Journal of Applied Data Sciences
Vol. 6, No. 2, May 2025, pp. 871-884

Improved Deep Learning Model for Prediction of Dermatitis in Infants

Debi Setiawan[1,*], Ramalia Noratama Putri[2], Imelda Fitri[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 17, 2024; Revised: December 13, 2024; Accepted: January 22, 2025; Available online: February 21, 2025)

Abstract

Indonesia's equatorial climate, characterized by summer and rainy seasons, presents environmental conditions that contribute to a high incidence of dermatitis in infants. Dermatitis, an inflammatory skin condition, can lead to significant discomfort in infants, affecting their sleep, growth, and development. Early diagnosis is crucial for effective treatment; however, conventional diagnostic methods in clinics and hospitals—such as physical observation and parental interviews—are often time-consuming, subjective, and may lack precision, creating a need for more efficient diagnostic tools. This study explores the application of deep learning models to enhance the accuracy and speed of dermatitis diagnosis in infants. Four convolutional neural network (CNN) models were evaluated: MobileNet, VGG16, ResNet, and a Custom CNN model specifically designed for this study. Using a dataset of 1,088 skin images collected from three regions in Riau Province, Indonesia, we conducted training and testing to assess each model's performance in distinguishing between dermatitis-affected and healthy skin. Results show that MobileNet and the Custom CNN outperformed other models, achieving accuracy rates of 97% and 85%, respectively. MobileNet's high accuracy and efficiency make it a viable option for mobile applications, enabling rapid, on-site diagnosis in resource-limited settings. The Custom CNN model, tailored to the unique features of infant skin, also showed promising results. These findings demonstrate the potential of automated, image-based diagnostic tools for assisting medical professionals in early dermatitis detection, improving patient outcomes. This study contributes a valuable diagnostic solution that leverages deep learning to support healthcare providers, particularly in areas with limited access to specialized medical resources.

Keywords: Deep Learning, Custom CNN, MobileNet, Infants, Dermatitis

1. Introduction

Like in many other countries, babies in Indonesia are also susceptible to infectious diseases such as respiratory infections, diarrhea, and non-communicable diseases, such as dermatitis in infants. Dermatitis is a medical term that refers to inflammation or irritation of the skin. It is a common condition that can affect people of all ages, including babies. Dermatitis can be caused by a variety of factors, including exposure to allergens, irritants, infections, genetics, or certain medical conditions. In addition, the body of a baby affected by skin diseases is really very uncomfortable. Baby's skin that is still smooth and soft can also become dry, scaly, yellowish patches or oily crusts, and red. This will disturb the baby such as being fussier and lack sleep due to the itching suffered, so early diagnosis is necessary. However, early diagnosis poses challenges in the world of medicine, both in central hospitals and clinics. Currently, doctors diagnose dermatitis in babies by observing the skin area and interviewing the baby's parents. Early diagnosis of dermatitis in babies is very important to do as soon as possible, because without proper treatment, dermatitis

*Corresponding author: Debi Setiawan (debisetiawan@univrab.ac.id)
DOI: https://doi.org/10.47738/jads.v6i2.542

symptoms such as redness, itching, and inflammation can increase to more severe [1], [2], [3]. This can make the baby feel very uncomfortable and it may be difficult for him to sleep or rest well. If severe dermatitis interferes with your baby's sleep and comfort significantly, this can affect his growth and development. So that it can have a negative impact on the physical and cognitive development of babies. However, early diagnosis of dermatitis poses a challenge in the world of medicine, both in central hospitals and clinics [4]. Currently, doctors diagnose dermatitis in babies by observing the skin area, interviewing the baby's parents [5], if necessary, a skin biopsy on the baby is performed [6]. Diagnosis in this way takes a long time and can produce a less accurate diagnosis [2]. Therefore, technology is needed that can help doctors diagnose dermatitis in babies [7]. The urgency of this study is that no researcher has classified atopic dermatitis in infants, many researchers have classified atopic dermatitis skin disease only in adults.

Atopic dermatitis occurs when the skin is exposed directly to chemicals or other substances that cause irritation or allergic reactions. Common examples of irritants include detergents, soaps, cosmetics, metals, and certain plants. Atopic dermatitis is a chronic condition that causes the skin to become dry, itchy, red, and scaly. Symptoms of dermatitis in babies can vary depending on the type of dermatitis and the causative factors. Some of the symptoms that appear in babies with dermatitis include: Skin Rash: The area of the baby's skin affected by dermatitis can be red, reddish, or dark in color. This rash may be scaly, mottled, or crusty, depending on the type of dermatitis. Itching: Itching is a common symptom in many types of dermatitis. The baby may seem uncomfortable, scratching, or rubbing the affected area of the skin to relieve the itching. Dry Skin: In some types of dermatitis, such as atopic dermatitis, the baby's skin may become dry and rough. The skin can look cracked or scaly. Swelling or Inflammation: Dermatitis can cause swelling or inflammation in the affected area of the skin. This can make the skin look thicker or larger than usual. Spots or Blisters: Areas of skin affected by dermatitis can have small spots, blisters, or open wounds from friction or scratching. Crust or Cortical Crust: In some types of dermatitis, such as atopic dermatitis, babies may experience yellow or white crust that sticks to the skin, especially in the head area (cradle cap). Secondary Infections: Severe or chronic dermatitis can cause damage to the skin and increase the risk of bacterial, fungal, or viral infections.

Current clinical methods for diagnosing dermatitis in infants primarily rely on physical observation of skin lesions and interviews with caregivers. These approaches are often subjective, as diagnoses heavily depend on the clinician's expertise and experience, leading to variability in accuracy. Additionally, conducting such evaluations is time-consuming, especially in busy clinical settings, and can delay the initiation of treatment. In rural or resource-constrained areas, the availability of dermatologists is limited, making it difficult for families to access timely and accurate diagnoses. These challenges are further compounded by the financial burden associated with repeated clinic visits or advanced diagnostic tests, such as biopsies, which are not feasible for infant patients. Together, these limitations highlight the urgent need for automated, image-based diagnostic tools that can offer consistent, quick, and cost-effective solutions for early detection and management of dermatitis in infants.

Several researchers have previously found solutions regarding the diagnosis of dermatitis with various methods from artificial intelligence [8], [9], [10]. The method offered by some previous researchers is the Convolutional Neural Network (CNN) method [11], [12]. The CNN method is one type of architecture in the deep learning which is specifically designed for pattern recognition on image-based data [13]. CNNs have become one of the most effective techniques in a variety of image processing tasks, including object classification and detection [14]. Researchers have previously proven that computational models of the CNN-based approach are capable of making a very accurate diagnosis of atopic dermatitis [15], [16], contact dermatitis, seborrheic dermatitis [17], and herpetiformis dermatitis [18]. Therefore, the CNN algorithm is able to perform image recognition by automatically extracting features from skin images and recognizing patterns of dermatitis in babies. The AI model developed uses Deep Learning with the Convolutional Neural Network architecture.

