Journal of Theoretical and Applied Information Technology
15th August 2024. Vol.102. No. 15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SKIN CANCER DIAGNOSIS SYSTEM ON OBJECT DETECTION USING VARIOUS CNN YOLOV5 IN ANDROID MOBILE

EVI FATIMATUR RUSYDIYAH[1], DIAN CANDRA RINI NOVITASARI[2], FARIS MUSHLIHUL AMIN[3], ANISA NUR AZIZAH[4], HUDA FEBRIANTO NURROHMAN[5], DINA ZATUSIVA HAQ[6], ELEN RISWANA SAFILA PUTRI[7]

[1] UIN Sunan Ampel Surabaya, Jl. Ahmad Yani No.117, Surabaya, 60237, Indonesia
[2,5,7]Mathematics Department, Faculty of Science and Technology UIN Sunan Ampel Surabaya, Jl. Ahmad Yani No.117, Surabaya, 60237, Indonesia
[3]Information System Department, Faculty of Science and Technology UIN Sunan Ampel Surabaya, Jl. Ahmad Yani No.117, Surabaya, 60237, Indonesia
[4,6]Department of Informatics, Faculty of Intelligent Electrical and Informations Technology, Institut Teknologi Sepuluh Nopember, Jl. Teknik Kimia, Sukolilo, Surabaya, 60111, Indonesia

E-mail: evifatimatur@uinsby.ac.id

ABSTRACT

Skin cancer is a type of cancer that is dangerous and causes death. In Indonesia, skin cancer is one of the most common health problems, with the number of sufferers ranging from 5.9% to 7.8% per year, making it one of the most common cancers besides cervical cancer and breast cancer. This research aims to create a skin cancer diagnosis system based on object detection using deep learning artificial intelligence based on Android mobile. The method used in this research is CNN YOLOv5, which focuses on real-time object detection. In this study, manual annotation was carried out according to 8 classes of skin lesion objects, of which 3 were skin cancers. Several experiments were carried out to obtain the best skin cancer detection system. In the trial, YOLOv5m had a small average evaluation value compared to YOLOv5l and YOLOv5x. However, the computing system performance on YOLOv5m is much lighter and faster than on YOLOv5l and YOLOv5x. The YOLOv5m model produces sensitivity values >80% in almost all classification classes. Therefore, YOLOv5m is the optimal model for an accurate and real-time skin lesion detection system, especially for skin cancer. The implementation on Android phones aims to make it easier for users to detect early and participate in controlling the growth of skin lesions.

eywords: *Skin Cancer, Diagnosis, Object Detection, YOLOv5, Android Mobile.*

1. INTRODUCTION

This Skin cancer is caused by the growth of Deoxyribose Nucleic Acid (DNA) which is damaged in skin cells and cannot be repaired, resulting in genetic damage or mutations in human skin tissue [1]. Skin cancer is one of the most dangerous types of cancer and occurs in many communities. In Indonesia, skin cancer is a disease that often suffers in addition to cervical cancer and breast cancer, with the number of sufferers from 5.9% to 7.8% for all types of skin cancer per year [2]. Various factors can cause skin cancer, such as UV exposure, allergies, infections, smoking, alcohol use, physical activity, and environmental changes [3].

There are two main types of skin cancer, namely melanoma and non-melanoma skin cancer [4]. Non-melanoma skin cancer is a type of skin cancer that grows quite slowly, but if not treated immediately, cancer cells will grow and increase the severity of the cancer. Melanoma skin cancer is a dangerous type of skin cancer and causes many deaths. This disease easily metastasizes early and spreads to other parts of the body [5]. Melanoma most often appears on parts of the body that are exposed to UV rays, such as the hands, face, neck, and lips. Melanoma cancer is a type of cancer that requires special attention, so it is necessary to do early detection to reduce the risk. Early detection of skin cancer is usually carried out by medical professionals using several criteria such as, irregular border asymmetry, uneven, color,

diameter, and evolving. The most common method of diagnosing skin cancer is by visual examinations by dermatologists, which have an accuracy of about 60% [6]. The detection technique based on these 5 criteria is called the "ABCDE" rule [7]. The "ABCDE" rule is a simple way to know the symptoms of cancer by paying attention to the shape, color, diameter, and changes of cancer cells after some time of infection. The examination level of accuracy of skin cancer is highly dependent on medical experts or doctors, the cost is quite high, and the examination time is quite long. Therefore, we need a system to help someone identify skin disorders experienced quickly and accurately. Doctors can use Computer-Aided Detection (CAD) as a clinical tool to minimize human error [8].

