2018 International Skin Imaging Collaboration (ISIC) challenge archive. In this study, we use 6705 benign melanocytic nevi cases and 1113 malignant melanoma cases. It contains an 8-bit RGB dermoscopy image with various image resolutions [14]. The sample of dermoscopy image can be seen in Figure 3 which shows the difference between benign melanocytic nevi lesion (a) and malignant melanoma lesion (b).

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)



(a) Benign melanocytic nevi lesion (b) Malignant melanoma lesion

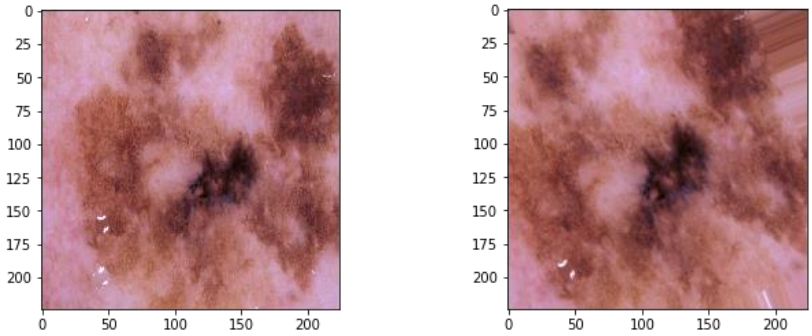
FIGURE 3. The Example of Skin Lesion [14]

3.2. PREPROCESSING

This stage is carried out pre-processing data to prepare images so that they are ready to be processed into the input of the network system. Input images are preprocessed by resizing the image to 224x224 pixels, augmenting the training data set using image processing and normalizing the pixel value.

The HAM 10000 dataset contains images with various resolutions. Therefore, it is necessary to resample the image by improving image resolution. The network input layer uses 224x224 pixel RGB images. This resolution is used because the default input image size on the VGGNet model is 224 x 224 pixels.

CNN training requires a large amount of data. This can be done with data augmentation techniques. It can increase the amount of data without removing the essence of the data. Data augmentation techniques used in this paper are rotation 40^{o}, horizontal flipping and vertical flipping. Also, the image will undergo minor changes by conducting random transformation for the data in each iteration in the training process. These changes include a random zoom, horizontal shift, vertical shift, and shear. Data augmentation process to improve CNN training is shown in Figure 4 which shows the difference between the original dermoscopy image (a) and the result lesion augmented image (b).



(a) the original dermoscopy image (b) the result lesion augmented image

FIGURE 4. Data Augmentation Process

Furthermore, we do scaling input values on the neural network model. We adjust the pixel values by dividing each RGB value from 0–255 by 255 so that it becomes a value in the range 0 and 1. This process is useful for improving the learning process.

4. EXPERIMENT

4.1. PROPOSED CNN ARCHITECTURE

In this section, the classification process of benign skin lesion (melanocytic nevi) and melanoma through dermoscopy images using a transfer learning technique in the pre-trained VGG-19 model as a feature extractor will be carried out. VGG-19 is a variation of the VGGNet architecture which has 19 layers (16 convolutional layers and 3 fully connected layers) [15]. The last fully-connected layer and the output layer of all pre-trained networks were removed and replaced by a new fully-connected layer with 1024 neurons to solve the binary classification problem. To prevent overfitting, a dropout layer is added. The architecture is summarized in Figure 5 and Table 1.

