Development of Mobile Skin Cancer Detection using Faster R-CNN and MobileNet v2 Model

Cahyo Adhi Hartanto
Department of Informatics
Faculty of Sciences and Mathematics,
Diponegoro University,
Semarang, Indonesia

Adi Wibowo
Department of Informatics
Faculty of Sciences and Mathematics,
Diponegoro University,
Semarang, Indonesia
Coresponding author: bowo.adi@live.undip.ac.id

***Abstract*— The development of cameras in smartphones is possible as a point of care for early detection of cancer. Early detection using smartphones is carried out by giving the smartphone the ability to recognize objects with skin cancer characteristics. The convolution neural network (CNN) is often used in disease detection and classification. However, the CNN method requires high computing capability and a large memory that is difficult to perform on smartphones. In this paper, the MobileNet v2 and Faster R-CNN methods are utilized and executed on an Android-based application that can detect skin cancer. Both proposed architectures were trained to recognize actinic keratosis and melanoma skin cancer targets on images. The dataset used was 600 images, divided into two classes, actinic keratosis images and melanoma images with no attention for gender, age, or additional factors. In this study, an android app was developed to utilize the smartphone camera for skin cancer detection. The Faster R-CNN and MobileNet v2 models were implemented as an intelligent system for the screening. Two testing methods were performed in this study, the Jupyter notebook and android camera. Based on the experiment result, Faster R-CNN obtained higher accuracy when testing using the Jupyter, and MobileNet v2 got the same high accuracy when applying on a smartphone.**

Keywords—Skin Cancer, Melanoma, Actinic keratosis, Faster R-CNN, MobileNet v2.

I. Introduction

Skin cancer is highly malignant and can prompt death. There were more than 60 thousand cases of skin cancer in America and caused more than eight thousand people deaths [1]. To decrease the mortality rate, an effort can be done by conducting early detection. Early detection can be done to find out the presence of cancer earlier before it develops into an advanced stage. One of the latest technologies that can be used to carry out cancer early detection is by using a smartphone camera [2].

Automatically detecting skin cancer is an essential function in the development of early detection technology. The rapid development of technology nowadays makes it possible to do early detection only through a smartphone. Various features available on smartphones are currently very supportive of early detection [2]. These features include the camera on a smartphone that not only able to capture images, but also able to process images and detect objects in the captured image. Another reason for choosing a smartphone as an early detection device is the connection, the picture-taking and the portability [3]. The detection of objects in imagery is currently growing more rapidly. Technology that is increasingly developing at this time is deep learning [4].

Deep learning is a new technology that comes as a development of machine learning. Deep learning has a higher level of sensitivity than other machine learning methods [5]. Deep learning is able to produce good results in image classification, object detection, and natural language processing. The success rate of deep learning depends on the resource used for the training process, which in this case is the GPU, and the number of datasets used [6].

Research related to skin cancer detection was carried out by Bourouis [2] using the Artificial Neural Network algorithm which was developed using Netbeans Java and produced an accuracy value of 96.50%. Yuan, et al [7] tried to overcome the challenges of automated lesion segmentation on dermoscopic images using Deep Fully Convolutional Networks with Jaccard distance. These challenges included the strong imbalance between background and foreground pixels, fuzzy and irregular borders, low contrast, and other noises. Esteva [8] developed deep learning technology by training CNN to detect two types of skin cancer, keratinocyte carcinomas and benign seborrheic. The dataset was 129,450 clinical images. In its testing, the CNN that being developed was able to classify skin cancer with a level of ability comparable to a dermatologist. In 2017, Burdick [9] conducted a segmentation to classify skin cancer using VGGNet. The dataset used was 1279 with details of 900 images included in the training data and 379 included in the testing data. Each picture was labeled benign or malignant. The accuracy obtained was 58.7% for perfect segmentation, 60.7% for partial segmentation and 51.3% for unsegmented. In the same year, Chang [10] developed the network using Inception v3 using two thousand dataset. Jin Qi [11] used 16 layer net VGG which consisted of 13 convolution layers, 3 fully connected layers, 15 RELU layers, 5 pooling layers, 2 dropout layers and a softmax layer to detect melanoma. The dataset used were 150 validation

978-1-7281-7226-2/20/$31.00 ©2020 IEEE

data, 600 test data, and 2000 training data. Xulei [12] also conducted a research related to skin cancer using CNN to detect the types of skin cancer, squamous cell carcinoma, benign seborrheic keratosis, malignant melanoma, nevi benign with a dataset of 130,000. In 2018, Haenssle [13] used Inception v4 CNN to detect melanoma. The results of the CNN model training were compared with dermatologists. The sensitivity result obtained during comparison with level 1 dermatologists was 86.6% and specificity was 71.3%, the sensitivity result obtained when compared with level 2 dermatologists was 88.9% and specificity was 75.7%. The difference in level 1 and level 2 was that in level 2, a close up image was provided. There are several Deep Learning models available including Faster R-CNN, R-CNN Mask, MobileNet and others. However, the deep learning model that can be used and works well on smartphones is MobileNet.

