JOURNAL MEDICAL INFORMATICS TECHNOLOGY
medinftech.org ISSN 2988-7003 (online)

Application of Deep Learning with ResNet50 for Early Detection of Melanoma Skin Cancer

Nurul Khasanah[1] and Monikka Nur Winnarto[2]

[1]*Universitas Nusa Mandiri, Indonesia*
[2]*Universitas Bina Sarana Informatika, Indonesia*

ARTICLE HISTORY
Received: 23 March 24
Final Revision: 30 March 24
Accepted: 30 March 24
Online Publication: 31 March 24

KEYWORDS
Classification, Deep Learning, InceptionV3, ResNet50, Skin Cancer

CORRESPONDING AUTHOR
nurul.nuk@nusamandiri.ac.id

DOI
10.37034/medinftech.v2i1.31

A B S T R A C T

Cancer is a type of disease that can be fatal. Some of the cancers with the highest death rates in Indonesia include uterine cancer, breast cancer and skin cancer. The most malignant types of skin cancer are melanoma, which has a high mortality rate, especially if not detected in the early stages, and non-melanoma skin cancer (NMS Cs). Management of this disease depends on whether the type of skin cancer is malignant (malignant) or non-malignant (benign). Therefore, we need a system that can classify types of skin cancer with high accuracy. In this research, the author will use deep learning with the InceptionV3 and ResNet50 algorithms to carry out classification. The aim of this research is to classify types of skin cancer using the InceptionV3 and ResNet50 architecture. The skin cancer dataset used consists of two classes, namely Benign and Malignant, with a total of 3297 data, consisting of 660 data for testing and 2637 data for training. Research stages include data acquisition, preprocessing, classification, and analysis of results. Experimental results show that ResNet-50 produces the best performance with an accuracy level of 0.87. Innovations from this research include using a larger dataset, testing two deep learning architectures, modifying hyperparameters, and using a different layer architecture, which produces better accuracy than previous research. It is hoped that the results of this research can be applied to classify skin cancer more accurately.

1. Introduction

Cancer is a disease that is included in the category of deadly diseases. Based on data on New Cancer Cases and Deaths based on Gender, United States, 2020, it is stated that several cancers that have high mortality rates are Oral cavity and pharynx, Digestive system, Respiratory system, Bones and joints, Soft tissue (including heart), Skin (excluding basal and squamous), and others [1].

Skin cancer is a disease that is ranked third after cervical cancer and breast cancer in Indonesia. The prevalence of skin cancer reaches around 5.9 to 7.8% of total cancer cases each year. The most common skin cancer in Indonesia is basal cell carcinoma (65.5%), followed by squamous cell carcinoma (23%), malignant melanoma (7.9%) and other skin cancers. The most invasive form of skin cancer is melanoma, has a high mortality rate, especially if not detected early and nonmelanoma skin cancers (NMS Cs), such as basal cell carcinoma and squamous cell carcinoma are more common but less metastatic, and only a small proportion lead to death [2].

Image processing technologies and artificial intelligence have emerged as potential tools to improve early detection of melanoma. The use of deep learning algorithms in medical image analysis has provided new hope in this regard. One of the prominent deep learning architectures is ResNet50, which is known for its ability to overcome the vanishing gradient problem and its deep neural network architecture. With its extraordinary ability to understand complex features in images, ResNet50 offers great potential in supporting the process of melanoma detection in dermatoscopic images.

In the study [3] titled "Classification of skin cancer images using local binary pattern and SVM classifier," utilizing 655 images (544 benign and 111 malignant), the obtained accuracy result was 76.1%. In another study [4], titled "Segmentation and Classification of Skin Cancer Melanoma from Skin Lesion Images," employing the IDIC dataset comprising 220 training data and 20 testing data, the highest accuracy of 78.2% was achieved with the SVM algorithm.

Arief Budhiman, et al, in 2019 carried out Melanoma skin cancer classification using CNN, namely using the ResNet architecture with ResNet 50, 40, 25, 10 and 7. This architecture was trained using augmented train data and under sampling. Validation Results for each model are calculated using F1 Score. The best architecture is ResNet 50 without augmentation which provides validation accuracy of 0.83 and f1 score of 0.46 [5].

Pratik Dubal et al (2017) used 463 skin cancer lesion datasets taken using a common camera to classify skin cancer benign lesions and malignant lesions using a Neural Network. This research uses segmented images and then features are extracted using ABCD and Neural Network rules as classification. The classification results produced six classes and an accuracy of 76.9% [6].