The research gap seen in this study is the limitation of data, there are still not many people who research about dermatitis in babies directly and get datasets about skin diseases in babies, this has been checked both in regular reviews and at the Kaggle and the University of Indonesia data center. Dermatitis is widely studied only on the skin of adults, but in babies it is still not present. It can be seen from the geographical side that Indonesia is located in the equatorial area which has two seasons, hot and rainy, this causes the tendency of babies to suffer from dermatitis due to the location of the region. Indonesia's equatorial climate, characterized by high humidity and temperatures, plays a significant role in the prevalence of dermatitis in infants. Data collection was conducted in three regions of Riau Province chosen for their distinct environmental conditions. These regions experience variations in air quality, allergen exposure, and hygiene practices, which were considered during data collection. For instance, areas with higher humidity levels were hypothesized to have a higher incidence of dermatitis due to increased skin irritation and bacterial growth. So that this can be categorized as GAP for countries adjacent to Indonesia, namely Malaysia and Brunei Darusalam. Seen from the demographic side of the tendency of unhealthy parental lifestyles, consuming fast food will certainly affect the quality of breast milk in children. Bad breast milk will have an impact on the baby's skin because the baby's skin is sensitive and soft. In clinical and medical techniques, there is a gap between some of the treatment methods for this dermatitis research with previous studies that have been patterned on adult skin. Adult skin can be biopsied, namely taking a sample of body tissue, one of which is the skin, but it is impossible to do a biopsy on the baby's skin because the baby still cannot talk about what is sick and must be quickly handled so that it does not spread to other skin tissues. The modeling used in the study was VGG16, MobileNet, Resnet, and CNN Costum, there were several comparisons to the measurement pattern, namely recall, precision, and accuracy.

The latest in this study is the CNN Custom pattern that the researcher designed himself with several layers. Then in this study, it can also be seen that the results of the comparison of models, patterns and appropriate measurement values by processing the real set of 1088 data taken based on the recording of images from the variables of the baby's hand which were used as data samples from 3 regions in Riau Province. Because the skin fibers of the baby's upper hands are softer, with minimal pigmentation, consisting of a moister layer of the epidermis and dermis so that it is more precise in taking pictures.

2. The Proposed Method

The steps in the first phase are: a) collecting images of patients infected with dermatitis and images of healthy infant patients as datasets; b) data labeling; c) image pre-processing such as image size setting (resize), normalization of color intensity, flip, rotation, and noise removal; d) image segmentation to make it easier to determine the location of objects and their limitations; e) feature map to make it easier to understand the features of the detected inputs or maintained by the model, f) the creation of a CNN classification model; g) the evaluation of the model using the confusion matrix. Figure 1 shows the research process.

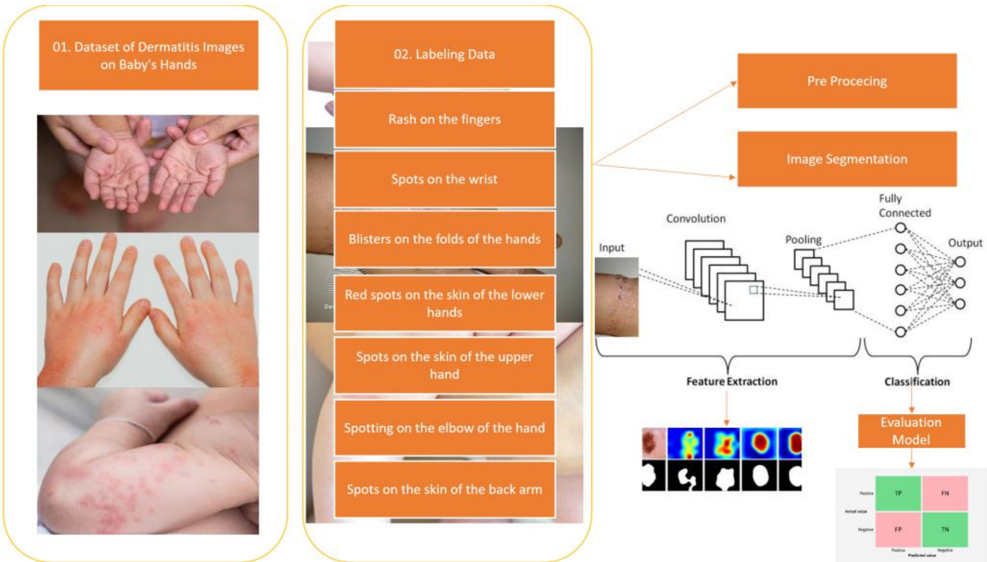

Figure 1. Dermatitis Detection Model in Infants

The figure illustrates the overall workflow of the study. It begins with a dataset comprising images of dermatitis on infants' hands, showcasing various conditions such as rashes on the fingers, spots on the wrist, and blisters on the folds of the hands. These images were labeled according to specific symptoms and locations, ensuring a systematic categorization process. The workflow proceeds with preprocessing steps, including resizing, noise reduction, and segmentation to enhance image quality. The processed images are then fed into CNN for feature extraction and classification. The final step involves model evaluation, where metrics like true positives, false positives, true

ISSN 2723-6471

negatives, and false negatives are analyzed to assess the model's performance. This structured workflow provides a clear and detailed pathway from raw data to model evaluation.

2.1. Collection of Images

The collection of images of patients infected with dermatitis and images of the skin of healthy baby patients was carried out from July to August. Data collection was carried out by collecting data via the internet from July 19 to August 31. Data collection was conducted by field observation in Pulau Gadang village, District XII Koto Kampar, Labuhan Tangga Hilir Village, Rokan Hilir Regency, and in Teluk Kenidai Village. These locations were chosen to capture variations in environmental conditions and the prevalence of dermatitis cases. To ensure data quality, all images underwent pre-screening, and those with blurriness, uneven lighting, or distortions were excluded. Additionally, efforts were made to reduce demographic and geographical bias by including participants from diverse backgrounds in terms of age, gender, and socioeconomic status. This diversity enhances the dataset's generalizability and ensures its robustness for model training and evaluation. Data collection is conducted using various methods to gather comprehensive information about the baby's skin condition. Parents or caregivers are interviewed to obtain details on disease history, symptoms, diet, hygiene, and environmental factors, aiming to identify patterns of symptoms and risk factors. Additionally, photos of the affected skin areas are taken to document the condition for diagnosis, monitoring progression, and comparison with medical references.

