Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

June 2026. Vol. 5. No. 3; e-ISSN: 2808-4519

Early Detection of Dermatitis Through Comparison of Image Size Variations Using the You Only Look Once (YOLO) Framework

Septora Ivanda Gabrani Agda[1*], Rudi Heriansyah[2], Zaid Romegar Mair[3]

[1,2,3] *Universitas Indo Global Mandiri*
2021110041@students.uigm.ac.id[1], rudi@uigm.ac.id[2], zaidromegar@uigm.ac.id[3]*

Abstract

Dermatitis is an inflammatory skin disease characterized by symptoms such as redness and itching, requiring early identification to prevent the development of more serious conditions. The use of image processing technology and deep learning is important as a supporting solution in the process of rapid and accurate skin disease detection. This study aims to compare the performance of the You Only Look Once (YOLO) model on several image size variations, evaluate the model's ability to detect types of dermatitis based on precision, recall, and mean Average Precision (mAP) metrics, and determine the optimal number of epochs to improve model performance. The dataset used consisted of 440 images of patients' skin obtained from Dr. Rivai Abdullah General Hospital and augmented to 1,320 images. The data was divided into training, validation, and test data. The YOLOv11 model was trained to detect four types of dermatitis, namely contact dermatitis, atopic dermatitis, static dermatitis, and circumscribed neurodermatitis. The results showed that image size and epoch number affected model performance. The best configuration was obtained with an image size of 640 × 640 pixels and 150 epochs, resulting in a precision value of 0.693 and a recall value of 0.674. These results indicate that the YOLO model has the potential to be used as an effective early identification support system for dermatitis.

***Keywords**: Dermatitis; Deep Learning, Early Disease Detection; Image Size Variation; YOLO*

1. Introduction

Dermatitis is a common inflammatory skin disorder characterized by erythema, pruritus, and irritation of the epidermal layer, affecting individuals across various age groups and representing a significant public health concern worldwide. According to the World Health Organization (WHO), atopic dermatitis affects approximately 204 million people globally, accounting for 2.6% of the world's population [1], while regional health reports consistently identify dermatitis as one of the most prevalent skin diseases treated in public healthcare facilities [2]. Although epidemiological data in Indonesia remain limited, particularly in South Sumatra Province, previous studies have reported that 9.6% of pineapple farmers experience contact dermatitis due to pesticide exposure [3], while atopic dermatitis affects approximately one in five children and one in ten adults in high-income countries [4]. The high prevalence of dermatitis, combined with risk factors such as irritant exposure, inadequate preventive knowledge, and poor personal hygiene, highlights the importance of early identification to prevent disease progression and reduce potential complications. However, the clinical manifestations of dermatitis often overlap with other skin disorders, including psoriasis and tinea, increasing the likelihood of misdiagnosis during initial examinations and emphasizing the need for more systematic and objective diagnostic support tools. In this context, advances in Artificial Intelligence (AI), Computer Vision, and Deep Learning have created new opportunities for automated medical image analysis, enabling computational models to recognize complex visual patterns from skin lesion images with high efficiency and consistency. The availability of digital skin images obtained from RSUP Dr. Rivai Abdullah further supports the development of image-based diagnostic systems, although variations in skin color, lighting conditions, and image resolution remain important challenges that can significantly affect feature extraction and model performance. Among various deep learning approaches, the You Only Look Once (YOLO) framework has emerged as one of the most effective object detection methods due to its ability to perform object localization and classification within a single computational process, resulting in faster inference times [5] than traditional multi-stage detection [6] approaches and enabling real-time implementation [7][8] in medical imaging applications. Previous studies have demonstrated the effectiveness of YOLO-based models for skin disease analysis, including the use of YOLOv5 for skin cancer classification with an accuracy of 89.1% [6], YOLOv8 for skin cancer detection achieving 89% accuracy, 0.975 precision, and 0.969 recall [5], as well as YOLO-based detection of pigmented dermoscopic lesions with an accuracy of 89.7% [7]. Despite these promising results, most existing studies have primarily focused on skin cancer and pigmented lesions, whereas research specifically targeting dermatitis detection remains relatively scarce. Furthermore, the influence of image size variation, a critical factor that directly affects feature representation and extraction within convolutional neural networks, has received limited attention, particularly in the context of the recently developed YOLOv11 architecture. To the best of our knowledge, studies comparing multiple image size configurations in YOLOv11 for dermatitis detection are still limited, creating a significant research gap in understanding how image resolution influences model performance. Therefore, this study investigates the impact of image size variations and training epochs on the performance of the YOLOv11 model for detecting four types of dermatitis, namely contact dermatitis, atopic dermatitis, stasis dermatitis, and circumscribed neurodermatitis. By evaluating model performance across different image resolutions and epoch settings, this research aims to identify the optimal configuration for dermatitis detection and provide practical insights for improving the effectiveness of AI-based skin disease identification systems in clinical settings.

