A Comparative Study of Deep Learning Algorithms for Detecting Fungal Infection Skin Diseases

Fajar Masya[1], Joko Triloka[* 2], Setia Wulandari[2]
[1]Mercu Buana University, Meruya Sel., Kembangan, Jakarta 11650, Indonesia
[2]Institute of Informatics and Business Darmajaya, Jl. Z.A. Pagar Alam No.93, Bandar Lampung 35141, Indonesia
E-mail: fajar.masya@mercubuana.ac.id, joko.triloka@darmajaya.ac.id, setiawulan.2121211001@mail.darmajaya.ac.id
*Corresponding author

Keywords: mask r-cnn, yolov5, image classification; skin fungal infection

Received: July 11, 2024

Many people place a high value on the health of their skin, frequently spending large sums of money on skincare products. Fungal infections are one of the most common skin conditions that can damage a person's self-esteem. When dealing with skin health issues, seeking advice from a knowledgeable dermatologist is essential. Deep learning is a contemporary technique that saves doctors time and helps them spot diseases early. Two deep learning algorithms that are useful in identifying patterns of skin illnesses are Mask R-CNN and YOLOv5. This paper explores using Mask R-CNN and YOLOv5 to recognize skin illnesses caused by fungal infections, going through several processing phases. The research results show that the YOLOv5 strategy performed best in accuracy, recall, precision, F1-Score, and AUC. This algorithm shows great potential and warrants further investigation in practical applications.

Povzetek: Primerjava algoritmov Mask R-CNN in YOLOv5 za zaznavanje glivičnih kožnih bolezni kaže, da YOLOv5 dosega najboljše rezultate, s čimer izkazuje velik praktični potencial.

1 Introduction

Skin covers the entire surface of the human body and is the largest organ, directly exposed to the external environment [1]. Various diseases affect the skin, ranging from mild, itchy conditions to serious, potentially fatal ones [2]. Despite the importance of skin health, it is often overlooked, and many underestimate skin conditions. Most skin diseases result from bacterial, fungal, or viral infections and allergies [3]. Several factors can directly or indirectly impact the skin, causing diseases that may be treatable with medications, while others necessitate consultation with a professional skin disease specialist [4,5]. Consultation with a specialist in dermatology is essential for individuals with skin health concerns. However, due to embarrassment and the high cost of treatment, many individuals with skin diseases remain silent, leading to decreased self-confidence and social withdrawal. This social isolation can contribute to depression. Therefore, dermatologists must engage in early detection and prevention of skin diseases, as these conditions can be easily transmitted.

In the modern era, nearly all sectors, including medicine, rely on computerized systems to replace conventional methods with automated technology [6]. Researchers, particularly in medical science, are actively seeking solutions to help doctors diagnose diseases early without excessive time expenditure [7]. This is where digital image processing becomes essential [8]. Digital image processing involves using computer algorithms to enhance images by extracting valuable information. Object detection algorithms, often employing machine learning or deep learning, automate relevant findings. In medical science, digital image processing is instrumental in automating diagnostic processes [9].

Several studies have applied popular object detection algorithms, such as the Mask Regional-based Convolutional Neural Network (Mask R-CNN) and You Only Look Once (YOLO) algorithms. One study using the Mask R-CNN algorithm for breast cancer detection reported an accuracy of 91% and a precision of 84% [10]. Another study implemented Mask R-CNN to find, detect, and classify objects in images or videos of the Ryze Tello drone, achieving an average accuracy of 95.6% [11]. Additionally, research using Mask R-CNN for automatically detecting and recognizing small magnetic targets in shallow underground layers demonstrated an average detection accuracy of 97%, a recall rate of 94%, and an average detection speed of 0.35 seconds per image on a GPU [12]. Studies employing the YOLOv5 algorithm have also shown significant results. One study detecting face masks with YOLOv5 after 300 epochs achieved an accuracy rate of approximately 96.6% [13]. Another study using YOLOv5 to determine whether a face mask is being worn reported an accuracy of 97.90% [14]. The application of popular object detection algorithms like Mask R-CNN and YOLOv5 has been widely successful across diverse fields. The specific accuracies and precision rates mentioned for different applications like breast cancer detection, drone imagery classification, underground magnetic target detection, and face mask detection highlight these algorithms' versatility and high performance in various domains.

While multiple studies have investigated the use of Mask R-CNN and YOLO for a variety of medical applications, including breast cancer detection, face mask recognition, and other skin illnesses, there has been a striking paucity of research focusing on fungal skin infections. Existing research focuses mostly on bacterial or viral skin disorders or non-specific skin diseases, leaving a vacuum in the early identification and categorization of fungal infections with advanced deep-learning models. This gap is crucial since fungal infections are common and sometimes misdiagnosed due to symptoms that overlap with other skin disorders.

This study intends to close the highlighted gap by thoroughly comparing two cutting-edge deep learning systems, Mask R-CNN, and YOLOv5, for identifying and categorizing fungal skin diseases. This is critical since fungal infections are among the most common skin disorders, affecting millions of people worldwide, and early detection is essential for avoiding consequences. This study enhances the application of deep learning in dermatology by comparing the performance of these algorithms. It also provides practical insights for real-time diagnostic tools in healthcare settings.