Amirreza Mahbod et al (2019) classified malignant and benign skin cancer using a Hybrid Deep Learning Network optimized from a number of CNNs and several different levels of abstraction. Three methods are used consisting of Alexnet, VGG16 and Resnet-18 as electricity generator features. This study used a dataset of 150 images consisting of 30 malignant melanoma, 42 orrheic seb-keratosis and 78 benign nevus images. This study yielded 83.83% for melanoma [7].

In 2019, A Murugan et al, classified skin cancer types into 2 classes, namely benign or melanoma. This research uses the watershed segmentation method as segmentation. The extracted segmented image will undergo feature extraction. The features extracted are shape, ABCD rule and GLCM. The feature files are then used for classification. The classification methods used are Random Forest, KNN, SVM. The classification results from 1000 sample datasets were 82.31% for SVM, 71.97 for Random Forest and 62.19 for KNN in shape feature classification [8].

Based on the background explanation above, the innovations of this research include the utilization of a larger skin cancer dataset, totaling 3297 data consisting of 660 for testing and 2637 for training, testing two deep learning architectures namely InceptionV3 and ResNet50 networks, modification of hyperparameters, and the implementation of different layer architectures. The framework of this paper comprises an introduction containing the research background, the methods used in section 2, section 3 consisting of the discussion or analysis of research results, and finally concluded with a conclusion.

2. Research Method

The following are the stages of research conducted by researchers presented in Figure 1.

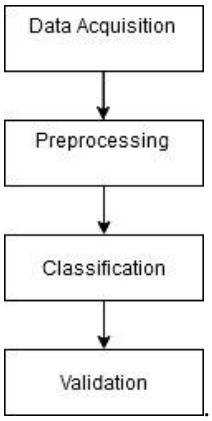

Figure 1. Research Methodology

2.1. Data Acquisition

The skin cancer dataset in this research was taken from Kaggle [9]. The dataset consists of two main classes, namely Benign and Malignant, with a total of 3297 data. The data is divided into 660 test data and 2637 training data. From the test data, 360 are Benign images and 300 are Malignant images. Meanwhile, from the training data, 1440 images are Benign and 1197 images are Malignant. The details of the dataset used are presented in Table 1.

Table 1. Details dataset

	Benign	Malignant
Test	360	300
Train	1440	1197

The visual representation of the sample dataset is as follows: Figure 2 depicts an image classified as Benign, while Figure 3 displays an image categorized as Malignant.

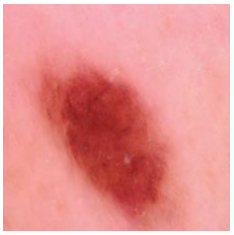

Figure 2. Benign image

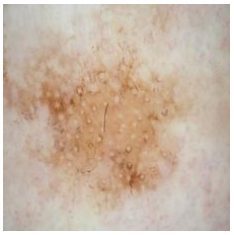

Figure 3. Malignant image

2.2. *Preprocessing*

After getting the dataset, the next stage is data preprocessing. In the preprocessing stage, the first step is to change the image size to 224 x 224 pixels. After that, the dataset is divided into two parts, namely the

training dataset which covers 90% of the total data and the testing dataset which covers the remaining 10%.

2.3. Classification

In the classification process, the algorithm used is deep learning. Deep learning is a part of machine learning that takes inspiration from the way the human brain works by applying multilevel learning [10]. In this context, multilevel refers to the various layers that make up deep learning. The initial layers produce simple features, while the final layers produce more complex features. Deep learning automatically performs feature extraction and classification [11]. The algorithm used for classification is Resnet50.

ResNet50 is a 50 layer Residual network and has other variants such as ResNet101 and ResNet152 [12]. Using ResNet as a trained model for medical image classification has provided good results [13]. Figure 4 displays the structure of the ResNet50 architecture.

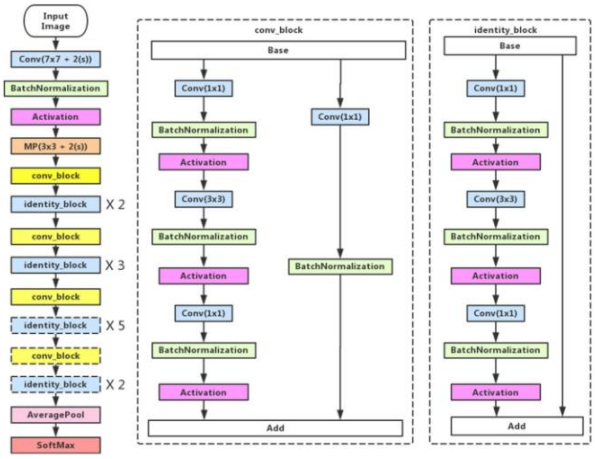

Figure 4. Architecture of ResNet50

In general, ResNet50 consists of Convolutional, followed by Average Pooling and ending with a Fully connected layer as a classification layer. There are several differences between the Resnet-50 architecture used in this research and the original architecture. This was done as needed in this research. Here are the modifications we made:

a. The number of outputs on the fully connected layer is 2 classes

b. The activation function at the fully connected layer becomes ReLU

After the model is formed, it then enters the data training process using the ResNet-50 algorithm. Some of the hyperparameters that we use include: batch size = 32, learning rate = 0.001, epoch = 14. Before entering the classification process, feature extraction is carried out using a combination of convolutional blocks and identity blocks. From these two blocks, there are several filter sizes used, including: 1x1, 3x3 and 1x1. After feature extraction is complete, enter the classification process. In this process, the reLU activation function is used. The total training parameter results are 23,591,810.

2.4. Validation of Results

Validating research results using a confusion matrix is an important step in measuring the performance of skin image classification models (such as in melanoma skin cancer detection research) built using deep learning such as ResNet-50. Confusion matrix is a table used to visualize model performance in classifying data. This table contains four main matrices: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [14].

a. True Positive (TP)

This is the number of images that were truly positive (melanoma skin cancer) and correctly classified by the model as positive. In this context, this means a skin image of a melanoma that is actually correctly detected as melanoma by the model.

b. True Negative (TN)

This is the number of images that were truly negative (non-melanoma) and correctly classified by the model as negative. In this context, this means non-melanoma images that are actually correctly detected as non-melanoma by the model.

c. False Positive (FP)

This is the number of images that were actually negative (non-melanoma) but incorrectly classified by the model as positive (melanoma). This was an error in which the model triggered a melanoma skin cancer warning when none actually existed.

d. False Negative (FN)

This is the number of images that were actually positive (melanoma) but incorrectly classified by the model as negative (non-melanoma). This is an error in which the model cannot detect melanoma that is actually present.

Using the TP, TN, FP, and FN values, it is possible to calculate a number of model performance evaluation metrics, such as:

a. Accuracy: Accuracy is the proportion of all correct predictions, namely (TP + TN) divided by the total images [15].

b. Sensitivity (Sensitivity or True Positive Rate): Sensitivity measures the extent to which the model can detect positive cases correctly, namely TP divided by (TP + FN) [15]. This is the model's ability to detect melanoma.

c. Specificity: Specificity measures the extent to which the model can avoid false positives, namely TN divided

by (TN + FP) [16]. This is the model's ability to identify non-melanoma.

d. Precision: Precision measures how many of the cases classified as positive by the model are truly positive, i.e. TP divided by (TP + FP) [15].

e. F1-Score: F1-score is the harmonic average of sensitivity and precision, and is useful when there is a trade-off between these two metrics [16].

3. Result and Discussion

From the results of the experiments we conducted using 3297 skin cancer data using the ResNet50 method, the results were as follows:

3.1. Model summary

Table 2 provides detailed information about the model summary of the proposed ResNet50 architecture in this study.

Table 2. Model Summary of ResNet50 Architecture

Layer	Pixel Size	Parameter
Input layer	224, 224, 3	0
ZeroPadding2D	230, 230, 3	1792
Conv2D	112, 112, 64	9472
BatchNormalization	112, 112, 64	256
Activation	112, 112, 64	0
ZeroPadding2D	114, 114, 64	0
MaxPooling2D	56, 56, 64	0
Conv2D_1	56, 56, 64	4160
BatchNormalization	56, 56, 64	256
Activation	56, 56, 64	0
Conv2D_2	56, 56, 64	36928
BatchNormalization	56, 56, 64	256
Activation	56, 56, 64	0
Conv2D_0	56, 56, 256	16640
Conv2D_3	56, 56, 256	16640
BatchNormalization	56, 56, 256	1024
BatchNormalization	56, 56, 256	1024
Add	56, 56, 256	0
Activation	56, 56, 256	0
Conv2D_1	56, 56, 64	16448
BatchNormalization	56, 56, 64	256
Activation	56, 56, 64	0
Conv2D_2	56, 56, 64	36928
BatchNormalization	56, 56, 64	256
Activation	56, 56, 64	0
Conv2D_3	56, 56, 256	16640
BatchNormalization	56, 56, 256	1024
Add	56, 56, 256	0
Activation	56, 56, 256	0
Conv2D_1	56, 56, 64	16448
BatchNormalization	56, 56, 64	256
Activation	56, 56, 64	0
Conv2D_2	56, 56, 64	36928
BatchNormalization	56, 56, 64	256
Activation	56, 56, 64	0
Conv2D_3	56, 56, 256	16640
BatchNormalization	56, 56, 256	1024
Add	56, 56, 256	0
Activation	56, 56, 256	0
Conv2D_1	28, 28, 128	32896
BatchNormalization	28, 28, 128	512
Activation	28, 28, 128	0
Conv2D_2	28, 28, 128	147584
BatchNormalization	28, 28, 128	512
Activation	28, 28, 128	0