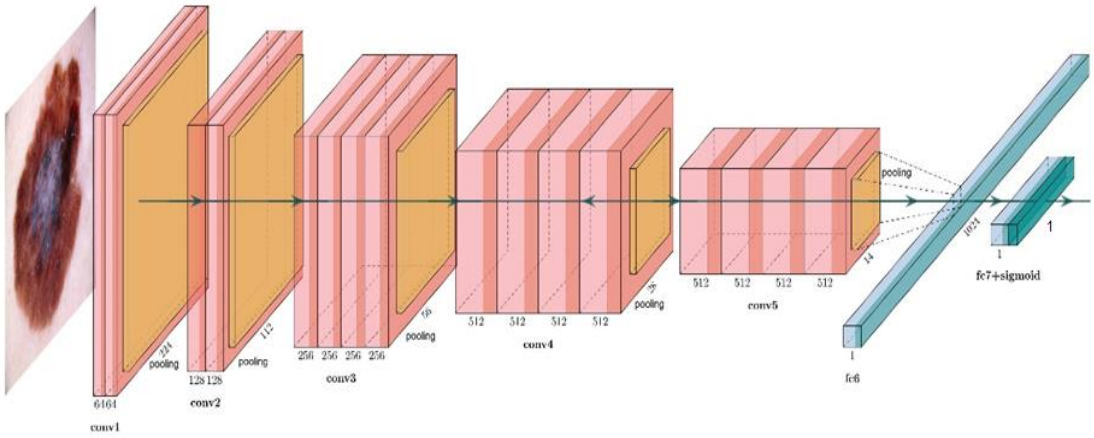


FIGURE 5. CNN Architecture for Classification of Skin Lesions using Transfer Learning in The Pretrained VGG-19 Model

The ReLU non-linearity is applied to every convolutional layer. Additionally, dropout optimization in the fully connected layer is applied with a value of 0.5. The output of the last fully-connected layer is fed to sigmoid which produces a distribution of 2 class labels. The sigmoid function is used for binary classification because this function takes the value of a real number and replaces its output value to be in the range of 0 to 1.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

TABLE 1.
Summaries of Architecture Layers of the Pretrained VGG-19 Model

Layer	Kernel size, Fearure map	Stride	Activation	Output Shape
Input data	-		-	224 x 224 x 3
Convolution Layer 1	3 x 3, 64	1	ReLu	224 x 224 x 64
Convolution Layer 2	3 x 3, 64	1	ReLu	224 x 224 x 64
Max Pooling	3 x 3, 64	2	-	112 x 112 x 64
Convolution Layer 3	3 x 3, 128	1	ReLu	112 x 112 x 128
Convolution Layer 4	3 x 3, 128	1	ReLu	112 x 112 x 128
Max Pooling Layer	3 x 3, 128	2	-	56 x 56 x 128
Convolution Layer 5	3 x 3, 256	1	ReLu	56 x 56 x 256
Convolution Layer 6	3 x 3, 256	1	ReLu	56 x 56 x 256
Convolution Layer 7	3 x 3, 256	1	ReLu	56 x 56 x 256
Convolution Layer 8	3 x 3, 256	1	ReLu	56 x 56 x 256
Max Pooling Layer	3 x 3, 256	2	-	28 x 28 x 256
Convolution Layer 9	3 x 3, 512	1	ReLu	28 x 28 x 512
Convolution Layer 10	3 x 3, 512	1	ReLu	28 x 28 x 512
Convolution Layer 11	3 x 3, 512	1	ReLu	28 x 28 x 512
Convolution Layer 12	3 x 3, 512	1	ReLu	28 x 28 x 512
Max Pooling Layer	3 x 3, 512	2	-	14 x 14 x 512
Convolution Layer 13	3 x 3, 512	1	ReLu	14 x 14 x 512
Convolution Layer 14	3 x 3, 512	1	ReLu	14 x 14 x 512
Convolution Layer 15	3 x 3, 512	1	ReLu	14 x 14 x 512
Convolution Layer 16	3 x 3, 512	1	ReLu	14 x 14 x 512
Max Pooling Layer	3 x 3, 512	2	-	7 x 7 x 512
Flatten	-		-	25088
Fully Connected Layer 1	-		ReLu	1024
Dropout	$p = 0.5$		-	1024
Output Layer	-		Sigmoid	1

4.2. TRAINING AND TESTING

The training process in the pre-trained VGG-19 model was carried out using the HAM 10000 dataset. The data used was a dermoscopy image from the augmentation process through the Image Data Generator. The amount of data produced was 17806 images including 8902 images for melanocytic nevi cases and 8904 images for melanoma cases. The augmented dataset is then randomly divided into 3 partitions; training data, test data, and validation data. The final dataset consisted of 10683 training images, 3561 testing images and 3562 validating images.