This study was conducted following approval from the institutional ethical review board (IRB) of Abdurrab University, ensuring compliance with ethical guidelines for research involving human participants. Informed consent was obtained from the parents or legal guardians of all infant participants prior to data collection. The consent process included a clear explanation of the study's purpose, the methods used, and the rights of participants, including the right to withdraw at any time without consequence. To protect patient privacy, all images and associated data were anonymized, with unique identifiers used in place of personal information. Data were securely stored on encrypted servers, and access was restricted to authorized personnel only. These measures ensured that the study met the highest ethical standards while safeguarding the rights and confidentiality of participants.

2.2. Labelling Data

Data labeling is the process of marking or labeling raw data, so that it can be used in machine learning algorithms (machine learning) [19], [20], [21]. Labeling is usually done for data to be used where the model needs the already labeled data to learn and make predictions [22],[23]. Labeling of Skin Disease Photo Images can be seen in the following figure 2:

Figure 2. Labeling of Skin Disease Photo Images

The data labeling process was conducted meticulously to ensure accuracy and consistency. Parents or caregivers were first interviewed to gather detailed information about the symptoms, potential risk factors, and medical history of the infants. Subsequently, trained medical professionals, including dermatologists, reviewed the images and assigned labels based on clinical observation guidelines. The labeling followed a standard protocol, which involved identifying key visual markers such as rashes, redness, or lesions, and cross-verifying the labels within a team of experts to minimize subjectivity. This systematic approach ensured that each image was accurately categorized as either "dermatitis" or "non-dermatitis" and further classified into specific types of dermatitis when applicable.

2.3. Image pre-processing

The preprocessing steps, including resizing, normalization, and noise removal, were essential in improving the accuracy and reliability of the models. In figure 3 is a pre-positioning of the photo image of skin diseases.

Figure 3. Image Pre-processing Results

Image pre-processing is the first step in digital image processing that aims to improve the quality of the image before conducting further analysis, such as feature detection, segmentation, or classification [24]. Preprocessing is essential to eliminate interference (noise) or distortion that can interfere with the pattern recognition process of machine learning models or other algorithms [25].

To assess the impact of each preprocessing step, we conducted experiments by training the models with and without specific steps. Results showed that resizing images to a consistent resolution (224 x 224 pixels) significantly improved model accuracy, with a 6% increase compared to models trained on unscaled images. Normalization of pixel values to a range of [0,1] further enhanced accuracy by 4%, likely by standardizing the input data and making it easier for the models to converge. Noise removal, achieved through Gaussian filtering, resulted in an additional 3% improvement in accuracy by reducing distortions that could hinder feature extraction. Combined, these preprocessing steps contributed to a cumulative improvement in model accuracy of approximately 13%, demonstrating their critical role in achieving robust performance.

2.4. Feature Map

Feature Map to make it easier to understand the features of the inputs detected or retained by the model. A feature map is a visual representation of the result of feature extraction from data, such as images or signals, in an artificial neural network [26], specifically in convolutional neural networks (CNNs). The Feature Map provides an overview of the patterns or significant attributes that have been identified by the filter (or kernel) applied to enter data on different layers in the network. In the context of image processing on CNN, the feature map is created as follows [27], [28], [29]: When an input image is passed through a convolution layer, the filter performs a convolution operation on the image to detect specific features such as specific edges, textures, or patterns. The result of this operation is a feature map. After the convolution layer, a pooling layer can be applied to reduce the dimension of the feature map. This helps reduce the number of parameters and computations required and strengthens the dominant features. In CNNs, there are usually many filters used in a single layer of convolution. Each filter generates its own feature map, which represents different features from the input image. In the early layers of CNNs, feature maps usually capture basic patterns such as lines or angles, while in deeper layers, feature maps become more complex and able to recognize more abstract patterns such as shapes or objects as a whole.

2.5. Classification Models

The models used in the study are MobileNetV1, VGG16, Resnet and CNN Custom. The MobileNet model is highly efficient for developing mobile applications. This research will be developed for mobile-based applications. The way MobileNet works relies on an efficient architecture in convoluting using special techniques such as depthwise separable convolutions, Input images with a size of 224 x 224 are inserted into the network. The first layer is used to change the dimensions of the input image, then a series of depthwise separable convolution layers is applied. Each layer performs depthwise convolution followed by pointwise convolution. After the convolution layer, the resulting features are summarized using global average pooling to further reduce the dimensions. This last iDir is passed to the fully connected layer that issues predictions based on the output category. The final prediction is obtained through the softmax activation function, which generates probabilities from various output classes. The VGG model is very popular, especially in image processing, VGG16 consists of several structures, namely convolution layers, has 13 convolution layers arranged in blocks, each block consists of two or three convolutional layers in a row with a 3×3 size filter. The Pooling layer reduces the spatial dimension of the feature, reducing the number of parameters and avoiding overfitting, and the Fully Connected layer provides probabilities for each class. The ResNet model is widely used for object detection, the way the ResNet model works is the Initial Convolution Layer layer with a large filter of 7×77 \times 77×7, followed by a pooling layer to reduce dimensions. After the initial layer, the tissue consists of several residual blocks, each containing multiple layers of convolution and shortcut connections. After the residual block, the final layer consists of an average pooling and a fully connected layer that produces a classification output. The Custom CNN model is a model designed by researchers starting from the number and type of layers, filter size, activation function, number of neurons, and many other parameters.

The Custom CNN model was specifically designed to capture the nuances of dermatitis detection in infants. The architecture consists of three convolutional layers with increasing numbers of filters (16, 32, and 64), each followed by a ReLU activation function and max-pooling layers to down-sample the feature maps while preserving critical spatial information. The number of filters was chosen to progressively extract low- to high-level features from the input images. After feature extraction, the architecture includes two fully connected layers with 128 and 64 neurons, respectively, to combine extracted features and improve classification precision. A SoftMax activation function is applied in the output layer for multi-class classification. Dropout layers were included after the pooling and fully connected layers to reduce overfitting, and L2 regularization was employed to penalize overly complex weight configurations. This tailored architecture allows the model to effectively detect fine-grained differences in infant skin conditions, providing robust performance in distinguishing between dermatitis and healthy skin.

2.6. Evaluation

Evaluating the effectiveness of a model is essential to ensure accuracy, efficiency, and reliability in theoretical frameworks, contributing to the advancement and refinement of knowledge in the field [30], [31]. The model evaluation in this study uses a confusion matrix. The Confusion Matrix is one of the most efficient ways to analyze the performance of classification models, especially for binary and multi-class classifications. The Confusion Matrix provides a detailed overview of how the model's predictions compare to actual classes.

