Improved Classification of Arsenic-Affected Skin Diseases Through Image Processing and Transfer Learning

Rudi Kurniawan[1], Samsuryadi[2*], and Fatma Susilawati Mohamad[3]
[1]Department of Computer System Engineering, Faculty of Engineering Science,
Universitas Bina Insan
Lubuklinggau, Indonesia 31629
[2]Department of Informatics Engineering, Faculty of Computer Science, Universitas Sriwijaya
Palembang, Indonesia 30129
[3]Department of Information Technology, Faculty of Informatics and Computing,
Universiti Sultan Zainal Abidin
Kuala Terengganu, Malaysia 22200
Email: [1]rudi.kurniawan@univbinainsan.ac.id, [2]samsuryadi@unsri.ac.id, [3]fatma@unisza.edu.my

***Abstract*—Arsenic contamination of groundwater is a global health concern, leading to adverse health effects, including skin diseases. Early detection is crucial for prevention and treatment, although manual methods are often time-consuming and error-prone. To address this issue, deep learning methods, specifically transfer learning, offer a promising solution for accurate and efficient skin disease detection. Therefore, the research aims to propose a comprehensive framework that uses ResNet152V2 architecture along with Gaussian smoothing methods to improve the classification accuracy of skin images exposed to arsenic. ResNet152V2 model is pre-trained on large-scale image datasets, providing powerful feature extraction fine-tuned on the ArsenicSkinBD dataset. The images are preprocessed using Gaussian smoothing to reduce noise and enhance feature clarity. Specifically, the research introduces the innovative application of Gaussian smoothing along with transfer learning for skin disease classification, which has not been extensively explored in previous studies. The results show a significant increase in classification accuracy, achieving approximately 0.9904 on the testing set compared to 0.9881 without enhancements. This improvement shows the effectiveness of the method in detecting skin diseases caused by arsenic exposure. The use of Gaussian smoothing also reduces loss values on the testing set, indicating that the model becomes more efficient in optimizing its predictions. The proposed framework not only enhances detection accuracy but also supports more efficient diagnostic processes, contributing to better prevention and treatment efforts.**

***Index Terms*—Arsenic-Affected Skin Diseases, Image Processing, Transfer Learning**

Received: July 14, 2024; received in revised form: Feb. 06, 2025; accepted: Feb. 24, 2025; available online: April 25, 2025.
*Corresponding Author

I. Introduction

ARSENIC is currently a global health concern, particularly in the public health community. High levels of arsenic in the soil can contaminate groundwater, which is an important source of drinking water in many areas [1, 2]. The impacts of long-term exposure to groundwater are considered a public health threat. It is because of the potential to increase the risk of skin cancer, diabetes, cardiovascular disease, bad pregnancy events, and reduced intellectual functioning in children [3].

Arsenic affects community through the skin in the form of skin lesions, patches, or unusual discoloration [4]. Using advanced medical imaging technology, healthcare professionals can identify the characteristic signs of skin disease, which is often an indicator of harmful arsenic exposure to human health. Hence, early detection through visual observations is important to initiate appropriate treatment and prevent further complications that are caused by arsenic exposure.

For the identification of skin diseases caused by arsenic, deep learning has been proven effective. Using deep learning to analyze medical images, such as processing and pattern recognition, the researchers can use the trained system to recognize typical patterns of skin diseases [5–8]. Since the use of deep learning in skin disease detection has shown great promise [9], adversarial and progressive transfer learning methods are used to improve the performance of cross-domain skin disease detection. These methods have high performance in the detection of melanoma and cancer

on various skin disease datasets. Skin cancer detection from digital medical images is still an active study area for more precise results, although many efforts have been made over the years. As reported in a previous study [10], a single and fairly lightweight deep learning model has achieved the best results in skin cancer detection. It is implemented in a mobile application after training and validation with HAM10000 dataset.

Arsenic exposure has been related to an increased risk of skin cancers, including squamous cell carcinoma, Bowen's disease, and basal cell carcinoma [11]. Understanding the detrimental effects of arsenic on skin health is essential for developing effective detection and treatment strategies. Chronic arsenic exposure often leads to characteristic skin changes such as hyperpigmentation and keratosis, which can serve as early warning signs before malignancy develops. Moreover, arsenic's interference with cellular repair mechanisms and immune response contributes to the carcinogenic process, making it crucial to monitor at-risk populations closely.

Several studies have explored advanced methods for skin lesion detection using deep learning methods. For instance, DermoExpert, a computerized dermoscopic framework, integrates preprocessing methods and hybrid Convolutional Neural Network (CNN) to enhance skin lesion detection accuracy, achieving an Area Under the Curve (AUC) of 0.97 [12]. Another previous research presents a transfer learning model for automatic classification of melanoma and nevus lesions, yielding accuracies of 0.9591, 0.9686, and 0.9770 across three datasets [13].

Several previous studies using AlexNet have shown high precision in classifying multi-class skin lesions, achieving accuracy rates of 0.9870 and above when validated against the International Skin Imaging Collaboration (ISIC) 2018 dataset [14]. Additionally, a framework combining deep learning with Internet of Health Things (IoHT) has shown promise in classifying skin lesions with an impressive accuracy of 0.99 [15].

The impact of image preprocessing methods on classification performance has been investigated. Previous research has showed that cropping images outperforms resizing in classification tasks, with a modified CNN fusion method achieving accuracy [16], Furthermore, a GoogleNet pre-trained model based on transfer learning successfully identified eight classes of skin lesions with accuracy of 0.9492 [17].

Automated skin cancer identification is crucial for enhancing accuracy and expertise of pathologists at an early stage. A previous research conducted [18] has proposed a Deep Convolutional Neural Network (DCNN) model using transfer learning for accurately classifying benign and malignant skin lesions. The model shows impressive training and testing accuracy of 0.9316 and 0.9193, respectively, after applying various preprocessing filters to eliminate noise and artifacts.

In another research [19], an optimal Deep Neural Network (DNN) based on probability is developed to detect skin diseases accurately. This classification tool achieves a significant 0.95 accuracy rate, with sensitivity at 0.91 and specificity at 0.97 by using the Whale optimization method. Then, a method for classifying skin images has been introduced using CNN optimized with the Spatial and Spectral Attention (SpaSA) meta-heuristic optimizer [20]. It achieves high accuracy rates across multiple datasets, including 0.9827 for ISIC Melanoma dataset (2019), 0.9883 for Human Against Machine with 10000 training images (HAM10000), and 0.8587 for the Skin Disease Images dataset. The significance of deep learning model in skin cancer detection is further emphasized [21], where a vision transformer is introduced with accuracy of 0.9615 using HAM10000 dataset. Additionally, deep learning model improves robustness through advanced preprocessing methods.

Then, previous studies have explored practical methods for detecting skin cancer through image analysis to enhance specialists' ability to distinguish between malignant tumors and benign ones [22]. Various methods such as Swarm Intelligence for locating lesions in areas of interest have shown promising results with classification accuracies reaching approximately 0.9852. According to another research [23], Canny edge detection combines with CNN to classify multi-class skin cancers using ISIC datasets, achieving an impressive accuracy of 0.99. Lastly, a deep-learning method trained on approximately 10,000 images shows over 0.95 accuracy for test datasets in distinguishing between malignant and benign lesions [24].

Although these studies have made significant contributions, the analysis often neglects the specific impact of preprocessing methods like Gaussian smoothing on classification performance. Therefore, the research aimed to investigate the effects of Gaussian smoothing combined with transfer learning on the classification accuracy of skin images affected by arsenic exposure. Moreover, the main contributions of the research are as follows:

1) Propose a framework and analyze the concept of transfer learning using ResNet152V2 architecture, which is previously trained for the detection and classification of arsenic-affected as well as unaffected skin.
2) Provide an analysis of comparative performance measures on the effect of using noise removal

Cite this article as: R. Kurniawan, Samsuryadi, and F. S. Mohamad, "Improved Classification of Arsenic-Affected Skin Diseases Through Image Processing and Transfer Learning", CommIT Journal 19(1), 71–88, 2025.

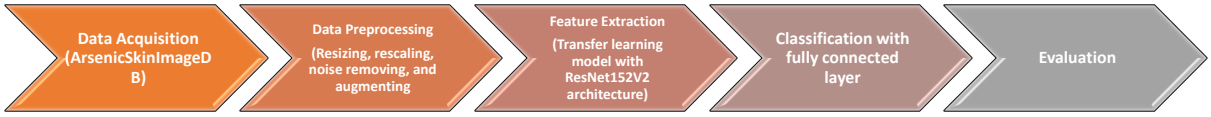

Fig. 1. Proposed framework.

(Gaussian blur) and data augmentation methods on ResNet152V2 architecture, which is related to the performance of increasing accuracy in the context of transfer learning

II. Research Method

In this section, the dataset used is introduced. Then, it is followed by a detailed description of the proposed method, which includes image preprocessing, the introduction of Residual Networks (ResNet), and the evaluation process, as shown in Fig. 1. The image preprocessing step involves techniques to enhance image quality and normalize data, ensuring better feature extraction by the model. ResNet, a deep convolutional neural network known for its residual learning capability, is employed to improve classification accuracy and mitigate the vanishing gradient problem. Finally, the evaluation process uses standard metrics to rigorously assess the model's performance and generalizability on unseen data.