In simplification of clinical practice in the health sector, several cases utilize the CAD system, such as mammography detection of breast lesions, colonography, CT polyp detection, and skin cancer detection [9]. CAD systems can help speed up the disease diagnosis process and are used as decision support by medical personnel. Until now, CAD has made a lot of use of Artificial Intelligence (AI) such as machine learning and deep learning because it processes data faster to provide diagnosis and is more reliable. In the recognition of skin cancer, some scientists have used the CAD system. However, among the different computer-aided methods, deep-learning-based methods gave promising results in segmenting and classifying skin lesions because of their ability to extract complex features from skin lesion images in much more detail. Deep learning algorithms can also learn task-specific characteristics and are much more efficient than other methods [10]. Gouda et al. [11] enhanced the dataset size for training the CNN network to classify skin lesion photos and generated synthetic images using the ESRGAN model. With the ISIC 2018 dataset as its training set, the CNN network attained an accuracy of 83.2%. Fraiwan and Faouri [12] tested thirteen pre-trained deep convolutional neural network models for diagnosing skin lesions from images acquired from the HAM1000 dataset to demonstrate an artificial intelligence-based method for classifying skin cancer. DenseNet201 gave the best accuracy of 82.9% and an F1-score of 0.744. Demir et al. [13] classified skin lesion images into two categories, benign and melanoma. The classification was performed using ResNet-101 and Inception-v3 trained on the ISIC archive. Accuracy of 84.09% and 87.42% was achieved using ResNet-101 and Inception-v3, respectively. Yali Nie, et al, detecting melanoma skin cancer based on object position and increasing accuracy using several YOLO methods, namely YOLOv1, YOLOv2, and YOLOv3. The results of the detection carried out get an mAP value of 0.82 in 200 datasets [14]. Unver and Ayan, segmenting skin cancer objects using a combination of the YOLOv3 method and the GrabCut Algorithm. The results of the skin cancer object segmentation research get 90.82% sensitivity value [15]. Hasya, et al, detecting skin cancer using CNN and YOLOv3. Real-time detection results get an accuracy of 80%. The system detects skin cancer in 7 classes, namely BKL, NV, DF, MEL, VASC, BCC, and AKIEC [16]. Based on several previous studies, the YOLO model has better performance than other models.

Recently, feature learning and object detection have used many deep learning architectures. YOLOv5 (You Only Look Once Version 5) is one of the architectures on the CNN method which is an implementation of Darknet-53 to produce a deeper feature network. This method

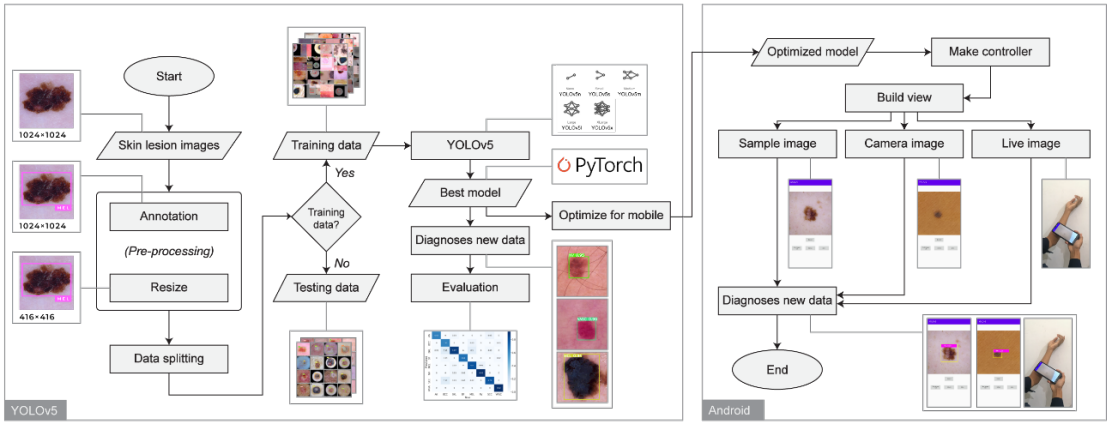

Figure 1. Skin Cancer Diagnosis Flowchart using YOLOv5

can adjust the size of the model flexibly. YOLOv5 has a similar network to YOLOv3 [17]. YOLOv5 processes the image as an input to the detection model network by dividing the image into squares with a certain size. YOLOv5 can detect objects in real-time with high accuracy [18]. Ramya A. compares the performance of YOLOv3, YOLOv4, and YOLOv5. Research shows that YOLOv5 is the most suitable method for detecting small objects [19]. Aqsa Mohiyuddin et al. investigated the detection and classification of breast cancer using YOLOv5, which resulted in good performance on YOLOv5x with an accuracy of 93.60% [20].

YOLOv5 is a deep learning CNN architecture that uses the depth levels described above. YOLOv5 has good pattern recognition capabilities. The YOLOv5 architecture can classify multi-class images. Therefore, this study aims to create an object detection-based skin cancer classification system using deep learning artificial intelligence based on android mobile. This classification system uses the YOLOv5 CNN (Convolutional Neural Network) model. The application of the YOLOv5 method can increase the effectiveness of the system because it can detect skin cancer objects in a relatively short time. This system application uses an android phone [21].

Android phones have practicality, making it easier for users to check for skin cancer automatically. This allows users to know the progress of skin cancer independently and in real-time so that users can treat skin cancer faster. Researchers hope that the proposed system can help the public detect skin cancer early and can help doctors in diagnosing skin cancer.

