Classification of Cervical Cancer Images Using Deep Residual Network Architecture

Hilman Fauzi*,[a,1], Revydo Bima Anshori [a,2], Thomhert S. Siadari [a,3], Ali Budi Harsono[b,4], Qisthi Nur Rahmah [a,5]

[a] Telkom University, Jl. Telekomunikasi 1 , Bandung, West Java, 40257, Indonesia
[b] Universitas Padjadjaran, Jl. Raya Bandung Sumedang, Sumedang, West Java, 45363, Indonesia
[1] hilmanfauzitsp@telkomuniveristy.ac.id; [2] rvydobima@student.telkomuniversity.ac.id;
[3] thomhert@telkomuniversity.ac.id; [4] alibudih@yahoo.com; [5] qisthinur@student.telkomuniversity.ac.id
* corresponding author

ARTICLE INFO

ABSTRACT

According to data from the World Health Organization (WHO), cervical cancer is ranked second, with a high mortality rate in women every year. Cervical cancer is caused by the presence of the Human Papilloma Virus (HPV), which directly attacks the cervix. Additionally, an unhealthy lifestyle can cause attacks of this disease. Several methods can be used to detect cervical cancer early, one of which is Visual Inspection with Acetic Acid (VIA). Through VIA, tests can determine whether patients are infected with the HPV virus. The results of the VIA test can be seen with the naked eye, but medical experts have different opinions about the diagnosis made using their vision. Therefore, to assist medical practitioners in diagnosing the results of VIA, an examination with a technological approach was carried out. Digital imagery was used for the analysis. A medical expert's Android camera was used with .jpg image format to capture pictures of the VIA test results. In this study, cervical cancer image classification was carried out from the results of the VIA test examination that had been carried out at Hasan Sadikin Hospital, Bandung, with as many as 255 data points for Negative VIA and 65 data points for Positive VIA. In the image processing of the VIA test results, CLAHE images and Canny Edge Detection images are used. Deep learning was used with the ResNet-50 and ResNet-101 architectural models to classify images, and different hyperparameter configurations, such as optimizers, learning rates, batch sizes, and input sizes, were tested. In this study, the best results were obtained using Canny Edge Detection images with hyperparameter configurations using the SGD optimizer with a learning rate of 0.1, a batch size of 32, and an input size of 224×224.

1. Introduction

Cancer is characterized by abnormal cell growth beyond normal limits, which can then invade adjoining parts of the body and/or spread to other organs (WHO, 2018). There are many types of malignant cancer, one of which is cervical cancer. Cervical cancer is a malignant cancer with a high mortality rate that affects the cervix of women. The Human Papilloma Virus (HPV), an oncogenic subtype that attacks cervical cells and causes abnormal cell growth, is what causes cervical cancer [1]. Based on GLOBOCAN data in 2020, cervical cancer in Indonesia ranks second among the 10 most common cancers with an incidence of 9.2%. According to the Indonesian Ministry of Health, the number of women suffering from early cervical cancer was approximately 23.4 per 100,000 women diagnosed with cervical cancer, and the mortality rate was 13.9 per 100,000 people in 2018

[2]. In order to reduce the risk for cervical cancer patients, the government conducts a cervical cancer screening program where counseling on cervical cancer is given and a Visual Inspection with Acetic Acid (VIA) examination is carried out. This cervical cancer screening program is provided free of charge through national health insurance [3]. Visual Inspection with Acetic Acid (VIA) is a simple screening method that uses vinegar acid solution at 2% and iosium lugol solution on the cervix and discolored cervical tissue to find pre-cancerous lesions and cervical cancer in women. Where if there is a well-defined white area with regular borders, then the result is positive for cervical cancer [4]. However, a lack of knowledge about cervical cancer, HPV vaccination, cervical cancer screening, concerns about side effects, and vaccine costs are obstacles to HPV vaccination among Indonesian women [5]. HPV vaccination planned by the government has shown an overall uptake of HPV vaccine in Indonesia, with an estimated uptake of more than 90%, and several cities in Indonesia have contributed to the high acceptance rate of HPV vaccine in some schools [6].