A. Dataset Acquisition

ArsenicSkinImageBD dataset consists of a total of 8,892 images, where each half represents healthy skin (4,466) and skin affected by arsenic exposure (4,466) [25, 26]. Each image has been converted to a resolution of 244 × 244 pixels for further analysis. All images in this dataset are saved in standard Portable Network Graphic (PNG) format.

B. Image Preprocessing

First, image resizing is intended to change the size of original image [27]. Resizing data related to image processing has several advantages. First, in memory and computational savings, data that are resized to a smaller size can save memory and reduce computational burden, specifically when working with large datasets. Second, it has more efficient processing. Resized data have the potential to be processed more quickly by the model because of smaller size and easy processing. Third, in size consistency, image resizing to a consistent size ensures that all input data have the same dimensions for identical model training process. Fourth, it has easier processing. Resized data are easily organized and manipulated, particularly when used in a series of data processing algorithms or model training.

Second, dataset rescaling is a normalization method commonly used in deep-learning image processing [28]. It is achieved by dividing every pixel value in the image by 255, which is the maximum value a color channel can provide on Red, Green, Blue(RGB) scale. Equation (1) shows the mathematical function of normalized pixels. The normalized pixel is the adjusted value, and real pixel ranges from 0 to 255. With the normalization, each pixel value is in the range between 0 and 1, which is easier for deep learning models to process.

$$\text{Normalized Pixel} = \frac{\text{Real Pixel}}{255}. \tag{1}$$

Third, Gaussian filtering is a spatial filter commonly used in image processing to decrease noise and extract important features [29, 30]. This filter works by averaging the pixel intensity around each pixel in the image using Gaussian function. Equation (2) shows the function of the Gaussian filtering. The $G(x, y)$ is the Gaussian value at coordinates (x, y). Then, σ is a parameter that controls the sharpness of the Gaussian curve. Meanwhile, x and y are the distances from the middle point of the filter to the pixel being processed. Figure 2 shows the process of removing noise from original input image.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}. \tag{2}$$

Fourth, data augmentation is a method used in deep learning to enhance the diversity and quantity of training data. It is done by generating various versions of existing images [31]. Augmentation aims to reduce overfitting, improve model generalization, and enhance the model robustness to variations in input data [32]. Several augmented methods have been used like rotation, mirroring, and flipping. A sample of image data augmented is shown in Fig. 3.

C. Transfer Learning for Feature Extraction

Transfer learning-based feature extraction is a powerful method in machine learning that uses a pre-trained model on a specific task to extract features from new data [33]. The method uses the knowledge embedded in a pre-trained model. Furthermore, it allows

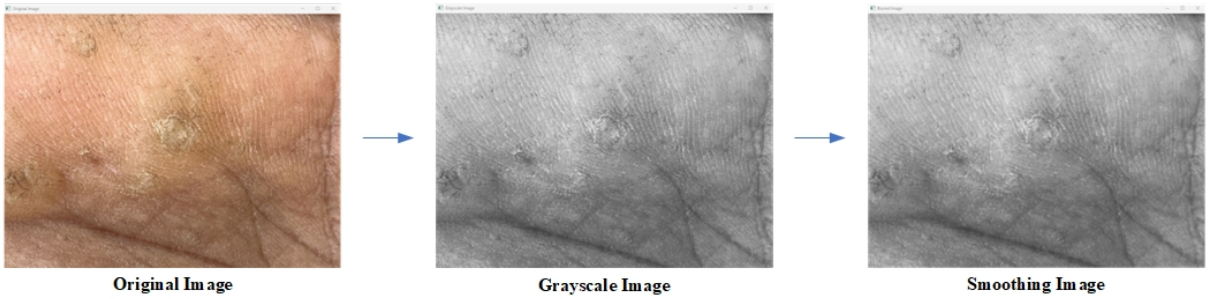

Fig. 2. Noise removal with Gaussian smoothing.

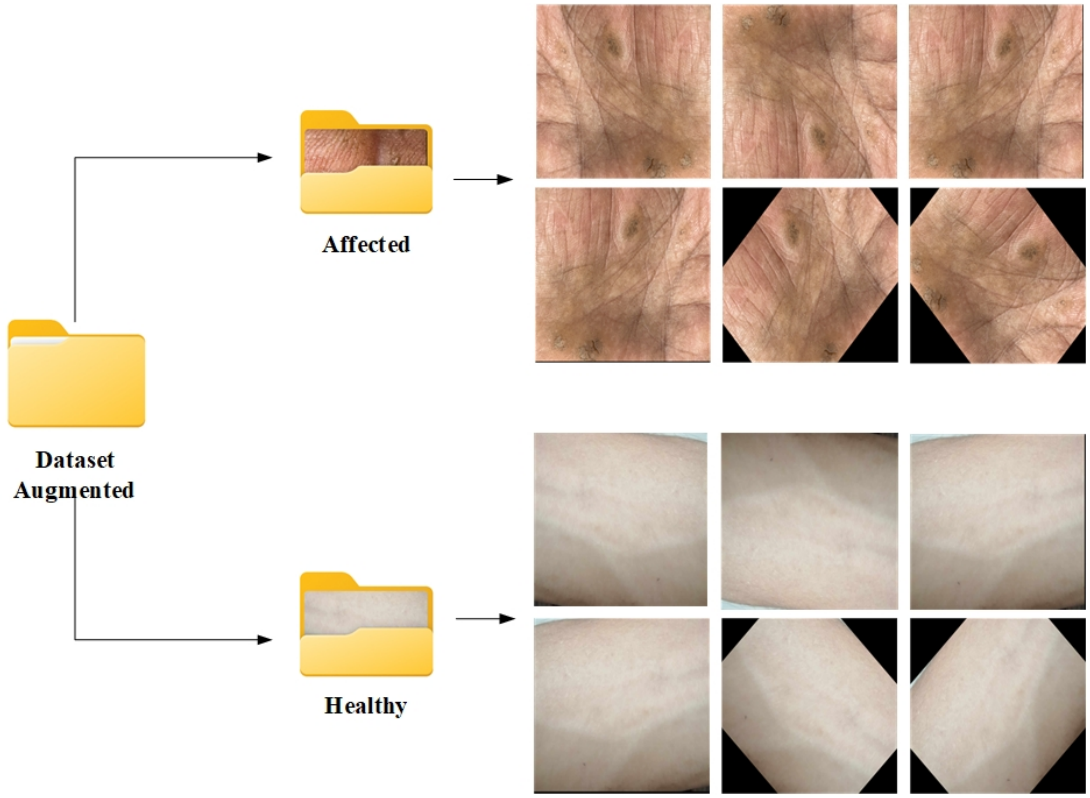

Fig. 3. Sample of augmented images.

to bypass the need for training a model from scratch, which can be resource-intensive and time-consuming. The process of transfer learning for feature extraction typically includes the following steps:

1) Select a pre-trained model: A model that has been pre-trained on a similar task or a large and diverse dataset needs to be selected. For the research, ResNet152V2 architecture is used, which shows strong performance in image classification tasks. Pre-trained models like ResNet152V2 have learned to identify general features such as shapes, textures, and patterns from extensive datasets.
2) Freeze convolutional layers: In this step, the initial convolutional layers of pre-trained model are frozen. It suggests that during subsequent training, the weights of the layers will not be updated. Freezing layers also preserve the general characteristics learned by the model, which are crucial for effective feature extraction.
3) Add a new classification layer: A new classification layer is appended to the model. The layer typically consists of one or more fully connected layers leading to an output with a softmax activation function. Specifically, it is designed to classify new datasets based on the features extracted from frozen convolutional layers.

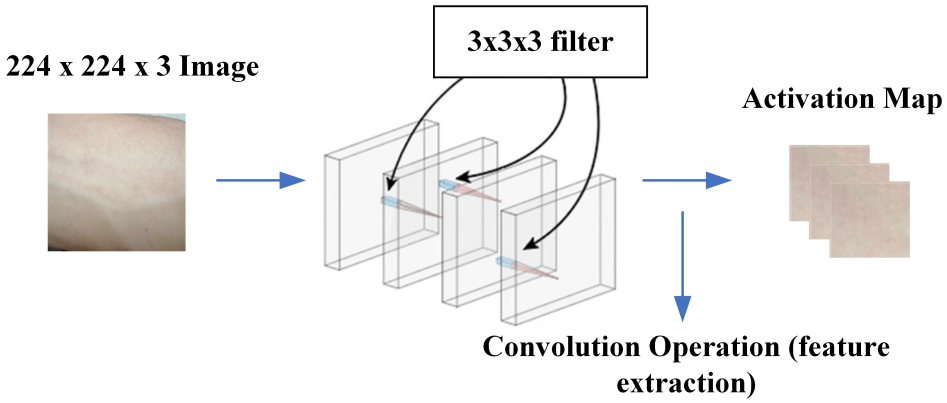

Fig. 4. Multi-layer convolutional process in the residual network.