Based on Table 2, the ResNet50 architecture follows a workflow consisting of several crucial steps in image processing. It starts with an input layer that receives images with a size of 224x224 pixels and 3 color channels. Then, ZeroPadding2D operation is performed to expand the size of the image, followed by the first Conv2D with 64 filters to extract initial features. After that, Batch Normalization and ReLu activation function are applied to accelerate model convergence and introduce non-linearity. This is followed by ZeroPadding2D and MaxPooling2D to extract important features from the image. The culmination is the repetition of residual blocks consisting of convolutional layers performed in each block, with the addition of shortcut connections to strengthen information flow and address the vanishing gradient problem. Finally, the last convolutional layer generates a feature vector used for image classification.

3.2. Accuracy and Loss

The following table presents the accuracy and loss values obtained from the conducted experiments:

Table 3. Accuracy and Loss

Method	Epoch	Accuracy	Loss
ResNet50	10	0.84	0.34
	15	0.86	0.30
	20	0.87	0.28

Table 3 presents the training results of the ResNet50 model across various epochs, namely 10, 15, and 20 epochs. Each epoch yields different accuracy and loss values. In epoch 10, the model achieves an accuracy of 0.84 with a loss of 0.34. Then, in epoch 15, there is an improvement in accuracy to 0.86 with a loss of 0.30. Finally, in epoch 20, the accuracy increases to 0.87 with a loss of 0.28. Thus, it can be concluded that as the number of epochs increases, the model's accuracy tends to improve, while its loss tends to decrease. This indicates that the model becomes better at learning patterns in the training data over time.

3.3. Accuracy and Loss Graph

Below are the graphical representations illustrating the accuracy and loss values acquired from experiments performed utilizing the optimal epoch (epoch 20):

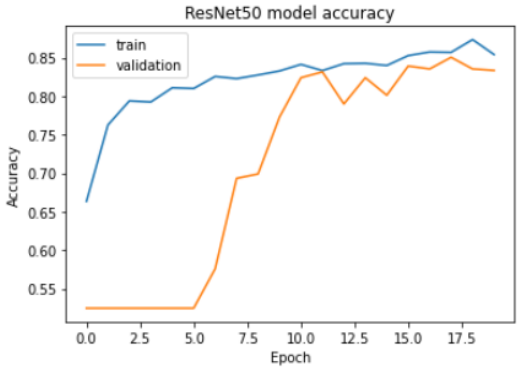

Figure 5. Accuracy Graph of ResNet50

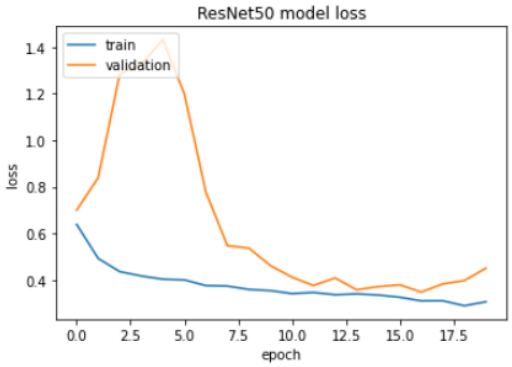

Figure 6. Loss Graph of ResNet50

The graphs in Figure 5 and Figure 6 illustrate the development of accuracy and loss values at each epoch for training and validation data. The yellow line indicates changes in accuracy and loss values in the validation data, while the blue line shows changes in accuracy and loss values in the training data. The graph indicates that the loss value decreased gradually until reaching the epoch 20, while the accuracy value continued to increase until reaching the epoch 20. At the end of the epoch, the training data reached an accuracy level of 0.87, while the validation data reached an accuracy level of 0.84.