2. Literature Review

Digital image processing is used to extract visual information from images so that it can be analyzed computationally. This process generally involves preprocessing, feature extraction, and object detection stages. The preprocessing stage aims to improve image quality and adapt the data format to the needs of the learning model. Commonly used techniques include noise removal, intensity normalization, and image resizing [9][10]. Variations in image size affect the amount of visual information processed by the model. Low resolution can obscure object details, while high resolution increases computational complexity during training [11]. Therefore, resizing is often combined with data augmentation to enrich image variations and improve the model's generalization ability.

Deep learning approaches are widely applied in image analysis because they can automatically extract features. Convolutional Neural Networks (CNNs) have become the primary architecture in various computer vision systems across medical applications [12][13] because they can learn spatial patterns in images through convolution and pooling operations [14]. Advances in CNN-based object detection methods have yielded two main approaches: two-stage detectors and one-stage detectors. The two-stage method separates the processes of region proposal generation and object classification, resulting in high accuracy but requiring greater computational time. Conversely, the one-stage approach performs detection and classification simultaneously, making it more efficient for real-time applications [15].

You Only Look Once (YOLO) is a one-stage object detection algorithm widely used in various computer vision applications [16]. This method divides the input image into several grids and predicts bounding boxes and object class probabilities for each grid. The agreement between predicted bounding boxes and ground truth is measured using Intersection over Union (IoU), while duplicate detections are filtered using Non-Maximum Suppression (NMS) [17]. The performance of object detection models is typically evaluated using the mean Average Precision (mAP) metric as well as other classification metrics such as accuracy, precision, recall, and F1-Score, which are calculated via the Confusion Matrix [18].

3. Methodology

3.1. Research Stages

This study develops an early detection system for dermatitis based on skin images using the YOLOv11 object detection model, which is built using a CNN architecture. This approach allows for simultaneous object localization and classification in a single detection stage, resulting in a more efficient inference process.The research workflow is illustrated in a flowchart, as shown in Fig. 1.

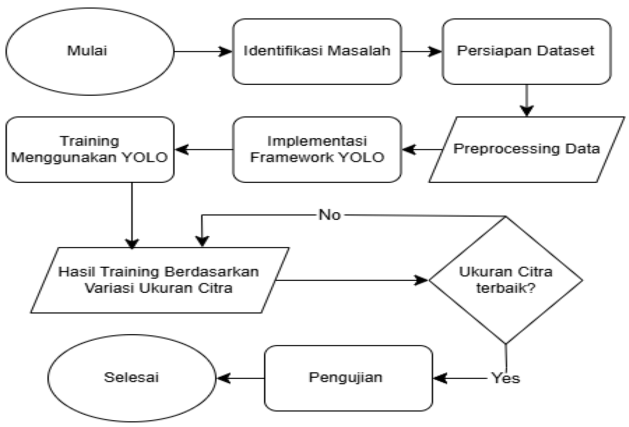

Fig. 1: Flowchart Research Phases

3.2. Problem Identification

Problem identification was conducted through a literature review by examining and analyzing various books, articles, and scientific journals related to image-based skin disease detection, particularly dermatitis, as well as the concepts of computer vision, deep learning, and CNNs. This study aims to identify the methodological approaches used in previous research and to understand the challenges in the process of automatic skin disease detection. The results of this literature review are used as a basis for formulating an appropriate methodological approach, including the application of a YOLO-based object detection model to develop an early dermatitis identification system.

3.3. Dataset Preparation

The research dataset was obtained from medical records at Dr. Rivai Abdullah General Hospital and consists of 440 images of dermatitis, including contact dermatitis, atopic dermatitis, static dermatitis, and circumscribed neurodermatitis. The dataset was then expanded using data augmentation techniques, bringing the total number of images to 1,320. Next, the data was divided into three parts with a ratio of 70:20:10, namely 924 images as training data, 264 images as validation data, and 132 images as test data. The training data was used to train the model to recognize visual patterns of the disease, the validation data was used to monitor the model's performance during the training process, while the test data was used to evaluate the model's ability to make predictions on new, previously unseen data.