Another challenge of skin cancer image detection is that mostly in an image, the ratio between skin cancer and normal skin areas is an imbalance. The size of normal skin usually larger than the cancer area. Therefore, In this paper, paramters tuning were performed to obtain the best parameters for the MobileNet v2 and Faster-R-CNN models for skin cancer detection on the mobile application. The android app is developed to detect skin cancer types of actinic keratosis and melanoma.

II. Materials and Methods

A. Skin Cancer

In this paper, two skin cancer types are use, Actinic Keratosis, and Melanoma. The first skin cancer type is triggered by ultraviolet(UV) radiation. The color properties of the Actinic Keratosis are pink, a brownish color or a compound of these colors [14]. The Actinic Keratosis originating is shown in Fig. 1. The second type of skin cancer is the most malignant skin cancer. Melanoma's case is identified to begin from two factors, i.e., UV light exposure and hereditary factors [15]. The color properties of melanoma skin cancer consist of more than one color [16]. Fig. 2 shows an example of melanoma skin cancer.

B. MobileNet

MobileNet model is developed for efficiency and is designed for possibly running embedded devices or mobile devices. The depthwise separable convolution architecture is the primary layer of this model to reduce the feature number. MobileNet v2 was released in April 2017. The updates from the previous version are the bottleneck layers and shortcut connections [17]. The bottleneck layer is shown in Fig. 3.

C. Faster R-CNN

Faster R-CNN is a deep learning model developed from Fast-R-CNN. In Faster R-CNN, there are additional Region Network Proposals as shown in Fig. 4. In the overall architecture, Faster R-CNN has a detector derived from Fast R-CNN and the addition of region proposers [18].

Faster R-CNN offers faster computing than its predecessors and other architectures such as YOLO and SSD [19]. The three core parts of the Faster R-CNN are convolutional layers for the feature extraction process, RPN as an object detector, and two fully connected layers that function for object classification and produce object position coordinates [20]. RPN accepts input in the form of features that have been extracted in the previous convolutional layer. The RPN determines the position of the object using sliding windows measuring n x n. RPN generates various bounding box sizes to recognize objects, even though the objects have different sizes [21]. According to the predetermined number, the RPN output will be processed by ROI Pooling to reduce the number of bounding boxes. The ROI Pooling layer output results will be forwarded to two fully connected layers, which aim to determine the object class using the softmax activation function and the object's coordinates using the linear activation function.

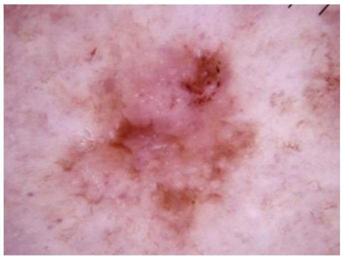


Fig. 1. Actinic keratosis skin cancer

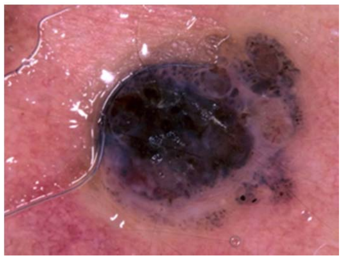


Fig. 2. Melanoma skin cancer

D. Data Collection

In this study, the dataset as many as 600 images were used, which divided into 2 classes, i.e. 300 images of actinic keratosis and 300 images of melanoma with no regard for gender, age, or other factors. The data is divided into 70% training data and 30% testing data. Training data is separated into 90% for training data and 10% for validation data. Also, there are 40 additional images used for testing using an Android-based smartphone camera. The size of the image that we used was still the same when downloading and there was no change in image size or resolution. The images were obtained from the website isic-archive. Region of Interest (ROI) for each image was determined to be section of the cancer. Every image that has been prepared by ROI would generate an XML file containing the coordinates of the target suspected of being cancer. Examples of images with ROI determination to the dataset are shown in Fig. 5.