In research [7], a new method was developed to automate cervical cancer screening using an object detection approach based on a Convolutional Neural Network (CNN), which provides reference information on the location and category of abnormal cells in more detail. A convolutional neural network (CNN) is a type of architecture in artificial neural networks that is used to process data in the form of images inspired by how visual sensory organs work in living things and uses convolution, pooling, and activation techniques to recognize objects in images [8]. The optimal accuracy results from the new development for cervical cancer screening automation obtained in the study [7] reached 89.3%. In this case, it can be known that the classification of cervical cancer images from VIA test results can use the CNN method. There are various types of CNN architecture models, one of which is the Deep Residual Network (ResNet). ResNet can solve the problem of vanishing gradients in very deep neural networks and introduce residual blocks that allow information to pass through multiple layers in the network [9]. Deep Learning training also takes a long time and is limited to a certain number of layers. To solve problems in Deep Learning training, the Deep Residual Network method uses skip connections or connection shortcuts [10]. This feature makes it possible to obtain a deeper network and reduce the error rate. A deep residual network is divided into several different types of architectures based on the number of layers used, ranging from 18 layers, 34 layers, 50 layers, 101 layers, and 152 layers [11].

In a study [12], a complex classification of cervical cancer images was performed using the Deep Residual Network (ResNet) method. The ResNet architecture model used in this study is ResNet-18. The optimal accuracy obtained in this study was 97%. The accuracy results obtained using the ResNet architecture model showed better accuracy in classifying cervical cancer images. In this study, a digital image classification system for cervical cancer was designed using a deep residual network architecture. Two types of Deep Residual Network architecture were used in this study: ResNet-50 and ResNet-101. The ReNet-50 architecture model is a type of residual network model with 48 convolutional layers, one maximum pooling layer, and one average pooling layer [13]. ResNet-50 introduces a new concept called shortcut connections. This concept is related to vanishing gradient problems produced by the deepening of a network [14]. ReNet-101 is a type of residual network that has a total of 101 layers, including 99 convolutional layers, one maximum pooling layer, and one average pooling layer [15]. The dataset used in this study was obtained from Hasan Sadikin Hospital, Bandung. This study aimed to assist medical experts in diagnosing white lesion patterns using the Deep Residual Network method in classifying cervical cancer images from VIA test results.

2. Method

A cervical cancer image classification system using a Deep Residual Network architecture will be designed in several stages in accordance with the objectives of this study. The stages of the cervical cancer image classification system are shown in Fig. 1.
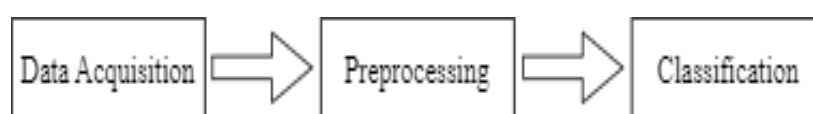

Fig. 1. Block System Design Diagram

The design of the cervical cancer image classification system starts at the first stage, namely data acquisition, which refers to the process of taking data directly using a camera to obtain datasets. After the data acquisition process was complete, a preprocessing step was carried out to obtain a better image. Then, the data that has been preprocessed will be classified with the aim of increasing the level of recognition and performance of image classification.

2.1. Data Acquisition

In this study, data were obtained from the results of the VIA test conducted at Hasan Sadikin Hospital, Bandung. The data were captured directly using a camera with a minimum resolution of 8 MP on Android. Taking pictures will be carried out with a distance of approximately 10–18 cm focused on the cervix and indoor lighting of 300–500 lux, and the images will be stored in .jpg format. For testing, this system uses data divided into two classes: positive and negative. The initial input data of the VIA test examination is shown in Fig. 2 as follows:


(a)   (b)