This experimental research wants to show the efficiency of utilizing the YOLOv5 method on mobile applications to detect skin cancer. By conducting this research, it is hoped that the proposed system can help doctors in diagnosing skin cancer. Other contributions of this research are; i) creating a mobile application to detect and diagnose skin cancer, and ii) optimizing the YOLOv5 method to increase the effectiveness of the system in detecting images (in the case of skin cancer).

2. METHODS AND MATERIAL

In this research, the detection of skin cancer in real-time using the YOLOv5 method. There are several stages in the skin cancer detection system, such as data collection, data preparation stage (pre-processing), making a skin cancer detection system, evaluation stage, selection of the best model, to making a mobile android-based detection system. The steps of this research can be seen in Figure 1.

2.1 Skin Cancer Dataset

This study uses skin lesion image data from the 2019 ISIC Skin Lesion Classification Challenge [22], [23], [24]. The data consists of 8 classes, namely Actinic Keratosis (AK), Basal Cell Carcinoma (BCC), Benign Keratosis (BKL), Dermatofibroma (DF), Melanocytic Nevus (NV), Melanoma (MEL), Squamous Cell Carcinoma (SCC), and Lesions Vascular (VASC). An example of skin lesion data is shown in Figure 2.

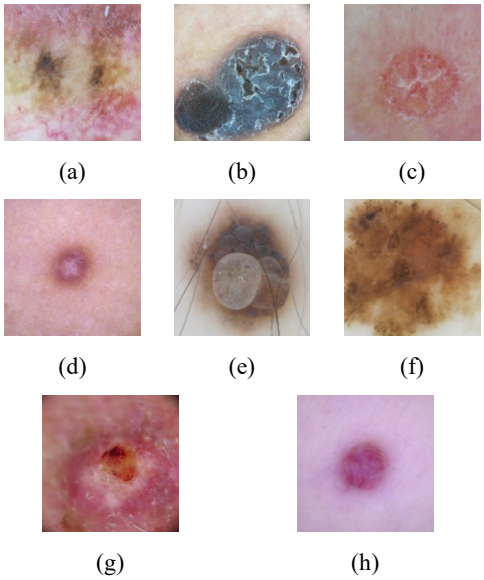

Figure 2. Sample skin lesion data (a) AK (b) BCC (c) BKL (d) DF (e) NV (f) MEL (g) SCC (h) VASC

2.2 Pre-Processing

Pre-processing stage can improve the performance of the classification system. In this study, the pre-processing stage consisted of 2 steps: annotation and resizing of each skin lesion image data. Annotation is the process of labeling the dataset according to the type of each object class. Resize is the process of resizing an image. The resize used in this study follows the size of the input image on the YOLOv5 architecture. In this study, resizing aims as an image reduction process so that it can reduce computational work and optimize the classification model. The image data results that are ready at the pre-processing stage are then divided into two parts: training data and validation data. Training data selection is used as input for feature learning using YOLOv5. While the selection of validation data is used to test the classification model and is included in the calculation of the evaluation model.

Journal of Theoretical and Applied Information Technology
15th August 2024. Vol.102. No. 15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.3 Object Detection using YOLOv5

YOLOv5 implements the DarkNet53 architecture so that the object recognition system becomes faster. Darknet 53 is a layer that includes 53 convolution layers and some residual blocks. Darknet53's deep layers make the architecture perform very well [17]. YOLOv5 produces three outputs to detect large, medium, and small objects. This algorithm has excellent performance and can perform data processing in real-time. YOLOv5 works by dividing image input into a grid of 614 x 614 cells. Each cell directly detects the bounding box and classifies objects based on the confidence or probability of the existence of an object in the bounding box. The YOLOv5 architecture comprises several convolution processes with the addition of residual blocks, concatenates, and up-sampling. These three features can study the image optimally. The YOLOv5 architecture can be seen in Figure 3.

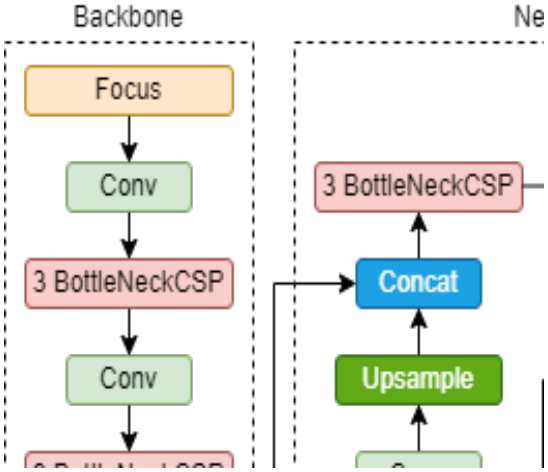

Figure 3. YOLOv5 architecture

2.2.1 Intersection over Union (IoU)

Intersection over Union is a feature to calculate the area that intersects with the combined area between bounding boxes. The IoU value is intended to determine the suitability of the detected bounding box with the object in the actual image. Calculation of IoU can use Equation 1.

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (1)$$

Where B is the area of the intersection of coordinates and the ratio of the size of the detection box results, while in the area of the intersection of the coordinates and the ratio of the size of the actual box.

