A Hybrid YOLO–CNN Model for Automatic Detection and Severity Assessment of Atopic Dermatitis in Infant Images

Debi Setiawan[1], Ramalia Noratama Putri[2,*], Sara Herlina[3], Achmad Nizar Hidayanto [4], Yuda irawan[5], Naohiro Hohashi[6]

[1]*Department of Informatics Engineering, Faculty of Engineering, Abdurrab University, Pekanbaru, Indonesia*

[2]*Department of information Systems, Faculty of computer science, Indonesian Pelita Institute of Business and Technology, Pekanbaru, Indonesia*

[3]*Department of midwifery, Faculty of Pharmacy and Health Science, Abdurrab University, Pekanbaru, Indonesia*

[4]*Department of Computer Science, Faculty of Computer Sciences, Universitas Indonesia, Jakarta, Indonesia*

[5]*Department of Computer Science, Faculty of Computer Sciences, Universitas Hang Tuah Pekanbaru, Pekanbaru, Indonesia*

[6]*Department of Family and Community Health Nursing, Graduate School of Health Sciences, Kobe University, Kobe, Japan*

(Received: November 5, 2025; Revised: January 5, 2026; Accepted: March 18, 2026; Available online: April 4, 2026)

Abstract

Atopic dermatitis is one of the most common skin diseases affecting infants and children worldwide and has a particularly high prevalence in tropical countries. Traditional diagnosis methods, which still rely on physical examinations and laboratory tests often face challenges such as delays, high costs, and limited facilities, thereby necessitating an artificial intelligence–based system that is more efficient and accurate. This study aims to develop a hybrid YOLO–CNN model for the automatic detection and severity classification of atopic dermatitis in infants. The dataset comprises 2,000 infant skin images, including lesions categorized as mild, moderate, and severe, obtained from an online repository and field observations conducted in three villages. The labeling process was performed by a specialist doctor to ensure clinical validity. In the first step, YOLO was used to detect the lesion area in real time by generating a bounding box. This produced a region of interest (ROI), which was subsequently analyzed by a CNN model employing transfer learning in the second step to determine the severity level. Experimental results indicate that YOLO achieved high detection performance, with an mAP@0.5 of 91.2% and an F1-score of 90.2%, while the CNN model attained an average accuracy of 85% and a macro-F1 score of 85% in classification. The visualization of predictions indicates that most lesions were detected with confidence levels ≥0.9, confirming the model's consistency. These findings highlight the potential of the hybrid YOLO–CNN framework as a supportive system for digital clinical diagnosis, applicable to both mobile applications and teledermatology services, particularly in regions with limited medical personnel. Future research should employ larger, multi-center datasets and integrate explainable AI approaches to promote broader clinical adoption.

Keywords: Atopic Dermatitis, Infant, CNN, Skin Lesion Detection, Severity Classification, YOLO

1. Introduction

Autopic Dermatitis is a common skin disease that affects infants worldwide [1]. It may result from multiple factors, including allergies, irritation, genetic predisposition, and environmental conditions. Among its various forms, atopic dermatitis is the most prevalent type in infants. The incidence of atopic dermatitis in infants is notably high, particularly in hot and humid tropical climates. Epidemiological studies report that the prevalence of atopic dermatitis among children in Southeast Asia ranges between 10% and 20%, with Indonesia exhibiting similar estimates according to regional pediatric dermatology surveys [2], [3]. These findings confirm that atopic dermatitis in infants and children is a significant global health concern [4]. Furthermore, atopic dermatitis remains among the twelve most frequently occurring skin diseases worldwide [5].

***Corresponding author: Ramalia Noratama Putri (ramalia.noratamaputri@lecturer.pelitaindonesia.ac.id)**
DOI: https://doi.org/10.47738/jads.v7i2.1212

The impact of atopic dermatitis in infants extends beyond physical symptoms such as itching, rashes, and inflammation, as it may also affect sleep quality and overall comfort. Infancy is often referred to as the "golden age," representing a critical developmental stage from birth to 12 months [6]. Therefore, if not properly managed, recurrent dermatitis can increase the risk of secondary infections [7], long-term skin disorders, and developmental delays [8], [9]. Early diagnosis of dermatitis in infants is crucial as a preventive measure for both parents and healthcare professionals [10], [11].

The diagnosis of dermatitis in infants is still predominantly performed using traditional methods, primarily through physical examinations conducted by physicians based on clinical symptoms, lesion distribution, family history, and daily habits [12], [13]. When necessary, physicians may conduct additional tests, such as allergy testing or immunoglobulin E (IgE) blood analysis [14]. However, traditional diagnostic approaches are often inaccurate or delayed. Dermatitis, for instance, is frequently misdiagnosed as other skin conditions such as miliaria (commonly known as heat rash) or skin infections [15], [16]. Furthermore, allergy tests such as the skin prick test and IgE blood test remain relatively expensive and are not available in all healthcare facilities, resulting in many infants in rural or remote areas failing to receive accurate diagnoses. Therefore, there is an urgent need to develop innovative and accurate diagnostic systems utilizing deep learning models, particularly hybrid YOLO and CNN architectures [17], [18], [19].

One of the most promising approaches in developing artificial intelligence, based diagnostic systems is the application of You Only Look Once (YOLO), a deep learning model designed for object detection [20], [21], [22], [23]. Several previous studies have demonstrated that YOLO achieves outstanding performance in the rapid and precise detection of skin lesions [24]. For example, research has shown that YOLO-based models can identify skin abnormalities in real time with high sensitivity [25], [26], while other studies have confirmed the model's effectiveness in detecting various types of dermatological conditions in clinical images [27]. These findings suggest that YOLO holds strong potential for adaptation in detecting atopic dermatitis in infants.

The YOLO model can be trained using an annotated dataset of infant skin images to automatically detect and localize areas affected by inflammation due to atopic dermatitis. The main advantages of YOLO lie in its remarkable speed and accuracy in real-time detection, making it an excellent option for mobile applications and digital clinical services. After the lesion area is successfully identified, a Convolutional Neural Network (CNN) can be employed for further analysis to extract visual features from the affected region and classify the severity of atopic dermatitis into mild, moderate, or severe categories. Previous studies have also demonstrated the effectiveness of CNN in the classification of dermatological conditions, showing that CNNs can achieve accuracy levels comparable to dermatologists in evaluating clinical skin images [28], [29].

