EfficientSkinCaSV2B3: An Efficient Framework Towards Improving Skin Classification and Segmentation

Quy Lu Thanh, Triet Minh Nguyen
FPT University, Can Tho, Viet Nam

***Abstract*—Ozone layer depletion has gained attention as a serious environmental issue. Because of its effects on human health especially skin cancer. Besides, Ultraviolet (UV) radiation is known to be a major risk factor for skin cancer. For instance, it can damage the DNA in skin cells leading to mutations that may eventually result in cancerous growth. Basal cell carcinoma, squamous cell carcinoma, and melanoma are the three primary forms of skin cancer linked to UV exposure. Additionally, it triggers associated illnesses including nevus, seborrheic keratosis, actinic keratosis, dermatofibroma, and vascular lesions. Many medical and computer studies were published as a result to address these disorders. Especially, using an aspect of deep learning that is transfer learning and fine-tuning for the classification of skin images. In this research, the EffecientSkinCaSV2B3 framework was proposed and applied to classify and segment the skin cancer dataset, which were collected and validated by The International Skin Imaging Collaboration (ISIC). In addition, Gradient-weighted Class Activation Mapping (Grad-CAM) is used in skin cancer classification to visually explain images, aiding in understanding model decisions and highlighting important areas. Based on color and texture, k-means clustering was used for the segmentation between portions that were healthy and those that were unhealthy. The study reached a surprising accuracy of 84.91% in nine classes of classifying skin cancer. In other experiments, the customized EfficientNetV2B3 model achieved 94.00% in classifying malign and benign. Moreover, scenarios pointed out that in classifying six classes (i.e., between benign skin diseases) and three classes (i.e., between malign skin diseases) the model earned a high accuracy of 89.56% and 96.74%, respectively.**

***Keywords*—Skin cancer; Convolutional Neural Network (CNN); transfer learning; fine tuning; classification; segmentation; EffecientNetB3V2**

I. INTRODUCTION

The progressive thinning of the ozone layer in the upper atmosphere as a result of chemicals released by businesses or other human activities is known as ozone layer depletion. Nowadays, the depletion of the ozone layer is a serious issue that releases various problems such as climate change, melting ice, and health issues. In particular, ozone layer depletion creates an increase in UV radiation on the surface of the earth. Moreover, UV radiation exposure has been the primary reason responsible for the development of skin cancer in recent decades [1]. The effects of skin cancer on health extend beyond the physical, often causing emotional damage. Patients may experience heightened anxiety, depression, and a diminished quality of life as they navigate the complexities of diagnosis, treatment, and potential recurrence. Moreover, the visible nature of skin cancer lesions can contribute to feelings of self-consciousness and social isolation, exacerbating the emotional damage of the disease.

During the research, 649,2 new melanoma skin cancer cases occurred in men, women, and both sexes per 100,000 persons in 2020 (i.e., the ratios for men, women, and both sexes are 173.8, 150.8, and 324.6, respectively) [2]. Moreover, according to the number of new cases and deaths from skin cancer in the USA (excluding dependent countries) and China (excluding the province of Taiwan) in 2022. In total, it is anticipated that in China and the USA, there will be roughly 8114 and 99.935 people newly diagnosed with melanoma skin cancer, and 4369 and 7530 people dying from melanoma skin cancer, respectively [3]. According to statistics on skin cancer at the National Hospital of Dermatology and Venereology from 2017 to 2021. Basal cell carcinoma was the most common type of skin cancer, followed by squamous cell carcinoma and melanoma. In addition, the majority of patients were over 60 years old, and there was an increase in the proportion of patients under 60 years old over the years [4].

Fortunately, advancements in medical science have led to a variety of treatment options for skin cancer, offering hope to those affected by this insidious disease. The choice of treatment depends on factors such as the type and stage of cancer, as well as the overall health and preferences of patients. Surgical interventions, such as excisional surgery and Mohs micrographic surgery, remain primary options for removing cancerous lesions while preserving as much healthy tissue as possible. In cases where surgery may not be feasible, other modalities such as radiation therapy, chemotherapy, immunotherapy, and targeted therapy may be used to combat the disease at its source. Additionally, early detection plays an important role in improving treatment outcomes and reducing the risk of complications. Regular skin examinations by dermatologists and self-checks at home can help identify suspicious growths or changes in existing moles, prompting timely medical intervention. However, with advances in medical technology, computer technology, and increased awareness of preventive measures, individuals can employ technology to minimize their risk of developing this disease and seek prompt treatment. Thus, applying artificial intelligence (I.e., AI) has become popular in recent years in classifying and detecting illnesses [5][6][7].