2 Related work

Numerous studies have explored the efficacy of various algorithms for classifying skin diseases caused by fungal infections. In 2017, [15] investigated the use of image processing techniques, including Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Singular Value Decomposition (SVD), achieving an impressive detection efficiency of up to 80%. The average training time across the three transformations and their parallel combinations was 2.066 seconds, with an average testing time of 0.7866 seconds. Subsequently, in 2018, [16] delved into the utilization of the K-Means and Fuzzy C-Means algorithms, providing valuable insights into skin disease detection. The adoption of these algorithms supported early diagnosis and disease-type identification.

In deep learning, [17] introduced Convolutional Neural Network (CNN) algorithms for skin disease detection in 2018, demonstrating enhanced accuracy and efficiency compared to traditional methods. The CNN approach yielded better results, paving the way for more advanced diagnostic tools. Building on this progress, [18] explored the application of the YOLOv3 algorithm in the medical field in 2019. Their investigation encompassed diverse tasks, such as white blood cell detection and identifying target strings of bananas and fruit stems. Notably, the YOLOv3 algorithm achieved impressive accuracy rates, showcasing its versatility and potential in medical imaging.

Further research by [19] focused on facial skin disease analysis using CNN algorithms based on clinical images. Their study encompassed the detection of five facial skin diseases, achieving notable accuracies for various conditions. Additionally, [20] investigated skin pathology detection using CNN algorithms, reaching a test accuracy of 89%.

Meanwhile, several studies have utilized YOLO in research on different tasks, including [21], which has achieved 92.20% accuracy in real-time face mask detection under multiple conditions. [22] calculating melanoma skin cancer using a web application integrated with the YOLOv5. The model evaluates if the stain is cancerous or benign. [23] applying YOLO for early skin cancer detection with the test results showed that the YOLOv5's model has an accuracy of 89.1% in detecting skin cancer types. Moreover, a proposed Yolo deep neural network which can classify 9 different classes of skin cancer was conducted by [24], their experimental analysis shows that the proposed method achieves the mean average precision score of 88.03% and 86.52% for Yolo V3 and Yolo V4 respectively.

3 System model

3.1 Mask R-CNN

Mask R-CNN, developed by the Facebook AI Research (FAIR) team in 2017, is a deep learning algorithm renowned for detecting objects in images while simultaneously generating a segmentation mask for each instance, a technique commonly referred to as instance segmentation [25]. As depicted in Figure 1. instance segmentation shares similarities with object detection, wherein individual objects are detected sequentially. However, it integrates semantic segmentation, enabling each object to be categorized, localized, and distinguished at the pixel level.

During the detection process, Mask R-CNN operates across three main components: the feature extraction network, region-proposal network, and instance detection and segmentation networks. Mask R-CNN employs various backbone architectures [26], including ResNet-101 and FPN for feature extraction. Through experimentation, the ResNet-101 backbone has demonstrated above-average accuracy and speed in feature extraction. In the Region Proposal Network (RPN) phase, Regions of Interest (ROIs) are generated, serving as input for the subsequent instance detection and segmentation networks stage.

a. ***Feature extraction***: Feature extraction aims to distill information from images and represent it in a lower-dimensional space, facilitating the classification of patterns. In the context of Mask R-CNN, feature extraction involves generating Region of Interest (RoI) features through the fusion of ResNet-101 architecture with FPN (Feature Pyramid Network). FPN plays a crucial role in recognition systems by enabling the identification of objects of various sizes within the same image. FPN enhances information quality by utilizing multiple feature maps. It adopts a pyramid design principle for feature extraction, offering superior speed and accuracy.

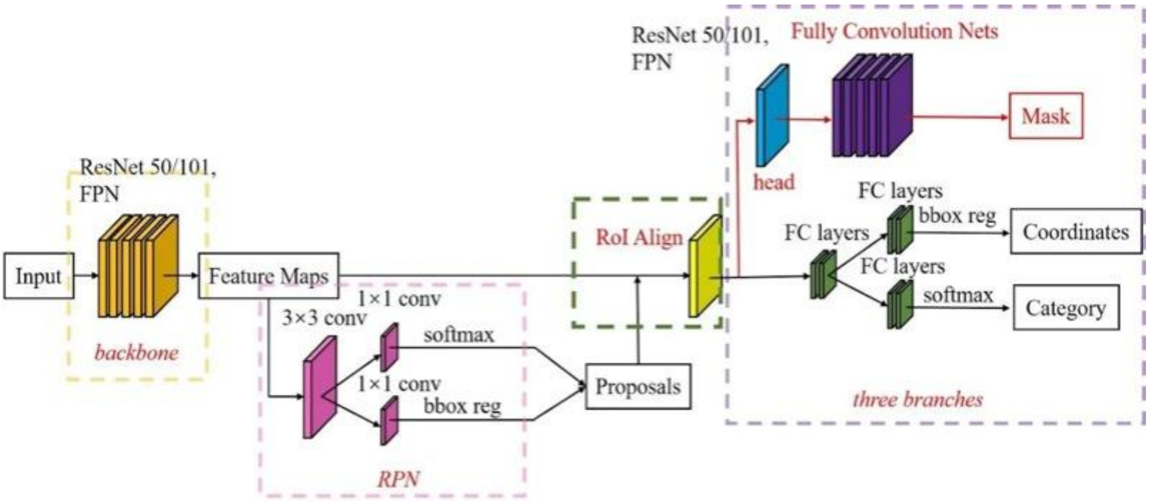

Figure 1: The Mask R-CNN framework for instance segmentation

FPN integrates both bottom-up and top-down information processing techniques to achieve comprehensive feature representation.