4) Conduct advanced training: The newly added classification layers are then trained on a smaller, task-specific dataset that pertains to the problem, namely classifying skin images affected by arsenic. During training phase, only the weights of the new classification layers are updated while keeping the weights of original convolutional layers frozen. It allows the model to adapt to the specifics of new tasks without losing the general features learned during pre-training.

By using transfer learning-based feature extraction with ResNet model, valuable features can be extracted from the images. It is trained on a large dataset using existing knowledge to build a better classification model through less training data. This approach leverages pre-trained weights from extensive datasets, allowing the model to recognize complex patterns and textures which are relevant to arsenic-related skin conditions. Consequently, transfer learning reduces the need for a large amount of labeled data and shortens training time while maintaining high accuracy and robustness in classification tasks.

D. ResNet152V2

ResNet152V2 is a variant of a very deep and complex CNN architecture consisting of a total of 152 layers. It belongs to ResNet family of models, which is known for its capabilities in deep network training [34, 35]. ResNet152V2 enables deeper learning without performance degradation by using residual and bottleneck blocks. Using batch normalization and dropout, ResNet152V2 can prevent overfitting during training. Furthermore, it is a powerful option for computer vision tasks where the representation of increasingly complex features from image data is a requirement. ResNet152V2 architecture consist of several parts.

First, it is convolutional block. ResNet152V2 consists of a set of basic blocks that contain multiple consecutive convolution layers. Each basic block comprises two or three convolutional layers, which are then followed by batch normalization and Rectified Linear Unit (ReLU) activation, enhancing clarity and flow. The formula for the convolution operation is shown in Eq. (3). It has x as input for this layer, W as the convolution weight matrix, b as a bias, $*$ as convolution operator, and σ as an activation function, such as ReLU.

$$y = \sigma(W * x + b). \tag{3}$$

Figure 4 shows the multi-layer convolutional process in the residual network. This network architecture employs a series of convolutional blocks, each containing multiple convolutional layers designed to extract increasingly complex features from the input images. Within these blocks, convolutional layers are followed by batch normalization and ReLU activation to improve clarity and flow of information, enhancing the network's overall performance. The successive application of these convolutional blocks enables the network to learn intricate patterns while maintaining efficient information propagation.

Second, generally, batch normalization is used to increase stability and speed up training artificial neural networks [36]. The formula for batch normalization is shown in Eq. (4). It has x as the input for the layer, μ as the batch mean, σ^2 as the batch variance, γ and β as the learned parameters, and ϵ as the stability factor to prevent division by zero.

$$\text{BN}(x) = \gamma \frac{(x - \mu)}{\sqrt{(\sigma^2 + \epsilon)}} + \beta. \tag{4}$$

Third, one of the key innovations of ResNet is the use of residual block, which enables deeper learning

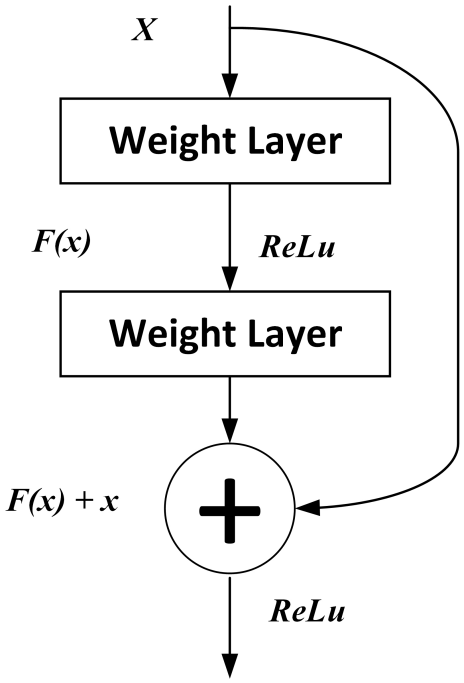

Fig. 5. Multi-layer convolutional process in the residual network.

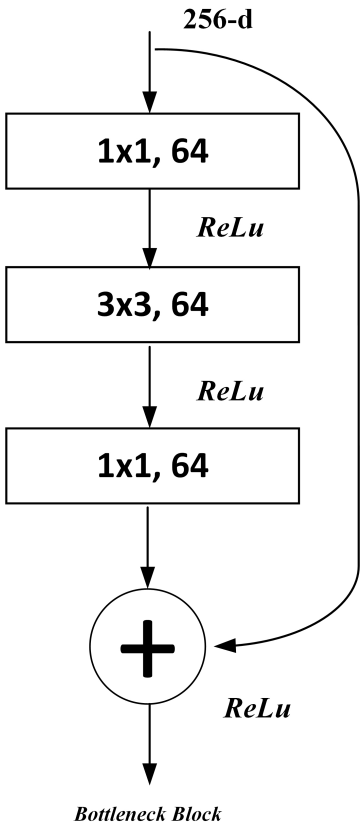

Fig. 6. Bottleneck block.

by mitigating the problem of performance degradation as the network becomes deeper [37]. Residual block allows direct flow from input to output, bypassing convolutional layers. It allows the network to easily learn identities, facilitating training of deep networks. The formula for residual block is shown in Eq. (5). It has x as the input to the residual block, $F(x, \{W_{ii}\}$ as the function of the learning in the residual block, and W_i as the learned parameter.

$$\text{Output} = F(x, \{W_{ii}\}) + x. \tag{5}$$

A building block of the residual process is shown in Fig. 5. It highlights a key element of ResNet models, designed to promote efficient learning in deep architectures. Each block typically incorporates multiple convolutional layers arranged in sequence, enabling the network to extract complex and hierarchical features. This design facilitates the network's ability to learn identity mappings, which eases the optimization of deeper networks.

Fourth, ResNet152V2 uses bottleneck block to decrease the number of parameters used by the network while preserving strong learning capabilities [38]. Bottleneck block comprises a 1×1 convolution layer that reduces the dimensions of the feature, succeeded by a more extensive convolution layer, and another 1×1 to enhance the dimensions. Figure 6 shows bottleneck block for ResNet152. The block reduces computational complexity, enabling deeper networks to be trained without excessive resource demands by compressing the feature maps initially. This approach maintains representational power by re-expanding the feature dimensions after processing, ensuring that crucial information is retained for subsequent layers. The bottleneck design effectively balances network depth and computational efficiency, making ResNet152 a powerful architecture for complex image recognition tasks.

Fifth, in deep architecture, ResNet152V2 has a total of 152 layers, making it a very deep network. This depth allows higher learning of complex feature representations from image data. Figure 7 shows the overall architecture of ResNet152. In summary, ResNet152 leverages its deep architecture and residual connections to enable effective learning of complex features from input images. The network employs bottleneck blocks to manage computational complexity while maintaining high representational power. This combination of depth, residual connections, and bottleneck structures allows ResNet152 to excel in various image recognition tasks.

Sixth, loss function is used to measure how well

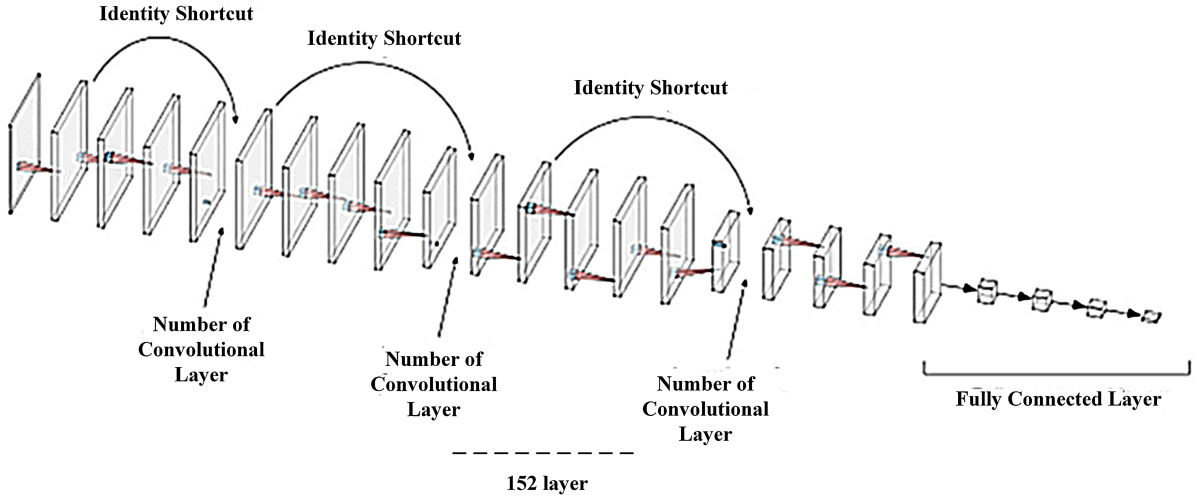

Fig. 7. ResNet152 Architecture.

the model predicts the target [39]. In classification cases, the cross-entropy loss is often used with the formulation in Eq. (6). It has N as the number of samples, y_i as the actual label, and $\hat{y}_i$ as the model prediction for the i sample.

$$\text{Loss} = -\sum_{i=1}^{N} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right]. \quad (6)$$

E. Data Evaluation

To assess the classification performance, the researchers obtain Confusion Matrix (CM). CM is an important performance evaluation tool in classification that allows visualization of model performance in more detail [40]. Figure 8 shows how well the predicted results of the model classification match the actual values of observed data. From these components, several important classification evaluation metrics can be calculated.

