Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 179 (2021) 423–431

Procedia Computer Science

www.elsevier.com/locate/procedia

5th International Conference on Computer Science and Computational Intelligence 2020

# Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer

Devvi Sarwinda[a,*], Radifa Hilya Paradisa[a], Alhadi Bustamam[a], Pinkie Anggia[b]

[a]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Kampus UI Depok, 16424, Depok, Indonesia*
[b]*Faculty of Computer Science, Universitas Indonesia, Kampus UI Depok, 16424, Depok, Indonesia*

**Abstract**

This paper investigates a deep learning method in image classification for the detection of colorectal cancer with ResNet architecture. The exceptional performance of a deep learning classification incites scholars to implement them in medical images. In this study, we trained ResNet-18 and ResNet-50 on colon glands images. The models trained to distinguish colorectal cancer into benign and malignant. We assessed our prototypes on three varieties of testing data (20%, 25%, and 40% of whole datasets). The empirical outcomes confirm that the application of ResNet-50 provides the most reliable performance for accuracy, sensitivity, and specificity value than ResNet-18 in three kinds of testing data. Upon three test assortments, we perceive the best performance value on 20% and 25% test sets with a classification accuracy of above 80%, the sensitivity of above 87%, and the specificity of above 83%. In this research, a deep learning method demonstrates the profoundly reliable and reproducible outcomes for biomedical image analysis.



*Keywords:* colon glands; deep learning; ResNet; image classification.

## 1. Introduction

The large intestine is a part of the digestive tract that includes colon, cecum, and rectum. In contrast, the colorectal divided into three types, namely ascending colon, transverse colon, and descending colon [1]. Colorectal cancer is a type of cancer that transpires in the colon of the large intestine. It is quite prevalent to befall nowadays. Furthermore, it has

* Corresponding author.
*E-mail address:* devvi@sci.ui.ac.id


10.1016/j.procs.2021.01.025

also confirmed throughout the investigation of colon biopsies collected during colonoscopy or surgery by the histopathologist. Some pathologists detect cancer and analyze the morphology of the biopsy images within the microscope.

Colorectal cancer treated with various methods such as chemotherapy, immunotherapy, radiotherapy, and others. Management of colorectal cancer efficiently is handled if known earlier about the level of development of abnormal cells, a cell that suspected of developing into cancer cells. In medicinal practice, the most conventional approach applied to obtain and define the level of colorectal cancer is by histopathological examination of tissue specimens. The pathologist will visually observe alterations in cell morphology and tissue components under a microscope and ascertain tissue specimens that have region of malignant cancer [2]. Nevertheless, expert pathologists often disagree with one another regarding network classifications, which can lead to the conclusion that solely the performance of expert pathologists for a histopathological appraisal is inadequate [3].

Medical imaging presently serves to distinguish the diseases in various components of the body. Machine learning devices are employed in medical image processing for pattern recognition and resolve medical images [4,5]. Deep learning has been outperformed conventional methods on the large-scale dataset [6]. Researchers invented numerous networks to attain tremendous precision and decrease parameters, such as VGGnet [7] in 2014, GoogLeNet [8,9] in 2015, ResNet [10]in 2016 and DenseNet [11] in 2017. The outstanding performance of deep learning invigorates scholars to implement them in medical images [12,13].

There were comprehensive studies that have been executed with numerous methods and focus on individual detection of colorectal cancer, such as colonic polyp classification [14,15,16,17,18,19], nuclei classification [20,21], colon gland classification [22,23,24], and classification of colon cancer grade [25,26]. In this research, we presented a deep learning approach to identify the colon gland (benign and malignant). We selected ResNet architecture for our proposed technique. ResNets has been confirmed that the residual mapping and shortcut connections lead to better results compared to very deep plain networks and the training is easier as well [11]. Images with a benign and malignant case have a form of pathology consisting of lumen, cytoplasm, nuclei, and stroma.