3.4. Data Pre-processing

The preprocessing stage was conducted to standardize the quality and format of the images before they were used in the model training process. This process included adjusting the image orientation (auto-orientation) and resizing the images so that all images had uniform orientation and dimensions, specifically 540×540, 640×640, and 740×740 pixels. Additionally, an annotation process is performed to mark skin areas affected by dermatitis using the bounding box technique. This study uses four disease classes: contact dermatitis, atopic dermatitis, static dermatitis, and circumscribed neurodermatitis. The labeling process was performed using the Roboflow platform, and the annotation results were exported in a format compatible with YOLOv11 as .txt files containing class identifiers, object center coordinates ((x,y)), and relative bounding box dimensions ((w,h)).

In addition to preprocessing, data augmentation techniques were also applied to increase the variation in training images. The augmentation techniques used included rotation, flipping, cropping, and shearing. This process generated new image variations from the original data, thereby improving the model's generalization ability. After augmentation, the dataset size increased from 440 images to 1,320 images, comprising training, validation, and test data.

4. Results and Discussion

4.1. Dataset

The dataset consists of 440 images, evenly divided into four classes: contact dermatitis, atopic dermatitis, circumscribed neurodermatitis, and static dermatitis. The dataset was then divided into three subsets: training data, validation data, and test data, with a split ratio of 70:20:10. The complete breakdown of the dataset for each class is presented in Table 1.

Table 1: Dataset Distribution for Each Class

Kelas	Total Data	Total Data (Augmentasi)	Data Latih	Data Valid	Data Uji
Dermatitis Kontak	110	330	231	66	33
Dermatitis Atopi	110	330	231	66	33
Neurodermatitis Sirkumskripta	110	330	231	66	33
Dermatitis Statis	110	330	231	66	33
Jumlah	**440**	**1320**	**924**	**264**	**132**

Based on the table above, Table 1 shows the distribution of the dataset before and after the augmentation process. Initially, each class had 110 images, resulting in a total dataset of 440 images. After augmentation, the number of images in each class increased to 330, for a total of 1,320 images. Next, the dataset was divided into three subsets: training data, validation data, and test data, in a 70:20:10 ratio, resulting in 924, 264, and 132 images, respectively. This division was performed to support the model training process and enable a more objective performance evaluation.

4.2. Model Training

Once the dataset was ready for use, the next step was to train the model using the YOLOv11 framework to identify the most optimal image size configuration. The training process was conducted online via the Google Colab platform, which utilizes a cloud-based computing environment. At this stage, an exploration of the number of epochs specifically 150 epochs was conducted to observe its impact on model performance. These image size variations were applied to obtain a training configuration capable of producing the best detection performance in identifying dermatitis in skin images. The model was trained using the entire augmented dataset, totaling 1,320 images.

After the training process was completed, a validation stage was conducted to evaluate the model's performance and determine the configuration yielding the most optimal results. Validation was performed using a separate dataset consisting of 264 images that were not used during the training process and had therefore never been learned by the model previously. This validation dataset included 284 objects divided into four categories of dermatitis on human skin. The details of the number of objects in each class within the validation data are presented in Table 2.

Table 2: Breakdown of Validation Data Object Counts by Class

No	Kelas	Jumlah Gambar	Jumlah Objek
1	Kontak	66	74
2	Atopik	66	69

3	Neurodermatitis Sirkumskripta	66	67
4	Statis	66	74
	Total	**264**	**284**

Table 2 shows the distribution of the valid data that will be used to evaluate the performance of the YOLOv11 model across three image size variations: 540×540 pixels, 640×640 pixels, and 740×740 pixels. Each dermatitis class has the same number of images, namely 66 images, but the number of objects contained varies between classes, ranging from 67 to 74 objects. This variation in the number of objects occurs because some images contain more than one dermatitis lesion, so even though the number of images is uniform, the representation of objects in each class is not entirely identical.

These results are further visualized in Fig 2, which displays a comparison of model performance based on evaluation metrics for each image size.