E. MobileNet v2 and Faster R-CNN Training

The Deep Learning model that we used in this training was MobileNet v2 and Faster R-CNN inception. The training process was run on Windows 10 with hardware specifications of Nvidia GTX 1070 TI, 8 GB RAM, and i5 Processor. The parameters we used in this study were the learning rate with the values of 0.001, 0.0005, 0.001 and 0.005 on the MobileNet v2 and Faster R-CNN models and step 30,000.

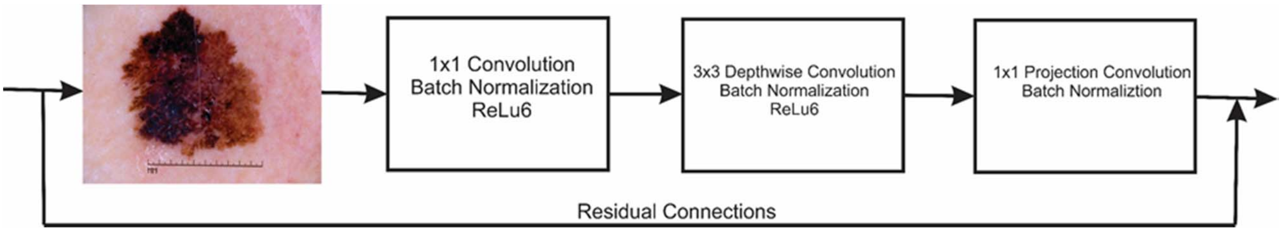


Fig. 3. Bottleneck layer of MobileNet v2

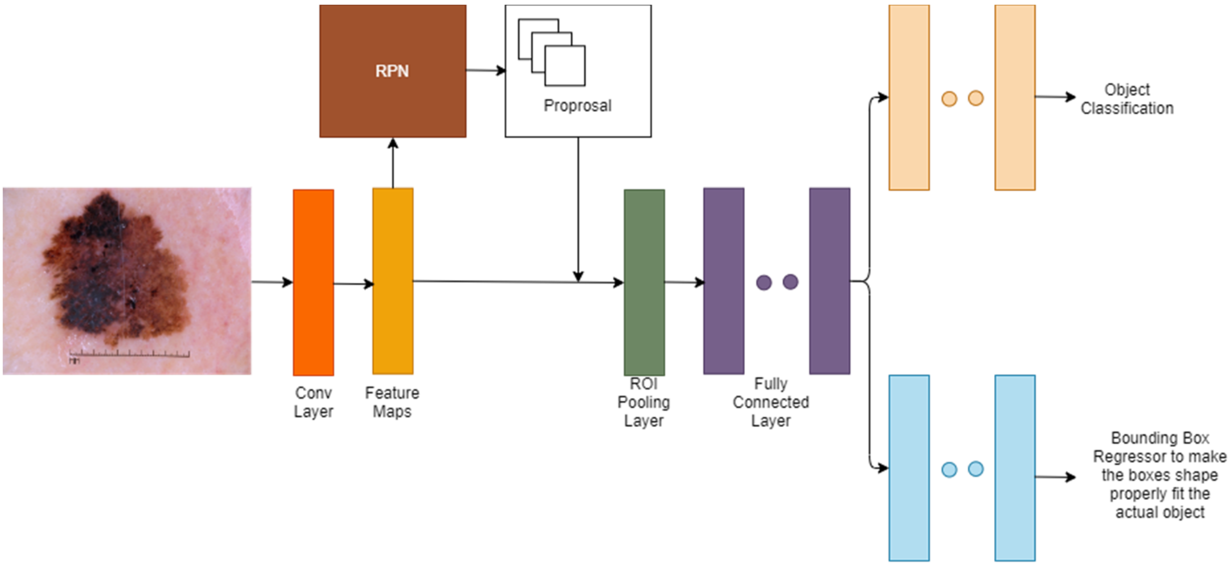


Fig. 4. Faster R-CNN architecture model

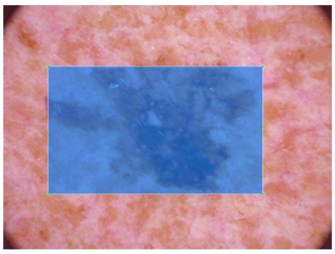


Fig. 5. ROI determination using LabelImg

III. Experiments and Results

The training process required different execution times for each model and used pre-trained model on MobileNet v2 and Faster R-CNN. In MobileNet v2 model, it took around 8 hours for the training process while only around 1.5 hours for Faster R-CNN model. At each learning rate, the decrease in the total loss was also different. Testing for each learning rate is carried out in two ways, testing with the original image that is loaded, and the original image is displayed on the monitor screen and then taken using an Android-based smartphone camera. We used our previous mobile app [22] to compare the perfomance of MobileNet v2 and Faster R-CNN on android devices for detection. The android app was tested using Realme 3 smartphone with 3 gigabyte RAM and 13-megapixel camera. The first test was carried out using 30% of the total dataset, while the second test was carried out using 40 images that were not in the dataset. Each of the data used in the two tests has a balanced comparison. For each testing conducted at each learning rate, we used the Jupyter notebooks by using testing images as many as 40 images, i.e., 20 melanoma class images and 20 actinic keratosis class images. The following is a discussion on the training and testing process for each model.

A. *MobileNet v2*

In this experiment, we used the MobileNet v2 model to detect melanoma and actinic keratosis skin cancer objects. In Fig. 6, the training process are presented. The accuracy values produced in the training process would be different in the testing process. The testing process was carried out using the Jupyter notebook and smartphone. Table 1 shows the results of the testing at each learning rate. The testing process was still using 30,000 steps.

TABLE I. Comparison of MobileNet v2 Accuracy Values based on Learning Rate in Testing Process Using Real Image and Additional Image for Smartphone Testing

Step	Learning Rate	Accuracy	Accuracy on Smartphone
30,000	0.005	86.1%	78.4%
30,000	0.001	85.5%	84.3%
30,000	0.0005	85.0%	82.4%
30,000	0.0001	85.0%	86.3%