2.2.2 Confidence Score

In YOLOv5, the confidence score has the task of finding the midpoint of an object contained in the cell grid. The confidence score formula is as shown in Equation 2, where the better the detection of objects with the specified class, the confidence score will be closer to 1 [25]. The confidence score must be zero when there are no objects in the grid of cells.

$$C_o = p(o) \cdot IoU\ (b, o) \quad (2)$$

$p(o)$ is the probability the box contains an object.

$IoU\ (b, o)$ is the IoU between the predicted box and the ground truth.

2.2.3 Non-Max Suppression

Non-Max Suppression aims to filter the bounding box if it appears excessively. This feature makes a selection based on IoU until one bounding box remains and has the highest confidence score. Non-max Suppression analyzes the results of several detections on objects in the image even though they have different sizes and shapes. Non-max Suppression only uses one bounding box with the best value from several detected bounding boxes.

2.4 Evaluation: mAP

The confusion matrix is an evaluation method that contains information between actual and predictive data from a classification system result. Evaluation calculations are useful for measuring the performance of a classification system. The confusion matrix is a matrix that contains how much the accuracy of the prediction results in the actual data in the classification system. The values in the confusion matrix are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). A multi-class confusion matrix measures a classification system with more than two classes. An example of the contents of a multi-class confusion matrix can be seen in Table 1.

Table 1. Multiclass confusion matrix

Actual Data	Predicted Data			
	x_1	x_2	…	x_n
x_1	TN	FP	…	TN
x_2	FN	TP	….	FN
…	…	…	…	…
x_n	TN	FP	…	TN

False positive (FP) is the number of redundant squares that detect the same ground fall. False negative (FN) is the number of ground truths

that were not detected [26]. Determination of the IoU threshold significantly affects the calculation results on the confusion matrix. TP calculated as one if IoU $\geq$ threshold. Meanwhile, if IoU $\leq$ threshold FP value is calculated as one. The equation formula for these values can be calculated at Equation 3, Equation 4, and Equation 5. The calculation results through the confusion matrix get several assessment categories such as accuracy, specificity, and sensitivity.

$$Accuracy = \frac{TP_{all}}{n_{all}} \quad (3)$$

$$Specificity = \frac{\Sigma\left(\frac{TN}{TN+TP}\right)}{n} \quad (4)$$

$$Sensitivity = \frac{\Sigma\left(\frac{TP}{TP+FN}\right)}{n} \quad (5)$$

The evaluation measurement of object detection problems uses the Mean Average Precision (mAP), where mAP is the value of the accuracy of the object detection results against the position and class of the object. mAP is the average value of the interpolated AP (Average Precision) for each class. The AP value is calculated based on recall and precision in each class, which can be seen in Equation 8. The mAP value is calculated using Equation 9.

The value of $x_{1,2,\ldots,n}$ is the first, second, up to the nth class. Table 1 shows the measurement of the predicted results and actual data in the 2nd class by determining the values of TP, TN, FP, and FN as in the table above. In the case of object detection, the value of TP is defined as the number of detection frames worth Intersection over Union (IoU).

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

$$AP = \sum_{k=0}^{k=n-1} (Recall_k - Recall_{k+1}) * Precision \quad (8)$$

$$mAP = \frac{1}{N}\sum_{k=0}^{k=N} AP_k \quad (9)$$

Where, k is the index of each object class. N is the total amount of data in the classification system. The AP value is calculated for each object class then the measurement of the whole system is determined from the mAP value. The higher the mAP value, the better the detection system.

2.5 Android Mobile Configuration

In this study, skin cancer detection was developed by making a detection system based on android mobile. This application aims to facilitate skin cancer detection systems in daily life. In making an android mobile application, detection and android system configuration are needed. The best YOLOv5 skin cancer detection model and the system configuration for android will be exported. Furthermore, the model will be modeled into Torchscript Pytorch Lite before implementing it on android phones. Next, the model processing uses YOLOv5 to get the object box and identify the class in the image based on the confidence score. After that process, cut the overlapping objects using non-max emphasis. Object pruning is based on selecting the highest confidence score and discarding boxes with IoU values above a certain threshold. The final result is object detection in the form of frames with class name identification that appears on Android Mobile.

3. RESULT AND ANALYSIS

This research conducted training on creating a skin cancer detection system using the ISIC 2019 dataset 8 classification classes: AK, BCC, BKL, DF, NV, MEL, SCC, and VASC. The initial pre-processing stage is labeling image data by knowing the position of the cancer object. This labeling produces an object annotation file according to each classification class. The visualization of the object resulting from the annotation process can be seen in Figure 4. The next stage is changing the image size according to the YOLOv5 input. The input image size is resized to 416x416. Next, separate the data into training and test data, with a ratio of 90:10. The total training data is 3252 images, and the total data for testing is 316. Then, the data is ready for the training process using YOLOv5.