4.3. Model Training Results

The YOLOv11 model was trained using three image resolution variations: 540×540, 640×640, and 740×740 pixels, with the dataset split into 70% training data and 20% validation data. A comparison of model performance across these three resolutions is presented in Fig 2, illustrating the impact of image size variations on the model's accuracy and consistency in object detection. Model performance was evaluated using the metrics precision, recall, mAP@50, mAP@50–95, and the confusion matrix, which are visualized in Fig 3 through 8.

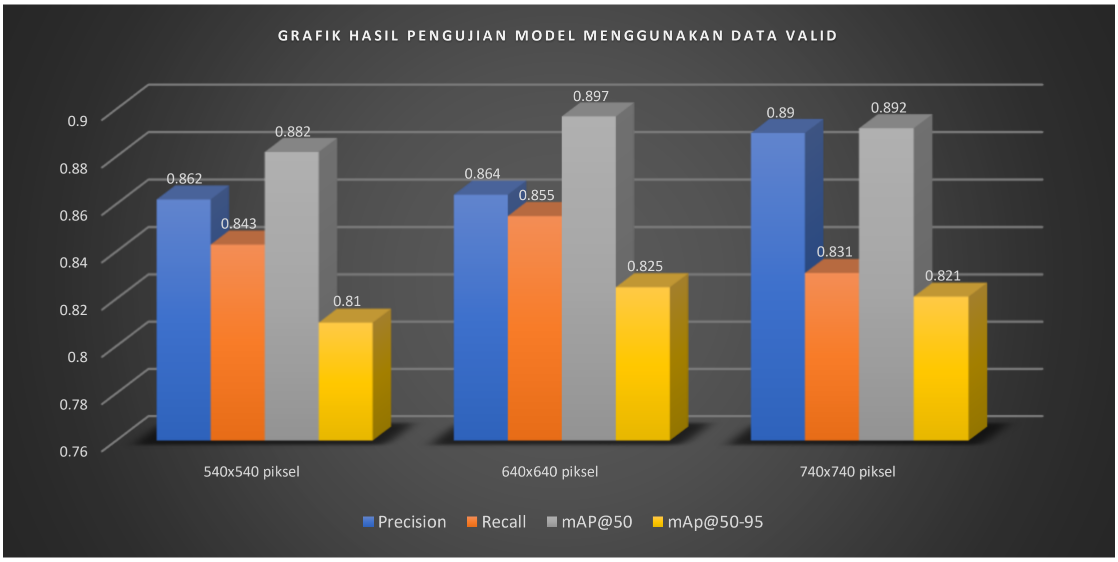

Fig 2: Model Testing Results Using Validation Data

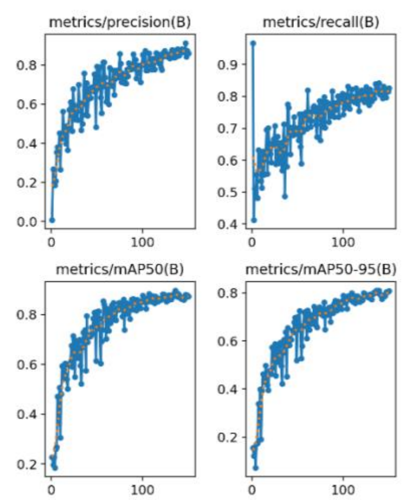

Fig.3: Training Results Graph at 540x540 Pixels

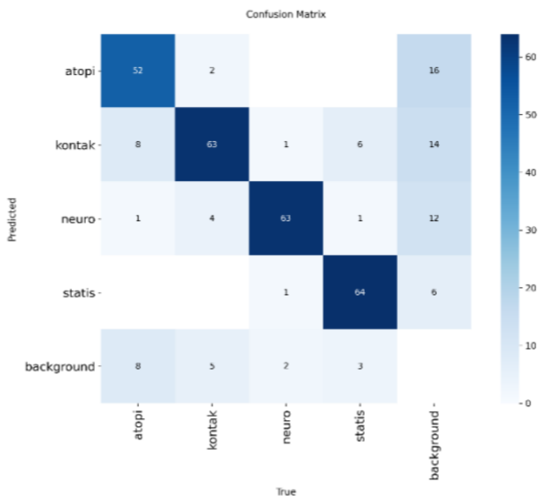

Fig.6: Confusion Matrix at 540x540 Pixels

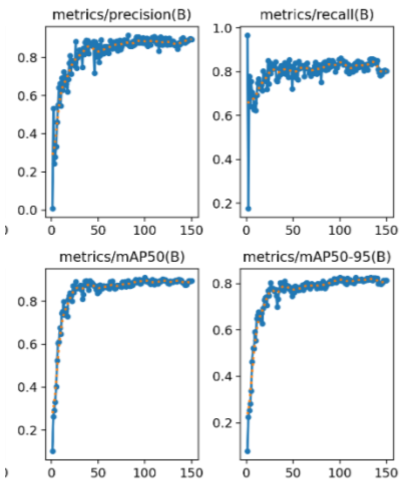

Fig.4: Training Results Graph at 640x640 Pixels

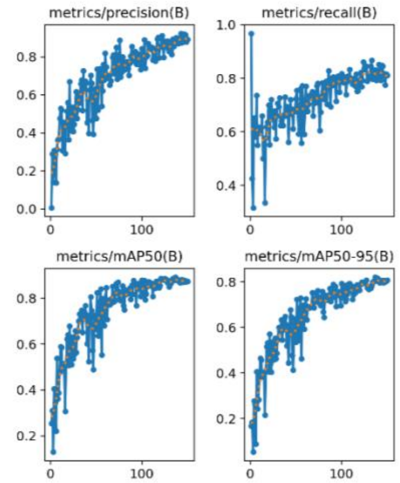

Fig.5: Training Results Graph at 740x740 Pixels

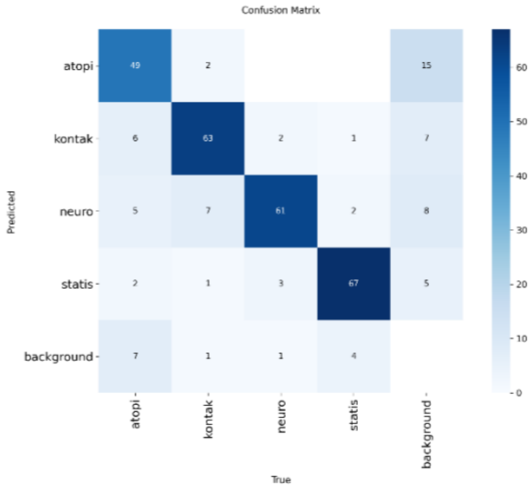

Fig.7: Confusion Matrix at 640x640 Pixels

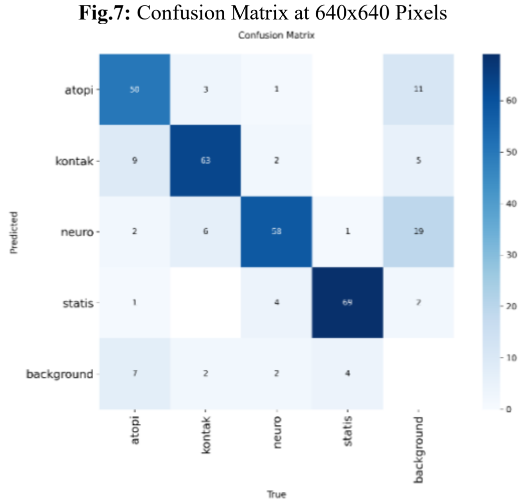

Fig.8: Confusion Matrix at 740x740 Pixels

4.4. Model Performance Comparison

As shown in Fig. 2, it is evident that an increase in image size does not always correlate directly with an improvement in overall performance. The 640×640-pixel model demonstrates an optimal balance between detection accuracy and training time efficiency, with a precision of 0.864, recall of 0.855, mAP@50 of 0.897, and mAP@50–95 of 0.825, although the training time is slightly longer compared to the 540×540-pixel size.

Meanwhile, the 540×540-pixel image size yields the fastest training time 0.794 hours with performance metrics of precision 0.862, recall 0.843, mAP@50 0.882, and mAP@50–95 0.810. Although the detection metrics are slightly lower than those for the 640×640-pixel size, this configuration still demonstrates the model's ability to recognize objects with reasonable accuracy.

For the 740×740-pixel image size, although the highest precision of 0.890 was achieved, the recall value of 0.831 decreased slightly compared to the 640×640-pixel configuration, indicating that the model's ability to detect all objects did not increase proportionally. The mAP@50 value of 0.892 and mAP@50–95 of 0.821 indicate good detection performance; however, the training time of 0.848 hours is relatively longer compared to the 540×540-pixel size and nearly equivalent to the 640×640-pixel size.