b. ***RPN (Region Proposal Network)***: Within the feature extraction process, a 3 x 3 convolution layer is applied to each generated feature map. Initially, the feature map undergoes scanning utilizing anchor boxes of various sizes and ratios. Subsequently, the output is bifurcated into two branches: one associated with the objectivity or confidence score, and the other with the bounding box regressor, as depicted in Figure 2.

c. ***Instance Detection and Semantic Segmentation***: During the instance segmentation process, objects, bounding boxes, class labels, and confidence values are detected through a fully connected network that takes the Region of Interest (RoI) as input. Semantic segmentation is then performed on the image using a Fully Convolutional Network (FCN), which predicts the semantic class of each pixel within the bounding box. As a result, distinct colors are assigned to each instance based on the bounding box delineation, facilitating visual differentiation of individual objects.

3.2 YOLO

You Only Look Once (YOLO) is an algorithm developed by the Facebook AI Research (FAIR) team to quickly and accurately detect various types of objects. YOLO addresses single regression problems directly by mapping image pixels to bounding box coordinates and class probabilities. It requires only one look at an image to predict what objects are present and where they are located. YOLO operates by using a single convolutional network that simultaneously predicts multiple bounding boxes and the probability of each class within those boxes. it has overall 24 convolutional layers, four max-pooling layers, and two fully connected layers as illustrated in Figure 3. It is trained on images to optimize detection performance. The architecture works as follows:

a. The input image is resized to 448x448 before being processed by the convolutional network.
b. A 1x1 convolution is initially applied to reduce the number of channels, followed by a 3x3 convolution to generate a cuboidal output.
c. The ReLU activation function is used throughout, except for the final layer, which uses a linear activation function.
d. Additional techniques, such as batch normalization and dropout, are employed to regularise the model and prevent overfitting.

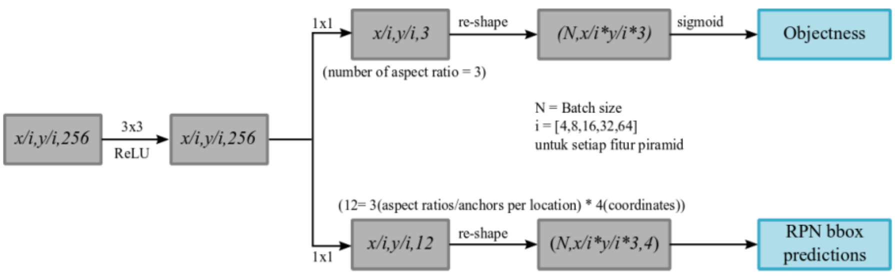

Figure 2: RPN processing

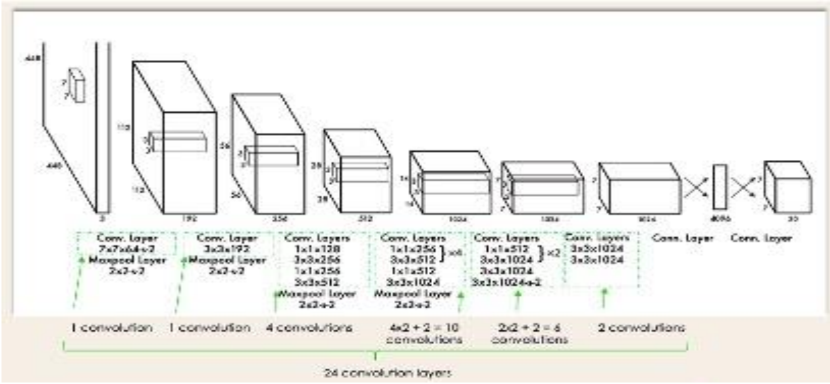

Figure 3: YOLOv5 architecture

4 Proposed procedures

Proposed Mask R-CNN comprises three primary stages. To begin with, it uses the darknet-53 architecture to extract features. Second, it uses the input image to derive the coordinates of Regions of Interest (RoI) using the Region Proposal Network (RPN) approach. Finally, it predicts the class of the discovered objects, revealing information about the ROI sites. This procedure yields a mask that highlights areas suggestive of fungal-induced skin disorders. A suggested Mask R-CNN technique based on edge detection is shown in Figure 4 to identify the skin conditions in the dataset.

The YOLOv5 algorithm utilized in this study has multiple phases for object detection. Using PyTorch as a feature extractor, YOLOv5 detects objects by classifying them and locating them depending on the features that are extracted. The goal of YOLOv5 feature extraction is to supply input variables for the classification procedure. The suggested YOLOv5 algorithm architecture is displayed in Figure 5.