Although previous studies have explored hybrid deep learning architectures combining detection and classification models, these approaches primarily focus on adult dermatological datasets or perform severity prediction using full-image classification without explicit lesion localization. In contrast, the present study is specifically designed for infant skin images and integrates lesion-level localization with severity assessment in a unified pipeline, addressing the unique visual and clinical characteristics of infant atopic dermatitis [29], which primarily focused on the detection or classification of skin diseases in general, this study presents a more comprehensive approach that not only automatically identifies lesion areas but also evaluates the severity level. The emphasis on infants, whose skin is more sensitive and clinically distinct, can add uniqueness and significance to this study, contributing to the advancement of more specific, accurate, and clinically relevant AI-based diagnostic systems. The integration of YOLO for lesion detection and CNN for severity classification provides a holistic diagnostic framework that can accelerate early diagnosis, assist healthcare professionals in timely intervention, and ultimately improve the quality of care for infants with atopic dermatitis.

2. Research Methodology

The method proposed in this study integrates two deep learning models: You Only Look Once (YOLO) for skin lesion detection and a Convolutional Neural Network (CNN) for classifying the severity of atopic dermatitis in infants. The overall system is structured into two main stages. The following is figure 1 of the proposed method in this study:

ISSN 2723-6471

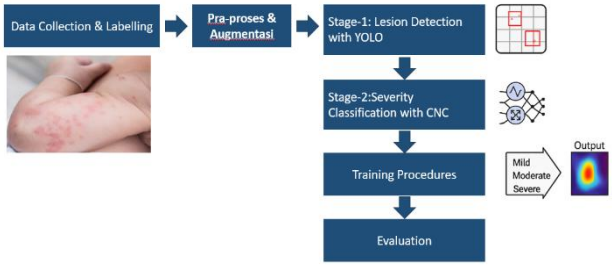


Figure 1. The method proposed in the research

In the first stage, YOLO was trained on a dataset of infant skin images annotated with bounding boxes by medical professionals. This training enabled YOLO to autonomously identify and label inflamed skin regions. The primary strength of YOLO lies in its high speed and accuracy in real-time detection, making it highly suitable for implementation in mobile applications and digital clinical services. Each detected lesion region was subsequently cropped (Region of Interest – ROI), padded, and normalized to prepare it for further analysis.

In the second stage, a Convolutional Neural Network (CNN) was employed to extract visual features from the Regions of Interest (ROIs) detected by YOLO. A transfer learning-based CNN architecture, such as EfficientNet or ResNet, was utilized due to its proven effectiveness in medical image classification. The CNN was trained to classify the severity of atopic dermatitis into three categories: mild, moderate, and severe. The class imbalance in the severity classification stage with address, a weighted cross-entropy loss function was consistently applied throughout all experiments. Preliminary comparisons with focal loss were conducted during pilot experiments; however, weighted loss demonstrated more stable convergence and comparable classification performance on the validation set, and was therefore selected as the final optimization strategy. This combined method offers novelty in its focus on infants, an age group with different skin characteristics than adults, as well as full integration between detection and severity classification. Unlike previous studies [30],[31] that primarily concentrated on a singular aspect, this approach offers more comprehensive information, ranging from the automatic identification of disease locations to the assessment of severity levels. The developed model not only accelerates the diagnostic process but also holds significant potential to support healthcare professionals in conducting early interventions and improving the quality of care for infants affected by atopic dermatitis.

2.1. Collecting and Labeling Data

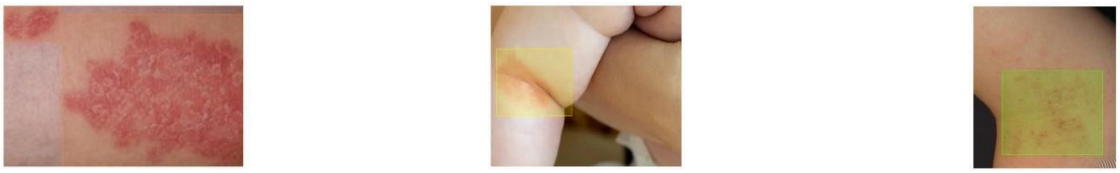
The data collection process in this study was conducted systematically to obtain a representative dataset for training and evaluating the YOLO–CNN model. A total of 2,000 images were collected, consisting of infant skin images diagnosed with atopic dermatitis and healthy skin images used as controls. The dataset was compiled from two primary sources, namely online repositories and direct field observations.

Data collection was carried out from two primary sources: online repositories and field observations, between July 19, 2024, and August 31, 2025, resulting in a total of 2,000 images. Online data were obtained from medical image repositories, research publications, and open-access dermatology databases, following strict selection criteria based on image resolution, lighting conditions, and skin area clarity. This process yielded approximately 1,000 images (50% of the total dataset). The remaining 1,000 images (50%) were gathered through field observations conducted in Desa Pulau Gadang (Kecamatan XII Koto Kampar), Desa Labuhan Tangga Hilir (Kabupaten Rokan Hilir), and Desa Teluk Kenidai in Indonesia. Specifically, 350 images were collected from Pulau Gadang, 350 from Labuhan Tangga Hilir, and 300 from Teluk Kenidai. Field images were collected from three rural villages in Indonesia—Pulau Gadang, Labuhan Tangga Hilir, and Teluk Kenidai—which share relatively similar socio-environmental characteristics. As a result, the dataset may reflect limited demographic variability, including relatively homogeneous skin tone distributions and comparable natural lighting conditions during image acquisition. The dataset includes images of infants with atopic dermatitis at varying levels of severity, as well as images of healthy infant skin for comparison. All field data collection was conducted with informed consent from parents or legal guardians and in accordance with established health research ethics standards.

 ISSN 2723-6471

The next step involved data labeling. Two independent annotators, a dermatologist and an obstetrician, reviewed and labeled each image. Labels for the detection stage consisted of bounding boxes marking the lesion areas on the infant's skin, while labels for the classification stage represented the severity levels of atopic dermatitis, categorized as mild, moderate, and severe. In cases where discrepancies occurred between the annotators, a resolution process was carried out through consensus between both specialists to ensure the reliability and clinical validity of the labeled data.

Through this procedure, the resulting dataset not only possesses high visual quality but also strong clinical validity, enabling its optimal use for training and evaluating the YOLO and CNN model in detecting and classifying the severity of atopic dermatitis in infants. Figure 2 illustrates an example of labeling the severity of atopic dermatitis in infants, categorized into mild, moderate, and severe cases used in this study.