A subset of machine learning in AI is deep learning, it has revolutionized the field of image analysis [8][9][10][11][12]. Deep learning models mimic the ability to process and recognize patterns of the human brain such as CNN. These models consist of multiple layers of interconnected neurons, each layer learning increasingly abstract features from the input data. Deep learning algorithms examine pictures of skin lesions and extract minute details that might not be visible to the human eye to classify skin cancer. Through the process of training on large datasets of labeled skin images, deep learning models become adept at distinguishing between benign and malignant lesions with high accuracy, providing valuable support to dermatologists in clinical decision-making. One of the leading methodologies used in skin cancer classification is transfer learning and fine-tuning. Transfer learning is the process of applying pre-trained neural network models on a large dataset for a different job to a particular classification problem [13][14][15], such as identifying malignant or benign skin lesions. Contrarily, fine-tuning is the process of retraining the previously trained model on a smaller dataset pertinent to the intended job [16][17][18], allowing its parameters to be optimized for optimal performance.

In general, the utilization of AI techniques in the diagnosis and treatment of medical become popular around the world. Especially, in image classification and segmentation by transfer learning combined with fine-tuning which created several successful promotions on both sides of computer and medical science. EfficientSkinCaSV2B3 framework provided computer vision technology for the classification and segmentation of skin cancer illnesses by employing transfer learning and fine-tuning in a customized CNN model. In addition, Grad-CAM was applied for visual explanation that helped create an overall vision for the final analysis. Furthermore, k-means clustering is a suitable technology used for image segmentation which provides extremely good results.

The contributions of this paper are as follows:

- In a classification of nine classes of skin cancer (i.e., includes actinic keratosis, basal cell carcinoma, dermatofibroma, melanoma, nevus, pigmented benign, keratosis, seborrheic keratosis, squamous cell carcinoma, and vascular lesion), our study demonstrated a custom CNN model based on EffecientNetV2B3 with successfully effective in multiple classes. Thus, it offers a time-saving and easy way for the dermatologist and patient when diagnose abnormal positions on the skin early.
- In the scenario of nine classes classification, our model reached outstanding validation accuracy, test accuracy, and F1 score (i.e., 85.13%, 84.91%, and 84.68%). Consequently, tables and confusion matrices were also created to show the effectiveness of the training and testing duration of the model.
- Grad-CAM is provided as a tool in skin cancer classification by elucidating pertinent features utilized by models for decision-making. It enables doctors and researchers to understand model predictions with increasing diagnostic confidence. By highlighting regions. Grad-CAM aids in the interpretation of model outputs ultimately facilitating accurate classification of skin lesions for improved patient care.
- In this article, K-means clustering was proposed in skin cancer segmentation which supports categorizing lesions based on features like color, texture, and size. This method assists in identifying distinct regions within an image and helps precise delineation of cancerous areas for diagnostic purposes, treatment planning, and monitoring disease progression.
- This research gathered a dataset consisting of 2357 images of malignant and benign oncological diseases, which were formed by the ISIC. This dataset is verified for the development of automated machine learning and deep learning algorithms for the classification and segmentation of skin diseases. In addition, it can also be used to instruct students studying medical.

The structure of the research paper is created by six principal sections. Firstly, Section I presents an overview providing a general introduction to the article. After that, Section II provides a comprehensive analysis of the body of literature that serves as the foundation for our study and identifies relevant studies. Subsequently, Section III illustrates the methodology employed which provides detailed insights into the methods used throughout the article. Section IV indicates experiments, including the procedures for their execution and the evaluation of each scenario. Moreover, Section V presents the results of the most important experiment and conducts a comparative analysis with existing scenarios. Finally, the article summarizes the key and analyzes the overview of our research in Section VI.

II. Related Work

Recent advancements in classification and segmentation research have witnessed a surge in deep learning approaches, particularly in the area of computer vision. CNN continues to control these fields due to their remarkable ability to extract features hierarchically from data. In addition. Techniques such as transfer learning and fine-tuned created for specific tasks have gained traction enabling effective classification and segmentation even with limited data. Moreover, Researchers are increasingly focusing on developing more robust architectures capable of handling diverse datasets with improved accuracy and efficiency. Ahmed Abdelhafeez et al proposed a customized CNN model to classify eight classes of skin cancer and reached a surprising accuracy of 85.74% when compared with GoogleNet and DarkNet[19]. Additionally, Pooja Nadiger et al developed a CNN for skin cancer detection and achieved an accuracy of 90% in classifying skin lesions as benign or malignant [20].