3. Result and Discussion

3.1. Dataset Description

In this study, 1088 datasets were used. With atopic dermatitis 748 and non-atopic dermatitis 340. The dataset is divided into training data and testing data with details of 80% training data of 748 and 20% of testing data of 136. The study categorized dermatitis into three main types: atopic dermatitis, contact dermatitis, and other forms of dermatitis. The majority of the dataset consisted of atopic dermatitis cases, reflecting its high prevalence in infants. Contact dermatitis and other less common forms were also included but constituted a smaller portion of the dataset. Each image was labeled based on clinical observation and caregiver interviews to ensure accurate categorization. This categorization allowed the models to learn and differentiate between these variations effectively, providing insights into the model's performance across different types of dermatitis. In figure 4(a) is normal skin in babies, and in figure 4(b) is skin with atopic dermatitis in babies.

(a) (b)

Figure 4. (a) Normal Skin (b)Skin Atopic Dermatitis

One of the primary limitations of this study is the relatively small dataset, consisting of 1,088 images, which may limit the generalizability of the model to broader populations. A small dataset can lead to overfitting, where the model performs well on the training data but struggles with unseen data. To address this issue, data augmentation techniques were employed during model training, including random flipping, rotation, and zooming. These techniques artificially expanded the dataset and introduced variability, which helped improve the model's robustness and accuracy. However, while augmentation mitigates some limitations, it does not fully replace the benefits of acquiring a larger and more diverse dataset. Future studies should focus on expanding the dataset by collecting images from multiple regions with varying environmental and demographic conditions. This would enhance the model's ability to generalize and improve its reliability in real-world applications.

3.2. Training Model

To measure the performance of the MobilNet, VGG16, Resnet and CNN Custom models in classifying, TensorFlow is used. TensorFlow measures the precision of object classification in aotpic dermatitis images [32], [33], as shown in table 1, the precision of the MobileNet model is 94.44% for Non Dermatitis and 100% for Dermatitis Atopic, with an overall weighted average precision of 97.22%. The VGG16 model achieves a precision of 83.87% for non-Dermatitis and 60.00% for Dermatitis Atopic, resulting in a weighted average precision of 71.94%. The ResNet model has a precision of 75.28% for non-Dermatitis and 97.87% for Dermatitis Atopic, with a weighted average precision of 86.58%. Lastly, the Custom CNN model records a precision of 34.91% for non-Dermatitis and 72.07% for Dermatitis Atopic, resulting in a weighted average precision of 60.43%.

Table 1. Model Evaluation

Model	Class	precision	accuracy	recall	f1-score
MobilenetV1	Non-Dermatitis	0.944444	0.97058	1.000000	0.971429
	Dermatitis Atopic	1.000000	0.970588	0.941176	0.969697
	accuracy	0.970588	0.970588	0.970588	0.970588
	macro avg	0.972222	0.970588	0.970588	0.970563
	weighted avg	0.972222	0.970588	0.970588	0.970563
VGG16	Non-Dermatitis	0.838710	0.654412	0.382353	0.525253
	Dermatitis Atopic	0.600000	0.654412	0.926471	0.728324
	accuracy	0.654412	0.654412	0.654412	0.654412
	macro avg	0.719355	0.654412	0.654412	0.626788
	weighted avg	0.719355	0.654412	0.654412	0.626788
ResNet	Non-Dermatitis	0.752809	0.830882	0.985294	0.853503
	Dermatitis Atopic	0.978723	0.830882	0.676471	0.800000
	accuracy	0.830882	0.830882	0.830882	0.830882
	macro avg	0.865766	0.830882	0.830882	0.826752
	weighted avg	0.865766	0.830882	0.830882	0.826752
CNN Custom	Non-Dermatitis	0.349057	0.539171	0.544118	0.425287
	Dermatitis Atopic	0.720721	0.539171	0.536913	0.615385
	accuracy	0.539171	0.539171	0.539171	0.539171
	macro avg	0.534889	0.539171	0.540515	0.520336
	weighted avg	0.604255	0.539171	0.529171	0.555815

In figure 5, you can see the graph of loss and validation loss in the MobileNetV1 and CNN Custom models, in the graph it can be seen that the values of training loss and validation loss decrease together and remain low, this indicates that the model is working well and is able to learn from the training data while making good generalizations to the validation data[34]. In contrast to the validation loss graph in the ResNet and VGG16 models, the graph shows that the training loss and validation loss values are not close.

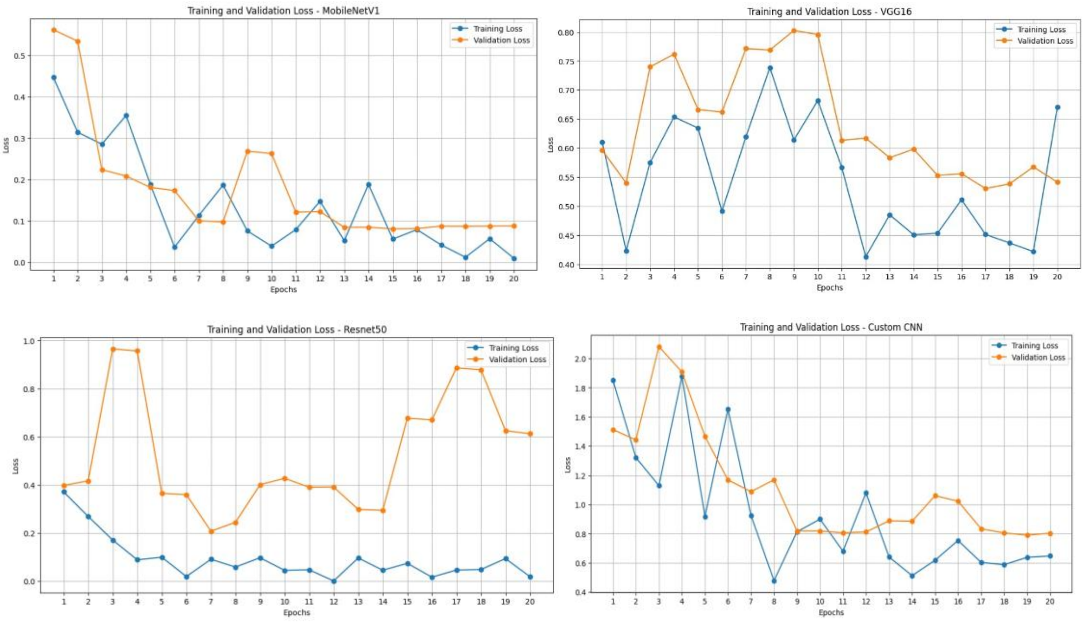

Figure 5. Loss Value and Validation Loss Graph

In figure 5, it can be seen that the training loss and validation loss values in the Custom CNN model remain low and close to it, this shows that the model is able to learn from the training data and can also generalize to the validation data, but in table 1 the accuracy value of the custom CNN model is 60%. So, the researcher raised the epoch value to find the best accuracy value on custom CNN. In figure 6, can see the values of the Loss and Validation Loss graphs on the custom CNN with 100 epochs, the graph shows that the values of Loss and Validation Loss on the Custom CNN model remain low and almost close. The accuracy value of the CNN Custom model with 100 epochs can be seen in table 2, which is 85%.