(a) (b) (c)

Figure 2. Labeling (a) Mild (b) Moderate (c) Severe

2.2. Pre-processing and Augmentation

Image pre-processing included resizing and pixel normalization to ensure consistent input distribution. During training, data augmentation was applied using geometric transformations and color jitter to simulate variations in orientation and illumination while preserving lesion characteristics. During the pre-processing stage, all images were adjusted in terms of color and illumination to minimize exposure differences that might occur during image capture. The images were then resized to a uniform resolution of 640×640 pixels for YOLO detection and 224×224 pixels for CNN classification. Pixel normalization was also applied by scaling the image intensity to a [0–1] range to stabilize the data distribution. After YOLO identified the lesion area, the region of interest (ROI) was cropped with an additional padding of approximately 10–15% to ensure that the inflamed area was fully represented.

Furthermore, the augmentation process was applied to increase data variability and prevent overfitting. The augmentation techniques included geometric transformations such as horizontal and vertical flipping, small rotations of ±10°, translation, and zooming in or out. Adjustments in brightness, contrast stretching, and color jitter were also applied to simulate variations in lighting conditions during image capture. In addition, mild Gaussian noise was added to make the model more robust against visual noise. Random cropping and scaling were used to introduce diversity in the framing of skin regions. During the experimental phase, advanced augmentation methods such as MixUp and CutMix were selectively implemented to enhance data diversity without compromising the clarity of the lesion areas. Through this series of procedures, the resulting dataset became cleaner, more balanced, and more representative, thereby improving the performance of YOLO in detecting lesions and CNN in classifying the severity of atopic dermatitis in infants. Below is figure 3 which shows the results of Preprocessing and Augmentation:

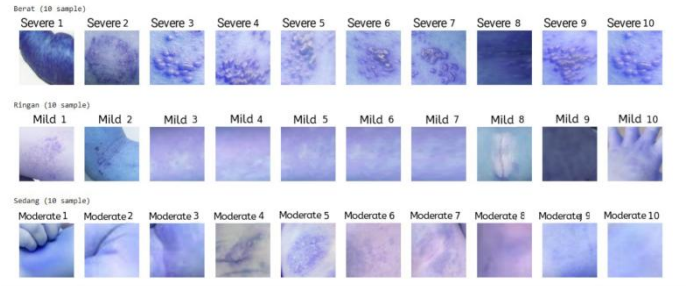


Figure 3. Pre-processing & Augmentation

2.3. Find the Lesion with YOLO

In this study, the role of YOLO is strictly limited to lesion localization. The model is trained to detect and delineate regions of inflamed skin by generating bounding boxes, without performing lesion severity or normal-versus-lesion classification. All subsequent classification tasks are handled exclusively by the convolutional neural network using the detected regions of interest. In this stage, YOLO architecture was employed to identify skin lesions. YOLO is capable of detecting objects in real time with high accuracy. The model was trained using a dataset of infant skin images annotated with bounding boxes around the inflamed areas. During training, YOLO extracted essential image features through its backbone network and generated bounding box predictions, confidence scores, and lesion classifications via the detection head. The optimization process utilized a composite loss function consisting of bounding box regression, objectness, and classification components. This procedure enabled the model to accurately locate and distinguish between infected and normal skin regions. The detection outputs, in the form of bounding box coordinates, were cropped with an additional 10–15% padding to obtain the Region of Interest (ROI), which was subsequently used in the severity classification stage with CNN.

2.4. Severity Classification with CNN

The classification of atopic dermatitis severity was performed using a Convolutional Neural Network (CNN) that received input from the Region of Interest (ROI) identified by YOLO. The CNN extracted key visual features such as texture, color distribution, and inflammation patterns on infant skin, and subsequently classified the images into three severity categories: mild, moderate, and severe. The CNN training process employed a transfer learning-based architecture, such as ResNet or EfficientNet, with weight adjustments applied to the final layer. To prevent overfitting, techniques including dropout, batch normalization, and data augmentation were implemented. To address class imbalance, the model was optimized using a weighted loss function or focal loss, resulting in more stable and consistent classification outcomes. The output of this stage was a severity label for atopic dermatitis, which was then integrated with YOLO detection results. This integration enabled the system to provide a comprehensive diagnosis of the infant's skin condition.

2.5. Training Procedures

The training process was conducted in stages, beginning with the separate training of the YOLO and CNN models, followed by their integration in the inference pipeline. The YOLO model was trained first using a dataset of infant images annotated with bounding boxes around the lesion areas, with a batch size of 16–32 and 150 training epochs. The AdamW or SGD optimizer was applied, along with an early stopping mechanism to prevent overfitting. Once the YOLO model achieved optimal detection performance, the detected Regions of Interest (ROIs) were used as input for CNN training. The CNN model was trained for 100 epochs with a batch size of 32–64 using transfer learning from a Conv2D-based architecture. To enhance generalization, cosine decay for the learning rate, dropout regularization, and image augmentation techniques were employed. To address class imbalance, class weighting or focal loss was applied, allowing the model to consistently classify the three severity levels of dermatitis: mild, moderate, and severe. Validation was performed using a patient-based k-fold cross-validation scheme to prevent data leakage, and the model was subsequently evaluated on a hold-out test set to assess overall system performance.

2.6. Evaluation

The model evaluation was performed to objectively measure how well the system could detect and classify atopic dermatitis in infants. During the detection phase with YOLO, performance was evaluated using the mean Average Precision (mAP) metric at various Intersection over Union (IoU) thresholds, as well as precision, recall, and F1-score to ensure a balance between the model's ability to detect lesions and minimize errors. Meanwhile, during the severity classification phase with CNN, performance was assessed using accuracy, precision, recall, macro-F1 score, and Area Under the Curve (AUC) for each category: mild, moderate, and severe. In addition, a confusion matrix was employed to provide a detailed analysis of classification performance by illustrating the distribution of true positives, false positives, true negatives, and false negatives for each severity class [32], [33].