Skin cancer is one of those deadly diseases where survival depends on early identification. In recent years, a lot of studies about deep learning models have been published. Mijwil et al selected and trained 24,000 skin cancer images between two classes by CNN model applying three architectures (i.e., InceptionV3, ResNet, and VGG19). Consequently, the best architecture InceptionV3 achieved a diagnostic accuracy of 86.90% [21]. Furthermore, Karar Ali et al trained and evaluated seven classes on EfficientNets B0 to B7 and achieved the best result in EfficientNet B4 with an accuracy of 87.91% [22].

Moreover, Solene Bechelli et al used fine-tuning in the VGG16 model to perform extremely well for skin tumor classification of 88% in two classes of classification [23].

Various techniques have been proposed to improve the accuracy of classification. In a comparative analysis, Krishna Mridha et al optimized CNN to identify the seven forms of skin cancer and reached a high accuracy of 82% [24]. Moreover, Duggani Keerthana et al proposed a DenseNet-201 and MobileNet model for skin cancer classification using the dataset of benign and malignant. The top-performing networks achieved accuracies of 88.02% [25]. In addition, Satin Jain et al pointed out that the XceptionNet model outperforms the rest of the transfer learning nets used for the study, with an accuracy of 90.48% for the classification of seven classes [26]. Besides, Ayesha Atta et al employed a customized CNN model with 3600 images of malignant and benign for classifying and gained an accuracy of 86.23% [27].

Advances in science and technology have promoted developments in the classification and segmentation of skin diseases. According to Vatsala Anand et al, one flattening layer, two dense layers with activation functions (LeakyReLU), and another dense layer with activation function (sigmoid) are added to a pre-trained VGG16 model to increase its performance. This model achieves an overall accuracy of 89.09% in identifying benign and malignant skin cancer [28]. Md Shahin Ali et al propose a deep convolutional neural network (DCNN) model based on a deep learning approach and compared it with transfer learning models such as AlexNet, ResNet, VGG-16, DenseNet, and MobileNet for the accurate classification between benign and malignant. Thus, the model obtained the highest 91.93% testing accuracy [29]. After several adjustments to the parameters and classification functions, Dipu Chandra Malo et al proposed VGG-16 model demonstrated a positive development and attained an accuracy of 87.6% [30].

The modern world is full of terrible diseases. Among them is skin cancer. Because skin cancer cells grow and spread like tumors in the human body. As a result, Mohammed Rakeibul Hasan et al compared several models in CNN and proposed the result that VGG16 provided the highest accuracy of 93.18% in classifying benign and malignant [31]. Additionally, Abdurrahim Yilmaz et al employed transfer learning and fine-tuning approaches and deep learning models in 3 different mobile deep learning models and 3 different batch sizes. Consequently, NASNetMobile gained the best outcome with an accuracy of 82% [32]. Furthermore, Chandran Kaushik Viknesh et al used convolutional neural networks, including AlexNet, LeNet, and VGG-16 models to gain a 91% accuracy rate after 100 compute epochs for classifying benign and malignant in ISIC datasets [33].

In conclusion, existing research on skin cancer classification demonstrates progress but faces challenges. Compared to human diagnosis, machine-learning models show lower accuracy, indicating the need for further refinement. Despite advancements, closing the gap between automated systems and human expertise remains a critical objective for enhancing diagnostic capabilities.

III. METHODOLOGY

A. The Research Implementation Procedure

12 steps of the pipeline this study suggested are depicted in Fig. 1. The following roles of the steps are displayed:

1) Collecting dataset: Curated meticulously by the International Skin Imaging Collaboration (ISIC), the dataset comprises 2357 high-resolution images encompassing a spectrum of skin cancer types, including Actinic keratosis, Basal cell carcinoma, Dermatofibroma, Melanoma, Nevus, Pigmented benign keratosis, Seborrheic keratosis, Squamous cell carcinoma, and Vascular lesion. Each image has undergone rigorous validation procedures to ensure accuracy and reliability. This compilation serves as an invaluable asset for scholarly investigations, providing comprehensive insights into the classification and management of skin cancer.

2) Pre-processing image and data augmentation: Image pre-processing techniques are crucial in refining input data for enhanced model performance. Key procedures like resizing and normalization are essential for standardizing images, and fostering consistency across datasets. Additionally, leveraging data augmentation methods such as rotation, flipping, and contrast enhancement diversifies the dataset, enriching the ability to generalize and learn from various skin lesion presentations of the model. These preprocessing steps contribute to improved accuracy and aid in the robustness and reliability of skin cancer classification models.

3) Dividing the dataset into three categories train, validation, and test: After being randomly chosen on an 8-1-1 scale, the datasets are organized into 8 training, 1 validation, and 1 testing folder. This ensures a balanced distribution, which is necessary for reliable model construction and assessment.