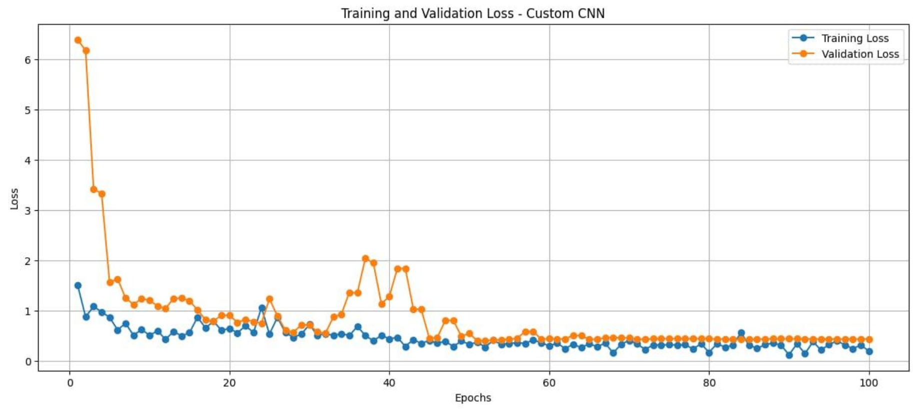

Figure 6. Custom CNN Loss Value and Loss Validation Graph

The loss value and loss validation graphs provide key insights into the models learning behavior. The staying low trend observed in these graphs indicates that the model is effectively minimizing error during training while maintaining a good generalization on the validation data. Specifically, a consistently low training loss suggests that the model is learning well from the training dataset, while a low and stable validation loss reflects that the model can generalize effectively to unseen data. Additionally, the convergence of training and validation loss values indicates that the model is not overfitting, as it performs well on both the training and validation datasets. These observations validate the robustness and reliability of the MobileNet and Custom CNN models for the task of infant dermatitis classification. Table 2 provides a detailed evaluation of the CNN Custom model after training for 100 epochs. This table highlights

 ISSN 2723-6471

key performance metrics that reflect the model's ability to classify between Non-Dermatitis and Dermatitis Atopic cases.

Table 2. Evaluation of the CNN Custom 100 epoch Model

Model	Class	precision	accuracy	recall	f1-score
CNN Custom	Non-Dermatitis	0.714286	0.85253	0.882353	0.789474
	Dermatitis Atopic	0.939850	0.85253	0.838926	0.886525
	accuracy	0.852535	0.85253	0.852535	0.852535
	macro avg	0.827068	0.85253	0.860640	0.837999
	weighted avg	0.869166	0.85253	0.852535	0.856112

Table 2 shows the performance metrics of the CNN Custom model, indicating its capability in classifying Non-Dermatitis and Dermatitis Atopic cases after 100 epochs of training. The model achieved an overall accuracy of 85.25%, reflecting a solid performance in distinguishing between the two classes. The precision, recall, and f1-score for Dermatitis Atopic are higher than for non-Dermatitis, suggesting that the model is more reliable in identifying dermatitis cases than non-dermatitis cases. This may be due to the distinctive features of dermatitis that the model can learn more effectively from the dataset. The macro average and weighted average scores further confirm the model's balanced performance across both classes, with weighted averages slightly higher, indicating that the model performs well even when considering class imbalances. Overall, these metrics demonstrate that the CNN Custom model, with its tailored structure, provides a robust approach for classifying skin conditions, making it a viable option for practical applications where reliable detection of dermatitis is essential.

3.3. Performance Analysis

To measure the accuracy of the model, a confusion matrix was used, with a value of 68 for the non-dermatitis classification and 64 for the dermatitis classification. Table 1 shows the accuracy value of the MobileNet model, which is 97%, which indicates that the classification model with MobileNet is included in the good classification[35]. In Figure 7, you can see the accuracy graph of the MobileNetV1 model with values that are almost close to and up, so that this model can be said to be good and the model learns well on the training data and is able to generalize well on the validation data[34]. The VGG16 model produces an accuracy value of 65% as seen in table 1, and when viewed in the graph, the accuracy value decreases. The ResNet model produces an accuracy score of 83%, and when viewed on the accuracy graph, the model is able to learn well as evidenced by the increasing accuracy value. In contrast to the CNN Custom model with 20 epochs, the accuracy value is 53%, and the graph value is decreasing. The following is Figure 7 (a) and figure 7 (b) which presents a graph of the test results using the confusion matrix and Accuracy Values.