 ISSN 2723-6471

3. Results and Discussion

3.1.Detection Performance using YOLO

This study employs a clinical image dataset comprising various manifestations of skin lesions in infants. This dataset was obtained from visual documentation with varying anatomical locations (hands, arms, legs, and torso), lighting conditions, and different levels of severity. To make it easier to train and test the deep learning model, each image has been given an annotation in the form of a bounding box that marks the area of the lesion, along with a category label (0, 1, 2) that shows how serious the condition is, from normal to severe. The visual representation of this dataset reflects the complexity of clinical variations and highlights the challenges faced by computer vision–based automated detection systems. Following is figure 4 of Annotated Clinical of Skin Lesions in Children:

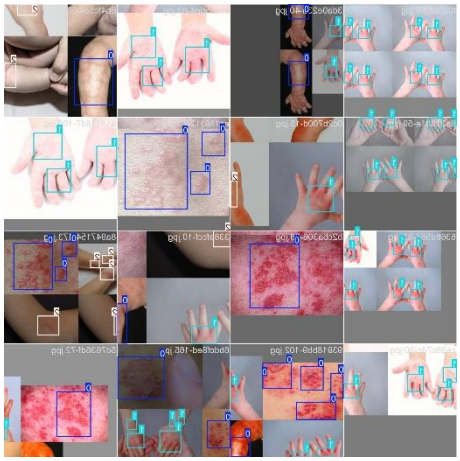


Figure 4. Annotated Clinical Images of Skin Lesions in Children

Figure 4 llustrates the diversity of lesion appearance and anatomical distribution in the infant dataset, high lighting variations in lesion size, color intensity, and body location that challenge automated detection models. The clinical conditions shown include erythema patches, diffuse rashes, papular lesions, and the distribution of lesions on different parts of the body, such as hands, arms, and torso. Double annotation on certain images is used to capture multi-location distribution, which helps train deep learning models to detect more than one lesion at a time. The diversity of data in terms of lighting, skin color, and anatomical position makes this dataset representative of real clinical conditions and serves as an important ground truth in the training, validation, and evaluation processes of computer vision-based automatic detection models. The evaluation graph can be seen in figure 5.

All four evaluation graphs show that the skin lesion detection model works very well for all categories (mild, moderate, and severe). The recall-confidence curve shows that the model's sensitivity stays high until the confidence level reaches about 0.8, when it drops significantly. The precision–recall curve, on the other hand, shows that the mAP@0.5 value reaches 0.991, with almost perfect precision and recall across all classes. The precision-confidence curve shows that predictions with a higher level of confidence have a consistent precision above 0.9, and the F1-confidence curve shows the best balance between precision and recall with an F1 value of 0.98 at the threshold of 0.57. Overall, these results confirm that the model is capable of classifying skin lesions with very high accuracy and consistency across a wide range of testing conditions. In terms of efficiency, the average inference speed of 16 milliseconds per image demonstrates that YOLO is not only accurate but also capable of real-time deployment, making it suitable for both digital clinical systems and mobile applications for infant health monitoring.

ISSN 2723-6471

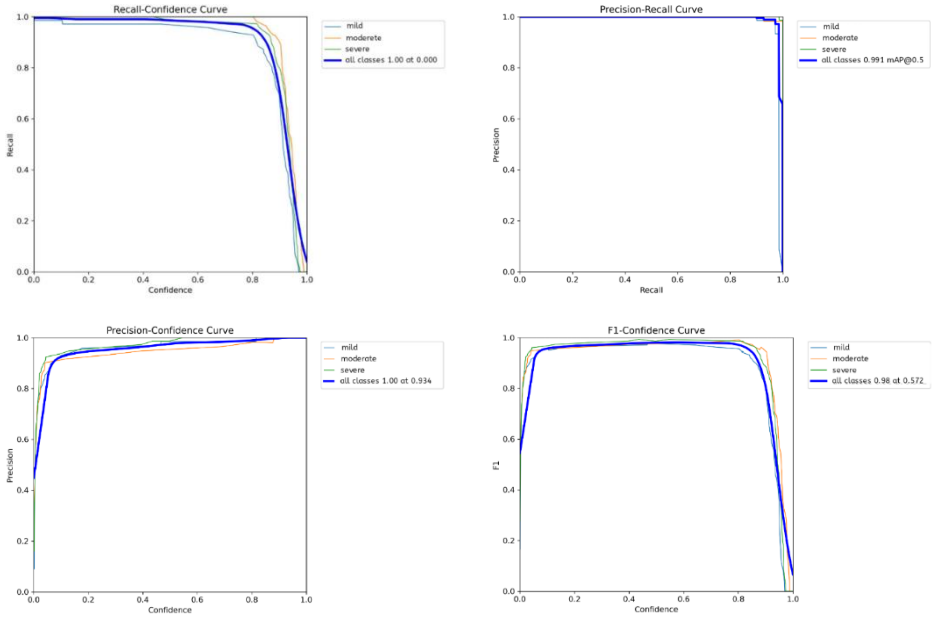


Figure 5. Evaluation Chart

The following is figure 6, the results of the evaluation using the confusion matrix. The results of the confusion matrix and normalized confusion matrix show that the model for classifying skin lesions can make very accurate predictions for all classes (mild, moderate, and severe). Figure 6 indicates that misclassifications occur predominantly between mild and moderate classes, suggesting visual overlap in intermediate severity levels rather than random prediction errors. The majority of the test data was successfully mapped to its target class, as evidenced by the dominant diagonal values in the matrix. In the Heavy class, the model achieved a prediction accuracy of up to 99%. In the mild and Medium classes, the accuracy was also very high, with only a few classification errors that were more likely to be moved to the background class. Normalizing the matrix makes the model even more consistent, with the ratio of correct predictions being close to 1 for almost all classes. This demonstrates that the classification system developed is not only accurate but also stable in differentiating the severity of skin lesions under various conditions.