1) Accuracy of the model predictions, which is calculated according to Eq. (7), is represented as a percentage of correctly predicted samples. It describes TP as True Positive, TN as True Negative, FP as False Positive, and FN as False Negative.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}. \quad (7)$$

2) Precision is calculated using Eq. (8) and shows the percentage of positive data samples accurately predicted by the model.

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (8)$$

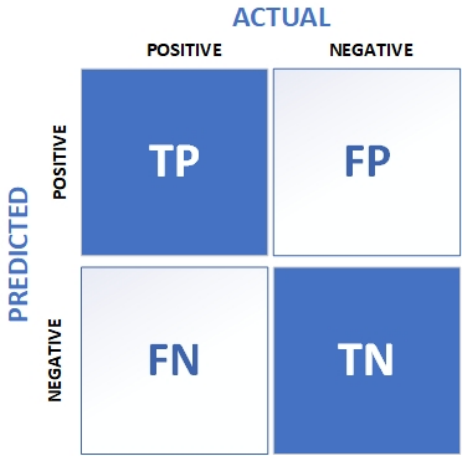

Fig. 8. Confusion matrix.

3) Recall is also known as sensitivity or true positive rate. It shows the percentage of positive samples that the model correctly predicts, which is calculated using Eq. (9).

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (9)$$

4) The F1-score is calculated by taking the harmonic mean of precision and recall, as represented by Eq. (10).

$$\text{F1-Score} = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}}. \quad (10)$$

F. Experiment Set-Up

In the research, the dataset was split into 80% for training and 20% for testing. Figure 9 presents

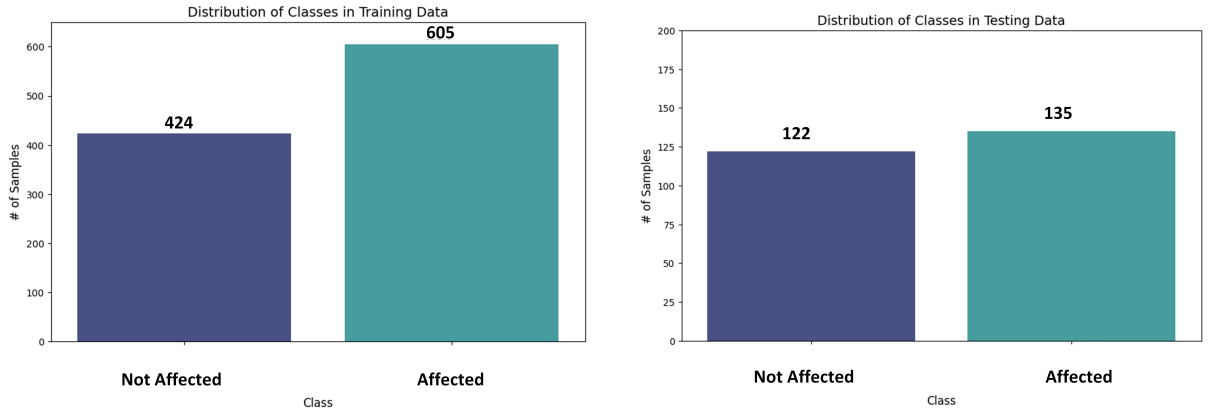

Fig. 9. Data distribution for each class (original data).

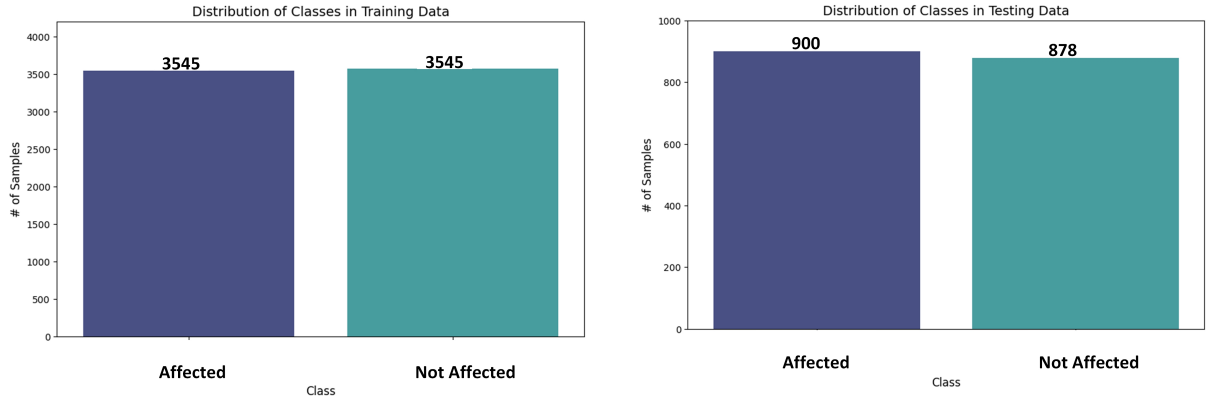

Fig. 10. Data distribution for each class (augmented data).

the data distribution for each class in the original dataset, showing an imbalance between the two classes. Data augmentation techniques are applied to balance the number of samples in each class, as shown in Fig. 10 to address the issue. This augmentation aims to improve the performance and generalization ability of the model. Furthermore, the research evaluates how Gaussian filtering influences the feature extraction process performed by the ResNet152V2 model, using both the original and augmented datasets for comparison.

Table I shows the same parameters used for both original and augmented data. These include the input image size (224×224 pixels), batch size (32), rescale applied to the data generator (1/255), zoom range (0.2), number of epochs (50), learning rate (0.001), loss function (binary_cross_entropy), and (Adam) optimizer. This same method for each data allows for a fair comparison of model performance on both types of data. There is also an accurate assessment of the effectiveness of data augmentation on model performance.

Table II shows an overview of the model architecture used, consisting of the form of output from each layer, as well as trainable and untrainable parameters. The model architecture uses ResNet152V2 with the final layer consisting of a global average pooling 2D and two dense layers. The output dimensions of the model are (None, 7, 7, 2048). It indicates that the output of ResNet152V2 layer is a 2D tensor with a size of 7×7 and a depth of 2048. All the parameters used in the model are 60,430,849, with 2,099,201 trainable and 58,331,648 untrainable parameters. The number of untrainable parameters mainly originates from pre-trained ResNet152V2 model, while trainable is mainly found in the dense layer. The weights of the dense layers are parameters that can be trained in the model and pre-trained ResNet152V2 layers are untrainable.

The research uses the Python programming language and the libraries such as OpenCV, Sci-kit Learn, TensorFlow, and Keras. The experiment is performed using a PC with the following specifications: CPU processor core i7 gen 9th, DDR4 16 GB, and GPU NVIDIA GeForce GTX 1660 Ti. These tools and

TABLE I
PARAMETER CONFIGURATION FOR EACH DATASET.

Parameters	Original Data	Augmented Data
Input size	224×224 pixel	224×224 pixel
Batch Size	32	32
Data generator	Rescale = 1/255, zoom range = 0.2	Rescale = 1/255, zoom range = 0.2
Learning rate	0.001	0.001
Epoch	50	50
Loss function	Binary cross-entropy	Binary cross-entropy
Optimizer	Adam	Adam

Note: All parameters are kept constant for both original and augmented data.

TABLE II
THE MODEL SUMMARY OF PROPOSED TRANSFER LEARNING.

Layer (type)	Output Shape	Parameter
Input (input_layer)	(None, 224, 224, 3)	0
ResNet152V2 (Functional)	(None, 7, 7, 2048)	58,331,648
Global_average_pooling2d	(None, 2048)	0
Dense (Dense)	(None, 1024)	2,098,176
Dense_1 (Dense)	(None, 1)	1,025
Total parameters		60,430,849
Trainable parameters		2,099,201
Non-trainable parameters		58,331,648

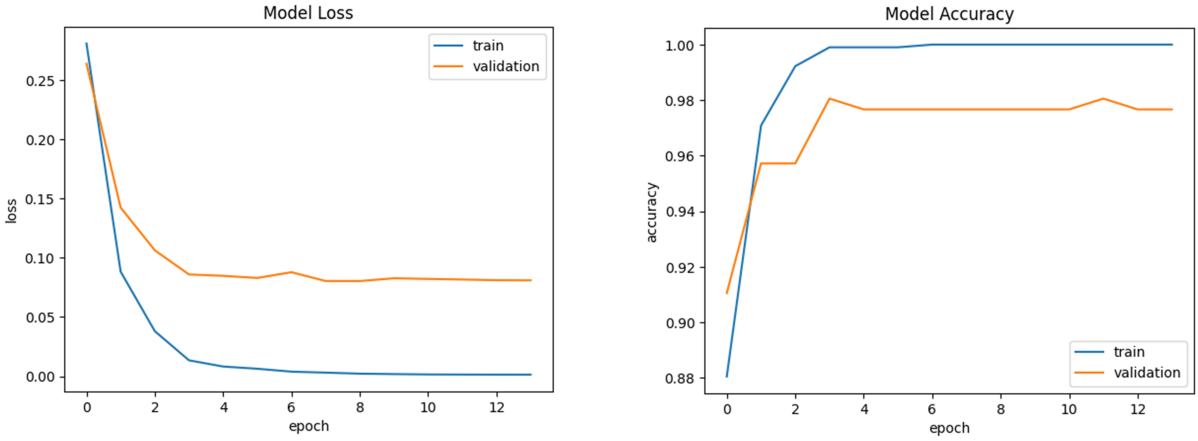

Fig. 11. Original dataset loss and accuracy function without Gaussian filtering.

libraries are selected due to their robustness and flexibility in handling image processing and deep learning tasks. OpenCV is utilized for pre-processing operations such as filtering and image enhancement, while TensorFlow and Keras are employed for designing and training deep learning models. Sci-kit Learn supports evaluation metrics and additional pre-processing steps. The hardware specifications ensure efficient training and testing processes, especially when handling large datasets and deep neural network architectures.