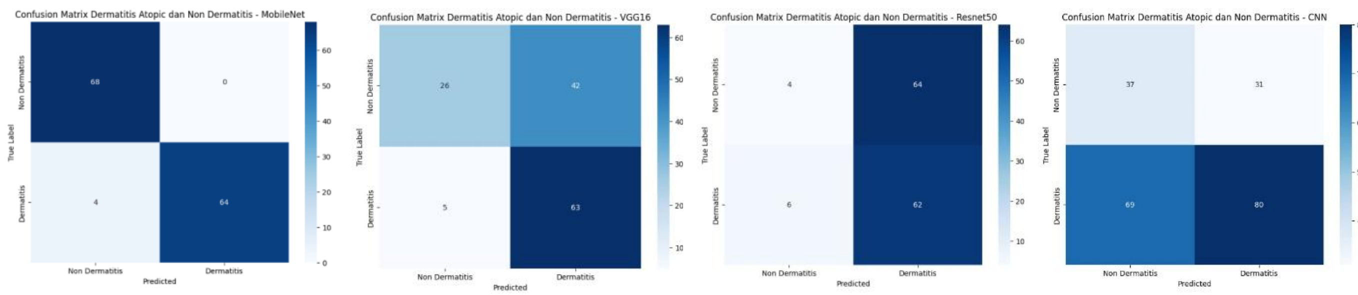

Figure 7. (a) Confusion Matrix 7

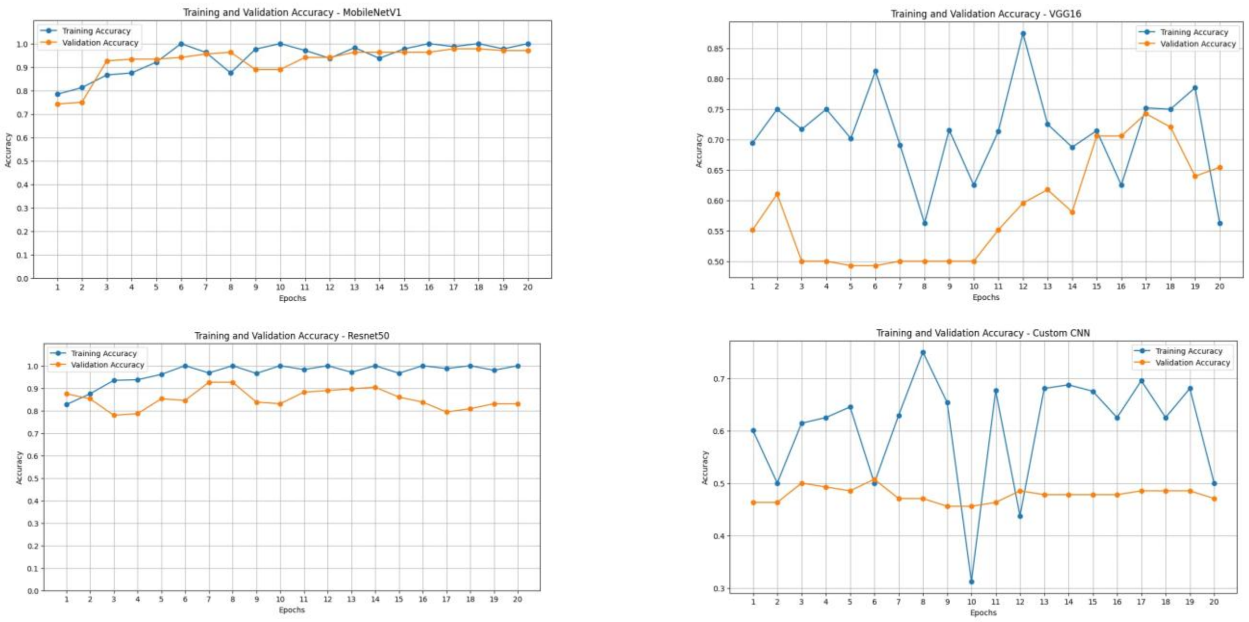

Figure 7. (b) Graph of Accuracy Values

However, when the epoch value is increased to 100 epochs in the CNN Custom model, the accuracy value increases with a value of 85% as seen in table 1. Likewise, the accuracy in figure 8 is increasing, this shows that the model learns well on the training data.

(a) (b)

Figure 8. (a) Confusion Matrix CNN Custom 100 epoch 9. (b) Accuracy Value Graph CNN Custom 100 epoch

The loss and accuracy metrics, along with validation curves, provide critical insights into the robustness of the models. For both MobileNet and Custom CNN, the training and validation loss curves converge and remain low, indicating that the models effectively generalize without significant overfitting. To further prevent overfitting, dropout layers were implemented, randomly deactivating neurons during training to reduce reliance on specific features. In contrast, models like VGG16 exhibited divergence between training and validation loss, a sign of overfitting, highlighting the need for additional regularization techniques or model tuning. These observations underscore the importance of monitoring these metrics and implementing regularization strategies to achieve robust model performance.

The table below presents a comparison of the performance of several deep learning models in detecting dermatitis in infants, including MobileNet and Custom CNN developed in this study. The purpose of this comparison is to evaluate the accuracy of the proposed model with other models commonly used in medical image classification. The following is table 3 which presents the Comparison of accuracy of deep learning models for detecting dermatitis in infants.

ISSN 2723-6471

Table 3. Accuracy Comparison of Deep Learning Models for Infant Dermatitis Detection

Model	Non-Dermatitis Precision	Dermatitis Precision	Non-Dermatitis Recall	Dermatitis Recall	Overall Accuracy
MobileNet (Our Model)	94.4%	100%	100%	94.1%	97%
Custom CNN (Our Model)	71.4%	93.9%	88.2%	83.8%	85%
ResNet	75.2%	97.8%	98.5%	67.6%	83%
VGG16	83.8%	60%	38.2%	92.6%	65%
EfficientNet	80%	88%	85%	80%	82%
AlexNet	70%	80%	75%	78%	76%
DenseNet	78%	85%	80%	82%	81%

This table provides a comparison of various deep learning models' performance in detecting dermatitis in infants, with a focus on MobileNet and Custom CNN developed in this study. All models were trained and tested on the same dataset, which was collected specifically for this study. The results show that the proposed MobileNet model achieves the highest overall accuracy at 97%, followed by the Custom CNN at 85%. Other models, including ResNet, VGG16, EfficientNet, AlexNet, and DenseNet, present lower accuracy levels, highlighting the superior predictive performance of the proposed models. This suggests that MobileNet and Custom CNN are promising solutions for efficient and accurate dermatitis detection in infants, suitable for practical applications in mobile and medical diagnostic tools.

The architectures of MobileNet, VGG16, and ResNet significantly differ in their approach to feature extraction and computational efficiency, which impacts their suitability for detecting skin conditions. MobileNet employs depthwise separable convolutions, reducing the number of parameters and computational costs while maintaining high accuracy. This makes it ideal for resource-constrained environments, such as mobile or real-time applications. In contrast, VGG16 uses a straightforward architecture with a stack of 13 convolutional layers and 3 fully connected layers, which excels in extracting detailed features but requires significantly more computational resources, making it less efficient for real-time use. ResNet introduces residual blocks that allow gradients to flow more easily during backpropagation, addressing the vanishing gradient problem and enabling the network to learn deeper representations. However, ResNet's higher complexity may lead to longer training times and increased computational demands. In the context of this study, MobileNet's lightweight architecture proved advantageous, achieving high accuracy with minimal computational overhead, while VGG16 and ResNet were better suited for scenarios where computational resources are less constrained but detailed feature extraction is prioritized.