Training was conducted using 30, 50 and 100 epochs with a batch size of 32. The optimization function based on the adaptive moment estimation (Adam) with a learning rate of 10^{-4} which is implemented in Keras artificial intelligence platforms. The test data used for the evaluation and optimization of the performance of the network is done individually.

4.3. VALIDATION AND EVALUATION

The confusion matrix is a table that is commonly used to performance measurement in the field of machine learning and specifically the problem of statistical classification. It describes the performance of classification on a set of test data for which the true values are known. Each row of a matrix describe the sample in a predicted class while each column describe the sample in the actual class [16]. Figure 6 shows a confusion matrix of binary classification.

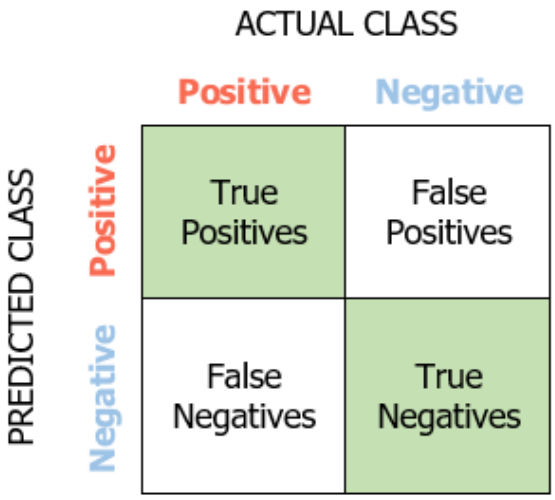


FIGURE 6. The Confusion Matrix of Binary Classification [17]

Binary classifiers have two data labels namely positive (P) and negative (N). This classification has four results with true positive (TP), true negative (TN), false positive (FP) and false-negative (FN) [17]. Evaluation using the confusion matrix produces equations (4) to (9) below.

$$ERR = \frac{FP+FN}{TP+TN+FP+FN} \quad (4)$$

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$SEN = \frac{TP}{TP+FN} \quad (6)$$

$$SPE = \frac{TN}{TN+FP} \quad (7)$$

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

$$PRE = \frac{TP}{TN+FP} \quad (8)$$

$$F1 = \frac{2 \times PRE \times SEN}{PRE+SEN} \quad (9)$$

Where accuracy (ACC) is the number of test results predicted correctly. Sensitivity (SEN) is the actual positive number that is correctly predicted. Specificity (SPE) is the actual negative number that is correctly predicted. Precision (PREC) is the number of positive predictions classified correctly divided by the total positive predictions. Error rate (ERR) is the number of test results predicted incorrectly. F-scores are averages of precision and sensitivity.

In the binary classification process between melanoma skin lesions and benign skin lesions (melanocytic nevi), the definition of TP is the number of patients suffering from skin cancer, TN is the number of patients with benign skin lesions, FP is the number of patients with benign skin lesions whose skin cancer prediction test results ,and FN is the number of patients suffering from skin cancer but the prediction results are benign skin lesions. A 2x2 confusion matrix table containing the sum of the 4 combinations of classification and actual disease status is shown in Table 2 [18].

TABLE 2.
Evaluation of Diagnosis Tests [18]

The Results of Diagnostic Tests	Status of Skin Lesions		Total
	Positive	Negative	
Positive	True Positive (TP)	False Positive (FP)	All Positive Tests (T+)
Negative	False Negative (FN)	True Negative (TN)	All Negative Tests (T-)
Total	Total with disease (D+)	Total without disease (D-)	Total Sample Size

5. RESULT AND DISCUSSION

This section shows the experiment of classification performance after training the pre-trained VGG-19 model against the deep learning on the dataset. We approach using 17806 dermoscopy images of the augmenting method. It includes 8902 images for melanocytic nevi cases and 8904 images for melanoma cases. The training process of the VGG19 pre-trained model uses a hyperparameter the rectification linear unit (ReLU) activation function at the fully connected layer, sigmoid activation function at the output layer, loss binary cross-entropy function, Adam optimizer with a learning rate of 10^{-4}, and the different number of epochs 30, 50 and 100 epochs. The plot graph of the results of accuracy and loss in training-testing the pre-trained VGG-19 model as a feature extractor with 100 epochs is shown in Figure 7 and 8.