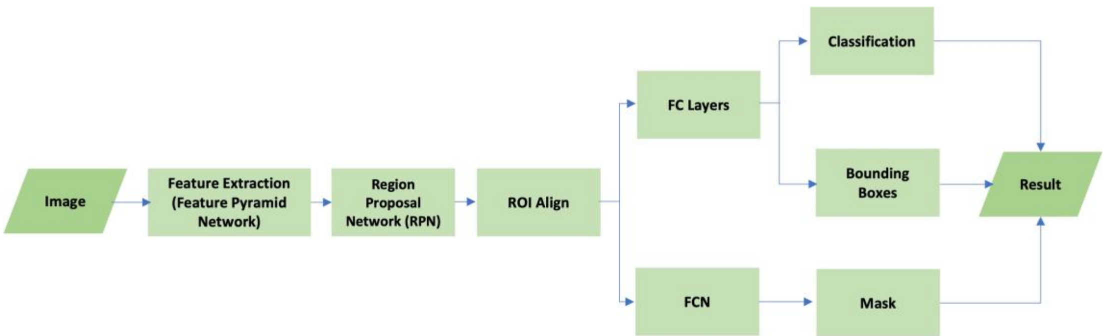

Figure 4: Proposed Mask R-CNN Architecture

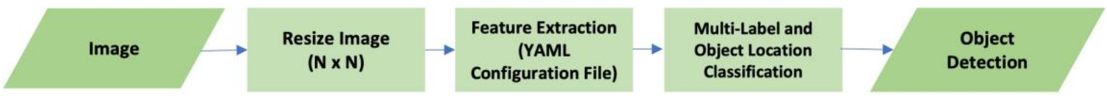

Figure 5: Proposed YOLOv5 architecture

5 Experiments on algorithm processing

5.1 Dataset description

This study's publicly accessible dataset is from http://www.dermnet.com/dermatology-pictures-skin-disease-pictures and consists of images of 1,473 data points and 3 class labels of skin diseases: Dermatomycosis, Mucocutaneous Candidiasis, and Pityriasis Versicolor. Before splitting the dataset, it is first preprocessed to ensure each image is appropriate for labeling. Figure 6 shows an example of a dataset with skin conditions brought on by fungus infections.

5.2 Data Pre-processing

Raw data needs to be treated first. Partitioning and labeling the dataset are preprocessing steps. Labeling gives each thing in the picture a name and ensures that it is part of the appropriate group. After that, the 1.473 image dataset is split into training and testing sets. The Mask R-CNN algorithm's labeling procedure entails object segmentation with the polygon tool. On the other hand, the YOLO algorithm utilizes bounding boxes created with a bounding box tool for labeling.

With 10% of the data for testing and the remaining 90% for training, the dataset was reduced to 1.136 following the labeling phase. A ratio of 80% for training and 20% for testing was also used in the experiment. Pre-processing procedures led to a minor decrease in the overall dataset size due to data cleaning procedures, much as the 90/10 split. The model's performance was likewise decreased to 1136 with this split, demonstrating how well it performs in comparison to the initial 90/10 split. Figure 7 provides a polygon tool example of data labeling, while Figure 8 provides a bounding box tool example.

5.3 Algorithm processing

Figure 9 illustrates the idea of processing both algorithms. Installing Python, TensorFlow, Keras, and other necessary software is part of the dependency installation process for the Mask R-CNN algorithm. Installing deep learning packages such as PyTorch, NumPy, and Pandas for YOLOv5 was required for YOLOv5. The dataset that will be utilized to train the object detection algorithms is prepared during the object detection data loading stage. The dataset needs to include pictures and properly formatted annotations (labels and bounding boxes) for the Mask R-CNN technique to function. Every object in the dataset needs to be labeled for detection for the YOLOv5 algorithm to work. The training configurations for both algorithms are established in the configuration setting.

Figure 7: Data labeling on images using the polygon tool

Figure 6: Sample images of skin diseases caused by fungal infections

The setup for the Mask RCNN algorithm contains details about the number of iterations, batch size, number of classes, and other pertinent parameters. Configuration options for the YOLOv5 include batch size, learning rate, and number of epochs. To improve object detection accuracy, Mask R-CNN performs gradient computations and modifies model weights throughout the training phase. Parameters in the YOLOv5 algorithm are optimized to improve the accuracy of object detection. During the testing phase, fresh photos are used to perform object detection. For every object that is recognized, the Mask RCNN algorithm produces bounding boxes and class labels. To evaluate the object detection performance of the YOLOv5 algorithm, a dataset that hasn't been seen before is used.

5.4 Algorithm evaluation

We assessed YOLOv5 and Mask R-CNN for object detection in this study because of their high accuracy and effectiveness in managing related tasks. These algorithms were selected due to their resilience and efficacy in object identification and classification, particularly in situations requiring quick and accurate detection. Although these techniques are the focus of our study, we acknowledge the possibility of expanding it to include other highly respected object detection algorithms, like SSD (Single Shot MultiBox Detector), EfficientDet, and Faster R-CNN. We consider these algorithms for future research initiatives because they may offer further insights into the comparative performance of our dataset. This procedure assesses how well the YOLOv5 and Mask R-CNN algorithms work. There are two phases to the evaluation: testing and training. During the training phase, both methods employed 1,023 photos from the dataset. During the testing phase, 113 different photos are used to assess the algorithms. At this stage, the algorithms' object detection performance is evaluated, and a confusion matrix is used to calculate the algorithms' accuracy. Numerous significant performance indicators, including accuracy, precision, recall, F1-score, mean average precision (MAP), and area under the curve (AUC), can be obtained from the confusion matrix. Five iterations and epochs of performance evaluation are granted for both algorithms.