Fig. 2. Class Image (a) Negative Class Image & (b) Positive Class Image
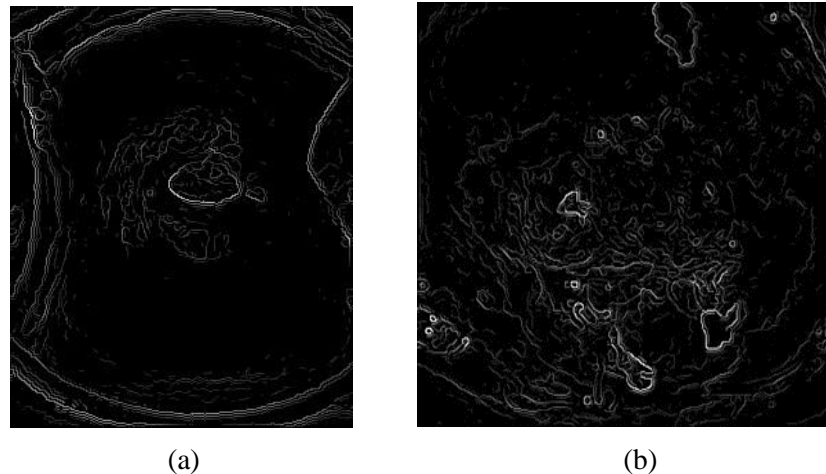
2.2. Preprocessing

Data preprocessing is the initial processing of data, which includes cleaning data, filling in missing values, smoothing noisy data, recognizing and eliminating outliers, and reducing large amounts of data [16]. In this study, data preprocessing was performed with the aim of obtaining a better image for the next stage.

In this preprocessing step, there are several stages, such as the resizing stage, which is used to resize the image. The size of the images entered is very diverse; therefore, the resizing stage is necessary so that the size of each image entered is the same. Furthermore, a data augmentation stage was used to manipulate the image data. Data Augmentation aims to add or subtract data if the amount of data is not equal to the number in each class. The Contrast Limited Adaptive Histogram Equalization (CLAHE) stage is then used to improve the contrast and image quality by fine-tuning the pixel intensity distribution in the image [17]. CLAHE can remove noise from images by limiting the contrast values. The CLAHE image is shown in Fig. 3.

(a)   (b)
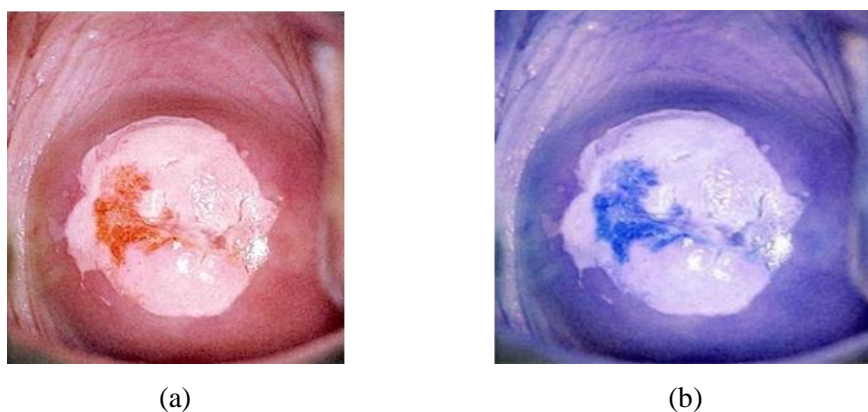
Fig. 3. CLAHE Image (a) Image Before CLAHE (b) Image After CLAHE

The final stage is Canny Edge Detection, one type of algorithm used in detecting edges in digital images by providing image results that show the location of the discontinuity of image intensity as an output of this process [18]. First, Canny Edge Detection will carry out the Gaussian Convolution process to smooth the input image and remove noise contained in the input. Furthermore, an application process will be carried out to highlight the location of the highest-rangking constituency. Various VIA test results in image input data will be divided into two groups with their respective portions, namely 90% training data and 10% test data. The training data and test data used will be randomly retrieved from the dataset used. Below is a flowchart of the preprocessing process in Fig. 4.

(a)   (b)

Fig. 4. Canny Edge Detection Image (a) Negative Class Canny Edge Detection Image (b) Positive Class Canny Edge Detection Image

2.3. Classification

At this classification stage, the preprocessed data are processed using the data split process. This data-splitting process was performed to divide the training, validation, and test data. After that, continue by configuring the architectural model. The architectural models used in this study are ResNet-50 and ResNet-101. The next stage is to configure hyperparameters before conducting the training and testing processes. The hyperparameter configurations used in this stage include the Optimizer, Learning rate, Batch size, and Input size. In this configuration process, the epoch value used is 40. The next stage was to conduct training and testing of the ResNet-50 and ResNet-101 architectural models.