There are numerous sections to elaborate on this study. Section 2 demonstrates the data and classifications adopted in this investigation, such as pre-processing approach, deep learning technique with ResNet-18 and ResNet-50 model, and evaluation of performance. Section 3 illustrates the empirical results of the proposed method. Subsequently, the conclusions articulated in Section 4.

## 2. Material and Methods

### *2.1. Material*

Colorectal gland images acquired from a Warwick-QU dataset via www.warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/download. The data consist of 165 images with 74 74 benign tumors images and 91 malignant tumors images. This data attained with the Zeiss MIRAX MIDI Scanner by applying an image data weight range of 716 kilobytes, 1.187 kilobytes, and an image data resolution range of 567x430 pixels to 775x522 pixels with each pixel has a distance of 0.6 μm from the actual distance. Figure 1 represents the model of glands colon images with benign and malignant colorectal cancer from the sample Warwick-QU dataset.

### *2.2. Research Method*

The classification technique adopted in this investigation presented in Figure 2. As the figure shows, our detection method extracts feature and divides into two classes of classifiers, such as ResNet-18 and Resnet-50.

### *2.3. Image Preprocessing*

In the preprocessing stage, the authors employed a grayscale image. After the grayscale image formation process is concluded, then gland colon images will enhance image features because this unprocessed image has low contrast.

The image contrast uniformity was completed with Contrast-Limited Adaptive Histogram Equalization (CLAHE). CLAHE is used to generate more solid images, and all areas of the image.

### *2.4. Deep Learning Model using ResNet*

Deep residual network or ResNet is one of the models developed by He et al. in 2016 [11]. This architecture formed to defeat quandaries in deep learning training because deep learning training, in general, takes quite a lot of time and is limited to a certain number of layers. The explication to the intricacy introduced by ResNet is to apply to skip connection or shortcut. The advantage of the ResNets model compared to other architectural models is that the performance of this model does not decrease even though the architecture is getting deeper. Besides, computation calculations are made lighter, and the ability to train networks is better. The ResNet model is implemented by skipping connections on two to three layers containing ReLU and batch normalization among the architectures [11]. He et al. showed that the ResNet model performs better in image classification than other models, indicating that the image features were extracted well by ResNet [11].

He et al. adopts residual learning to be applied to multiple layers of layers. The residual block on ResNet is defined as follows [11]:

$$y = F(x, W + x) \tag{1}$$

where $x$ is input layer; $y$ is output layer; and $F$ function is represented by the residual map.

Residual block on ResNet can be accomplished if the input data dimensions are identical to the output data dimensions. In addition, each ResNet block consists of two layers (for ResNet-18 and ResNet-34 networks) or three layers (for ResNet-50 and ResNet-101 networks). The two initial layers of the ResNet architecture resemble GoogleNet by doing convolution 7 × 7 and max-pooling with size 3 × 3 with stride number 227. In this study, we used ResNet-18 and ResNet-50 models. The authors resize the fundus image into a 224 x 224 grid. The weights of ResNet are initialized using Stochastics Gradient Descent's (SGD) with standard momentum parameters. The structure of the ResNet representation explicated in Figure 3.

### *2.5. Stochastics Gradient Descents with Momentum*

Stochastic Gradient Descent (SGD) is one of the most common algorithms applied to a logical network to optimize. This algorithm renews the weight and bias parameters by diminishing the primary weight value to the gradient value obtained. SGD aims to minimize the loss function [27]. SGD is an algorithm ordinarily used in machine learning and structural networks; and deep learning. Calculation of SGD value is expressed in equation (2) with η as a parameter size that stewarding the learning level of neural networks.

$$w_{ij} = w_{ij} + \Delta w_{ij} \tag{2}$$

where

$$\Delta w_{ij} = -\eta\, \delta D/(\delta w_{ij})$$

Several types of SGD algorithm construction are employed to obtain the optimal neural network performance, one of which is the stochastic gradient descent with momentum. The algorithm can achieve convergence more promptly, so it is conceivable to bypass low local minimums. This renewal method is regarded faster than SGD in the presence of the learning rate parameter, which determines how fast the gradient was lost before. The formula of SGD with momentum can be formulated as follows:

$$\Delta w_{ij}(t+1) = -\eta \delta E/\left(\delta w_{ij}\right) + \alpha \Delta w_{ij}(t) \tag{3}$$