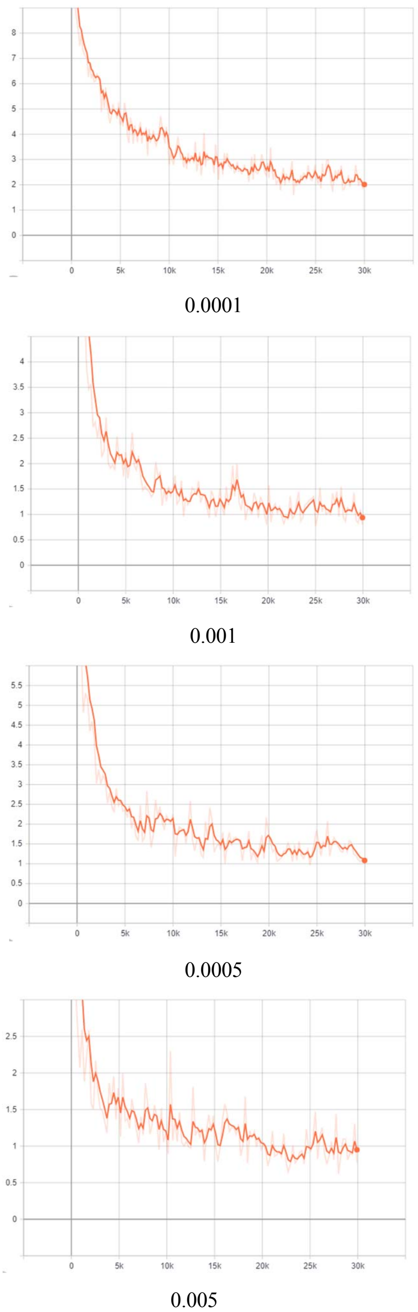


Fig. 6. Graph of the MobileNet-v2 Model Training Process

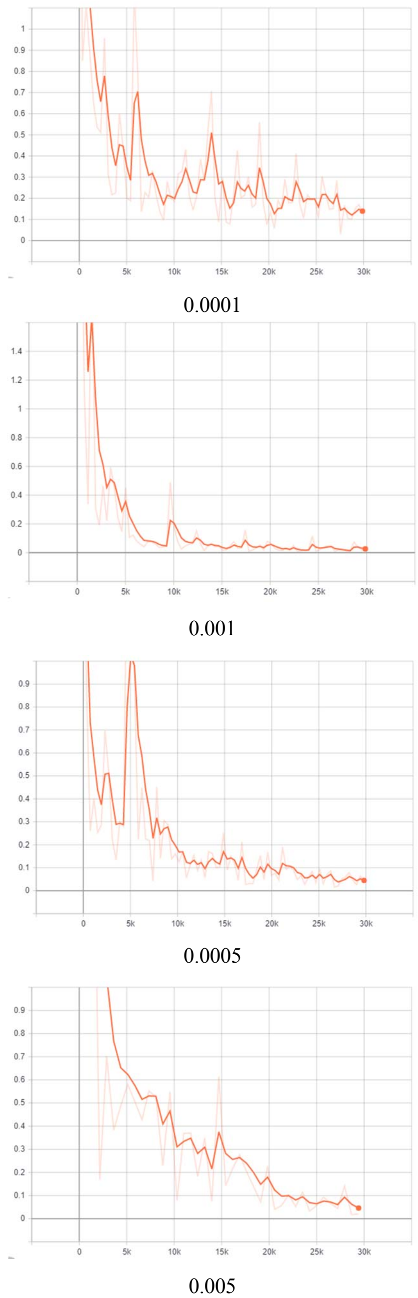


Fig. 7. Graph of the Faster R-CNN Model Training Process

B. Faster R-CNN

The Deep Learning model that we used here was the training result from the use of the Faster R-CNN model to detect melanoma and actinic keratosis skin cancer objects. In Figure 7, training process are presented. The testing process was carried out using the Jupyter notebook and using smartphone. Table 2 shows the results of the testing at each learning rate. The testing process was still using 30,000 steps.

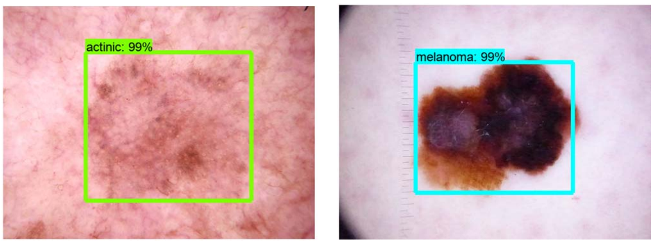
TABLE II. COMPARISON OF FASTER R-CNN ACCURACY VALUES BASED ON LEARNING RATE IN THE TESTING PROCESS USING REAL IMAGE AND ADDITIONAL IMAGE FOR SMARTPHONE TESTING

Step	**Learning Rate**	**Accuracy on The Jupyter Notebook**	**Accuracy on Smartphone**
30,000	0.005	85%	86.3%
30,000	0.001	86.6%	78.4%
30,000	0.0005	87.2%	76.5%
30,000	0.0001	83.8%	84.3%

C. *MobileNet v2 vs Faster R-CNN*

The Faster R-CNN and MobileNet v2 models produced different levels of accuracy. The difference between the two models not only laid in the accuracy but also in the length of the training process. The length of training time for each learning rate in the MobileNet v2 model was 8 hours with a time of around 0.9 sec/step and the training time in the Faster R-CNN model was 1.5 hours with a time of about 0.2 sec/step.

The loss values produced on each model were also different. In the MobileNet v2 model, the loss value generated in the training process was still quite high, reaching 2% while the Faster R-CNN model produced a very small loss value of only 0.3%. The testing and training processes yield different accuracy. This difference due to the model detects an entirely new image, then the possibility of detection errors can occur. In addition, the accuracy when