4) Dividing dataset for scenarios: The dataset was partitioned into four scenarios. In the initial scenario, nine classes including actinic keratosis, basal cell carcinoma, dermatofibroma, melanoma, nevus, pigmented benign keratosis, seborrheic keratosis, squamous cell carcinoma, and vascular lesion were chosen due to their distinguishability through surface observation. Following this, the second scenario comprised two classes: benign and malignant, focusing on internal characteristics. The third scenario encompassed six classes dedicated to the classification of benign conditions. Finally, the fourth scenario involved three classes specifically targeting malignant cases.

5) Building the model: The study employed transfer learning with the EfficientNetV2B3 model, a pre-trained convolutional neural network architecture for conducting experiments. During fine-tuning, external layers were utilized to adapt the pre-trained model to the specific data of the skin cancer classification task. The evaluation of training results indicates that the EfficientNetV2B3 model achieved excellent performance, particularly in skin cancer classification.

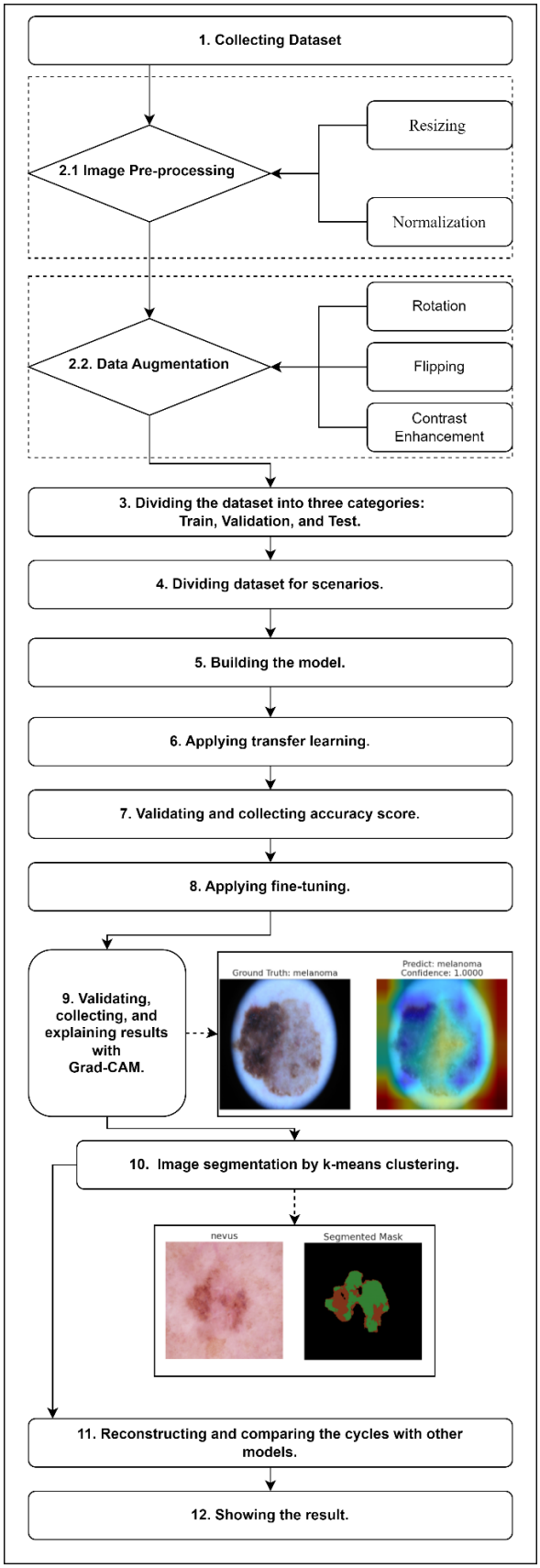

Fig. 1. The EfficientSkinCaSV2B3 framework.

6) Applying transfer learning: Transfer learning in skin cancer detection works by utilizing pre-trained models that have been trained on large datasets, often of general images. These models have already learned features that are useful for image recognition tasks. Instead of training a model from scratch, transfer learning involves taking these pre-trained models and adapting them to the specific task of skin cancer detection.

7) Validating and collecting accuracy score: Once the model finished training, its efficacy was evaluated based on its training accuracy and other performance metrics. Subsequently, the validity of the test was assessed using the initially separated testing set.

8) Applying fine-tuning: Fine-tuning includes taking a pre-trained model and adjusting its parameters to specialize in a specific task, such as skin cancer detection. This process optimizes the model's performance for the new task by adapting its learned features and weights. It improves accuracy without requiring extensive training on a new dataset.