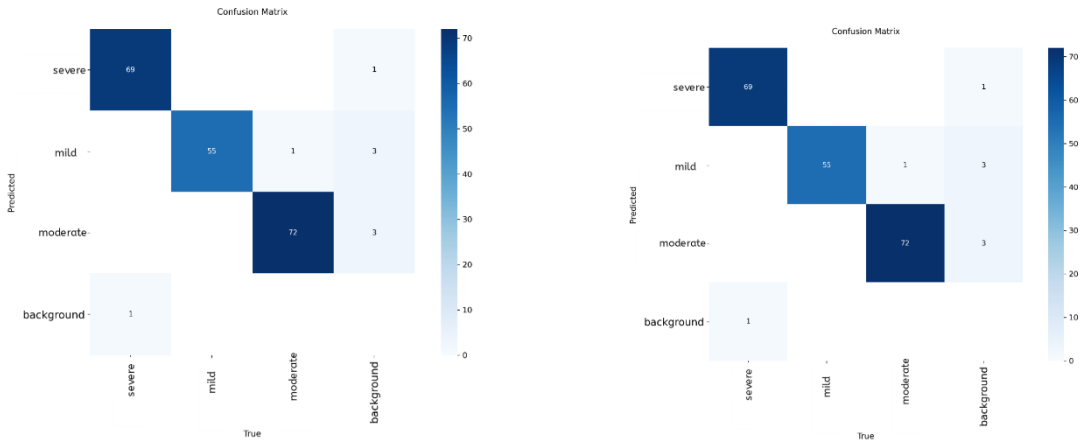


Figure 6. Confusion Matrix

This finding is consistent with previous studies such as[34], [20], which demonstrate that YOLO excels in detecting dermatological abnormalities in adults; however, this research expands the scope by focusing on the infant population. This is important because babies' skin is softer, more sensitive to color, and often has environmental artifacts (like skin folds, diaper rash, or uneven lighting) that make it harder to detect than in other age groups.

ISSN 2723-6471

3.2. Classification Performance with CNN

Headings To assess the performance of the CNN model in classifying the severity level of skin lesions, a training process was conducted over 50 epochs, monitoring accuracy and loss metrics on both training and validation data. The accuracy curve analysis is used to see how well the model can learn patterns from the training data and apply them to the validation data. The loss curve shows how well the model can reduce prediction errors as the number of epochs increases. The figure 7 below displays the results of this training.

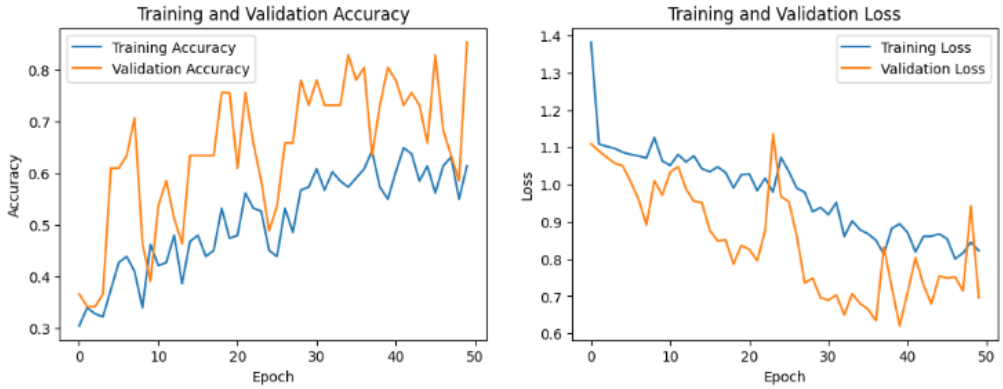


Figure 7. Evaluation Chart

The figure above shows the accuracy and loss curves for the CNN model training process for classifying the severity of skin lesions over 50 epochs. Figure 7 shows stable convergence of the classification model, with decreasing validation loss despite minor accuracy fluctuations, indicating limited overfitting and reasonable generalization. The accuracy graph indicates that the model's performance gradually improved on the training data, while on the validation data, the accuracy value tended to fluctuate but was able to reach more than 0.8 on some epochs, which shows that the model can generalize to test data. The loss graph shows a consistent downward trend for both the training and validation data. The validation loss drops to about 0.6, which means that the model is getting better at making fewer mistakes in its predictions. This pattern shows that CNN can learn to tell the difference between different levels of skin lesions, even though the performance changes because of the different validation data.

The results of training the CNN model, as shown in table 1, show that it is effective at classifying the severity of atopic dermatitis in babies based on the Region of Interest (ROI) detected by YOLO. Using the transfer learning-based EfficientNet-B0 architecture (g pakai), the model achieved an average accuracy of 85% and a macro-F1 score of 85%, indicating consistent performance across all categories. Table 1 provides a more detailed analysis, demonstrating a commendable overall accuracy of 85%. In the mild class, the model achieved a precision score of 75% and a perfect recall score of 100%, resulting in an F1-score of 86%. The result shows that the model can find all mild cases, even though some of its predictions are still not quite right. In the Medium class, the model had a high precision of 100% but a low recall of only 60%, which gave it an F1-score of 75%. This outcome means that even though the predictions were very accurate, some Medium cases were not found. For the Heavy class, on the other hand, the model did better overall, with a precision of 88% and a recall of 100%, which gave it an F1-score of 93%. Overall, the macro-average scores showed a precision of 88%, a recall of 87%, and an F1-score of 85%, while the weighted-average values were 88%, 85%, and 84%, respectively. These results suggest that the model performs consistently in classifying the severity of skin lesions, although further improvement is still needed to enhance detection accuracy in the moderate class. The results of the CNN model evaluation can be seen in the following table 1:

Table 1. Results of the CNN Model Evaluation

Classification	Precision	Recall	F1-Score
Mild	75%	100 %	86 %
Keep	100 %	60 %	75 %
Heavy	88 %	100 %	93%
Accuracy			85 %

Macro avg	88 %	87 %	85 %
Weighted avg	88 %	85%	84 %

Although the overall classification performance is satisfactory, the moderate class exhibits a notably lower recall of 60%, indicating that a substantial portion of moderate cases are misclassified as either mild or severe. This behavior suggests significant visual overlap between intermediate severity levels, where lesion characteristics such as erythema intensity and texture transition gradually rather than forming distinct boundaries. The results of this study are consistent with findings[35], [36], demonstrating that CNN can achieve accuracy comparable to dermatologists in classifying dermatological diseases. However, this study offers a novel contribution by concentrating on the infant population, which possesses distinct anatomical and physiological skin conditions compared to other age groups. The different textures of skin, how sensitive it is to mild, and environmental factors like tropical humidity make classification harder, but the CNN model was able to do an impressive job. The following is figure 8 about the search results and classification of skin lesions using the CNN-based YOLO model.