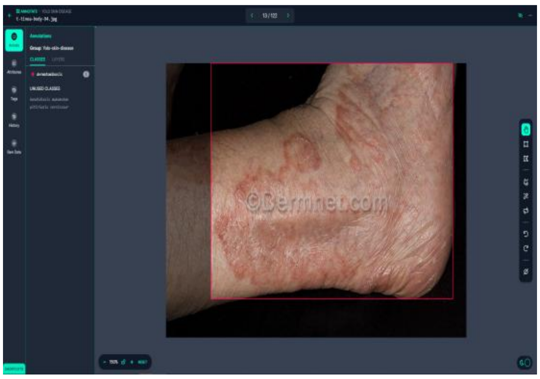

Figure 8: Data labeling on images using the bounding box tool

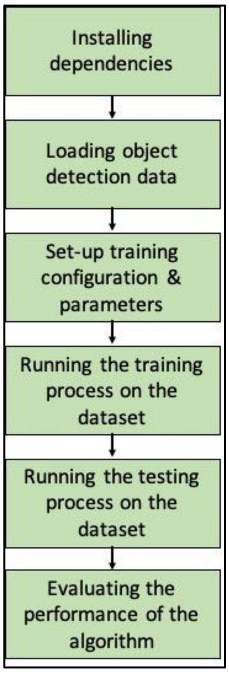

Figure 9: The notion of algorithmic processing

6 Results and discussion

Once training and testing are finished, the assessment step is carried out to gauge the effectiveness of the YOLOv5 and Mask R-CNN algorithms. We test the Mask RCNN algorithm with various iteration settings and thresholds. In contrast, the Mask RCNN algorithm is evaluated using 1000, 1500, 2000, 2500, and 3000 iteration values, respectively, with threshold values varying between 0.1 and 0.9 for every iteration. The YOLOv5 algorithm, on the other hand, makes use of distinct threshold values and epochs. 50, 75, 100, 125, and 150 are the employed epoch values, and each epoch's threshold values range from 0.1 to 0.9.

a. ***Mask R-CNN Algorithm:*** The Mask R-CNN algorithm identified 80 data labels for Dermatomycosis (D), 19 for Mucocutaneous Candidiasis (MC), and 0 for Pityriasis Versicolor (PV) after five tests. A total of 113 photos were positively detected. Table 1 presents the interpretation of the performance calculation for the Mask RCNN method used to treat skin diseases. The algorithm uses 3000 iterations and varies the threshold (T) from 0.1 to 0.9. The F1-score is the method for identifying the optimal model, with a threshold value for model evaluation between 0.1 and 0.9. When assessing the binary model, the harmonic mean of precision and recall is employed using the F1-score. According to Table 1, the maximum F1-score of 0.28 is attained at the 0.1 level. The precision is 49%, the recall is 19%, and the accuracy is 67% at this threshold.

Table 1: Performance of Mask R-CNN with 3000 Iterations

T	Accuracy				Recall				Precision				F1-Score			
	D	MC	PV	MAP	D	MC	PV	MAP	D	MC	PV	MAP	D	MC	PV	MAP
0.1	0,39	0,7	0,92	0,67	0,35	0,23	0	0,19	0,87	0,59	0	0,49	0,5	0,33	0	0,28
0.2	0,4	0,72	0,93	0,68	0,35	0,19	0	0,18	0,91	0,63	0	0,51	0,51	0,29	0	0,27
0.3	0,4	0,73	0,94	0,69	0,34	0,15	0	0,16	0,92	0,58	0	0,5	0,5	0,24	0	0,25
0.4	0,4	0,73	0,94	0,69	0,33	0,14	0	0,16	0,95	0,69	0	0,55	0,49	0,23	0	0,24
0.5	0,4	0,74	0,95	0,7	0,33	0,15	0	0,16	0,96	0,73	0	0,56	0,49	0,25	0	0,25
0.6	0,4	0,75	0,96	0,7	0,32	0,14	0	0,15	0,96	0,71	0	0,56	0,48	0,23	0	0,24
0.7	0,39	0,75	0,96	0,7	0,31	0,12	0	0,14	0,96	0,67	0	0,54	0,47	0,2	0	0,22
0.8	0,38	0,77	0,97	0,71	0,31	0,13	0	0,15	0,96	0,67	0	0,54	0,47	0,22	0	0,23
0.9	0,33	0,81	0,98	0,71	0,26	0,12	0	0,13	0,91	0,5	0	0,47	0,4	0,19	0	0,2

The area under the ROC graph is then computed using the AUC value. It is employed as a performance evaluation statistic to gauge a classification model's effectiveness. A higher AUC score indicates better model performance in differentiating between positive and negative classes. In the fifth test, an AUC value of 0.55 is displayed on the ROC graph of the Mask R-CNN method, as depicted in Figure 10.