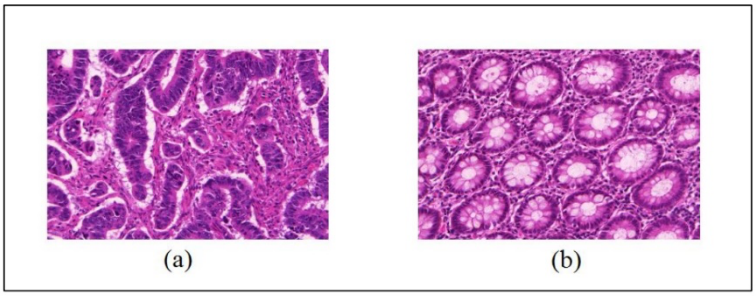


Fig. 1. Colorectal gland (a) Malignant; (b) Benign [21].

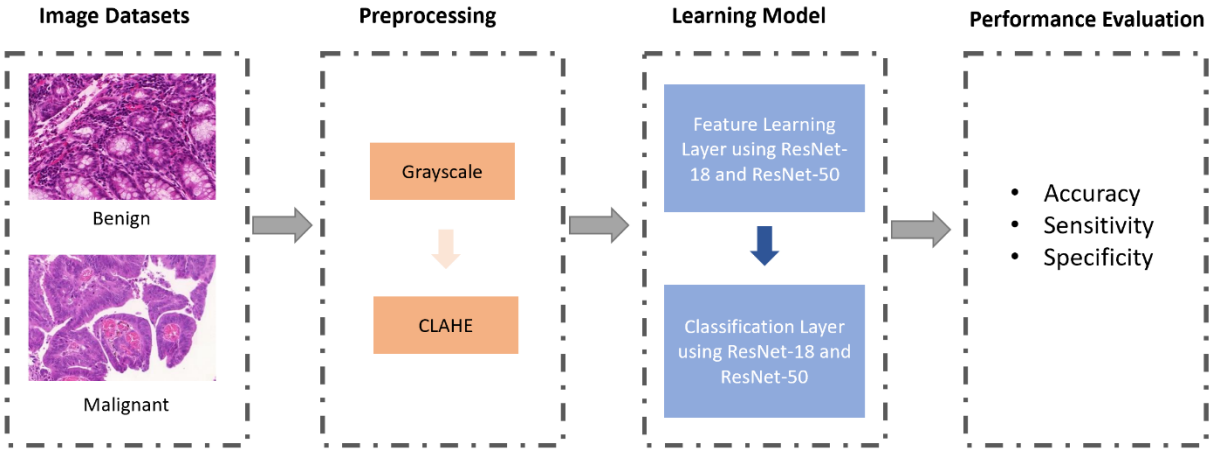


Fig. 2. Research Flow for Proposed Method

| layer name | output size | 18-layer | 34-layer | 50-layer | 101-layer | 152-layer |
|---|---|---|---|---|---|---|
| conv1 | 112×112 | 7×7, 64, stride 2 | | | | |
| conv2_x | 56×56 | 3×3 max pool, stride 2 | | | | |
| | | $\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$ | $\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$ |
| conv3_x | 28×28 | $\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$ | $\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$ | $\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$ | $\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$ | $\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$ |
| conv4_x | 14×14 | $\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$ | $\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$ | $\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$ | $\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$ | $\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$ |
| conv5_x | 7×7 | $\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$ | $\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$ |
| | 1×1 | average pool, 1000-d fc, softmax | | | | |
| FLOPs | | $1.8\times10^9$ | $3.6\times10^9$ | $3.8\times10^9$ | $7.6\times10^9$ | $11.3\times10^9$ |

Fig. 3. ResNet Architecture [27]

### 2.6. Performance Evaluation

The computations applied for the experiment results consist of accuracy, sensitivity, and specificity. The theorems toward accuracy, precision, and recall imply as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