2.4. System Test Parameters

System testing parameters are required to determine how preprocessing affects the architectural models used in testing, such as ResNet-50 and ResNet-101. Testing of this system was carried out using input data from CLAHE images and Canny Edge Detection images. This system was tested for 40 epochs using early stopping. In this test, a comparison was made against the results of hyperparameters such as the optimizer, learning rate, batch size, and input size. An optimizer is an algorithm used to correct the weight and bias contained in the model learning process, which can reduce the value of the loss function by equalizing the desired output with the predicted output [19]. The optimizers used for comparison are Adam, Adamax, and SGD.

Adam's optimizer is a combination of momentum and RMSProp optimizer types used to calculate the average of both gradients and leverage the previous gradient to accelerate learning [20]. The Adamax optimizer is a development of the Adam optimizer, which has a simpler parameter update to its rules that can maximize the average exponential value of the previous gradient into a more stable value [21]. A stochastic gradient descent (SGD) optimizer is one of the simplest types of gradient descent optimizers used to reduce workload in deep learning models [22]. The learning rate is one of the system testing parameters used to calculate the value of weight correction in model learning. Learning rate values vary greatly, but generally the values used range from 0.1 to 1 [23]. The learning rate values used for the comparison were 0.1, 0.01, and 0.001. The batch size is a parameter of system testing in deep learning that regulates the number of samples during the

optimization process [24]. The batch sizes used for comparison are 32, 64, and 128. The input size is a measure of the dimensions of the image used in the computational process [25]. The input size values used for comparison are 128×128, 224×224, and 256×256.

3. Results and Discussion

Cervical cancer image data from VIA test results was divided into two classes of data, namely positive VIA and negative VIA. In this system test, the results of the CLAHE input image and the Canny Edge Detection input image were compared using two architectures, namely ResNet-50 and ResNet-101. There are four stages of testing the optimizer: learning rate, batch size, and input size on CLAHE images and Canny Edge Detection images. Tests done on these two types of preprocessing show that the Canny Edge Detection image works well because it can learn the pattern of white lesions from VIA inspection. Meanwhile, CLAHE images can only improve the quality of the image. Then, in testing two architectural models, namely ResNet-50 and ResNet-101, on images from CLAHE preprocessing and Canny Edge Detection. The best test results for cervical cancer image classification from VIA test results are using ResNet-50 architecture with Canny Edge Detection preprocessing image types and hyperparameters in the form of SGD optimizers, learning rates of 0.1, batch sizes of 32, and input sizes of 224×224, with test accuracy results of 98.55%, as seen in Table 1.

Table. 1. Test Results

Architecture	Types of Imagery	Optimizer	Learning Rate	Batch Size	Input Size	Accuracy Tests
ResNet-50	CLAHE	SGD	0,1	32	224✕224	95,45 %
	Canny Edge Detection	SGD	0,1	32	224✕224	98,55 %
ResNet-101	CLAHE	SGD	0,1	32	224✕224	92,90 %
	Canny Edge Detection	SGD	0,1	32	224✕224	89,28 %