Figure 8. The results of finding and classifying skin lesions using the CNN-based YOLO model

Figure 8 demonstrates that the proposed framework can localize multiple lesions within a single image and assign consistent severity predictions with high confidence, supporting its applicability in real-world screening scenarios. Shows the results of skin lesion detection and classification using the CNN-based YOLO model on the validation data. Each lesion area was successfully identified with a colored bounding box annotation accompanied by a class label (Mild, Moderate, and Severe) and a confidence score indicating the model's confidence in the prediction. It is evident that the model can identify multiple lesions within the same image while simultaneously differentiating the severity based on the emerging visual characteristics. The prediction results indicate that most of the lesions were classified with a high confidence value (≥0.9), which shows that the model is consistent in finding patterns in skin lesions. This visualization shows that YOLO can not only classify images globally but also segment them based on objects with a high level of accuracy. This functionality makes it useful for clinical applications that need to diagnose the severity of dermatitis in children.

3.3. Discussion

Integrating YOLO and CNN into a two-step pipeline has been shown to create a complete system. Although several previous studies have demonstrated the effectiveness of YOLO-based models for dermatological applications, most of these works rely on adult datasets. Infant skin presents distinct challenges, including thinner epidermal layers, less pronounced texture patterns, higher color homogeneity, and greater sensitivity to illumination changes. These characteristics can reduce the contrast between lesions and surrounding skin, making lesion localization and severity assessment more difficult compared to adult skin images. YOLO is a fast way to find areas of damage, while CNN is a more accurate way to judge how bad the damage is. The stable detection performance of YOLO across various lesion sizes supports its reliability in clinical practice, while CNN can provide a relatively balanced assessment across all categories, although the heavy class still requires additional data to enhance accuracy.

ISSN 2723-6471

The main advantage of this research is its application to the population of infants, which has rarely been the focus of AI dermatology studies. The system has the potential to be applied in digital clinic and mobile health applications, especially in areas with limited medical specialists, due to its high level of accuracy. The term real-time in this study refers to near real-time inference suitable for point-of-care and mobile health applications rather than high-throughput clinical systems. The reported inference time of approximately 16 milliseconds per image corresponds to a processing rate of over 60 frames per second on standard GPU hardware, which is sufficient for interactive clinical screening and teledermatology workflows. Future work may also explore multi-scale feature extraction and attention-based mechanisms to improve robustness against noise, illumination variation, and subtle lesion boundaries commonly observed in infant skin images. This research, however, has limitations, such as an imbalance in data distribution among classes and variations in image quality due to differences in camera equipment. Therefore, further research needs to include larger, multi-flashlight datasets, as well as integration with richer explainable AI systems so that the results can be more widely adopted in children's healthcare. The following table 2 presents a comparative evaluation of different deep learning models tested on the same dataset to assess their performance in classifying the severity of atopic dermatitis in infants:

Table 2. Performance Comparison of Different Models Using the Same Dataset

Model	Architecture	Accuracy (%)	Precision (%)	Recall (%)	Macro-F1 (%)
CNN (Baseline)	Custom CNN	78	76	79	77
VGG16 (TL)	VGG16 + FC	82	81	83	82
ResNet50 (TL)	ResNet50 + FC	84	85	82	83
EfficientNet-B0 (TL)	EfficientNet-B0 + FC	85	88	87	85
Proposed Model	YOLO + CNN (EfficientNet-B0)	85	88	87	85

The performance comparison presented in table 2 is intended to provide a descriptive benchmark across different model architectures. No statistical significance testing or confidence intervals were computed; therefore, the reported differences in accuracy and macro-F1 score should be interpreted as indicative rather than conclusive. The comparison demonstrates that transfer learning–based CNN architectures generally outperform the baseline CNN model in terms of accuracy and macro-F1 score, indicating their superior ability to extract discriminative features from infant skin images. Among the evaluated models, EfficientNet-B0 achieved the highest classification performance, reflecting its balanced depth and parameter efficiency. However, these models operate solely on full-image classification without explicit lesion localization. In contrast, the proposed YOLO–CNN framework provides comparable classification performance while simultaneously enabling precise lesion detection and severity assessment. Hyperparameter values, including batch size and number of training epochs, were selected based on preliminary experiments to balance model convergence stability, generalization performance, and computational constraints. Larger batch sizes resulted in slower convergence, while smaller batches led to unstable gradients. Similarly, training beyond the selected number of epochs did not yield measurable performance improvements and increased the risk of overfitting. This dual-stage capability offers a more comprehensive and clinically meaningful diagnostic approach, particularly for real-world pediatric applications where both lesion localization and severity evaluation are required. While recent literature reports strong performance of hybrid deep learning models for skin disease detection, most existing studies target adult datasets and emphasize disease classification rather than severity grading at the lesion level. This study extends prior work by adapting these techniques to infant skin images and by explicitly linking lesion localization to severity assessment, addressing a gap not covered by recent adult-focused studies

4. Conclusion

The conclusion of this study indicates that the integration of YOLO for lesion detection and CNN for severity classification provides an effective, accurate, and efficient two-stage solution for diagnosing atopic dermatitis in infants. From a clinical perspective, an accuracy of 85% may be sufficient for preliminary screening and risk stratification in pediatric settings, particularly in resource-limited environments. However, it may not yet meet the reliability requirements for standalone diagnostic use, underscoring the importance of human-in-the-loop clinical

validation. The YOLO model successfully detected areas of inflammation with high performance (mAP@0.5 of 99% and F1-score of 98%) as well as real-time inference speed, while CNN managed to classify severity with an average accuracy of 85% and a macro-F1 score of 85%. These results confirm that AI-based systems can be a promising clinical diagnostic tool. This study contributes to the development of a more adaptive and applicable digital health system, particularly for pediatric healthcare services in areas with limited specialized medical personnel.

5. Declarations

5.1. Author Contributions

Conceptualization: D.S., R.N.P., S.H., A.N.H., Y.I., and N.H.; Methodology: S.H.; Software: D.S.; Validation: D.S., S.H., and N.H.; Formal Analysis: D.S., S.H., and N.H.; Investigation: D.S.; Resources: S.H.; Data Curation: S.H.; Writing Original Draft Preparation: D.S., S.H., and N.H.; Writing Review and Editing: S.H., D.S., and N.H.; Visualization: D.S.; All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

5.3. Funding

Thank you to the Directorate of Research, Technology and Community Service (DRTPM), Ministry of Education, Culture, Research and Technology in the 2025 Fundamental Research Program with contract number DT.05.00/PL/2025,003/LL17/DT.00.05.00/PL/2025,006/LPPM/KH-DIKTI/PEN/2025.

5.4. Institutional Review Board Statement

Not applicable.

5.5. Informed Consent Statement

Not applicable.