III. RESULTS AND DISCUSSION

A. Results with Original Data

Figure 11 shows plots of loss and accuracy functions without using Gaussian filtering method. During training, loss value is 0.0013, and accuracy value is 1.0000. Meanwhile, the loss value is 0.0811 with accuracy of 0.9767 during validation. It can be observed that loss value on training and validation set has decreased significantly from epoch to epoch. Accuracy during training and validation also increases as epoch progresses. It shows the model effectiveness in learning from training data and its ability to generalize to unseen data. Training process is halted at the 14^{th} epoch because learning rate decreases. The process is caused by using the ReduceLROnPlateau callback, which aims to reduce learning rate when loss reduction on validation set does not occur in several consecutive epochs.

Next, Fig. 12 shows plots of loss and accuracy func-

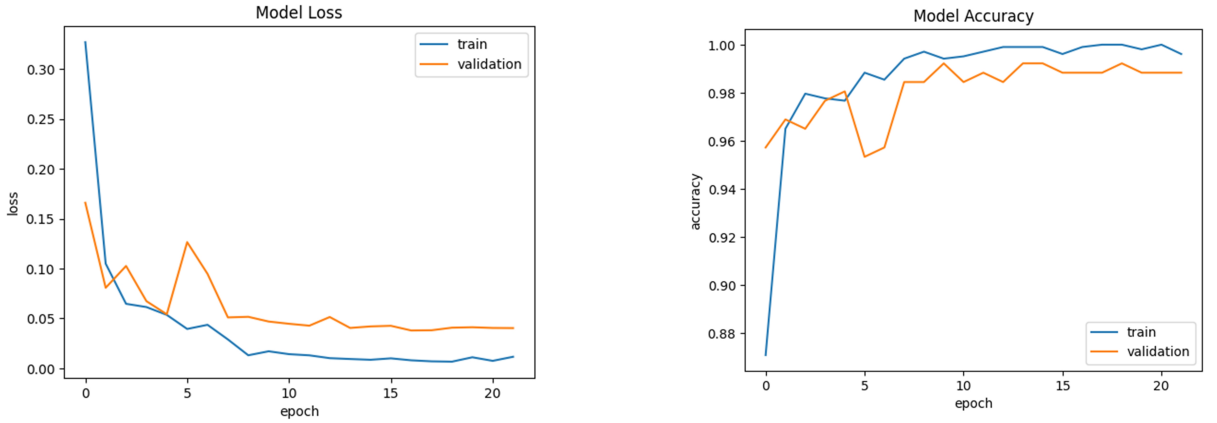

Fig. 12. Original dataset loss and accuracy function with Gaussian filtering.

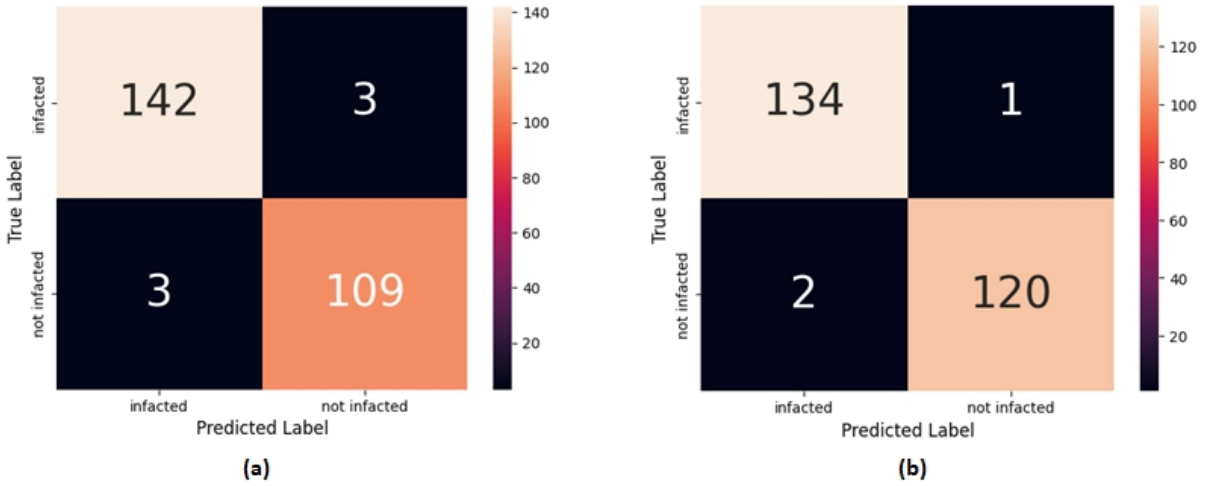

Fig. 13. Confusion Matrix (CM) of the original dataset: (a) unfiltered and (b) filtered.

tions using Gaussian filtering. During training, loss and accuracy values are 0.0116 and 0.996, respectively. In validation phase, loss value is 0.0404, while accuracy reaches 0.9883. It can be seen that loss on training set continues to decrease as epoch progresses, while accuracy shows a consistent increase. Loss value in validation also decreases as the epoch progresses but accuracy increases.

Figure 13 provides CM of original dataset ((a) unfiltered, and (b) filtered). From CM, the results of the model can be observed in two conditions, namely without and with Gaussian smoothing. In the “affected” category, the model without Gaussian smoothing has 142 correct and 3 incorrect predictions. Meanwhile, the model with Gaussian smoothing has 134 correct and 1 incorrect prediction. In the “not affected” category, the model without Gaussian smoothing has 109 correct and 3 incorrect predictions, while with Gaussian smoothing shows 120 correct and 2 incorrect predictions. These results show that the use of Gaussian smoothing tends to improve the model performance in classifying both categories, particularly in reducing the number of incorrect predictions for the “affected” category.

Figure 14 shows the prediction error with original dataset. The prediction error results are accompanied by a confidence value for each image, which indicates how confident the model is in the prediction. Figure 14(a) shows that the highest confidence values are found in the two image prediction results, namely 0.87 and 0.86. This result is followed by the next highest confidence value of 0.54. The remaining confidence values for the next three predictions are 0.21, 0.01, and 0.00. Figure 14(b) shows the highest confidence value of 0.56. Meanwhile, the rest are 0.41 and 0.02.

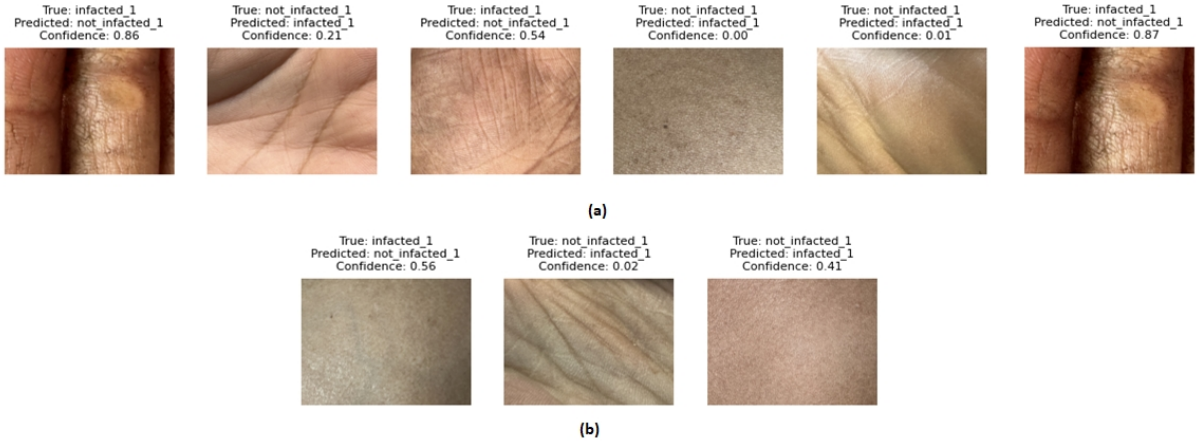

Fig. 14. Wrong prediction of the original dataset: (a) unfiltered and (b) filtered.