testing using a smartphone camera is also diverse compare to the jupyter testing when using the original image loaded on a computer—this problem due to several factors that affect image acquisition. The aspect of light, camera capability, and the distance between the camera and the object are required to normalize.



(a) (b)

Fig. 8. Testing on New Images. a) Actinic Keratosis, b) Melanoma

In the testing process, the results of the accuracy of the Faster R-CNN model were higher. At the same learning rate, the results of accuracy had significant difference. Based on the experiment results, MobileNet v2 produces the highest accuracy of 86.1% at the learning rate of 0.005 when the original image is loaded into the Jupyter. While testing using a smartphone, the highest learning rate parameter was 0.0001 and obtaining an accuracy of 86.3%. The best learning rate for the Faster R-CNN model is 0.0005, reaching an accuracy of 87.2% in testing on the Jupyter, while the testing process using a smartphone at a learning rate of 0.005 and obtains an accuracy of 86.3%.

IV. Conclusion

The study was conducted to perform mobile skin cancer detection using Faster R-CNN and MobileNet v2 Model. Mobile device was used to get the benefits of the smartphone's camera and its technology. The result showed that the Faster R-CNN model produced a better performance than MobileNet v2 on the training and the Jupter notebook test with the highest test accuracy values of 87.2%. The MobileNet v2 obtained same high accuracy with Faster R-CNN when using android app for skin cancer detection with 86.3% accuracy. This high accuracy proved that the both model could overcome the challenge related to the significant difference between the number of normal skin pixels and cancer skin pixels on the sample. The future work, multi classes of skin disease detection can be developed and blended with segmentation method to improve accuracy.