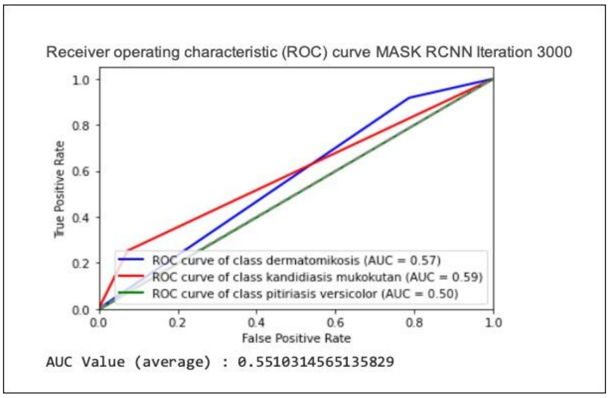

Figure 10: ROC following the fifth Mask R-CNN algorithm test

b. ***YOLOv5 Algorithm:*** By the time the fifth test was reached, the YOLOv5 algorithm had accurately identified 113 images; however, it had only identified 86 data labels for Pityriasis Versicolor, 39 for Mucocutaneous Candidiasis, and 86 for Dermatomycosis. The performance testing at epoch 150 with a threshold value of 0.1 is displayed in Table 2. According to Table 2, the greatest F1-Score value is 0.81, at the 0.1 threshold, with 86% accuracy, 8% recall, and 94% precision. With an AUC value of 0.88. Figure 11. displays a graph for the ROC of the YOLOv5 algorithm in the fifth test.

c. ***Evaluation of Proposed Algorithms:*** The Mask R-CNN and the YOLOv5 algorithm can be compared to the calculation results obtained from the algorithm testing technique for detecting fungal infections-caused skin problems, which contained 113 data points from three different skin conditions. Table 3 displays the comparison values.

In every metric that is examined, YOLOv5 outperforms Mask R-CNN, including accuracy (0.87), recall (0.80), precision (0.85), F1-Score (0.81), and AUC (0.88). The variety of fungal infections in appearance, size, form, and texture makes it particularly difficult to diagnose skin illnesses caused by these infections. Algorithms that process medical pictures accurately and efficiently are necessary for effective detection. Here, the effectiveness of two widely used object detection algorithms, YOLOv5 and Mask R-CNN, is compared.

This can be beneficial when precise infection borders are critical in medical imaging. Dermatologists may find the capacity to create segmentation masks especially helpful in diagnosing and treating infections since they offer comprehensive details about the affected regions. However, because of its multi-stage processing, Mask R-CNN requires a lot of computing power. This may lead to slower inference and longer training times, which could be problematic for real-time applications or for handling big datasets.

By adding a branch for predicting segmentation masks on each Region of Interest (RoI) in parallel with the current branch for classification and bounding box regression, Mask R-CNN expands upon Faster R-CNN. The two-stage method of Mask R-CNN, which includes region proposal and refining, enables very accurate object identification and segmentation.

Table 2: Performance of YOLOv5 with 150 epochs

T	Accuracy				Recall				Precision				F1-Score			
	D	MC	PV	MAP	D	MC	PV	MAP	D	MC	PV	MAP	D	MC	PV	MAP
0.1	0,77	0,83	0,99	0,86	0,77	0,72	0,92	0,8	0,71	0,92	0,8	0,81	0,8	0,71	0,92	0,81
0.2	0,74	0,86	0,99	0,86	0,68	0,72	0,92	0,77	0,78	0,92	0,84	0,85	0,76	0,75	0,92	0,81
0.3	0,72	0,86	0,98	0,85	0,62	0,69	0,83	0,71	0,8	0,91	0,85	0,85	0,72	0,74	0,87	0,78
0.4	0,67	0,85	0,98	0,83	0,52	0,62	0,82	0,65	0,83	0,9	0,86	0,86	0,65	0,71	0,86	0,74
0.5	0,62	0,83	0,97	0,81	0,42	0,53	0,5	0,48	0,85	1	0,89	0,91	0,57	0,65	0,67	0,63
0.6	0,58	0,8	0,96	0,78	0,33	0,42	0,33	0,36	0,89	1	0,92	0,94	0,49	0,57	0,5	0,52
0.7	0,53	0,75	0,93	0,74	0,23	0,25	0	0,16	1	0	1	0,67	0,37	0,4	0	0,26
0.8	0,44	0,69	0,93	0,69	0,08	0,07	0	0,05	1	0	1	0,67	0,15	0,13	0	0,09
0.9	0,4	0,67	0,93	0,67	0,02	0	0	0,01	0	0	1	0,33	0,04	0	0	0,01

Table 3. Performance comparison of proposed algorithms

Iteration/ Epochs		Threshold		ΣPrediction		Accuracy		Recall		Precision		F1-Score		AUC	
M	Y	M	Y	M	Y	M	Y	M	Y	M	Y	M	Y	M	Y
1000	50	0,6	0,1	176	163	0,79	0,84	0,34	0,76	0,6	0,76	0,43	0,75	0,61	0,81
1500	75	0,2	0,1	179	179	0,73	0,87	0,42	0,76	0,36	0,85	0,38	0,78	0,61	0,88
2000	100	0,1	0,1	220	183	0,67	0,86	0,22	0,78	0,45	0,82	0,3	0,8	0,55	0,87
2500	125	0,7	0,1	199	161	0,8	0,87	0,4	0,75	0,67	0,85	0,45	0,79	0,57	0,84
3000	150	0,1	0,1	260	183	0,67	0,86	0,19	0,8	0,49	0,81	0,28	0,81	0,55	0,88