Overall, this comparison demonstrates that the 640×640-pixel image size offers the best balance between prediction accuracy, comprehensive object detection capability, and training time efficiency. These results underscore the importance of selecting a balanced image size to optimize the performance of object detection models in the context of automated dermatitis identification.

To support the analysis, a confusion matrix at the 640×640-pixel resolution was used to evaluate the patterns of successful and incorrect predictions during the training process. This matrix illustrates the model's ability to distinguish and classify the four types of dermatitis, including the number of correct predictions as well as classification errors for each class and the background class. Based on the visualization in Fig 7, it is evident that the model can recognize all classes with adequate accuracy. Specifically, the Static Dermatitis and Circumscribed Neurodermatitis classes exhibit relatively higher performance compared to other classes, with the number of true positives reaching 67 and 61 objects, respectively.

4.5. Model Evaluation on the Test Dataset

After undergoing the training and validation processes, the model demonstrated the best performance when using a 640×640-pixel image size compared to other image size configurations. Based on this configuration, the subsequent testing phase was conducted to evaluate the model's ability to detect and classify types of dermatitis on the test dataset, which had not been used during training or validation. The test dataset consists of 132 images obtained from Dr. Rivai Abdullah General Hospital, with a total of 139 objects divided into four dermatitis classes, consistent with the class division in the training and validation data. The detailed number of objects in each class for the test dataset is presented in Table 3

Table 3: Breakdown of the Number of Test Data Objects by Class

No	Kelas	Jumlah Gambar	Jumlah Objek
1	Kontak	33	38
2	Atopik	33	34
3	Neurodermatitis Sirkumskripta	33	33
4	Statis	33	34
	Total	**132**	**139**

Although the number of images in Table 4 is consistent for each class, the number of successfully detected objects varies. This variation arises because some test images contain more than one lesion, so the number of objects per image is not always the same. The evaluation metrics are presented in Fig 11, while the Confusion Matrix and examples of predictions on the test data are visualized in Fig 10 and 9 to provide a comprehensive overview of the model's performance.