Acknowledgment

This work was supported by the RISTEKDIKTI, The Republic of Indonesia for PENELITIAN DASAR Research Grant 2020 under Grant No. 257-22/UN7.6.1/PP/2020.

References

[1] M. C. F. Simões, J. J. S. Sousa, and A. A. C. C. Pais, "Skin cancer and new treatment perspectives: A review," *Cancer Lett.*, vol. 357, no. 1, pp. 8–42, 2015.

[2] A. Bourouis, A. Zerdazi, M. Feham, and A. Bouchachia, "M-health: Skin disease analysis system using smartphone's camera," *Procedia Comput. Sci.*, vol. 19, pp. 1116–1120, 2013.

[3] W.-K. Tam and H.-J. Lee, "Accurate shade image matching by using a smartphone camera," *J. Prosthodont. Res.*, vol. 61, no. 2, pp. 168–176, 2017.

[4] E. Chao, C. K. Meenan, and L. K. Ferris, "Smartphone-Based Applications for Skin Monitoring and Melanoma Detection," *Dermatol. Clin.*, vol. 35, no. 4, pp. 551–557, 2017.

[5] G. Wang and F. P. Tomasella, "Ion-pairing HPLC methods to determine EDTA and DTPA in small molecule and biological pharmaceutical formulations," *J. Pharm. Anal.*, vol. 6, no. 3, pp. 150–156, 2016.

[6] A. R. Pathak, M. Pandey, and S. Rautaray, "Application of Deep Learning for Object Detection,‖ Procedia Comput." Sci, 2018.

[7] Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE Trans. Med. Imaging*, vol. 36, no. 9, pp. 1876–1886, 2017.

[8] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[9] J. Burdick, O. Marques, A. Romero-Lopez, X. Giró Nieto, and J. Weinthal, "The impact of segmentation on the accuracy and sensitivity of a melanoma classifier based on skin lesion images," in *SIIM 2017 scientific program: Pittsburgh, PA, June 1-June 3, 2017, David L. Lawrence Convention Center*, 2017, pp. 1–6.

[10] H. Chang, "Skin cancer reorganization and classification with deep neural network," 2017.

[11] J. Qi, M. Le, C. Li, and P. Zhou, "Global and Local Information Based Deep Network for Skin Lesion Segmentation," 2017.

[12] X. Yang, Z. Zeng, S. Y. Yeo, C. Tan, H. L. Tey, and Y. Su, "A novel multi-task deep learning model for skin lesion segmentation and classification," *arXiv Prepr. arXiv1703.01025*, 2017.

[13] 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.

[14] W. Jerjes, Z. Hamdoon, A. A. Abdulkareem, and C. Hopper, "Photodynamic therapy in the management of actinic keratosis: Retrospective evaluation of outcome," *Photodiagnosis Photodyn. Ther.*, vol. 17, pp. 200–204, 2017.

[15] M. Rastrelli, S. Tropea, C. R. Rossi, and M. Alaibac, "Melanoma : Epidemiology , Risk Factors , Pathogenesis , Diagnosis and Classification," *In Vivo (Brooklyn).*, vol. 28, pp. 1005–1012, 2014.

[16] N. Z. Tajeddin and B. M. Asl, "Melanoma recognition in dermoscopy images using lesion's peripheral region information,"

Comput. Methods Programs Biomed., vol. 163, pp. 143–153, 2018.

[17] T. S. Cohen, M. Geiger, J. Köhler, and M. Welling, "Spherical cnns," *arXiv Prepr. arXiv1801.10130*, 2018.

[18] and J. S. Shaoqing Ren, Kaiming He, Ross Girshick, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017.

[19] Y. Kawazoe, K. S. Id, and R. Yamaguchi, "Faster R-CNN-Based Glomerular Detection in Multistained Human Whole Slide Images."

[20] Y. Chen, "Domain Adaptive Faster R-CNN for Object Detection in the Wild."

[21] H. Nguyen, "Improving Faster R-CNN Framework for Fast Vehicle Detection," vol. 2019, 2019.

[22] Wibowo, Adi; Hartanto, Cahyo Adhi; Wirawan, Panji Wisnu. Android skin cancer detection and classification based on MobileNet v2 model. International Journal of Advances in Intelligent Informatics, , 6.2: p.135-148. 2020.