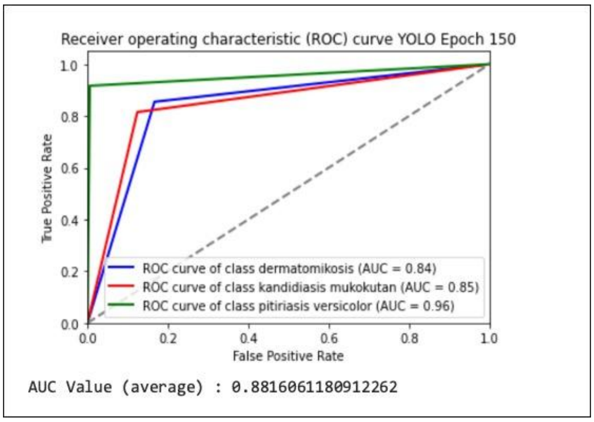

Figure 11: ROC following the fifth of the YOLOv5 Algorithm test

Our proposed technique, YOLOv5, is a quick and efficient single-stage object detection algorithm that outperforms two-stage detectors such as Mask R-CNN. Large-scale image analysis and real-time detection scenarios benefit greatly from this efficiency. Although it might not offer as much segmentation depth as Mask R-CNN, YOLOv5 has a very high degree of object detection and classification accuracy. It captures items of different sizes with the aid of anchor boxes and multi-scale

Additionally, YOLOv5 produced higher recall and precision metrics, demonstrating its efficacy in reducing false positives and false negatives. This is important for medical diagnosis since misdiagnosing a healthy area as sick (false positive) or failing to detect an infection (false negative) can have serious repercussions. This implies that YOLOv5 has a higher degree of accuracy when it comes to recognizing contaminated regions in the pictures.

Furthermore, the high AUC suggests that YOLOv5 performs better across a range of threshold values in differentiating between infected and non-infected areas. Although Mask R-CNN provides thorough segmentation, the comparison research indicates that for this specific task, the benefits of segmentation are not greater than those of YOLOv5's higher detection accuracy and efficiency. However, in some clinical situations where precise infection boundaries are required, Mask R-CNN's segmentation function might still be useful.

Because YOLOv5 processes information more quickly, it is more suited for real-world uses where timely findings are crucial, like automated screening systems in healthcare settings. According to the comparison analysis, YOLOv5 is a more sensible option for large-scale screens and real-time applications. Nonetheless, the needs of the application, such as the necessity for segmentation against the requirement for quick and precise identification, should be considered while choosing between the two algorithms.

7 Conclusion

There are notable variations in performance parameters including accuracy, recall, precision, F1-Score, and AUC when YOLOv5 and Mask R-CNN algorithms are compared to identify fungal diseases on the skin. The disparities arise from variations in iteration (or epoch) values, which affect the algorithms' capacity to acquire knowledge and generalize from the training set. The results of performance tests indicate that algorithm performance is influenced by epoch or iteration values from the Mask R-CNN and YOLOv5 algorithms' first to fifth tests. The second test had the highest AUC value with 1500 iterations of the Mask R-CNN method and 75 epochs of the YOLOv5 algorithm. With an AUC value of 66% and an F1-score of 38%, the Mask R-CNN algorithm is less successful in 1500 iterations at identifying diseases caused by fungal infections of the skin. On the other hand, with an AUC value of 88% and an F1-Score value of 78%, the YOLO algorithm's test results demonstrate its good ability to identify diseases brought on by skin infections, as evidenced by its ability to forecast 179 disorders in epoch 75.

The thorough investigation demonstrates that YOLOv5 performs better than Mask R-CNN in terms of accuracy, recall, precision, F1-Score, and AUC when it comes to identifying fungal diseases on the skin. Iteration and epoch settings have a significant impact on performance; YOLOv5 shows the best performance at 75 epochs. Mask R-CNN is less suited for this application due to its computational intensity and lower detection accuracy, even though it has segmentation capabilities. As a result, in both clinical and real-world contexts, YOLOv5 is the recommended algorithm for identifying fungal-induced skin disorders due to its exceptional efficiency and accuracy. Future research can continue to enhance the detection accuracy and practical applicability of deep learning models for diagnosing skin fungal infections.

References

[1] Majaranta, P. et al. 2019. Eye Movements and Human-Computer Interaction In: Klein, C., Ettinger, U. (eds) Eye Movement Research. Studies in Neuroscience, Psychology, and Behavioral Economics. Springer, Cham, pp.971-1015. https://doi.org/10.1007/978-3-030-20085-5_23

[2] Al Bshabshe, A. et al. 2023. An Overview of Clinical Manifestations of Dermatological Disorders in Intensive Care Units: What Should Intensivists Be Aware of? *Diagnostics* 13, no. 7: 1290. https://doi.org/10.3390/diagnostics13071290

[3] Nigat, T.D.; Sitote, T.M. 2023. Gedefaw, B.M. Fungal Skin Disease Classification Using the Convolutional Neural Network. *J. Healthc. Eng*. 6370416, pp.1-9. https://doi.org/10.1155/2023/6370416

[4] Badia, M. et al. 2020. Dermatological Manifestations in the Intensive Care Unit: A Practical Approach. *Crit Care Res Pract*. 2020:9729814. https://doi.org/10.1155/2020/9729814

[5] Altammami, G. S et al. 2024. Dermatological Conditions in The Intensive Care Unit at A Tertiary Care Hospital in Riyadh, Saudi Arabia." *Saudi Medical Journal* vol. 45, 8. pp. 834-839. https://doi.org/10.15537/smj.2024.45.8.20240479

[6] Khavandi, S. et al. 2023. Investigating the Impact of Automation on the Health Care Workforce Through Autonomous Telemedicine in the Cataract Pathway: Protocol for a Multicenter Study." *JMIR research protocols* vol.12. e49374. pp.1-10. https://doi.org/10.2196/49374