9) Validating, collecting, and explaining results with Grad-CAM: Validating results with Grad-CAM highlights regions important for classification, and researchers gain insight into the decision-making process of the model. This method helps explain model predictions by visually indicating which parts of the image contribute most significantly. By validating, collecting, and explaining results with Grad-CAM, this research enhances transparency and confidence in the model's performance, aiding in the development of more accurate and interpretable skin cancer detection systems.

10) Image segmentation by k-means clustering: By iteratively assigning pixels to clusters with similar characteristics, k-means effectively separates skin lesions from healthy tissue. This method aids in identifying the boundaries of lesions, facilitating accurate diagnosis and treatment planning. By segmenting skin cancer images with k-means clustering, dermatologists can efficiently analyze lesion morphology and texture, improving the precision of diagnostic assessments and enhancing patient care.

11) Reconstructing and comparing the cycles with other models: To arrive at the final outcome, the process was revised and compared with another model, which included ResNet50V2, MobileNetV2, MobileNet, EffecientNetB3, and ResNet50.

12) Showing the result: Following established procedures, the data will be meticulously organized into tables and graphs, allowing for precise and pertinent comparisons to be made with ease, thereby enhancing the depth of analysis and understanding.

B. Pre-processing Image and Data Augmentation

In the region of classifying skin cancer using transfer learning and fine-tuning techniques, pre-processing and data augmentation play important roles in increasing the

effectiveness of the model. Pre-processing means preparing the raw data to make it suitable for training, while data augmentation aims to increase the diversity of the training data to improve the robustness and generalization of the model.

a) Pre-processing: Pre-processing in this research includes two key steps: resizing (1) and normalization (2). Resizing (1) is a crucial step to ensure that all input images are of the same dimensions, which is necessary for feeding them into the neural network. This step is crucial because neural networks require fixed-size inputs. Let I (1) be the original image, $I_{resized}$ (1) represents the resized image, and $D_{desired}$ (1) denotes the desired dimensions. The resizing process can be represented mathematically as:

$$I_{resized} = resize(I, D_{desired}) \quad (1)$$

Normalization (2) means scaling the pixel values of the images to a standard range, often between 0 and 1 or -1 and 1. This step helps in stabilizing and speeding up the training process by ensuring that all input features have a similar scale. Let $I_{normalized}$ (2) indicates the normalized image, and $I_{resized}$ (2) represents the resized image, min and max show the minimum and maximum pixel values respectively. The normalization process can be expressed mathematically as:

$$I_{normalized} = \frac{I_{resized} - \min(I_{resized})}{\max(I_{resized}) - \min(I_{resized})} \# \quad (2)$$

b) Data augmentation: Data augmentation connects creating new training samples by applying various transformations to the existing data. This technique helps in increasing the variability and diversity of the dataset, thereby reducing overfitting and improving the ability to generalize to unseen data. Three common augmentation techniques include rotation (3), flipping (4), and contrast enhancement (5).

Rotation involves rotating the images by a certain angle. Let I (3) be the original image, θ (3) presents the rotation angle, and $I_{rotated}$ (3) represent the rotated image. The rotation process can be mathematically expressed as:

$$I_{rotated} = rotate(I, \theta) \quad (3)$$

Flipping horizontally or vertically involves flipping the images along the horizontal or vertical axis. Let I (4) denote the original image, and $I_{flipped}$ (4) represent the flipped image. The flipping process can be represented as:

$$I_{flipped} = flip(I) \quad (4)$$

Contrast enhancement involves adjusting the contrast of the images to make features more discernible. Let I (5) presents the original image, and $I_{enchanced}$ (5) represent the contrast-enhanced image. The contrast enhancement process can be expressed as:

$$I_{enchanced} = enchance_{contrast(I)} \quad (5)$$

In summary, pre-processing and data augmentation are important steps in the classification of skin cancer. Pre-processing ensures that the input data is standardized and ready for training, while data augmentation increases the diversity of the dataset, leading to more robust and generalized models. By carefully applying these techniques, researchers and practitioners can improve the performance of skin cancer classification models and contribute to more accurate diagnosis and treatment decisions.

C. Transfer Learning and Fine-tuning of EffecientNetV2B3

Transfer learning means using a pre-trained model that has been trained on a large dataset, and applying it to a different but related task, such as classifying skin cancer images. Instead of training a model from scratch, transfer learning utilizes the knowledge gained from solving one problem and applying it to a different but related problem [13][14][15]. On the other hand, fine-tuning means using a pre-trained model and further training it on a new dataset specific to the task [16][17][18]. This allows the model to adapt to the nuances of the new dataset while retaining the general knowledge learned during pre-training.