Fig.9: Sample Training Data

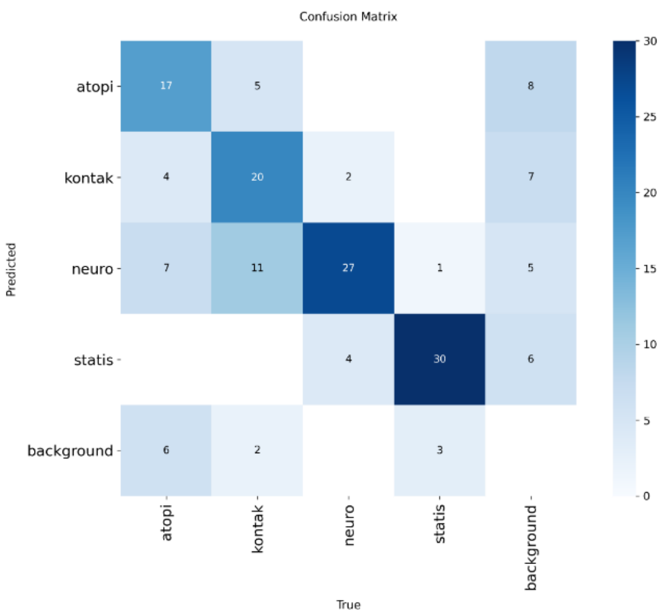

Fig.10: Confusion Matrix Using Test Data

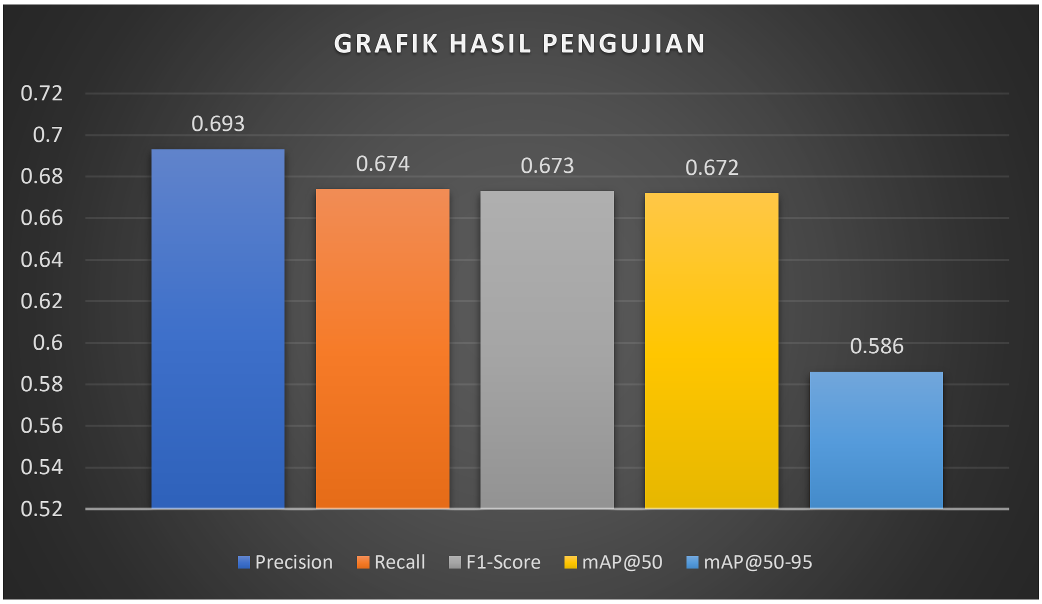

Fig.11: Model Testing Results Using Test Data

As shown in Fig 11, YOLOv11 achieved an average precision of 0.693, a recall of 0.674, and an F1-score of 0.673, with an mAP@50–95 of 0.586, indicating that the bounding boxes lack precision at strict IoU thresholds. Confusion matrix analysis in Fig 10 reveals that Static Dermatitis and Circumscribed Neurodermatitis are detected most optimally, while Contact Dermatitis and Atopic Dermatitis have low recall due to visual similarity between classes, an unbalanced training data distribution, and image variability. These results confirm that the model is capable of detecting all four types of dermatitis with reasonably good average metrics; however, performance varies across classes, necessitating an expanded dataset, class-based data augmentation, and training optimization.

5. Conclusion

Based on the results of the study, it can be concluded that variations in image size affect the performance of the YOLO model, where the use of 640×640-pixel images yields the most optimal and consistent performance compared to 540×540-pixel and 740×740-pixel images, as evidenced by the balance between precision and recall values and the highest Mean Average Precision (mAP) score of 0.897.

Additionally, the evaluation results for detecting types of dermatitis on the test data show that the YOLO model is capable of delivering fairly good performance with an average precision of 0.693, recall of 0.674, mAP@50 of 0.672, and mAP@50–95 of 0.586, although there are still performance differences across dermatitis classes.

Based on testing with varying numbers of epochs, it was found that using 150 epochs yields the most optimal number of epochs for improving average evaluation metrics. However, the resulting model performance is not yet fully optimal for application in real-world medical settings, likely due to differences in distribution between the training and test data, as well as variations in lighting conditions, backgrounds, and image capture angles. Therefore, further development is required, particularly regarding the improvement of data quality and diversity, so that the YOLO-based dermatitis detection model can be used more effectively and reliably in diverse medical settings.

References

[1] S. S. Malik *et al.*, “Adolopment Of Atopic Dermatitis Management Guidelines For Pakistan,” *East. Mediterr. Heal. J.*, vol. 31, no. 9–10, pp. 552–559, 2025, doi: 10.26719/2025.31.10.552.

[2] E. Heryanto, S. Sarwoko, F. Meliyanti, D. Prodi, K. Masyarakat, and S. Al-Ma’arif Baturaja, “Faktor Resiko Dermatitis Pada Anak,” *J. Kesehat. Abdurahman Palembang*, vol. 11, no. 1, pp. 10–16, 2022.

[3] P. Asih, “Faktor- Faktor Yang Berhubungan Dengan Kejadian Gejala Dermatitis Kontak Iritan Pada Petani Nanas Pengguna Pestisida Di Desa Tangkit Baru Kecamatan Sungai Gelam Kabupaten Muaro Jambi Tahun 2023,” Universitas Jambi, 2024. [Online]. Available: https://repository.unja.ac.id/id/eprint/62378

[4] S. M. Abuabara, K., Yu, A. M., Okhovat, J. P., Allen, I. E., & Langan, “The Prevalence Of Atopic Dermatitis Beyond Childhood: A Systematic Review And Meta-Analysis Of Longitudinal Studies,” *Allergy*, vol. 73, pp. 696–704, 2018, doi: 10.1111/all.13320.