Using the same hyperparameters, these two architectural models have different results. When using a combination of the ResNet-50 model with Canny Edge Detection preprocessing images, the accuracy results obtained reached 98.55%. While the accuracy results from the combination of the ResNet-101 model with the same preprocessing image data get 89.28% results. It can be seen that the accuracy level of the test between ResNet-50 and ResNet-101 with the same input image has decreased because the Canny Edge Detection preprocessing image is not good when used on ResNet-101. To evaluate the quality and performance of the system that has been created, a confusion matrix is used that can make it easier to analyze the results of accuracy, precision, recall, and F1-Score. The results of the confusion matrix in the best test, where the system can learn quite well in classifying images, can be seen in Fig. 5.
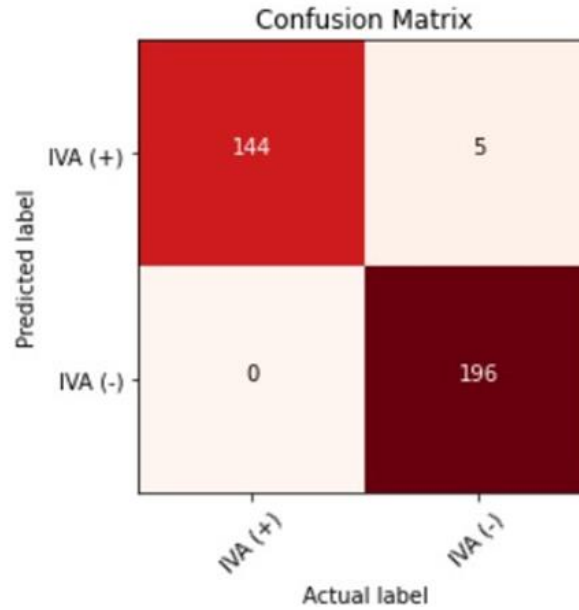
Fig. 5. Best Confusion Matrix Testing

The performance parameters obtained from the best tests in each class have precision, recall, specificity, and F1-Score values with an accuracy rate of 99%. The high accuracy value of this study was obtained to show that the performance of the model used is good and able to classify positive VIA and negative VIA, as can be seen in Table 2 below.

Table. 2. Performance Parameter Results

Class	Precision	Recall	Specificity	F1-Score
VIA (+)	100 %	97 %	97 %	98 %
VIA (-)	98 %	100 %	97 %	99 %
Accuracy				99 %

4. Conclusion

In this study, cervical cancer image classification from VIA examinations with Deep Residual Network architecture was applied. In the preprocessing process, tests are carried out on two types of preprocessing: CLAHE input images and Canny Edge Detection input images. From this preprocessing test, it was found that the input image of Canny Edge Detection results had a high performance value in studying white lesion patterns from VIA test results. Subsequently, a comparison of the accuracy results against two Deep Residual Network architecture models was performed. The architecture models used for comparison were ResNet-50 and ResNet-101, with each model tested on input images from CLAHE and Canny Edge Detection. From the test results, it was found that the best test result was obtained by the ResNet-50 architecture model using the Canny Edge Detection input image. In this test, the hyperparameters used were SGD optimizer, learning rate with a value of 0.1, batch size with a value of 32, and input size with a value of 224 × 224. By combining the ResNet-50 architecture model with Canny Edge Detection input images and hyperparameters, the best results were obtained, with an accuracy reaching 98.55%. Meanwhile, from the test results using the ResNet-101 architecture model, accuracy reached 89.28% using the same input image data and the same hyperparameters.

References

[1]   R. Elakkiya, V. Subramaniyaswamy, V. Vijayakumar, and A. Mahanti, "Cervical Cancer Diagnostics Healthcare System Using Hybrid Object Detection Adversarial Networks," *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 4, pp. 1464–1471, 2022.

[2]   Juwitasari, R. Harini, and A. A. Rosyad, "Husband Support Mediates the Association between Self-Efficacy and Cervical Cancer Screening among Women in the Rural Area of Indonesia," *Asia-Pacific J. Oncol. Nurs.*, vol. 8, no. 5, pp. 560–564, 2021.

[3]   E. S. Aoki *et al.*, "National screening programs for cervical cancer in Asian countries," *J. Gynecol. Oncol.*, vol. 31, no. 3, pp. 1–9, 2020.

[4]   G. Namale *et al.*, "Visual inspection with acetic acid (VIA) positivity among female sex workers: a cross-sectional study highlighting one-year experiences in early detection of pre-cancerous and cancerous cervical lesions in Kampala, Uganda," *Infect. Agent. Cancer*, vol. 16, no. 1, pp. 1–11, 2021.

[5]   D. Endarti, Satibi, S. A. Kristina, M. A. Farida, Y. Rahmawanti, and T. Andriani, "Knowledge, perception, and acceptance of HPV vaccination and screening for cervical cancer among women in Yogyakarta Province, Indonesia," *Asian Pacific J. Cancer Prev.*, vol. 19, no. 4, pp. 1105–1111, 2018.