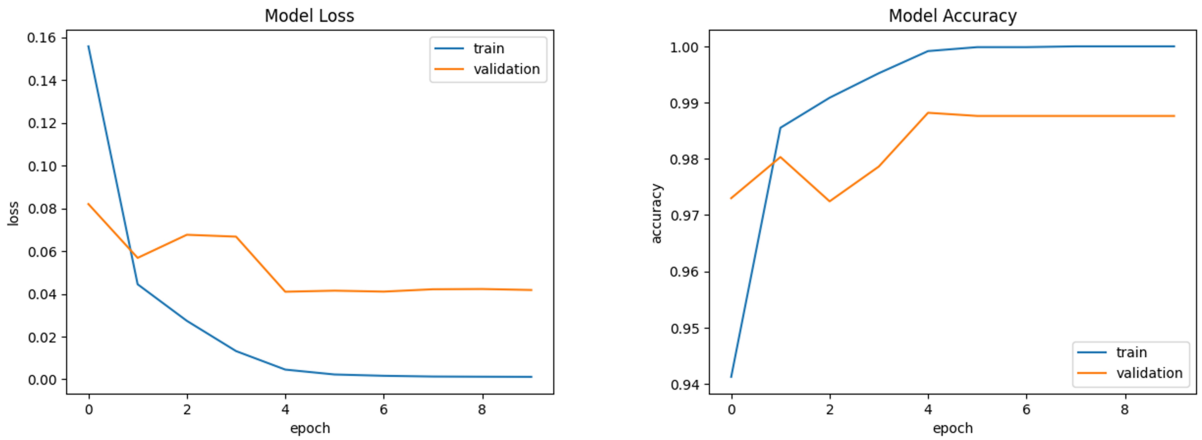

Fig. 15. Augmented dataset loss and accuracy function without Gaussian filtering.

B. Results with Augmented Data

Figure 15 shows plots of the loss and accuracy curves during the training process without applying the Gaussian filtering method. The training loss and accuracy reach 0.0012 and 1.0000, indicating that the model fits the training data very well. On the other hand, the validation loss and accuracy are 0.0410 and 0.9881, which are slightly lower than the training performance. This slight gap suggests that the model generalizes well to unseen data, with no significant signs of overfitting despite the absence of Gaussian filtering.

Figure 16 shows plots of the loss and accuracy curves when the Gaussian filtering method is applied before training. The training process results in a loss of 0.0062 and an accuracy of 0.9987, still indicating a strong fit to the training data. Meanwhile, the validation loss decreases to 0.0269, and the accuracy improves slightly to 0.9904. These results suggest that applying Gaussian filtering enhances the model’s ability to generalize by reducing noise in the input data and improving feature extraction during training.

Figure 17 provides CM of augmented dataset ((a) unfiltered, and (b) filtered). From the two CMs, the performance of the model is shown in two conditions, namely without and with Gaussian smoothing. In the “affected” category, the model without Gaussian smoothing has 891 correct and 9 incorrect predictions. Meanwhile, the model with Gaussian smoothing has 892 correct and 8 incorrect predictions. In the “not affected” category, the model without Gaussian smoothing has 866 correct and 12 incorrect predictions. The model with Gaussian smoothing has 869 correct and 9 incorrect predictions. From these results, the use of Gaussian smoothing is considered to slightly improve the model performance in classifying both categories, particularly in reducing the number of in-

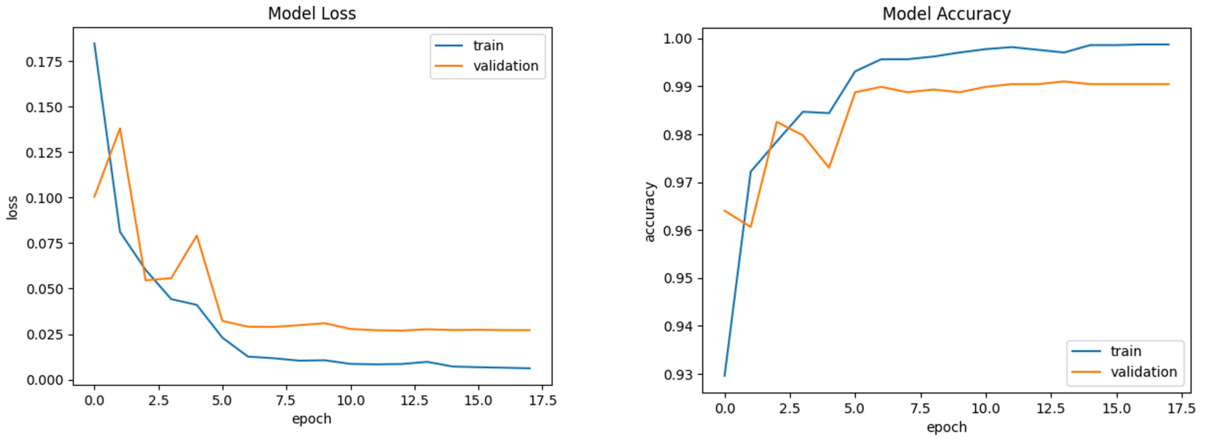

Fig. 16. Augmented dataset loss and accuracy function with Gaussian filtering.

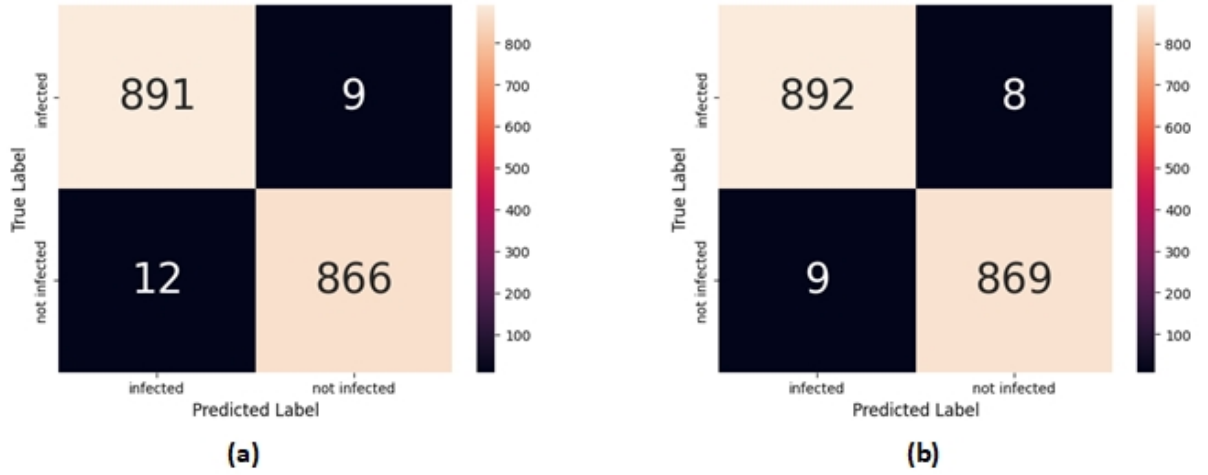

Fig. 17. Confusion Matrix (CM) of the augmented dataset: (a) unfiltered and (b) filtered.

correct predictions for the “affected” category.

In Fig. 18(a), the highest confidence value is 0.99 and the lowest is 0.01. Meanwhile, in Fig. 18(b), the highest confidence value is 0.96, and the lowest is 0.01. It suggests that a higher confidence score of a prediction correlates with a greater model certainty in prediction. This difference also indicates that the model produces more confident predictions on certain samples compared to others, possibly due to clearer distinguishing features. Additionally, the consistently low minimum values highlight that the model can differentiate between classes with high certainty when it is confident, but may still show uncertainty in more ambiguous cases.

C. Discussion

There are several main contributions of the research. First, it classifies good accuracy in the effects of arsenic exposure on the skin using an image processing-based framework and ResNet152V2 architecture. Second, the research aims to investigate the use of Gaussian smoothing method both on original and improved data. Tables III shows the performance comparison results of each experiment.

The results from Table III show that ResNet152V model effectively classifies images in ArsenicSkinBD dataset, achieving high accuracy and low loss values across experiments. The proposed framework not only enhances detection accuracy but also provides a robust method for generalizing new data, which is crucial for practical applications in clinical settings. Furthermore, the model successfully learns patterns from the data. Since accuracy of training set reaches 1.0000, the model is able to correctly classify all training samples. Accuracy on the test set is also high, with a value above 0.99 (ResNet+Gaussian). This result indicates

Fig. 18. Wrong prediction of augmented dataset: (a) unfiltered and (b) filtered.

that the model has a strong capability to generalize new data. Precision, recall, and F1-scores have high values, close to 0.99 for approximately all experiments. The results show that the model has a good balance between

TABLE III
PERFORMANCE COMPARISON RESULTS.