[5] Z. R. Mair and M. A. Rahmanda, “Perbandingan Versi Terbaik YOLO Dalam Mendeteksi Jarak Spasi Antar Baris Tulisan Tangan,” *J. Sains, Nalar, dan Apl. Teknol. Inf.*, vol. 4, no. 2, pp. 103–110, 2025, doi: 10.20885/snati.v4.i2.40414.

[6] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” 2018, [Online]. Available: http://arxiv.org/abs/1804.02767

[7] Tanadi, Kartika, and Najaf, “Sistem Pendeteksi Penyakit Kanker Kulit Menggunakan Convolutional Neural Network Arsitektur YOLOv8 Berbasis Website,” *Repeater Publ. Tek. Inform. dan Jar.*, vol. 2, no. 3, pp. 166–177, 2024, doi: 10.62951/repeater.v2i3.124.

[8] 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, 2021, doi: 10.1016/j.procs.2022.01.135.

[9] M. Widyaningsih and A. Harjoko, “Identifikasi Gejala Penyakit Tanaman Jeruk Melalui Pengolahan Citra,” *J. Sains Komput. dan Teknol. Inf.*, vol. 3, no. 2, pp. 104–113, 2021, doi: 10.33084/jsakti.v3i2.2294.

[10] Supiyandi Supiyandi, Mona Donaon, and Muhammad Yusuf Azmi, “Analisis Sistem Aplkasi Pengolahan Citra Pada Pertanian Cerdas Untuk Pemantauan Tanaman,” *SABER J. Tek. Inform. Sains dan Ilmu Komun.*, vol. 2, no. 3, pp. 221–228, 2024, doi: 10.59841/saber.v2i3.1443.

[11] M. M. Sebatubun and C. Haryawan, "Implementasi Algoritma Convolutional Neural Network untuk Klasifikasi Jenis Keris," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 3, pp. 595–602, 2024, doi: 10.25126/jtiik.937260.

[12] Nurohman, Rudi Heriansyah, Dwi Asa Verano, and Zaid Romegar Mair, "Deteksi Penyakit Diabetes Retinopathy Menggunakan Citra Digital Dengan Metode Convolutional Neural Network (Cnn)," *Pros. Snast*, no. November, pp. 311–320, 2024, doi: 10.34151/prosidingsnast.v1i1.5120.

[13] Z. R. Mair *et al.*, "An Enhanced Deep Learning Framework for Diabetic Retinopathy Classification Using Multiple Convolutional Neural Network Architectures," *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 11, pp. 769–776, 2025, doi: 10.14569/IJACSA.2025.0161176.

[14] T. B. Sasongko, H. Haryoko, and A. Amrullah, "Analisis Efek Augmentasi Dataset dan Fine Tune pada Algoritma Pre-Trained Convolutional Neural Network (CNN)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 4, pp. 763–768, 2023, doi: 10.25126/jtiik.20241046583.

[15] D. Nafis Alfarizi, R. Agung Pangestu, D. Aditya, M. Adi Setiawan, and P. Rosyani, "Penggunaan Metode YOLO Pada Deteksi Objek: Sebuah Tinjauan Literatur Sistematis," *J. Artif. Intel. dan Sist. Penunjang Keputusan*, vol. 1, no. 1, pp. 54–63, 2023, [Online]. Available: https://jurnalmahasiswa.com/index.php/aidanspk

[16] R. Supriyadi *et al.*, "Pengembangan Aplikasi Estimasi Kalori Makanan Berbasis Citra Dengan Pendekatan Deteksi Objek Menggunakan Yolo," *J. Inform. dan Tek. Elektro Terap.*, vol. 14, no. 1, pp. 1393–1404, 2026, doi: 10.23960/jitet.v14i1.8545.

[17] S. Nugroho, M. Kahfi, M. Alamsyah, A. Natanael, and P. Rosyani, "MENDETEKSI JENIS KENDARAAN di JALAN," vol. 2, no. 2, pp. 127–131, 2024.

[18] F. T. Hidayat and A. K. Whardana, "Deteksi Pelanggaran Sepeda Motor Menggunakan Algoritma Yolo Dan Mean Average Precision," vol. VIII, no. September, pp. 71–79, 2024.