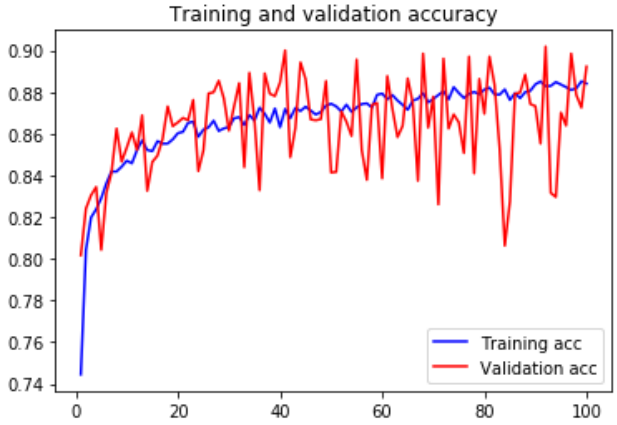


FIGURE 7. The Result of Overall Accuracy of the Pretrained VGG-19 Model

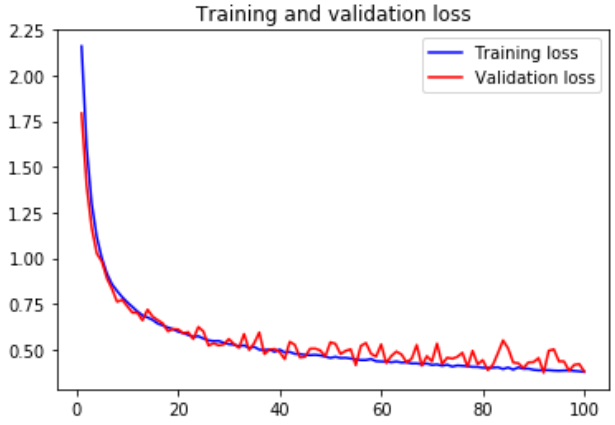


FIGURE 8. The Result of Overall Loss of the Pretrained VGG-19 Model

Also, Table 3 summarizes the binary classification performance results of the pre-trained VGG-19 model as a feature extractor in terms of the confusion matrix value.

TABLE 3.
Confusion Matrix for 100 Epoch of the Pretrained VGG-19 Model

	Malignant Melanoma	Benign Melanocytic Nevi
Malignant Melanoma	**1323**	458
Benign Melanocytic Nevi	58	**1722**

From Table 3, it can be concluded that if the number of cases of skin lesions correctly identified as malignant melanoma is considered as TP, but if not then it is considered FN. Conversely, the number of cases of skin lesions is considered as TN if the prediction is correctly classified as benign melanocytic nevi; if not, they are

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

FP. Also, the receiver operator characteristic curve (ROC) is used for the assessment of classification of skin lesions. Finally, Figures 9 and 10 illustrated the ROC curve and the precision-recall curve with the pre-trained VGG-19 model.