[6]   K. E. Wijayanti, H. Schütze, and C. MacPhail, "Parents' attitudes, beliefs and uptake of the school-based human papillomavirus (HPV) vaccination program in Jakarta, Indonesia – A quantitative study," *Prev. Med. Reports*, vol. 24, p. 101651, 2021.

[7]   Y. Xiang, W. Sun, C. Pan, M. Yan, Z. Yin, and Y. Liang, "ScienceDirect A novel automation-assisted cervical cancer reading method based on convolutional neural network," *Integr. Med. Res.*, vol. 40, no. 2, pp. 611–623, 2020.

[8]   A. Ajit and A. C. Layer, "A Review of Convolutional," pp. 1–5, 2020.

[9]   A. P. Coding, "ResNet-Like Belief-Propagation Decoding for Polar Codes," vol. 10, no. 5, pp. 2021–2024, 2021.

[10]   T. Alaoui *et al.*, "Classification of chest pneumonia from x-ray images using new architecture based on ResNet," 2021.

[11]   A. Mahajan, "Categorical Image Classification Based On Representational Deep Network (RESNET)," *2019 3rd Int. Conf. Electron. Commun. Aerosp. Technol.*, pp. 327–330, 2019.

[12]   K. M. A. Adweb, N. Cavus, and B. Sekeroglu, "Cervical Cancer Diagnosis Using Very Deep Networks over Different Activation Functions," *IEEE Access*, vol. 9, pp. 46612–46625, 2021.

[13]   A. Çınar, M. Yıldırım, and Y. Eroğlu, "Classification of pneumonia cell images using improved ResNet50 model," *Trait. du Signal*, vol. 38, no. 1, pp. 165–173, 2021.

[14]   Faiz Nashrullah, Suryo Adhi Wibowo, and Gelar Budiman, "The Investigation of Epoch Parameters in ResNet-50 Architecture for Pornographic Classification," *J. Comput. Electron. Telecommun.*, vol. 1, no. 1, pp. 1–8, 2020.

[15]   F. Ri *et al.*, "/ Deho , Pdjh & Odvvlilfdwlrq % Dvhg Rq 7Udqvihu," pp. 354–358, 2020.

[16]   O. Data, "Technical counseling of data processing with spss 25 for some employees of the bkad office of majene district," vol. 9, no. 3, pp. 184–187, 2020.

[17]   B. K. Umri, M. W. Akhyari, and K. Kusrini, "Detection of Covid-19 in Chest X-ray Image using CLAHE and Convolutional Neural Network," pp. 14–18, 2021.

[18]   G. J. H. Hwhfwlrq *et al.*, "7KH 5HVHDUFK RI , PSOHPHQWDWLRQ 0HWKRG RI & DQQ \," pp. 20–23, 2020.

[19]   S. Bera and V. K. Shrivastava, "Analysis of various optimizers on deep convolutional

neural network model in the application of hyperspectral remote sensing image classification," *Int. J. Remote Sens.*, vol. 41, no. 7, pp. 2664–2683, 2020.

[20]   Z. Zhang, "Improved Adam Optimizer for Deep Neural Networks," *2018 IEEE/ACM 26th Int. Symp. Qual. Serv.*, pp. 1–2, 2018.

[21]   B. I. N. Xiao, Y. Liu, and B. Xiao, "Accurate State-of-Charge Estimation Approach for Lithium-Ion Batteries by Gated Recurrent Unit With Ensemble Optimizer," *IEEE Access*, vol. 7, pp. 54192–54202, 2019.

[22]   N. Shirish and K. Richard, "Improving Generalization Performance by Switching from Adam to SGD," no. 1, 2017.

[23]   N. O. Attoh-okine, "Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance," vol. 30, pp. 291–302, 1999.

[24]   S. Edition, *Deep Learning with Python*.

[25]   S. Rajbhandari and J. Rasley, "ZeRO : Memory Optimizations Toward Training Trillion Parameter Models," pp. 1–24.