## 3. Results and Discussion

In this study, the experiments were carried out with three models of training and testing data, namely model 1 carried out 60% of training data and 40% of testing data; model 2 is done 75% training and 25% testing data, and; model 3 is carried out 80% training data distribution and 20% testing data. In the initial stages of the experiment, the preprocessing step is done using grayscale images and contrast enhancement with CLAHE. Figure 4 displays the outcomes of preprocessing images from initial input to CLAHE. After going through the preprocessing stage, learning with the ResNet model is applied. The formation of each architecture layer on ResNet has its conditions. The feature learning model is generally done by using the convolution layer. In addition to the convolution layer, there is a batch normalization, activation layer, and pooling layer in the ResNet architecture. Batch normalization is employed following the convolution layer, and then the activation layer uses the ReLu activation function. In each ResNet architecture, there are residual blocks that divide the convolution layer into five layers. Pooling layers are only applied at the beginning of feature learning or after the first convolution and the last stage of feature learning or after the previous convolution before being included in the classification layer. A description of the preparation of ResNet-18 and ResNet-50 architecture layers can be seen in Figures 5 and 6.

Our model uses tuning hyperparameter with SGD with momentum as optimal function, learning rate, and binary cross-entropy as the loss function. Based on the results of the experiments, a measurement of the training results of each model explicated in Table 1 and Table 2. Table 1 confers the performance of the ResNet-18 architecture model with three dataset sharing models. The highest accuracy results obtained from 80% of training data and 20% of testing data equal to 85%, while the highest sensitivity value of 96% is obtained from 75% of training data and 25% of testing data.

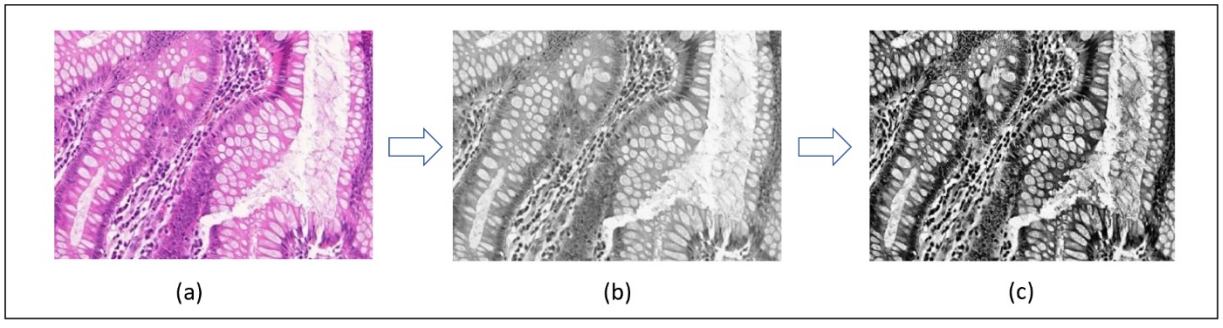


Fig. 4. The pre-processing process: a) input dataset; b) grayscale image; (c) and uniformity of contrast performed with CLAHE