5.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. Pagliaro, L. Pecoraro, C. Stefani, S. Pieropan, G. Piacentini, and A. Pietrobelli, "Bathing in Atopic Dermatitis in Pediatric Age: Why, How and When," *Pediatr*. Rep., vol. 16, no. 1, pp. 57–68, 2024, doi: 10.3390/pediatric16010006.

[2] P. A. Alayyannur, M. M. Al Hakim, and R. S. R. E. P. Sari, "Dermatitis among Workers and Its Relation with Personal Protective Equipment," *Indones. J. Occup. Saf. Heal*., vol. 13, no. 2, pp. 261–267, 2024, doi: 10.20473/ijosh.v13i2.2024.261-267.

[3] R. Melati, E. Aryanti, M. S. Alfaqih, and I. W. Ferianasari, "Prevalence and characteristics of dermatitis cases in Dermatology Clinics (2020 – 2024): A retrospective analysis," *Science Midwifery,* vol. 12, no. 6, pp. 1-9, 2025, DOI:10.35335/midwifery.v12i6.1846

[4] Faye O, Flohr C, Kabashima K, Ma L, Paller AS, Rapelanoro FR, Steinhoff M, Su JC, Takaoka R, Wollenberg A, Yew YW, Postigo JAR, Schmid-Grendelmeier P, Taïeb A, "Atopic dermatitis: A global health perspective," *J. Eur. Acad. Dermatology Venereol.*, vol. 38, no. 5, pp. 801–811, 2024, doi: 10.1111/jdv.19723.

[5] Laughter MR, Maymone MBC, Mashayekhi S, Arents BWM, Karimkhani C, Langan SM, Dellavalle RP, Flohr C, "The global burden of atopic dermatitis: lessons from the Global Burden of Disease Study 1990–2017," *Br. J. Dermatol*., vol. 184, no. 2, pp. 304–309, 2021, doi: 10.1111/bjd.19580.

[6] R. Tupsila, W. Siritaratiwat, S. Bennett, and L. Mato, "Intra-Individual Variability in Gross Motor Development in Healthy Full-Term Infants Aged 0 – 13 Months and Associated Factors during Child Rearing," *Child. Multidiscip. Digit. Publ. Inst*., vol. 9, no. 6, pp. 1-10, 2022, DOI: 10.3390/children9060801

[7] V. Wang, J. Boguniewicz, M. Boguniewicz, and P. Y. Ong, "The infectious complications of atopic dermatitis," *Ann. Allergy, Asthma Immunol.*, vol. 126, no. 1, pp. 3–12, 2021, doi: 10.1016/j.anai.2020.08.002.

[8] Napolitano M, Esposito M, Fargnoli MC, Girolomoni G, Romita P, Nicoli E, Matruglio P, Foti C, "Infections in Patients with Atopic Dermatitis and the Influence of Treatment," *Am. J. Clin. Dermatol.*, vol. 26, no. 2, pp. 183–197, 2025, doi: 10.1007/s40257-025-00917-z.

[9] Liang, S., Liu, S., Long, Z, "A self-powered hydration-monitoring and drug-delivery skin patch for closed-loop treatment of atopic dermatitis," *Microsystems Nanoeng.,* vol. 11, no. 1, pp. 1-12, 2025, doi: 10.1038/s41378-025-01000-7.

[10] G. Færk, E. Søndergaard, L. Skov, K. S. Hansen, and S. Reventlow, "A 'normal' life: a qualitative study exploring parents' experiences of everyday life with a child diagnosed with atopic dermatitis and atopic comorbidities," *Scand. J. Prim. Health Care*, vol. 43, no. 2, pp. 347–358, 2025, doi: 10.1080/02813432.2024.2440777.

[11] A. G. Kelbore, W. Enbiale, J. M. van Wyk, and A. Mosam, "Atopic dermatitis in Ethiopian children: a multicenter study of clinical severity, characteristics, and sociodemographic factors," *Front. Med*., vol. 11, no. July, pp. 1-10, 2024, doi: 10.3389/fmed.2024.1410310.

[12] B. Duarte, P. Mendes, B. Joana, A. Filomena, and A. Margarida, "The APOLO Study : A Cross - Sectional Analysis of Disease Characteristics and Patient Burden in Moderate - to - Severe Atopic Dermatitis in Portugal," *Dermatol. Ther. (Heidelb).,* vol. 15, no. 3, pp. 647–662, 2025, doi: 10.1007/s13555-025-01347-9.

[13] A. Ameen, A. Al, D. Ashraf, M. R. Ayman, A. Fatima, and A. Marzooqi, "Consensus Recommendations for the Management of Atopic Dermatitis in the United Arab Emirates," *Springer Healthcare,* vol. 14, no. 9, pp. 1-11, 2024. doi: 10.1007/s13555-024-01247-4.

[14] M. Raulf and S. Kespohl, "Skin tests, serological IgE detection, basophil test—what is available, useful and helps to clarify a mold allergy?," *Allergo J. Int.*, vol. 33, no. 4, pp. 133–139, 2024, doi: 10.1007/s40629-024-00287-8.

[15] N. Katoh, "Atopic Dermatitis, beyond the Skin Surface," *Pediatr. Allergy, Immunol. Pulmonol.*, vol. 29, no. 4, pp. 1-12, 2021, doi: 10.1089/ped.2016.29014.nk.

[16] Brănișteanu, D.E.; Boțoc, C.C.; Huțanu, A.E.; Munteanu, C.A.; Ciobanu, R.P.; Brănișteanu, D.C.; Colac, A.G.; Onu-Brănișteanu, C.I.; Brănișteanu, G.; Manolache, "Diagnostic and Therapeutic Challenges in Psoriasis–Atopic Dermatitis Overlap: A Retrospective Observational Cohort Study," *Diagnostics,* vol. 15, no. 11, pp. 1–16, 2025, doi: 10.3390/diagnostics15111381.

[17] M. G. Ragab, Abdulkadir, SJ, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access,* vol. 12, no. April, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.

[18] K. Khan and A. Aleem, "A Hybrid Approach Combining YOLO and Deep Neural Networks for Efficient Lung Disease Detection," *2025 3rd Int. Conf. Commun. Secur. Artif. Intell. ICCSAI 2025*, vol. 1, no. July, pp. 1204–1211, 2025, doi: 10.1109/ICCSAI64074.2025.11064147.