The superior performance of MobileNet and Custom CNN in this study can be attributed to their architectural designs and computational efficiencies. MobileNet, designed with depthwise separable convolutions, significantly reduces the number of parameters and computational requirements while maintaining high accuracy. This makes it ideal for real-time applications, especially on mobile devices or in resource-limited environments. On the other hand, the Custom CNN was tailored specifically for the unique features of infant skin images in this dataset. By optimizing the number of layers and parameters, the Custom CNN effectively balances accuracy and computational cost, achieving strong performance metrics without overfitting. In contrast, VGG16 and ResNet, though robust for general image classification tasks, are computationally intensive and less suitable for deployment in real-time diagnostic tools due to their complex architectures and higher inference times. These findings underscore the importance of selecting models not only based on accuracy but also on their suitability for specific operational contexts, such as scalability and processing speed.

The performance of the models, while promising, is not without limitations. Several potential sources of error and areas for improvement were identified through error analysis. The dataset primarily focuses on dermatitis cases from specific regions in Riau Province, which may limit the generalizability of the model to other populations with different environmental or genetic factors. Additionally, the smaller proportion of non-dermatitis images may have introduced a class imbalance affecting the model's generalization ability. Variability in image quality, such as differences in lighting, resolution, or focus, could have impacted the model's ability to extract consistent features, with noise from poorly lit or shadowed images reducing prediction accuracy. The models also struggled with borderline cases, where mild or ambiguous dermatitis symptoms presented subtle visual differences, as reflected in lower precision and recall, particularly for the Custom CNN model. Finally, the dataset's geographic specificity may have introduced demographic or environmental biases, further limiting the broader applicability of the model.

4. Conclusion

With the best accuracy and in the MobileNet and CNN Custom models with epochs of 100 i.e. with values of 97% and 85% as well as the training loss and validation loss values decreasing together and remaining low, this shows that this model has strong potential to be used in practical applications to detect dermatitis in infants. MobileNet is known as an efficient and lightweight model for image classification applications, especially on resource-constrained devices such as mobile phones. If MobileNet shows the best accuracy value among other models, it means that it is making correct predictions more often than other models. This can be caused by. MobileNet uses special convolution blocks such as depthwise separable convolution, which reduces the number of parameters and computations without sacrificing accuracy. MobileNet is designed to maintain a balance between model complexity and speed, which makes it more adaptable to classification tasks such as the detection of dermatitis in infants. Custom CNN models can be designed specifically and tailored for specific tasks or needs, usually with a unique structure or architecture and are different from standard CNN models, especially in medical image object recognition, this is proven by increasing accuracy values along with lower training loss and validation loss values. If the training loss and validation loss decrease consistently from the beginning to the end of the training, this is a good sign, it shows that the model is able to learn from the training data and can also generalize on the validation data.

5. Declarations

5.1. Author Contributions

Conceptualization: D.S., R.N.P., I.F., A.N.H., Y.I., N.H.; Methodology: N.H.; Software: D.S.; Validation: D.S., N.H., and Y.I.; Formal Analysis: D.S., N.H., and Y.I.; Investigation: D.S.; Resources: N.H.; Data Curation: N.H.; Writing Original Draft Preparation: D.S., N.H., and Y.I.; Writing Review and Editing: N.H., D.S., and Y.I.; 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 2024 Fundamental Research Program with contract number 112/E5/PG.02.00.PL./2024

5.4. Institutional Review Board Statement

Not applicable.

5.5. Informed Consent Statement

Not applicable.

5.6. Declaration of Competing Interest

The authors accepted and declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The research solely for the development of scienceth

References

[1] O. Mahmoud, G. Yosipovitch, and E. Attia, "Burden of Disease and Unmet Needs in the Diagnosis and Management of Atopic Dermatitis in the Arabic Population of the Middle East," *J. Clin. Med.*, vol. 12, no. 14, pp. 1-17, 2023, doi: 10.3390/jcm12144675.

[2] C. Gu, X. Yao, and W. Li, “Burden of Disease; the Current Status of the Diagnosis and Management of Atopic Dermatitis in China,” *J. Clin. Med.*, vol. 12, no. 16, pp. 1-12, 2023, doi: 10.3390/jcm12165370.

[3] A. Wollenberg, T. Werfel, J. Ring, H. Ott, U. Gieler, and S. Weidinger, “Atopic Dermatitis in Children and Adults Diagnosis and Treatment,” *Dtsch. Arztebl. Int.*, vol. 120, no. 13, pp. 224–234, 2023, doi: 10.3238/arztebl.m2023.0011.

[4] J. I. Silverberg, “Atopic dermatitis in the pediatric population: A cross-sectional, international epidemiologic study,” *Ann. Allergy, Asthma Immunol.*, vol. 126, no. 4, pp. 417–428, 2021, doi: 10.1016/j.anai.2020.12.020.

[5] J. Chittock, Linda Kay, “Association between skin barrier development and early-onset atopic dermatitis: A longitudinal birth cohort study,” *J. Allergy Clin. Immunol.*, vol. 153, no. 3, pp. 732-741.e8, 2024, doi: 10.1016/j.jaci.2023.10.017.

[6] A. B. Pavel, Jianni Wu., “Tape strips from early-onset pediatric atopic dermatitis highlight disease abnormalities in nonlesional skin,” *Allergy Eur. J. Allergy Clin. Immunol.*, vol. 76, no. 1, pp. 314–325, 2021, doi: 10.1111/all.14490.

[7] S. Wu, Li Lei, Yibo Hu, “Machine learning-based prediction models for atopic dermatitis diagnosis and evaluation,” *Fundam. Res.*, vol. 2, no. April, pp. 1-10, 2023, doi: 10.1016/j.fmre.2023.02.021.

[8] D. M. R. Sari, S. Nurmaini, D. P. Rini, and A. I. Sapitri, “Dermatitis Atopic and Psoriasis Skin Disease Classification by using Convolutional Neural Network,” *Comput. Eng. Appl. J.*, vol. 12, no. 1, pp. 1–10, 2023, doi: 10.18495/comengapp.v12i1.419.

[9] M. Hammad, P. Pławiak, M. ElAffendi, A. A. A. El-Latif, and A. A. A. Latif, “Enhanced Deep Learning Approach for Accurate Eczema and Psoriasis Skin Detection,” *Sensors*, vol. 23, no. 16, pp. 1-17, 2023, doi: 10.3390/s23167295.

[10] Y. I. Park, S H Coi, M C Han., “A New Approach to Quantify and Grade Radiation Dermatitis Using Deep-Learning Segmentation in Skin Photographs,” *Clin. Oncol.*, vol. 35, no. 1, pp. e10–e19, 2023, doi: 10.1016/j.clon.2022.07.001.

[11] S. S. Noronha, M. A. Mehta, D. Garg, K. Kotecha, and A. Abraham, “Deep Learning-Based Dermatological Condition Detection: A Systematic Review With Recent Methods, Datasets, Challenges, and Future Directions,” *IEEE Access*, vol. 11, no. December, pp. 140348–140381, 2023, doi: 10.1109/ACCESS.2023.3339635.