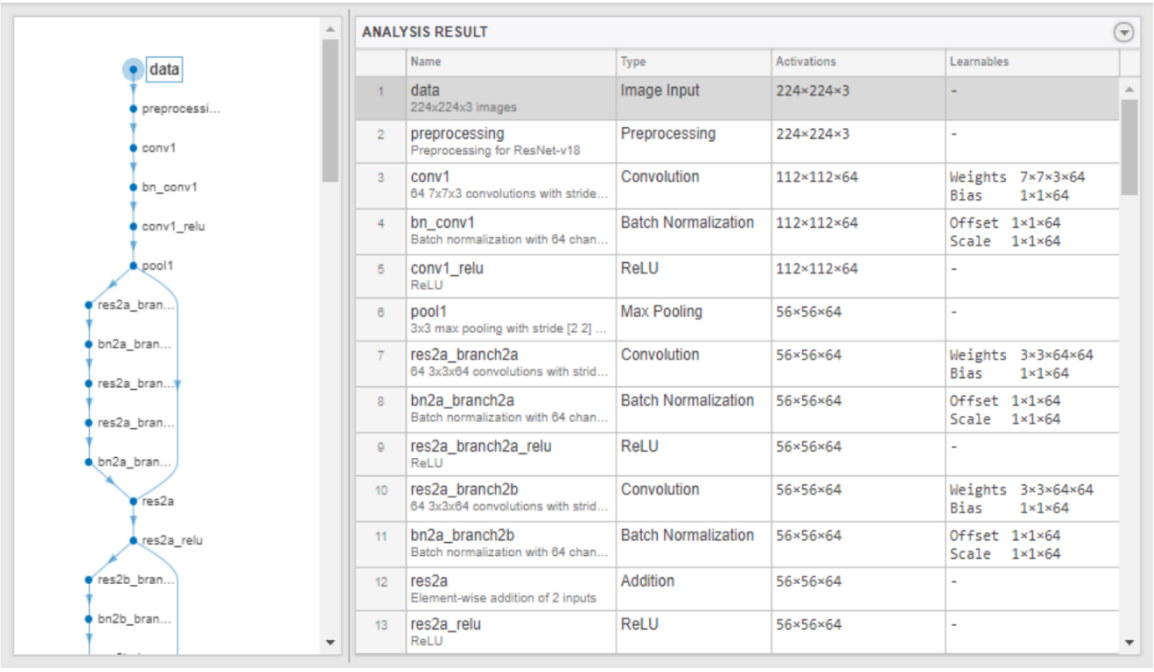


Fig. 5. Design of ResNet-18 architecture

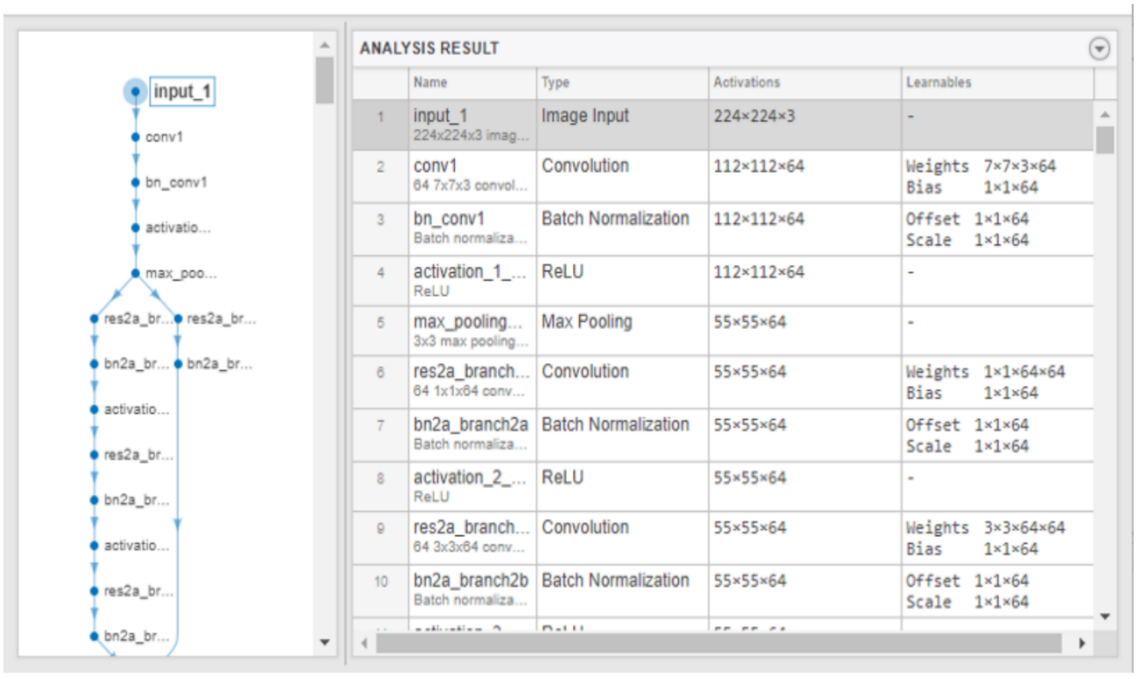


Fig. 6. Design of ResNet-50 architecture

Table 1. Comparison of Accuracy, Sensitivity, Specificity Value in Several Testing Datasets for ResNet-18