[19] B. Kurniawan, R. Wahyuni, Y. Irawan, and M. H. Yuhandri, "Multimodal Deep Learning and IoT Sensor Fusion for Real-Time Beef Freshness Detection," *J. Appl. Data Sci.*, vol. 6, no. 4, pp. 2921–2937, 2025. DOI: 10.47738/jads.v6i4.977

[20] N. C. Huang, A. Mukundan, R. Karmakar, S. Syna, W. Y. Chang, and H. C. Wang, "Novel Snapshot-Based Hyperspectral Conversion for Dermatological Lesion Detection via YOLO Object Detection Models," *Bioengineering,* vol. 12, no. 7, pp. 1–21, 2025, doi: 10.3390/bioengineering12070714.

[21] S. Kukreti, R. V. S. Praveen, S. Bansal, H. Raju, N. Singh, and R. Begum, "Object Detection in Real-Time Surveillance Using Deep Learning-Based YOLO Framework," IEEE Int. Conf. "*Computational, Commun. Inf. Technol. ICCCIT* 2025, vol. 1, no. March, pp. 601–606, 2025, doi: 10.1109/ICCCIT62592.2025.10928084.

[22] A. Febriani, R. Wahyuni, Y. Irawan, and R. Melyanti, "Improved Hybrid Machine and Deep Learning Model for Optimization of Smart Egg Incubator," *J. Appl. Data Sci*., vol. 5, no. 3, pp. 1052–1068, 2024. DOI: 10.47738/jads.v5i3.304

[23] A. Muhaimin, Edriyansyah, Wahyat, Y. Irawan, and R. Wahyuni, "Optimized IoT-Based Multimodal Fusion for Early Forest Fire Detection and Prediction," *ECTI Trans. Comput. Inf. Technol*., vol. 19, no. 4 SE-Research Article, pp. 569–582, Sep. 2025, doi: 10.37936/ecti-cit.2025194.262839.

[24] J. Huang, W. Ren, and L. Ru, "Intelligent Detection of Skin Diseases Based on The Optimized YOLO V8," *Proc. 2024 IEEE 4th Int. Conf. Inf. Technol. Big Data Artif. Intell. ICIBA* 2024, vol. 4, no. 1, pp. 1277–1289, 2024, doi: 10.1109/ICIBA62489.2024.10868331.

[25] Z. Chen, Q. Zhu, X. Zhou, J. Deng, and W. Song, "Experimental Study on YOLO-Based Leather Surface Defect Detection," *IEEE Access*, vol. 12, no. February, pp. 32830–32848, 2024, doi: 10.1109/ACCESS.2024.3369705.

[26] D. Jepisah, H. Octaria, Muhamadiah, and Y. Irawan, "YOLOv12 Model Optimization for Monitoring Occupational Health and Safety in Hospital Archive Rooms," *J. Appl. Data Sci.*, vol. 6, no. 4, pp. 2666–2681, 2025, doi: 10.47738/jads.v6i4.936.

[27] S. Pavithra, K. Tamilarasi, S. Shashidhar, and M. Shahzil, "Developing an Advanced Skin Disease Detection System by Integrating Hybrid Neural Network with Yolo V8 Model," *2024 3rd Int. Conf. Smart Technol. Syst. Next Gener. Comput. ICSTSN 2024*, vol. 1, no. September, pp. 1–6, 2024, doi: 10.1109/ICSTSN61422.2024.10671324.

[28] M. M. Musthafa, M. T R, V. K. V, and S. Guluwadi, "Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification," *BMC Med. Imaging*, vol. 24, no. 1, pp. 1–20, 2024, doi: 10.1186/s12880-024-01356-8.

[29] D. Setiawan, R. N. Putri, I. Fitri, A. N. Hidayanto, Y. Irawan, and N. Hohashi, "Improved Deep Learning Model for Prediction of Dermatitis in Infants," *J. Appl. Data Sci*., vol. 6, no. 2, pp. 871–884, 2025, doi: 10.47738/jads.v6i2.542.

[30] A. Ansari, A. Singh, M. Singh, and V. Kukreja, "Enhancing Skin Disease Classification: A Hybrid CNN-SVM Model Approach," *2024 Int. Conf. Autom. Comput. AUTOCOM 2024*, vol 1. no. 1, pp. 29–32, 2024, doi: 10.1109/AUTOCOM60220.2024.10486133.

[31] P. Gupta, J. Nirmal, and N. Mehendale, "Custom CNN architectures for skin disease classification: binary and multi-class performance," *Multimed. Tools Appl*., vol 84. no. December, pp. 32505–32532, 2024, doi: 10.1007/s11042-024-20503-5.

[32] Y. Devis, Muhamadiah, Yulanda, Y. Irawan, and R. Wahyuni, "Optimization of Machine Learning Models for Risk Prediction of DHF Spread to Support Management Strategies in Urban Areas," *J. Appl. Data Sci*., vol. 6, no. 4, pp. 2407–2420, 2025, doi: 10.47738/jads.v6i4.898.

[33] H. Fonda, Y. Irawan, R. Melyanti, R. Wahyuni, and A. Muhaimin, "A Comprehensive Stacking Ensemble Approach for Stress Level Classification in Higher Education," *J. Appl. Data Sci*., vol. 5, no. 4, pp. 1701–1714, 2024. DOI: https://doi.org/10.47738/jads.v5i4.388

[34] W. T. D. Tjahjono, C. Paramita, C. Supriyanto, A. J. Savicevic, S. Rakasiwi, and A. A. Haresta, "YOLOv10x Model for Accurate First Detection of Skin Diseases from Dermoscopic Objects," *2025 Int. Conf. Smart Comput. IoT Mach. Learn.,* vol. 11, no. 1, pp. 1–6, 2025, doi: 10.1109/siml65326.2025.11080864.

[35] A. Kumar and K. K. Tiwari, "Disease Classification in Dermatology: A CNN-RF Hybrid Approach," *2024 3rd Int. Conf. Artif. Intell. Internet Things, AIIoT 2024*, vol. 1, no. 1, pp. 1–5, 2024, doi: 10.1109/AIIoT58432.2024.10574607.

[36] M. S. Rashid, G. Gilanie, S. Naveed, S. Cheema, and M. Sajid, "Automated detection and classification of psoriasis types using deep neural networks from dermatology images," *Signal, Image Video Process*., vol. 18, no. 1, pp. 163–172, 2024, doi: 10.1007/s11760-023-02722-9.