[7] Roy, K. S. et al. 2019. *Skin Disease Detection Based on Different Segmentation Techniques.* 2019 Int. Conf. Opto-Electronics Appl. Opt. Optronix. pp. 1–5, 2019. https://doi.org/10.1109/OPTRONIX.2019.8862403

[8] Archana, R. and Jeevaraj, P.S.E. 2024. Deep learning models for digital image processing: a review. *Artif Intell Rev* 57, 11. https://doi.org/10.1007/s10462-023-10631-z

[9] Thakur, G. K. et al. 2024. Deep Learning Approaches for Medical Image Analysis and Diagnosis. *Cureus* vol. 16,5 e59507. pp. 1-8. https://doi.org/10.7759/cureus.59507

[10] Bhatti, H.M.A. et al. 2020. *Multi-detection and Segmentation of Breast Lesions Based on Mask RCNN-FPN*. Proc. - 2020 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2020, pp. 2698–2704. https://doi.org/10.1109/BIBM49941.2020.9313170

[11] Subash, K.V.V. et al. 2020. *Object Detection Using Ryze Tello Drone With Help of Mask-RCNN*. 2nd Int. Conf. Innov. Mech. Ind. Appl. ICIMIA 2020 - Conf. Proc., no. Icimia, pp. 484-490, 2020. https://doi.org/10.1109/ICIMIA48430.2020.9074881

[12] Zhou, Zhijian. et al. 2020. Detection and Classification of Multi-Magnetic Targets Using Mask-RCNN. *IEEE Access*. 8. 187202-187207. https://doi.org/10.1109/access.2020.3030676

[13] Ieamsaard, J. et al. 2021. Deep *Learning-based Face Mask Detection Using YoloV5*. In Proceeding of the 2021 9th International Electrical Engineering Congress, iEECON 2021, Mar. 2021, pp. 428–431. https://doi.org/10.1109/iEECON51072.2021.9440346

[14] Yang, G. et al. *Face Mask Recognition System with YOLOV5 Based on Image Recognition*. 2020. In 2020 IEEE 6th International Conference on Computer and Communications, ICCC 2020, Dec. 2020, pp. 1398-1404. https://doi.org/10.1109/ICCC51575.2020.9345042

[15] Ajith, A. et al. 2017. *Digital Dermatology Skin Disease Detection Model using Image Processing*. Int. Conf. Intell. Comput. Control Syst. ICICCS 2017 Digit., vol. 24, no. 7, pp. 168-173. https://doi.org/10.1109/iccons.2017.8250703

[16] Haddad, A and Hameed, S.A. 2018. *Image Analysis Model for Skin Disease Detection: Framework*. Proc. 2018 7th Int. Conf. Comput. Commun. Eng. ICCCE 2018, no. c, pp. 280-283, 2018. https://doi.org/10.1109/ICCCE.2018.8539270

[17] Rathod, J. et al. 2018. *Diagnosis of skin diseases using Convolutional Neural Networks.* Proc. 2nd Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2018, no. Iceca, pp. 1048-1051, 2018. https://doi.org/10.1109/ICECA.2018.8474593

[18] Rohaziat, N. et al. 2020. White Blood Cells Detection using YOLOv3 with CNN Feature Extraction Models. *International Journal of Advanced Computer Science and Applications.* 11. https://doi.org/10.14569/IJACSA.2020.0111058

[19] Wu, Z. et al. 2019. Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images. *IEEE Access*. vol. 7, no. c, pp. 66505-66511. https://doi.org/10.1109/ACCESS.2019.2918221

[20] Li, L. F. et al. 2020. Deep Learning in Skin Disease Image Recognition: A Review. *IEEE Access*. vol. 8, pp. 208264-208280. https://doi.org/10.1109/ACCESS.2020.3037258

[21] Salama, A. M. et al. 2024. A YOLO-based Deep Learning Model for Real-Time Face Mask Detection via Drone Surveillance in Public Spaces, *Information Sciences*, vol. 676, 120865. https://doi.org/10.1016/j.ins.2024.120865

[22] Chhatlani, J. et al. 2022. *DermaGenics - Early Detection of Melanoma using YOLOv5 Deep Convolutional Neural Networks*. 2022 IEEE Delhi Section Conference (DELCON). pp. 1-6. https://doi.org/10.1109/DELCON54057.2022.9753227

[23] Wiliani, N., et al. 2023. Identifying Skin Cancer Disease Types with You Only Look Once (YOLO) Algorithm. *Jurnal Riset Informatika*. 5(3), 455-464. https://doi.org/10.34288/jri.v5i3.241

[24] Aishwarya, N., et al. 2023. Skin Cancer diagnosis with Yolo Deep Neural Network. *Procedia Computer Science*. vol 220, pp. 651-658. https://doi.org/10.1016/j.procs.2023.03.083

[25] He, K., et al. 2020. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell*. vol. 42, no. 2, pp. 386-397. https://doi.org/10.1109/TPAMI.2018.2844175

[26] Shi, W., et al. 2019. Plant-part segmentation using deep learning and multi-view vision. *Biosyst. Eng*. vol. 187, no. September, pp. 81–95. https://doi.org/10.1016/j.biosystemseng.2019.08.014