| Training Data: Testing Data | *Accuracy* | *Sensitivity* | *Specificity* |
|---|---|---|---|
| 60%:40% | 73% | 64% | 83% |
| 75%:25% | 81% | 96%/ | 63% |
| 80%:20% | 85% | 83% | 87% |

Table 2. Comparison of Accuracy, Sensitivity, Specificity Value in Several Testing Datasets for ResNet-50

| Training Data: Testing Data | *Accuracy* | *Sensitivity* | *Specificity* |
|---|---|---|---|
| 60%:40% | 77% | 60% | 92% |
| 75%:25% | 88% | 89 | 87% |
| 80%:20% | 88% | 93% | 83% |

Table 3. Comparison of Running Time between ResNet-18 and ResNet-50 for each epoch

| Training Data: Testing Data | Running Time per epoch for ResNet-18 (second) | Running Time per epoch for ResNet-50 (second) |
|---|---|---|
| 60%:40% | 0.46 | 2.39 |
| 75%:25% | 0.53 | 2.89 |
| 80%:20% | 0.55 | 3.09 |

Table 2 confers the effectiveness of the ResNet-50 architecture model, where the highest accuracy value is obtained from 75% and 80% of training data with the same value equal to 88%. While the highest sensitivity value is obtained from sharing data with 60% training data and 40% testing data, which is 92%. Based on Table 1 and Table 2, it can be seen that ResNet-50 has a higher level of accuracy than ResNet-18. The sensitivity value on the same data sets shows ResNet-18 has a minimum sensitivity value of 64% and Resnet-50 has a minimum sensitivity value of 83%. The outcomes of these estimations signify that the more stack convolution layers in the ResNet architecture, the better a model is in learning features.

In this analysis, the extraction and classification sequences for both procedures estimated. The data processing stage was performed on Matlab tools version R2018a on an Intel Core i7 2.5 GHz with 16 GB of RAM. Table 3 shows that ResNet-18 is faster the ResNet-50 architecture. It occurred because the quantity of layers exhibited in the architecture impacted the length of training and testing times where ResNet-50 has more layers than ResNet-18.

## 4. Conclusion

Based on the results of the implementation of the Deep Residual Network (ResNet) method to identify colorectal cancer, the authors generated several outcomes. There are two ResNet architectural models, particularly ResNet-18 and ResNet-50, with three dataset distribution models designed to test accuracy, sensitivity, and specificity values. ResNet-50 has higher accuracy than ResNet-18. According to our analysis, the ResNet variants can be detected colorectal cancer with accuracy between 73% -88% and sensitivity value between 64%-96%. The best classification model that is built can identify colorectal cancer. In the forthcoming works, we will purpose to incorporate added architectures with their variants using the corresponding datasets.

**Acknowledgements**

The authors greatly appreciated the PUTI Proceedings 2020 Grant from Directorate of Research and Human Engagement Universitas Indonesia with a contract number of NKB-955/UN2.RST/HKP.05.00/2020 that has been supporting and furthering our research.