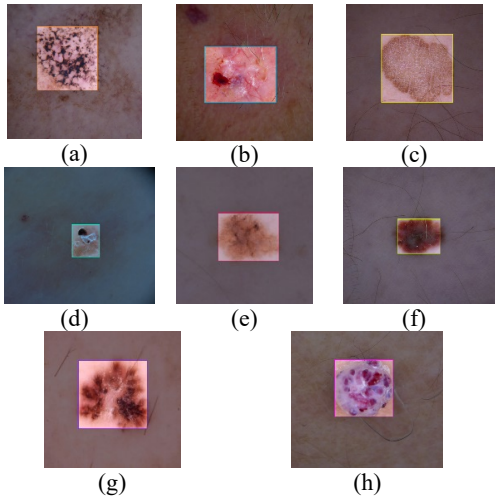

Figure 5. Anotated skin lesion data (a) AK (b) BCC (c) BKL (d) DF (e) NV (f) MEL (g) SCC (h) VASC

The pre-processing data then becomes input for training the YOLOv5 detection system. This research compares several types of models in YOLOv5, namely YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. This training process uses epoch 300, 600, and 1000 parameters. This research uses automatic batch size, where the batch size value depends on the extensive use of the model architecture and GPU memory. Models with mini architecture can get a maximum batch size of 360, and models with jumbo architecture can only get a maximum batch size of 38. Using automatic batch sizes affects accuracy because small batch sizes are used in models with jumbo architecture. The performance of each model is in Table 2.

Table 2. Training result of YOLOv5

Model Type	Epochs	Computation Time (h)	mAP@.5	Accuracy
YOLOv5n	300	0.578	0.798	0.815
	600	1.350	0.814	0.862
	1000	1.889	0.804	0.836
YOLOv5s	300	0.919	0.816	0.843
	600	1.625	0.813	0.846
	1000	2.723	0.815	0.846
YOLOv5m	300	1.318	0.834	0.872
	600	2.634	0.827	0.862
	1000	4.219	0.831	0.890
YOLOv5l	300	2.084	0.832	0.868
	600	6.953	0.844	0.861
	1000	6.690	0.833	0.884
YOLOv5x	300	7.045	0.833	0.862
	600	13.823	0.831	0.890
	1000	12.174	0.838	0.888

The results of the experimental tests in Table 2 show that the entire experiment has a range of values that are not so far away from the others. Based on the mAP value, the best performance on YOLOv5l with 600 epochs was 0.844, and the computation time was 6.953 hours. Meanwhile, the best accuracy value was obtained using YOLOv5x with 600 epochs to get 0.890 results and 13,823 hours of computation time. Meanwhile, when computing time is compared, the deeper the YOLOv5 network, the more computational time is required. The fastest computation time is YOLOv5n, with 270 layers and 300 epochs; it takes 0.578 hours.

Based on Table 2, the average mAP value, accuracy, and computation time were calculated to determine the best model for skin cancer detection systems. Figure 5 shows the results of comparing the average evaluation value and computation time in each YOLOv5 model. It is clear that the deeper the network of YOLOv5, the more computational time is needed to train the skin cancer detection system. The time difference between YOLOv5l and YOLOv5x is significant, with a difference of 5,772 hours. The best average mAP value is YOLOv5l, 0.836 for all epochs. Furthermore, the average best accuracy value for all epochs is YOLOv5x, with a value of 0.880.

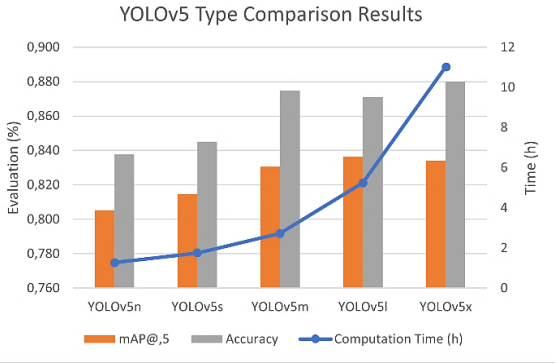

Figure 4. YOLOv5 Type Comparison Results

Based on the evaluation of the object detection system using mAP, the best results are on the YOLOv5l architecture with an mAP value of 0.844. In contrast, the best accuracy results for the classification system use the YOLOv5x and YOLOv5m architectures with an accuracy of 0.890. Evaluation of the object detection system of YOLOv5x and YOLOv5m architectures has a non-significant difference in the mAP value compared to the YOLOv5l mAP value, with a difference in the mAP value of only 0.01. It shows that the

YOLOv5x and YOLOv5m architectures perform reasonably well in object detection and classification systems. Regarding training time duration, YOLOv5x and YOLOv5m have a significant difference. YOLOv5x takes about 13 hours for the training process, while YOLOv5m only takes about 3 hours. Therefore, YOLOv5m is considered to have a good detection system performance in cases of skin cancer. The small number of layers and parameters on YOLOv5m has been able to get system performance results almost the same as the YOLOv5l and YOLOv5x versions with much different computation times. During YOLOv5 training, the time spent on each iteration tends to vary, thereby influencing the time of each epoch exponentially. Therefore, there is training of the YOLOv5 model with relatively the same time but almost twice as many different epochs. Accuracy cannot be improved due to small batch sizes, which result in relatively short iteration times for epochs.

The training results show the best mAP performance value on YOLOv5m with an mAP value of 0.831 and an accuracy of 0.890 with 1000 epochs and a computing time of 4,219 hours. The model results produce an overall precision (P) value of 0.811 and a recall (R) of 0.795 for object detection. Based on the parameters that have been determined, get the best training results in Table 3.