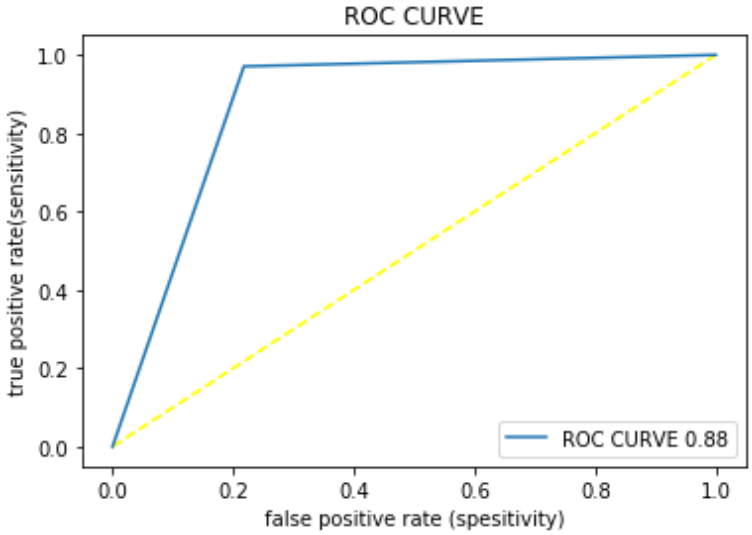


FIGURE 9. The ROC Curve of the Tested CNN Method for Skin Lesion Classification

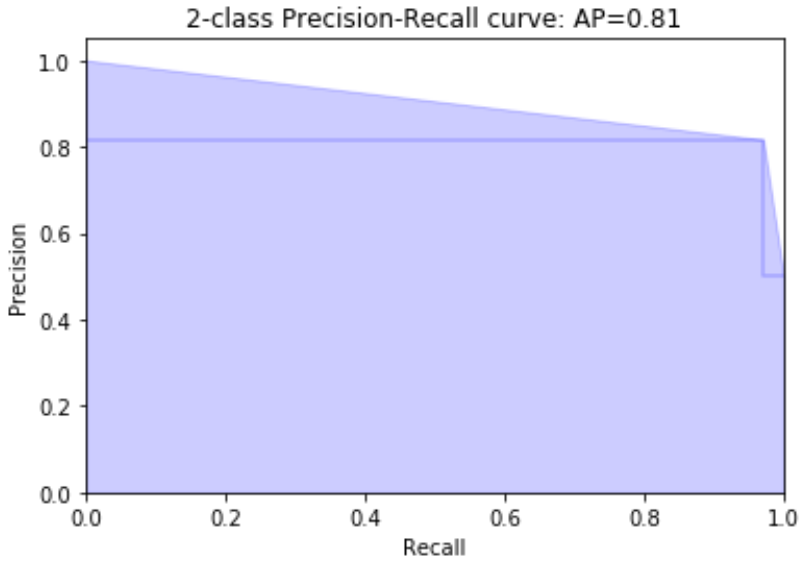


FIGURE 10. The Precision-Recall Curve of the Tested CNN Method for Skin Lesion Classification

Quantitatively, Table 4 shows the performances of the experiment in terms of sensitivity, specificity, precision, F1 score, and overall accuracy, then the best results from each comparison of models with different number of epochs.

TABLE 4.
Classification Performance of the Proposed CNN Method

Number of Epochs	Sensitivity	Specificity	Precision	F1 Score	Accuracy
30	95.80 %	78.99 %	74.28 %	83.68 %	85.45 %
50	92.21 %	83.64 %	81.80 %	86.70 %	87.46 %
100	**96.40 %**	81.66 %	78.21 %	86.36 %	**87.64 %**

The results of skin lesion classification of benign melanocytic nevi and malignant melanoma based on CNN pre-trained VGG-19 models with 100 epochs produce good classification with diagnostic accuracy is 87.64%, sensitivity is 96.40%, specificity is 81.66%, precision is 78.21% and F1 score is 86.36% as shown in Table 3. The highest sensitivity value is to be achieved in medical applications because the skin lesions were misdiagnosed cancers can seriously affect the health of the patient. Analysis of the overall results of the binary classification test shows that all networks have obtained good classification results. The best performance in terms of accuracy and sensitivity is obtained by a system that uses 100 epochs. In this study, the sensitivity coefficient reaches a value higher than 90% for all networks.

Table 5 summarizes the performance of the approach of the proposed method (ie the VGG19 model) and compares it with the previous approach regarding the binary classification of skin lesions using dermoscopy images.

TABLE 5.
Comparison of Performance using the CNN Algorithm in the Binary Classification of Skin Lesion

Author	Dataset	Approach	Sensitivity	Accuracy
Haenssle et al (2018) [19]	Image library of the Department of Dermatology, University of Heidelberg, Germany	Inception-v4 CNN	86.6 %	79 %
Esteva et al (2017) [20]	Image from dermatoscopic devices	Inception-v3 CNN	-	94 %
Hosny et al (2019) [21]	MedNod Dataset	AlexNet CNN	91.43 %	-
Proposed approach	The HAM10000 Dataset	VGG-19 CNN	**96.4 %**	87.64 %