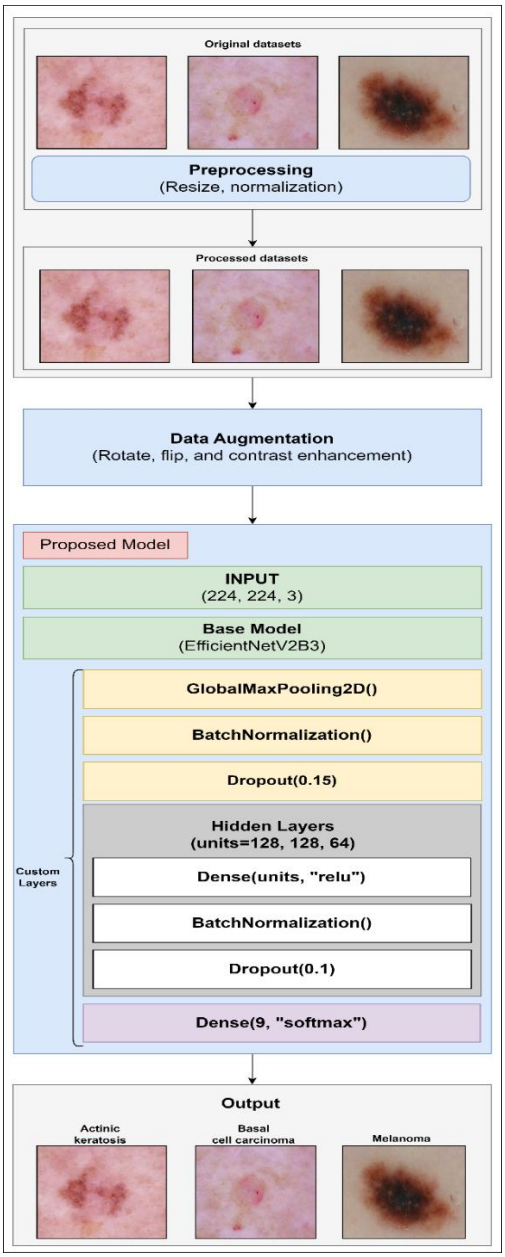

Fig. 2. Procedure of transfer learning and fine-tuning in our model with custom layers.

EfficientNetV2B3 is a convolutional neural network architecture known for its efficiency and effectiveness in image classification tasks. The additional layers mentioned in Fig. 2, such as GlobalMaxPooling2D, Batch Normalization, and Dropout (0.15). In addition, hidden layers consisting of Dense units with ReLU activation, Batch Normalization, and Dropout (0.2), followed by a Dense layer with 9 units and softmax activation, are commonly used to enhance the performance of the model.

GlobalMaxPooling2D reduces the spatial dimensions of the feature maps, summarizing them into a single vector. Batch Normalization normalizes the activations of the previous layer, helping to speed up training and improve generalization. Dropout randomly drops a fraction of neurons during training, reducing overfitting. The hidden layer with Dense units and ReLU activation adds non-linearity to the model, while Batch Normalization and Dropout further regularize it. Finally, the Dense layer with softmax activation produces probabilities for each class of skin cancer.

Combining transfer learning with fine-tuning using EfficientNetB3V2 as a base model with additional layers can lead to a powerful classifier for skin cancer images, leveraging both the general knowledge from pre-training and the specific features of the new dataset.

D. Visual Explanation with Grad-CAM

Grad-CAM is a technique used for visualizing the regions of an image that are key for the prediction of the CNN model. In the context of classifying skin cancer, Grad-CAM can help us understand which parts of the skin image are being attended to by the model when making a classification decision.

Given an image I (6) and a target class y, the final convolutional layer's feature map A (6) is extracted. The gradients of the target class score y_c (6) with respect to the feature map activations are computed using backpropagation:

$$\frac{\partial y_c}{\partial A^k} \tag{6}$$

Then, these gradients are global average pooled to obtain the neuron importance weights:

$$a_k^c = \frac{1}{Z} \sum i \sum j \frac{\partial y_c}{\partial A_{ij}^k} \tag{7}$$

Where Z is the normalization factor to ensure that the importance weights sum up to 1. For a particular neuron k (7), the gradients are summed across all spatial locations (i, j) (7) within the feature map A(7). Finally, the class-discriminative localization map is computed as a weighted combination of the feature maps:

$$L_{Grad-Cam}^c = ReLU(\sum_k a_k^c A^k) \tag{8}$$

For each neuron activation map A^k (8)in the final convolutional layer, it is multiplied element-wise by its corresponding importance weight a_k^c (8). This operation amplifies the activations of neurons that are deemed important for predicting the target class and suppresses the activations of less relevant neurons. Next, these weighted feature maps are summed up across all neurons $\sum_k a_k^c A^k$ (8). Finally, a ReLU (Rectified Linear Unit) (8) activation function is applied to the summed feature map to ensure that only positive values are retained