Table 3. Details Evaluation Training of Each Class in YOLOv5m with 1000 Epochs

Class	Precision	Recall	mAP@.5	mAP.5:.95
AK	0.793	0.62	0.667	0.416
BCC	0.792	0.711	0.794	0.569
BKL	0.728	0.897	0.912	0.741
DF	0.806	0.667	0.667	0.424
MEL	0.829	0.85	0.916	0.763
NV	0.898	0.924	0.974	0.834
SCC	0.735	0.795	0.795	0.605
VASC	0.905	0.9	0.922	0.723
All Classess	0.811	0.795	0.831	0.634

When training YOLOv5, the time used in each iteration tends to be different, so it affects the time of each epoch exponentially. Therefore, there is training of the YOLOv5 model with relatively the same time but almost twice as many different epochs. Accuracy cannot be improved due to the use of small batch sizes resulting in relatively short iteration times for epochs. The table also shows detailed information about the evaluation values (Precision, Recall, mAP@.5, and mAP.5:95) for each skin lesion classification class. In sequence, the mAP@.5 values for the best classes are VASC at 92%, NV at 97%, MEL at 92%, and BKL at 91%, while for other classes, the mAP@.5 values are still <80%.

Based on eight object detection classes, there are three types of cancer, namely BCC (slow growing), SCC (faster growing), and MEL (serious skin cancer). Based on the results of the YOLOv5m model, the MEL class is good, achieving a mAP@.5 value of 92%, while BCC and SCC both have a mAP@.5 value of 79%. The detection results of this research are quite satisfactory, especially in the most dangerous class of cancer, namely the MEL class. Meanwhile, the AK class can become candidates for the SCC class if they are not immediately recognized and treated. Based on

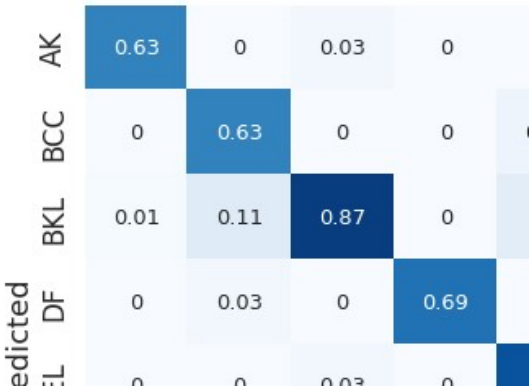

Figure 6. Confusion Matrix of Skin Cancer Classification in YOLOv5m with 1000 Epochs

the results of the confusion matrix in Figure 6, some of the data that produces AK class predictions turns out that the actual class is 3% BKL class and 5% SCC class. This is not good because as much as 5% of the actual data for the SCC class is still predicted as AK class. Furthermore, as many as 1% of the actual AK class is predicted to be BKL; this event is called False Negative (FN). In health analysis, the FN value is strictly avoided and must be minimized so that the sensitivity value is high. Different things happened in the actual data for the NV class, as many as 5% of the NV class were predicted to be MEL class. Where the NV class can become a candidate for the MEL class if it is not immediately recognized and treated. The development of this health prediction system is considered better because the patient produces the possibility that it is worse than the actual one, so the patient will have their condition checked early and get fast treatment. Meanwhile, the other three classes, BKL, DF, and VASC, are skin lesions that are not dangerous and do not develop into serious diseases.

Table 4. Classification result using YOLOv5m

Class	Sensitivity	Specificity	Accuracy
AK	0.984	0.888	0.875
BCC	0.716	0.889	0.663
BKL	0.897	0.828	0.702
DF	0.920	0.883	0.862
MEL	0.825	0.861	0.762
NV	0.948	0.845	0.893
SCC	0.908	0.848	0.745
VASC	0.958	0.853	0.958
Mean	**0.892**	**0.862**	**0.808**

In the case of object detection, the evaluation value uses precision and recall. However, researchers realize that in the case of health data, values such as sensitivity, specificity, and accuracy are also very necessary in determining the success of a data classification system, as shown in Table 4. Based on this table, the average accuracy value for all classes is 0.808, sensitivity 0.892, and specificity 0.862. The results of skin cancer classification showed the best accuracy in the VASC class with a value of 0.958, sensitivity of 0.958 and specificity of 0.853. If we look at the sensitivity value, there is 1 class out of 5 important classes, the BCC class, which has poor results, namely 72%. If analyzed further in Figure 6, the actual BCC class data has many prediction errors. As much as 11% of the actual BCC data was predicted as BKL class, 3% was predicted as DF class, and 11% was predicted as SCC class. Therefore, the sensitivity value of the BCC class is the worst compared to other classes.

Researchers also provide an Android-based application to make it easier for users to detect skin cancer. This application aims to detect various types of skin diseases in real-time, including the most dangerous disease, namely skin cancer. If dangerous skin lesion classes are detected, further examination and treatment must be immediately carried out, especially in the five classes (MEL, SCC, BCC, AK, and NV). The prediction accuracy value also accompanies the detection information, so you need to pay attention if the resulting value is not good. However, suppose the object detection value is >70% (which is recommended). In that case, it is necessary to carry out a medical examination to ensure the accuracy of the skin disease and immediately get appropriate treatment. The hope is that this early detection of skin cancer can help users carry out routine and easy monitoring and examinations (just by using a cell phone).