**References**

1. Campbell NA,RJB,ULA,CML,WSA,MPV,&JRB. Biologi (Damaring Tyas Wulandari, Penerjemah). In ; 2010; Jakarta: Penerbit Erlangga.
2. Iftikhar MA,HM,&AH. A colon cancer grade prediction model using texture and statistical features, SMOTE and mRMR. In in 2016 19th International Multi-Topic Conference (INMIC); 2016. p. 1-7.
3. Aeffner F,WK,MNT,BJC,HCL,BBN,RDG,GR,KSR,KJ,&YGD. The gold standard paradox in digital image analysis: manual versus automated scoring as ground truth. In Archives of Pathology & Laboratory Medicine; 2017. p. 1267-1275.
4. M. B. Patwari RRMaYMR. Detection and counting the microaneurysms using image processing techniques. In International Journal of Applied Information Systems; 2013. p. 11-17.
5. J. Lachure AVDSLSGRJ. Diabetic Retinopathy using morphological operation and machine learning. In in International Advance Computing Conference; 2015. p. 617-622.
6. Deng J,DW,SR,LLJ,LK,FFL. ImageNet: A large-scale hierarchical image. In in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2009. p. 248–255.
7. Krizhevsky A,SI,HGE. Imagenet classification with deep convolutional neural. In in Proceedings of the Neural Information Processing Systems Conference; 2012. p. 1097-1105.
8. Simonyan K,ZA. Very Deep Convolutional Networks for Large-Scale Image Recognition. In in Proceedings of the International Conference on Learning Representations; 2015. p. 730-734.
9. Szegedy C,LW,JY,SP,RS,AD,ED,VV,R. A Going deeper with convolutions. In in Proceedings of the IEEE Conference on Computer Vision and Pattern; 2016. p. 1-9.
10. Szegedy C,VV,IS,SJ,WZ. Rethinking the inception architecture for computer vision. In in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 2818–2826.
11. He K,ZX,RS,SJ. Deep residual learning for image recognition. In in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 770– 778.
12. Huang G,LZ,VDML,WKQ. Densely Connected Convolutional Networks. In in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017. p. 2261-2269.
13. Esteva A,KB,NRA,KJ,SSM,BHM,TS. Dermatologist-level classification of skin cancer with deep neural networks. In Nature; 2017. p. 115-118.
14. Tajbakhsh NSRGaJL. Automatic polyp detection in colonoscopy videos using an ensemble of convolutional neural networks. In in 12th International Symposium on Biomedical Imaging (ISBI); 2015. p. 79-83.
15. Ribeiro EAUaMH. Colonic polyp classification with convolutional neural networks. In in 29th International Symposiu on Computer-Based Medical Systems (CBMS); 2016. p. 253-258.
16. Komeda YHHTWTNMKTSAOea. Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience. In Oncology; 2017. p. 30-34.
17. Korbar BAMOAPMCMNMASLTAASaSH. Deep learning for classification of colorectal polyps on whole-slide images. Journal of Pathology Informatics. 2017; VIII.
18. Zhang RYZTWCMRYSHWJYLaCCP. Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. In IEEE J. Biomedical and Health Informatics; 2017. p. 41-47.
19. Chen PJMCLMJLJCLHHSLaVST. Accurate classification of diminutive colorectal polyps using computer-aided analysis. Gastroenterology. 2018; CLIV(3): p. 568-575.
20. Bayramoglu NaJH. Transfer learning for cell nuclei classification in histopathology images. In in European Conference on Computer Vision: Springer.
21. Sirinukunwattana KSEARYWTDRSIACaNMR. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. In IEEE transactions on medical imaging; 2016. p. 1196-1206.
22. Kainz PMPaMU. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. In ; 2017.

23. Xu YZJLBWYAFZMLIEaCC. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. BMC Bioinformatics. 2017 281-290; LXXIX(1).
24. Bychkov DNLRTSNPEKCVMWMLCHaJL. Deep learning based tissue analysis predicts outcome in colorectal cancer. Scientific reports. 2018 3395-3402; VIII(1).
25. Tomczak JM,MIMWMJHGCMLKdLea. Tomczak, Jakub M., Maximilian Ilse, Max Welling, Marnix Jansen, Helen G. Coleman, Marit Lucas, Kikki de Laat et al. 2018.
26. Manju Dabass RVaSV. Five-Grade Cancer Classification of Colon Histology Images via Deep Learning. In in Proceedings of the 2nd International Conference on Commuincation and Computing System; 2018.
27. A. G. Prostate Cancer Classification using Convolutional Neural Networks. Master Theses in Mathematical Sciences. 2016.