ArsenicSkinBD Dataset	Amount of Images	Experiments	Train Loss	Train Accuracy	Test Loss	Test Accuracy	Precision	Recall	F1-Score	Time
Original data	1287	ResNet152V	0.0013	1.0000	0.0811	0.9767	0.98	0.98	0.98	4 m 41 s
		ResNet152V + Gaussian	0.0116	0.9961	0.0404	0.9881	0.99	0.99	0.99	12 m 1 s
Augmented data	8892	ResNet152V	0.0012	1.0000	0.0410	0.9881	0.99	0.99	0.99	26 m 6 s
		ResNet152V + Gaussian	0.0062	0.9987	0.0269	0.9904	0.99	0.99	0.99	56 m 8 s

TABLE IV
COMPARISON OF SKIN DISEASE CLASSIFICATION.

Author	Year	Dataset	Methodology	Accuracy
[41]	2022	The International Skin Imaging Collaboration (ISIC) 2019, Human Against Machine with 10000 training images (HAM10000)	Digital hair removal, Gaussian filtering, Gray-Level Co-occurrence Matrix (GLCM) features, Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbors (KNN)	SVM: 95%
[42]	2023	Human Against Machine with 10000 training images (HAM10000)	16 Convolutional Neural Network (CNN) models for classification	Up to 99%
[43]	2023	Human Against Machine with 10000 training images (HAM10000)	Model fusion with attention module, data augmentation	95.29%
[44]	2024	Human Against Machine with 10000 training images (HAM10000)	Hybrid Feature-Based (HFB)-Convolutional Neural Network (CNN)-Bidirectional Long Short-Term Memory (BiLSTM) approach	96.3%
[45]	2024	Human Against Machine with 10000 training images (HAM10000), Programa de Assistência Dermatológica e Cirurgica - (PAD) at the Universidade Federal do Espírito Santo (UFES-Brazil)/Pigmented Skin Lesions Dataset from the Federal University of Espírito Santo – 2020 (PAD-UFES-20)	Hybrid Convolutional Neural Network (CNN)-DenseNet model	95.7%
[46]	2024	Human Against Machine with 10000 training images (HAM10000)	Self-Transfer Generative Adversarial Network (STGAN) for data synthesis and classification	98.23%
Proposed work	2025	ArsenicSkinImageBD	ResNet152V2 with Gaussian smoothing and transfer learning	99.04%

positive prediction accuracy. In line with the analysis, the model also shows the ability to find all positive samples and enhanced precision in classifying.

Using Gaussian method on ResNet152V model improves accuracy of the test set from approximately 0.9881 to 0.9904. It means that data augmentation with Gaussian method helps the model to better classify complex samples. The use of data augmentation with Gaussian method also causes a decrease in loss values on the set of tests. The model becomes increasingly efficient in optimizing its predictions. Although there is an increase in accuracy, recall, precision, and F1-scores either remain the same or slightly increase. The results show that the implementation of Gaussian smoothing method not only raises accuracy from approximately 0.9881 to 0.9904 but also improves and enhances the model reliability in detecting complex samples, contributing to more effective early diagnosis and treatment.

By using data augmentation with Gaussian method, there is an increase in training time from 26 to 56 minutes. It is reasonable considering the additional processing required to apply Gaussian augmentation to each batch of data. Overall, the method has a positive impact on model performance by increasing accuracy and consistency of prediction results. Although training time increases slightly, the improved prediction quality enhances trade-off, particularly in cases where high accuracy is preferred.

Table IV shows the comparison of previous studies for skin disease classification. Based on Table IV, several studies have explored various methods for skin disease classification, each with unique strategies and datasets. A previous research [41] uses traditional machine learning methods such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree on ISIC 2019 and HAM10000 datasets, achieving a maximum accuracy of 97% with SVM. Another previous research [42] utilizes 16 CNN models on HAM10000 dataset, achieving approximately 99% accuracy, and showing the potential of deep learning in large-scale image analysis. A model fusion method is also introduced with attention modules and data augmentation, achieving 95.29% accuracy on HAM10000 dataset [43]. The previous research [44] proposes a Hybrid Feature-Based (HFB)-CNN-Bidirectional Long

Short-Term Memory (BiLSTM) model and achieves 96.3% accuracy on HAM10000 dataset by addressing imbalanced data issues through augmentation. An integrated hybrid CNN-DenseNet model is also observed, showing accuracy of 95.7% on both HAM10000 and Programa de Assistência Dermatológica e Cirurgica - (PAD) at the Universidade Federal do Espírito Santo (UFES-Brazil)/Pigmented Skin Lesions Dataset from the Federal University of Espírito Santo – 2020 (PAD-UFES-20) datasets, emphasizing the potential for effective diagnosis across various skin diseases [45]. Furthermore, Self-Transfer Generative Adversarial Network (STGAN) is introduced for data synthesis and classification, achieving significant results with accuracy of 98.23% [46].

In comparison, the research leverages transfer learning using ResNet152V2 architecture combined with Gaussian smoothing method. The results show accuracy of approximately 99.04% on ArsenicSkinImageBD dataset. The method is competitive with state-of-the-art methods while addressing a unique problem, namely classifying arsenic-affected skin diseases. Compared to most previous reports on generic skin disease datasets like HAM10000, the research targets a specific dataset designed to analyze arsenic-related skin conditions. The innovative application of Gaussian smoothing for noise reduction and feature clarity further distinguishes from previous studies. This comparison shows the effectiveness and novelty of the framework in improving classification performance for a specialized application area.

IV. Conclusion

In conclusion, the research significantly contributes to the improved classification of arsenic-affected skin diseases using image processing methods and ResNet152V2 architecture. Specifically, the impact of Gaussian smoothing method is investigated on both original and augmented data to enhance the model classification performance. The results show that applying Gaussian smoothing improves the testing set accuracy from approximately 0.9881 to 0.9904. This improvement shows the effectiveness of Gaussian smoothing in enabling the model to better classify complex skin samples affected by arsenic exposure. Moreover, the use of Gaussian smoothing also reduces loss values on the testing set, indicating that the model becomes more efficient in optimizing its predictions. The high precision, recall, and F1-score (close to 0.99) across experiments further confirm the model consistency and reliability in performing accurate classifications. These results show the potential of the framework to support early detection and diagnosis of arsenic-induced skin diseases.

Despite the significant contribution, there are limitations in the research that should be acknowledged. The application of Gaussian smoothing significantly increases training time from 26 minutes to 56 minutes, making it less ideal for real-time applications or resource-constrained environments. Moreover, the research focuses on a specific augmentation method, leaving the exploration of other models for future analysis. Expanding the dataset to include more diverse skin conditions or environmental factors can also improve the model generalizability. Although training time increases, the improved classification accuracy and consistency achieved through Gaussian smoothing make it a valuable method for enhancing detection capabilities. Additionally, the framework shows significant promise for early detection and management of arsenic-affected skin diseases, potentially leading to better health outcomes and reduced disease burden.

Acknowledgement

The authors are grateful to Universitas Bina Insan for the support provided during the research. Moreover, the research was conducted independently and was not funded by any study scheme. The encouragement and resources provided by the university have been instrumental in the successful completion.

Data Availability

The data used in the research is obtained from ArsenicSkinBD public dataset which can be accessed on https://data.mendeley.com/datasets/x4hgnjj5gv/2.

Author Contribution

Conceived and designed the analysis, R. K.; Collected the data, R. K.; Contributed data or analysis tools, R. K.; Performed the analysis, S. and F. S. M.; Wrote the paper, R. K.; and Contributed others, S. and F. S. M.

References

[1] Y. Chen, F. Parvez, M. Gamble, T. Islam, A. Ahmed, M. Argos, J. H. Graziano, and H. Ahsan, "Arsenic exposure at low-to-moderate levels and skin lesions, arsenic metabolism, neurological functions, and biomarkers for respiratory and cardiovascular diseases: Review of recent findings from the Health Effects of Arsenic Longitudinal Study (HEALS) in Bangladesh," *Toxicology and Applied Pharmacology*, vol. 239, no. 2, pp. 184–192, 2009.

[2] W. K. Chowdhury, A. Tisha, S. Akter, S. Zahur, N. Hasan, A. S. Mahadi, S. M. F. Rabby, M. M.

Mohib, M. A. T. Sagor, and S. Mohiuddin, "The role of arsenic on skin diseases, hair fall and inflammation: An immunological review and case studies," *Journal of Clinical & Experimental Dermatology Research*, vol. 8, no. 2, pp. 1–9, 2017.

[3] Y. Zhang, B. Xu, Z. Guo, J. Han, H. Li, L. Jin, F. Chen, and Y. Xiong, "Human health risk assessment of groundwater arsenic contamination in Jinghui irrigation district, China," *Journal of Environmental Management*, vol. 237, pp. 163–169, 2019.

[4] S. Prakash and A. K. Verma, "Arsenic: It's toxicity and impact on human health," *International Journal of Biological Innovations (IJBI)*, vol. 3, no. 1, pp. 38–47, 2021.

[5] R. Ashraf, S. Afzal, A. U. Rehman, S. Gul, J. Baber, M. Bakhtyar, I. Mehmood, O. Y. Song, and M. Maqsood, "Region-of-interest based transfer learning assisted framework for skin cancer detection," *IEEE Access*, vol. 8, pp. 147 858–147 871, 2020.