4. Conclusion

Based on the results of the research that has been carried out, it can be concluded that an increase in accuracy in classification using the ResNet50 method with variations in the number of epochs has been observed. In the case of the ResNet50 method, accuracy increased from 0.84 at epoch 10 to 0.87 at epoch 20. In addition, the loss value at epoch 20 was 0.28. These results show that the ResNet50 method with 20 epochs provides the best results in terms of accuracy and loss value. The novelty in this research lies in the use of a larger dataset, namely 2637 datasets, as well as modifications to the layer architecture and hyperparameter settings used. The better accuracy results than previous studies indicate that this research is a valuable contribution to the development of image classification techniques. In conclusion, the ResNet50 method with epoch 20 is the best choice for image classification in this research, with an accuracy of 0.87 and a loss of 0.28.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA. Cancer J. Clin.*, vol. 70, no. 1, pp. 7–30, 2020.

[2] N. Alyyu, R. Fuadah, and N. Pratiwi, "Klasifikasi Kanker Kulit Ganas Dan Jinak Menggunakan Metode Convolutional Neural Network," *e-Proceeding od Eng.*, vol. 8, no. 6, pp. 3200–3206, 2022.

[3] F. Adjed, I. Faye, F. Ababsa, S. J. Gardezi, and S. C. Dass, "Classification of skin cancer images using local binary pattern and SVM classifier," *AIP Conf. Proc.*, vol. 1787, 2016.

[4] N. C. Lynn and Z. M. Kyu, "Segmentation and classification of skin cancer Melanoma from skin lesion images," *Parallel Distrib. Comput. Appl. Technol. PDCAT Proc.*, vol. 2017-Decem, pp. 117–122, 2018.

[5] A. Budhiman, S. Suyanto, and A. Arifianto, "Melanoma Cancer Classification Using ResNet with Data Augmentation," *2019 2nd Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2019*, pp. 17–20, 2019.

[6] P. Dubai, S. Bhatt, C. Joglekar, and S. Patii, "Skin cancer detection and classification," *Proc. 2017 6th Int. Conf. Electr. Eng. Informatics Sustain. Soc. Through Digit. Innov. ICEEI 2017*, vol. 2017-Novem, pp. 1–6, 2018.

[7] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinger, "Institute for Pathophysiology and Allergy Research , Medical University of Vienna , Austria Department of Research and Development , TissueGnostics GmbH , Austria Department of Computer Science , Loughborough University , U . K . Department of Biomedical," *Ieee*, pp. 1229–1233, 2019.

[8] A. Murugan, S. A. H. Nair, and K. P. S. Kumar, "Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers," *J. Med. Syst.*, vol. 43, no. 8, 2019.

[9] C. Fanconi, "Skin Cancer: Malignant vs. Benign," Skin Cancer: Malignant vs. Benign. Accessed: Mar. 10, 2024. [Online]. Available: https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign

[10] W. Zhang, X. Gu, L. Tang, Y. Yin, D. Liu, and Y. Zhang, "Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge," *Gondwana Research*, vol. 109, pp. 1–17, Sep. 2022, doi: 10.1016/j.gr.2022.03.015.

[11] A. W. Shandi Noris, "Implementasi deep learning untuk klasifikasi tanaman hias beracun menggunakan algoritma convolutional neural network (cnn)," vol. 6, no. 1, pp. 39–46, 2023.

[12] A. S. B. Reddy and D. S. Juliet, "Transfer Learning with ResNet-50 for Malaria Cell-Image Classification," *2019 International Conference on Communication and Signal Processing (ICCSP)*, Apr. 2019, doi: 10.1109/iccsp.2019.8697909.

[13] A. Sai Bharadwaj Reddy and D. Sujitha Juliet, "Transfer learning with RESNET-50 for malaria cell-image classification," *Proc. 2019 IEEE Int. Conf. Commun. Signal Process. ICCSP 2019*, pp. 945–949, 2019.

[14] Daniati Uki Eka Saputri, Nurul Khasanah, F. Aziz, and Taopik Hidayat, "Enhancing Skin Cancer Classification Using Optimized InceptionV3 Model," *Journal Medical Informatics Technology*, pp. 65–69, Sep. 2023, doi: 10.37034/medinftech.v1i3.14.

[15] S. Hadianti and W. A. G. Kodri, "Optimization of The Machine Learning Approach using Optuna in Heart Disease Prediction," *Journal Medical Informatics Technology*, pp. 59–64, Sep. 2023, doi: 10.37034/medinftech.v1i3.15.

[16] I. M. De Diego, A. R. Redondo, R. R. Fernández, J. Navarro, and J. M. Moguerza, "General Performance Score for classification problems," *Applied Intelligence*, vol. 52, no. 10, pp. 12049–12063, Jan. 2022, doi: 10.1007/s10489-021-03041-7.