(a) MEL detected *(b) VASC detected* *(c) NV detected*

Figure 7. Skin Cancer Diagnosis System on Android Mobile

The following are some explanations for using the application. First, download and install the application on your Android phone first. If the installation process is complete, then open the application. You can use several available features, such as pressing "Image Test" and then pressing the "Detection" button to find the results of skin cancer detection using image samples. Press the "Select" button to detect skin cancer from the gallery, or click the "Live" button to detect skin cancer in real-time. Next, the application will produce output as detection results as in Figure 7. Results in class name information with location markers or object boxes and confidence values according to each detection class. As additional information in using this application, users are expected first to clean the skin lesions they want to detect from fine hairs so that detection can be more visible and focused on the object of the lesions. Lighting also needs to be considered when using the application. In order to produce good detection, it is recommended that the camera flash is turned on. Researchers also suggest adjusting the distance between the camera and the object you want to detect. Please set the distance between 5-10 cm for optimal results to avoid detection errors.

Several suggestions can be made in further research to optimize the work of the detection system. First, an automatic object zoom system can be added [27] [28]. If the camera captures a lesion object without needing to test the proximity of the camera several times, the system will automatically zoom in on the object according to the recommended object size level so that detection can be carried out quickly and precisely. Second, it can increase the detection accuracy of dangerous lesion classes such as BCC, SCC, and MEL by using a geometric [29] and color [30] image augmentation approach. In terms of detection, various augmentation behaviors can improve the optimal operation of the detection system. Third, you can use the latest version of the YOLO method [31] [32], which can improve the detection system's performance optimally by paying attention to the level of computing and test image inference.

4. CONCLUSIONS

The method in this research is CNN YOLOv5 which focuses on real-time object detection. This research compares the performance of several YOLOv5 architectures, namely YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. Based on the evaluation of object detection systems using mAP, the best results are found in the YOLOv5l architecture with a mAP value of 0.844. In contrast, the best accuracy results for the classification system use the YOLOv5x and YOLOv5m architectures with an accuracy of 0.890. The evaluation of the YOLOv5x and YOLOv5m architectural object detection systems has an insignificant difference in mAP values compared to the YOLOv5l mAP value, with a difference in mAP values of only 0.01. This shows that the YOLOv5x and YOLOv5m architectures work well in object detection and classification systems. Regarding the duration of training time, YOLOv5x and YOLOv5m have significant differences. YOLOv5x takes around 13 hours for the training process, while YOLOv5m only takes around 3 hours. So, YOLOv5m is the most optimal architecture for detecting skin cancer in real-time and can be implemented on Android phones.

Other contributions of this research are; i) creating a mobile application to detect and diagnose skin cancer, and ii) optimizing the YOLOv5 method to increase the effectiveness of the system in detecting images (in the case of skin cancer). In addition, this research can be a reference for future research, either by adapting the YOLOv5 model or using the latest model. In addition, it can use more recent datasets and have more diverse types such as ISIC 2020, Dermophytes, BCN20000 to PAD-UFES-20.

REFERENCES:

[1] E. M. Senan and M. E. Jadhav, “Analysis of dermoscopy images by using ABCD rule for early detection of skin cancer,” *Glob. Transitions Proc.*, vol. 2, no. 1, pp. 1–7, 2021.

[2] Y. N. Fu’adah, N. K. C. Pratiwi, M. A. Pramudito, and N. Ibrahim, “Convolutional neural network (cnn) for automatic skin cancer classification system,” in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 982, no. 1, p. 12005.

[3] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. K. Islam, “An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models,” *Mach. Learn. with Appl.*, vol. 5, p. 100036, 2021.

[4] N. H. Khan *et al.*, “Skin cancer biology and barriers to treatment: Recent applications of polymeric micro/nanostructures,” *J. Adv. Res.*, 2021.

[5] D. S. Davis, C. Robinson, and V. D. Callender, “Skin cancer in women of color:

Epidemiology, pathogenesis and clinical manifestations," *Int. J. women's dermatology*, vol. 7, no. 2, pp. 127–134, 2021.

[6] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic accuracy of dermoscopy," *Lancet Oncol.*, vol. 3, no. 3, pp. 159–165, 2002, doi: 10.1016/S1470-2045(02)00679-4.

[7] D. A. Gavrilov, E. I. Zakirov, E. V Gameeva, V. Y. Semenov, and O. Y. Aleksandrova, "Automated skin melanoma diagnostics based on mathematical model of artificial convolutional neural network," *Res. Pract. Med. J.*, vol. 5, no. 3, pp. 110–116, 2018.

[8] A. A. Alshehri, T. Daws, and S. Ezekiel, "Medical image segmentation using multifractal analysis," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 10, no. 2, pp. 420–429, 2020.