In skin cancer classification, Grad-CAM can provide insights into which parts of the skin lesion image the model is focusing on to make its decision. For example, if the model correctly classifies a malignant lesion, Grad-CAM in Fig. 3might highlight irregular borders or asymmetric color distribution as important features. Conversely, if the model misclassifies a benign lesion, Grad-CAM might reveal that it is focusing on features that are typically indicative of malignancy, leading to further investigation and refinement of the model.

By visually interpreting the Grad-CAM heatmaps generated for different skin lesion images, dermatologists and researchers can gain valuable insights into the decision-making process of the model and potentially improve the interpretability and trustworthiness of the classification system.

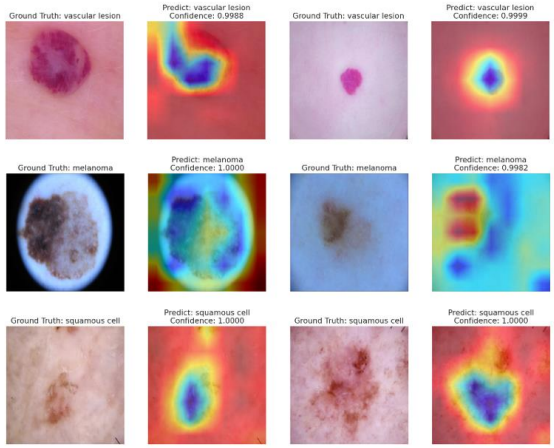

Fig. 3. Visual explanation by Grad_CAM of skin cancer.

E. Image Segmentation by k-means Clustering

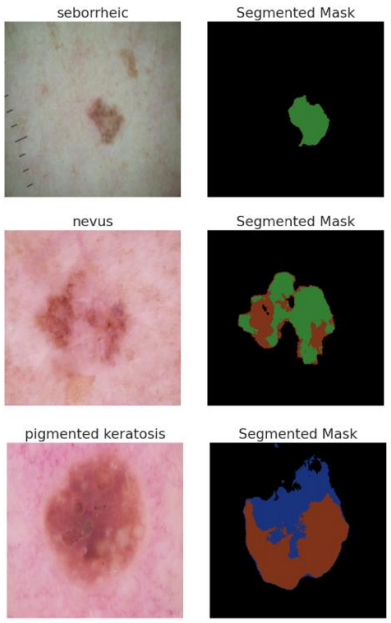

Fig. 4. Image segmentation in skin cancer by k-mean clustering.

Image segmentation using k-means clustering for classifying skin cancer means partitioning the image into different clusters based on pixel intensity values. K-means clustering is a popular unsupervised learning algorithm used for clustering tasks. In this research, it can help identify different regions within an image that may correspond to different types of skin lesions or healthy skin inFig. 4.

Given an input image I of size $m \times n$, the output is to partition the image into k (9) where each cluster represents a distinct region of the image clusters (i.e., $k = 3$ for the normal surface, abnormal surface, and background). The algorithm iteratively assigns each pixel in the image to the cluster with the nearest mean value, minimizing the within-cluster sum of squares. The objective function of k-means clustering is defined as:

$$J = \sum_{i=1}^{k=3} \sum_{x_j \in C_i} \left\| x_j - \mu_i \right\|^2 \quad (9)$$

Algorithm 1 Image Segmentation using K-means Clustering with $k = 3$

1: **Initialization:** Randomly initialize 3 cluster centroids.
2: **Assignment Step:**
3: Assign each pixel to the nearest cluster centroid based on Euclidean distance:
4: **for** each pixel x_j **do**
5: Find the nearest centroid μ_i:
6: $C_i = \{x_j : ||x_j - \mu_i|| \leq ||x_j - \mu_l||, \forall l, 1 \leq l \leq 3\}$
7: **end for**
8: **Update Step:**
9: Update the cluster centroids by computing the mean of all pixels assigned to each cluster:
10: **for** each cluster C_i **do**
11: $\mu_i = \frac{1}{|C_i|} \sum_{x_j \in C_i} x_j$
12: **end for**
13: **Repeat Step:** Iterate steps 2 and 3 until convergence or maximum iterations reached.
14: **Segmentation Step:** Assign each pixel to one of the 3 clusters, resulting in the segmentation of the image into the normal surface, abnormal surface, and background.