As the comparison in Table 5 shows, our proposed approach outperforms all other algorithms proposed at the sensitivity value. Although the accuracy is not good because it requires a larger training dataset to achieve higher evaluation results in deep learning architecture. The results of the performance evaluation in the training and testing process show good classification results, thus proving that the CNN method is a good method used for 2D image classification problems.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

6. CONCLUSION

This section shows the result of classification performance the pre-trained VGG-19 model as a feature extractor. Specifically, the pre-trained VGG-19 model as a feature extractor of skin lesion classification of benign melanocytic nevi and malignant melanoma produces good classification with accuracy is 87.64%, sensitivity is 96.40%, specificity is 81.66%, precision is 78.21% and F1 score is 86.36%. Although the pre-trained VGG-19 model classification method has improved result as we have trained it longer, it still needs enhancement, in particular about its specificity, precision, and F1 score.

REFERENCES

[1] R. I. Hartman and J. Y. Lin, "Cutaneous Melanoma—A Review in Detection, Staging, and Management," *Hematol. Oncol. Clin. North Am.*, vol. 33, no. 1, pp. 25–38, 2019.

[2] T. A. C. Society, "Key Statistics for Melanoma Skin Cancer," 2016. [Online]. Available: https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html. [Accessed: 14-Dec-2018].

[3] T. Y. Tan, L. Zhang, S. C. Neoh, and C. P. Lim, "Intelligent skin cancer detection using enhanced particle swarm optimization," *Knowledge-Based Syst.*, vol. 158, pp. 118–135, 2018.

[4] N. L. Husni, A. Silvia, E. Prihatini, S. Nurmaini, and I. Yani, "Swarm Intelligent in Bio-Inspired Perspective : A Summary," *Comput. Eng. Appl.*, vol. 7, no. 2, pp. 105–120, 2018.

[5] F. M. C´ıcero, A. H. M. Oliveira, and G. M. Botelho, "Deep Learning and Convolutional Neural Networks in the Aid of the Classification of Melanoma," *SIBGRAPI*, pp. 1–9, 2016.

[6] A. J. Pathiranage, "Convolutional Neural Networks for Predicting Skin Lesions of Melanoma," 2017.

[7] A. Romero Lopez, X. Giro-i-Nieto, J. Burdick, and O. Marques, "Skin Lesion Classification from Dermoscopic Images Using Deep Learning Techniques," *Biomed. Eng. (NY).*, 2017.

[8] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, 2018.

[9] Z. Zhao and A. Kumar, "Accurate Periocular Recognition under Less Constrained Environment Using Semantics-Assisted Convolutional Neural Network," *IEEE Trans. Inf. Forensics Secur.*, vol. 12, no. 5, pp. 1017–1030, 2017.

[10] Stanford computer science class, "Convolutional Neural Networks for Visual Recognition," *final course project reports online*, 2016. [Online]. Available:

http://cs231n.github.io/convolutional-networks/. [Accessed: 08-Oct-2018].

[11] M. A. Al-masni, M. A. Al-antari, M. T. Choi, S. M. Han, and T. S. Kim, "Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks," *Comput. Methods Programs Biomed.*, vol. 162, pp. 221–231, 2018.

[12] Y.-L. Boureau, J. Ponce, J. P. Fr, and Y. Lecun, "Icml2010B.Pdf," *Icml*, pp. 111–118, 2010.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Proceeding NIPS'12 Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, vol. 12, p. 04015009, 2015.

[14] P. Tschandl, C. Rosendahl, and H. Kittler, "Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, pp. 1–9, 2018.

[15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *ICLR*, pp. 1–14, 2015.

[16] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, pp. 861–874, 2006.

[17] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation David," *Tech. Rep. SIE-07-001*, no. December, 2007.

[18] K. H. Zou, A. J. O'Malley, and L. Mauri, "Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models," *Circulation*, vol. 115, no. 5, pp. 654–657, 2007.

[19] H. A. Haenssle *et al.*, "Man against Machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Ann. Oncol.*, vol. 29, no. 8, pp. 1836–1842, 2018.

[20] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, p. 115, Jan. 2017.

[21] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-net," *PLoS One*, vol. 14, no. 5, pp. 1–17, 2019.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)