[6] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: A literature review," *BMC Medical Imaging*, vol. 22, no. 1, pp. 1–13, 2022.

[7] L. Alzubaidi, M. Al-Amidie, A. Al-Asadi, A. J. Humaidi, O. Al-Shamma, M. A. Fadhel, J. Zhang, J. Santamaría, and Y. Duan, "Novel transfer learning approach for medical imaging with limited labeled data," *Cancers*, vol. 13, no. 7, pp. 1–22, 2021.

[8] S. A. AlDera and M. T. B. Othman, "A model for classification and diagnosis of skin disease using machine learning and image processing techniques," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, pp. 252–259, 2022.

[9] Y. Gu, Z. Ge, C. P. Bonnington, and J. Zhou, "Progressive transfer learning and adversarial domain adaptation for cross-domain skin disease classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1379–1393, 2019.

[10] I. Kousis, I. Perikos, I. Hatzilygeroudis, and M. Virvou, "Deep learning methods for accurate skin cancer recognition and mobile application," *Electronics*, vol. 11, no. 9, pp. 1–19, 2022.

[11] Q. Zeng and A. Zhang, "Assessing potential mechanisms of arsenic-induced skin lesions and cancers: Human and in vitro evidence," *Environmental Pollution*, vol. 260, 2020.

[12] M. K. Hasan, M. T. E. Elahi, M. A. Alam, M. T. Jawad, and R. Martí, "DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation," *Informatics in Medicine Unlocked*, vol. 28, pp. 1–17, 2022.

[13] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-Net," *PLOS ONE*, vol. 14, no. 5, pp. 1–17, 2019.

[14] K. M. Hosny, M. A. Kassem, and M. M. Fouad, "Classification of skin lesions into seven classes using transfer learning with AlexNet," *Journal of Digital Imaging*, vol. 33, pp. 1325–1334, 2020.

[15] A. Khamparia, P. K. Singh, P. Rani, D. Samanta, A. Khanna, and B. Bhushan, "An Internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 7, 2021.

[16] A. Mahbod, G. Schaefer, C. Wang, G. Dorffner, R. Ecker, and I. Ellinger, "Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification," *Computer Methods and Programs in Biomedicine*, vol. 193, pp. 1–9, 2020.

[17] M. A. Kassem, K. M. Hosny, and M. M. Fouad, "Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning," *IEEE Access*, vol. 8, pp. 114 822–114 832, 2020.

[18] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. K. Islam, "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models," *Machine Learning with Applications*, vol. 5, pp. 1–8, 2021.

[19] A. Jain, A. C. S. Rao, P. K. Jain, and A. Abraham, "Multi-type skin diseases classification using OP-DNN based feature extraction approach," *Multimedia Tools and Applications*, vol. 81, pp. 6451–6476, 2022.

[20] H. M. Balaha and A. E. S. Hassan, "Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm," *Neural Computing and Applications*, vol. 35, no. 1, pp. 815–853, 2023.

[21] G. M. S. Himel, M. M. Islam, K. A. Al-Aff, S. I. Karim, and M. K. U. Sikder, "Skin cancer segmentation and classification using vision transformer for automatic analysis in dermatoscopy-based noninvasive digital system," *International Journal of Biomedical Imaging*, vol. 2024, no. 1,

pp. 1–18, 2024.

[22] P. Thapar, M. Rakhra, M. Alsaadi, A. Quraishi, A. Deka, and J. V. N. Ramesh, "A hybrid grasshopper optimization algorithm for skin lesion segmentation and melanoma classification using deep learning," *Healthcare Analytics*, vol. 5, pp. 1–17, 2024.

[23] J. S. T. Purni and R. Vedhapriyavadhana, "EOSA-Net: A deep learning framework for enhanced multi-class skin cancer classification using optimized convolutional neural networks," *Journal of King Saud University-Computer and Information Sciences*, vol. 36, no. 3, pp. 1–25, 2024.

[24] N. V. Y. Lankadasu, D. B. Pesarlanka, A. Sharma, S. Sharma, and S. Gochhait, "Skin cancer classification using a convolutional neural network: An exploration into deep learning," in *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*. Manama, Bahrain: IEEE, Jan. 28–29, 2024, pp. 1047–1052.

[25] I. A. Emu, N. T. Niloy, B. M. A. Karim, A. Chowdhury, F. T. Johora, M. Hasan, T. Mittra, M. R. A. Rashid, T. Jabid, M. Islam, and M. S. Ali, "ArsenicSkinImageBD: A comprehensive image dataset to classify affected and healthy skin of arsenic-affected people," *Data in Brief*, vol. 52, pp. 1–11, 2024.

[26] N. T. Niloy, I. A. Emu, A. Chowdhury, B. M. A. Karim, F. T. Johora, M. R. A. Hasan, Mahamudul Rashid, T. Mittra, T. Jabid, M. Islam, and S. Ali, "ArsenicSkinImageBD," 2023. [Online]. Available: https://data.mendeley.com/datasets/x4hgnjj5gv/2

[27] S. Saponara and A. Elhanashi, "Impact of image resizing on deep learning detectors for training time and model performance," in *International Conference on Applications in Electronics Pervading Industry, Environment and Society*. Springer, 2021, pp. 10–17.

[28] M. Xiao, S. Zheng, C. Liu, Y. Wang, D. He, G. Ke, J. Bian, Z. Lin, and T. Y. Liu, "Invertible image rescaling," in *Computer Vision–ECCV 2020: 16th European Conference*. Glasgow, UK: Springer, Aug. 23–28, 2020, pp. 126–144.

[29] B. Desai, M. Paliwal, and K. K. Nagwanshi, "Study on image filtering—Techniques, algorithm and applications," 2022. [Online]. Available: https://arxiv.org/abs/2207.06481

[30] B. H. Chen, Y. S. Tseng, and J. L. Yin, "Gaussian-adaptive bilateral filter," *IEEE Signal Processing Letters*, vol. 27, pp. 1670–1674, 2020.

[31] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.

[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[33] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[34] S. Athisayamani, R. S. Antonyswamy, V. Sarveshwaran, M. Almeshari, Y. Alzamil, and V. Ravi, "Feature extraction using a residual deep convolutional neural network (ResNet-152) and optimized feature dimension reduction for MRI brain tumor classification," *Diagnostics*, vol. 13, no. 4, pp. 1–20, 2023.

[35] H. Yu, X. Miao, and H. Wang, "Bearing fault reconstruction diagnosis method based on ResNet-152 with multi-scale stacked receptive field," *Sensors*, vol. 22, no. 5, pp. 1–14, 2022.

[36] V. Thakkar, S. Tewary, and C. Chakraborty, "Batch normalization in convolutional neural networks—A comparative study with CIFAR-10 data," in *2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT)*. Kolkata, India: IEEE, Jan. 12–13 2018, pp. 1–5.

[37] X. Xu, W. Li, and Q. Duan, "Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification," *Computers and Electronics in Agriculture*, vol. 180, 2021.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, June 26–July 1, 2016, pp. 770–778.

[39] Y. Cui, M. Jia, T. Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, June 16–20, 2019, pp. 9268–9277.

[40] L. S. Bernardo, R. Damaševičius, V. H. C. De Albuquerque, and R. Maskeliūnas, "A hybrid two-stage SqueezeNet and support vector machine system for Parkinson's disease detection based on handwritten spiral patterns," *International Journal of Applied Mathematics and Computer Science*, vol. 31, no. 4, pp. 549–561, 2021.

[41] M. Ahammed, M. Al Mamun, and M. S. Uddin, "A machine learning approach for skin disease

detection and classification using image segmentation," *Healthcare Analytics*, vol. 2, pp. 1–15, 2022.

[42] M. Shahin, F. F. Chen, A. Hosseinzadeh, H. K. Koodiani, A. Shahin, and O. A. Nafi, "A smartphone-based application for an early skin disease prognosis: Towards a lean healthcare system via computer-based vision," *Advanced Engineering Informatics*, vol. 57, 2023.

[43] M. Wei, Q. Wu, H. Ji, J. Wang, T. Lyu, J. Liu, and L. Zhao, "A skin disease classification model based on DenseNet and ConvNeXt fusion," *Electronics*, vol. 12, no. 2, pp. 1–19, 2023.

[44] A. M. Vidhyalakshmi and M. Kanchana, "Classification of skin disease using a novel hybrid flash butterfly optimization from dermoscopic images," *Neural Computing and Applications*, vol. 36, no. 8, pp. 4311–4324, 2024.

[45] A. De, N. Mishra, and H. T. Chang, "An approach to the dermatological classification of histopathological skin images using a hybridized CNN-DenseNet model," *PeerJ Computer Science*, vol. 10, pp. 1–33, 2024.

[46] Q. Su, H. N. A. Hamed, M. A. Isa, X. Hao, and X. Dai, "A GAN-based data augmentation method for imbalanced multi-class skin lesion classification," *IEEE Access*, vol. 12, pp. 16 498–16 513, 2024.