[12] 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.

[13] M. R. Hall, Weston D A., “An Automated Approach for Diagnosing Allergic Contact Dermatitis Using Deep Learning to Support Democratization of Patch Testing,” *Mayo Clin. Proc. Digit. Heal.*, vol. 2, no. 1, pp. 131–138, 2024, doi: 10.1016/j.mcpdig.2024.01.006.

[14] S. Hao, Xiong Ying, Guo Sisi, “Development and performance validation of a low-cost algorithms-based hyperspectral imaging system for radiodermatitis assessment,” *Biomed. Opt. Express*, vol. 14, no. 9, pp. 4990-5004, 2023, doi: 10.1364/boe.500067.

[15] P. Guimarães, A. Batista, M. Zieger, M. Kaatz, and K. Koenig, “Artificial Intelligence in Multiphoton Tomography: Atopic Dermatitis Diagnosis,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–9, 2020, doi: 10.1038/s41598-020-64937-x.

[16] A. Dautovic, Dondras B, “Diagnosis of Atopic dermatitis Using Artificial Neural Network,” *IFAC-PapersOnLine*, vol. 55, no. 4, pp. 51–55, 2022, doi: 10.1016/j.ifacol.2022.06.008.

[17] H. Fonda, Y. Irawan, R. Melyanti, R. Wahyuni, and A. Muhaimin, "A Comprehensive Stacking Ensemble Approach for Stress Level Classification in Higher Education," Journal of Applied Data Sciences, vol. 5, no. 4, pp. 1701–1714, 2024.

[18] J. Hocke, J. Krauth, “Computer-aided classification of indirect immunofluorescence patterns on esophagus and split skin for the detection of autoimmune dermatoses,” *Front. Immunol.*, vol. 14, no. February, pp. 1–11, 2023, doi: 10.3389/fimmu.2023.1111172.

[19] S. K. Jangir, “Functional link convolutional neural network for the classification of diabetes mellitus,” *Int. j. numer. method. biomed. eng.*, vol. 37, no. 8, pp. 1-17, 2021, doi: 10.1002/cnm.3496.

[20] A. Laishram, “Automatic Classification of Oral Pathologies Using Orthopantomogram Radiography Images Based on Convolutional Neural Network,” *Int. J. Interact. Multimed. Artif. Intell.*, vol. 7, no. 4, pp. 69–77, 2022, doi: 10.9781/ijimai.2021.10.009.

[21] A. Lubis, Y. Irawan, J. Junadhi, and S. Defit, "Leveraging K-Nearest Neighbors with SMOTE and Boosting Techniques for Data Imbalance and Accuracy Improvement," Journal of Applied Data Sciences, vol. 5, no. 4, pp. 1625–1638, 2024.

[22] I. Kandel, “Comparative Study of First Order Optimizers for Image Classification Using Convolutional Neural Networks on Histopathology Images,” *J. Imaging*, vol. 6, no. 9, pp. 1-17, 2020, doi: 10.3390/JIMAGING6090092.

ISSN 2723-6471

[23] K. A. Alqahtani, “Deep convolutional neural network-based automated segmentation and classification of teeth with orthodontic brackets on cone-beam computed-Tomographic images: A validation study,” *Eur. J. Orthod.*, vol. 45, no. 2, pp. 169–174, 2023, doi: 10.1093/ejo/cjac047.

[24] J. Xia, “Analysis and classification of oral tongue squamous cell carcinoma based on Raman spectroscopy and convolutional neural networks,” *J. Mod. Opt.*, vol. 67, no. 6, pp. 481–489, 2020, doi: 10.1080/09500340.2020.1742395.

[25] H. Herianto, B. Kurniawan, Z. H. Hartomi, Y. Irawan, and M. K. Anam, "Machine Learning Algorithm Optimization using Stacking Technique for Graduation Prediction," Journal of Applied Data Sciences, vol. 5, no. 3, pp. 1272–1285, 2024.

[26] R. Yakkundimath, “Classification of Rice Diseases using Convolutional Neural Network Models,” *J. Inst. Eng. Ser. B*, vol. 103, no. 4, pp. 1047–1059, 2022, doi: 10.1007/s40031-021-00704-4.

[27] G. Li, “Diagonal-kernel convolutional neural networks for image classification,” *Digit. Signal Process. A Rev. J.*, vol. 108, no. January, pp. 1-17, 2021, doi: 10.1016/j.dsp.2020.102898.

[28] P. Akhenia, “Fault severity classification of ball bearing using SinGAN and deep convolutional neural network,” *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 236, no. 7, pp. 3864–3877, 2022, doi: 10.1177/09544062211043132.

[29] M. Wisniewski, “Drone Model Classification Using Convolutional Neural Network Trained on Synthetic Data,” *J. Imaging*, vol. 8, no. 8, pp. 1-17, 2022, doi: 10.3390/jimaging8080218.

[30] Y. Irawan, “Decision Support System For Employee Bonus Determination With Web-Based Simple Additive Weighting (SAW) Method In PT. Mayatama Solusindo,” *J. Appl. Eng. Technol. Sci.*, vol. 2, no. 1, pp. 7–13, 2020.

[31] U. Rahmalisa, A. Febriani, and Y. Irawan, “Detector leakage gas LPG based on telegram notification using wemos D1 and MQ-6 sensor,” *J. Robot. Control*, vol. 2, no. 4, pp. 287–290, 2021, doi: 10.18196/jrc.2493.

[32] G. Temidayo Adekunle and A. C. Aladeyelu, “Image Classification of Automobiles Using Deep Learning in Tensorflow,” *J. Multidiscip. Eng. Sci. Technol.*, vol. 10, no. 3, pp. 15818–15822, 2023.

[33] M. Naufal, A. Saputro, F. Liantoni, and D. Maryono, “Application of Convolutional Neural Network (CNN) Using TensorFlow as a Learning Medium for Spice Classification,” vol. 16, no. 1, pp. 8-15, 2024.

[34] M. M. Islam, P. Barua, M. Rahman, T. Ahammed, L. Akter, and J. Uddin, “Transfer learning architectures with fine-tuning for brain tumor classification using magnetic resonance imaging,” *Healthc. Anal.*, vol. 4, no. May, pp. 100270-1000280, 2023, doi: 10.1016/j.health.2023.100270.

[35] A. Maulana, T R Noviandy., “Evaluation of atopic dermatitis severity using artificial intelligence,” *Narra J*, vol. 3, no. 3, pp. 1-11, 2023, doi: 10.52225/narra.v3i3.511.