[9] E. L. Henriksen, J. F. Carlsen, I. M. M. Vejborg, M. B. Nielsen, and C. A. Lauridsen, "The efficacy of using computer-aided detection (CAD) for detection of breast cancer in mammography screening: a systematic review," *Acta radiol.*, vol. 60, no. 1, pp. 13–18, 2019.

[10] S. Rajarajeswari, J. Prassanna, M. Abdul Quadir, J. Christy Jackson, S. Sharma, and B. Rajesh, "Skin Cancer Detection using Deep Learning," *Res. J. Pharm. Technol.*, vol. 15, no. 10, pp. 4519–4525, 2022, doi: 10.52711/0974-360X.2022.00758.

[11] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning," *Healthc.*, vol. 10, no. 7, 2022, doi: 10.3390/healthcare10071183.

[12] M. Fraiwan and E. Faouri, "On the Automatic Detection and Classification of Skin Cancer Using Deep Transfer Learning," *Sensors*, vol. 22, no. 13, 2022, doi: 10.3390/s22134963.

[13] A. Demir, F. Yilmaz, and O. Kose, "Early detection of skin cancer using deep learning architectures: Resnet-101 and inception-v3," *TIPTEKNO 2019 - Tip Teknol. Kongresi*, vol. 2019-January, pp. 283–286, 2019, doi: 10.1109/TIPTEKNO47231.2019.8972045.

[14] Y. Nie, P. Sommella, M. O'Nils, C. Liguori, and J. Lundgren, "Automatic detection of melanoma with yolo deep convolutional neural networks," in *2019 E-Health and Bioengineering Conference (EHB)*, 2019, pp. 1–4.

[15] H. M. Unver and E. Ayan, "Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm," *Diagnostics*, vol. 9, no. 3, p. 72, 2019.

[16] H. F. Hasya, H. H. Nuha, and M. Abdurohman, "Real Time-based Skin Cancer Detection System using Convolutional Neural Network and YOLO," in *2021 4th International Conference of Computer and Informatics Engineering (IC2IE)*, 2021, pp. 152–157.

[17] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo algorithm developments," *Procedia Comput. Sci.*, vol. 199, pp. 1066–1073, 2022.

[18] Z. Qiu, Z. Zhao, S. Chen, J. Zeng, Y. Huang, and B. Xiang, "Application of an Improved YOLOv5 Algorithm in Real-Time Detection of Foreign Objects by Ground Penetrating Radar," *Remote Sens.*, vol. 14, no. 8, p. 1895, 2022.

[19] V. Gupta Pola, A. Bhavya Vaishnavi, and S. Suraj Karra, "Comparison of YOLOv3, YOLOv4 and YOLOv5 Performance for Detection of Blood Cells," *Int. Res. J. Eng. Technol.*, pp. 4225–4229, 2021.

[20] A. Mohiyuddin, A. Basharat, U. Ghani, S. Abbas, O. Bin Naeem, and M. Rizwan, "Breast tumor detection and classification in mammogram images using modified YOLOv5 network," *Comput. Math. Methods Med.*, vol. 2022, 2022.

[21] R. Kraleva, "ChilDiBu-A mobile application for Bulgarian children with special educational needs," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 7, no. 6, pp. 2085–2091, 2017.

[22] N. C. F. Codella *et al.*, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic)," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*, 2018, pp. 168–172.

[23] M. Combalia *et al.*, "Bcn20000: Dermoscopic lesions in the wild," *arXiv Prepr. arXiv1908.02288*, 2019.

[24] P. Tschandl, C. Rosendahl, and H. Kittler,

"The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. data*, vol. 5, no. 1, pp. 1–9, 2018.

[25] A. Horzyk and E. Ergun, "YOLOv3 precision improvement by the weighted centers of confidence selection," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–8.

[26] J. Kim, N. Kim, and Y. W. Park, "Object Detection and Classification Based on YOLO-V5 with Improved Maritime Dataset," 2022.

[27] S.-C. Hsia, S.-H. Wang, C.-M. Wei, and C.-Y. Chang, "Intelligent Object Tracking with an Automatic Image Zoom Algorithm for a Camera Sensing Surveillance System," *Sensors*, vol. 22, no. 22, p. 8791, 2022.

[28] Z. Jiang, D. Wang, Y. Zheng, C. Liu, and Q.-H. Wang, "Continuous optical zoom microscopy imaging system based on liquid lenses," *Opt. Express*, vol. 29, no. 13, pp. 20322–20335, 2021.

[29] M. Elgendi *et al.*, "The effectiveness of image augmentation in deep learning networks for detecting COVID-19: A geometric transformation perspective," *Front. Med.*, vol. 8, p. 629134, 2021.

[30] A. N. Azizah and C. Fatichah, "Tajweed-YOLO: Object detection method for Tajweed by applying HSV color model augmentation on Mushaf images," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 2, pp. 236–245, 2023.

[31] Q. J. Glenn Jocher, Ayush Chaurasia, "Ultralytics YOLO." Ultralytics, 2023.

[32] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "Yolov9: Learning what you want to learn using programmable gradient information," *arXiv Prepr. arXiv2402.13616*, 2024.