Fig. 5. Algorithm of skin cancer segmentation by k-means clustering.

Let $X = \{x_1, x_2, \ldots, x_{mn}\}$ be the set of pixels in the image, where x_i (9) represents the i-th pixel with its corresponding feature vector. Each feature vector typically consists of color intensity values or texture features. Where C_i (9) represents the i-th cluster, μ_i (9) is the mean (centroid) of cluster C_i (9), and $\left\| x_j - \mu_i \right\|$ (9) denotes the Euclidean distance. The steps involved in image segmentation using k-means clustering for skin cancer classification are as follows in Fig. 5.

By segmenting the skin lesion regions using k-means clustering, dermatologists can efficiently analyze and classify skin cancer from dermatological images, aiding in early detection and diagnosis.

IV. Experiments

A. Dataset and Performance Metrics

The dataset comprises 2,357 images sourced from the International Skin Imaging Collaboration (ISIC), encompassing various oncological conditions, with a focus on melanoma, a potentially fatal form of skin cancer constituting 75% of skin cancer-related deaths. This dataset is a critical resource for developing solutions to automate melanoma detection processes, thus aiding dermatologists in early diagnosis. Fig. 6includes images depicting malignant and benign conditions such as actinic keratosis, basal cell carcinoma, dermatofibroma, melanoma, nevus, pigmented benign keratosis, seborrheic keratosis, squamous cell carcinoma, and vascular lesions which were described in Fig. 7. the dataset offers a substantial and diverse collection for training machine learning algorithms or developing image analysis tools.

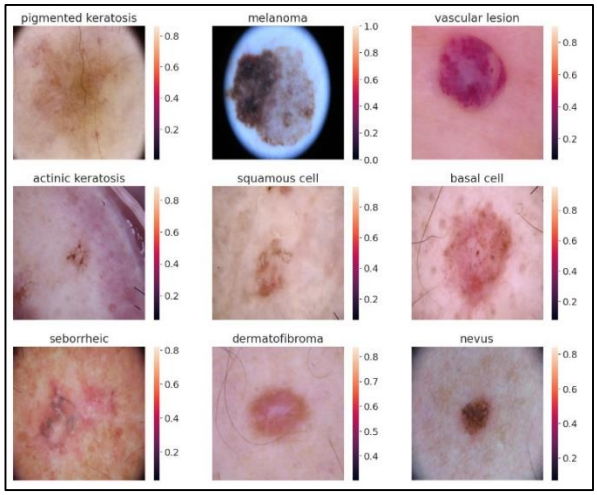

Fig. 6. Dataset about the skin diseases.

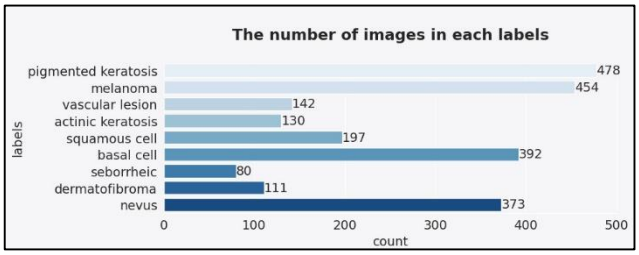

Fig. 7. The amount of data in nine classes.

To evaluate the effectiveness of classification models in classifying skin illnesses, various performance metrics are employed, among which accuracy, recall, precision, and F1 score stand as fundamental measures. These metrics provide quantitative insights into the ability of the model to correctly classify instances of malignant and benign skin lesions.

Accuracy (10) is the most intuitive metric which calculates the ratio of correctly predicted cases to the total number of cases evaluated. It is represented by the formula:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

True positives (TP) are instances of correctly classified skin disease, true negatives (TN) are correctly classified instances of absence of skin disease, false positives (FP) are samples incorrectly classified as having skin disease, and false negatives (FN) are cases incorrectly classified as not having skin disease are all represented by the equation (10). While accuracy provides a general overview of the performance, it may not be sufficient when dealing with imbalanced datasets, such as in skin cancer classification, where benign cases often outnumber

malignant ones. Hence, recall (11) and (12) precision metrics offer additional insights.

Recall (11) also known as sensitivity or true positive rate, measures the proportion of actual positive cases that are correctly identified by the model. It is calculated as:

$$Recall = \frac{TP}{TP+FN} \quad (11)$$

On the other hand, precision (12) quantifies the model's ability to correctly identify positive cases among all cases predicted as positive. It is expressed as:

$$Precision = \frac{TP}{TP+FP} \quad (12)$$

While recall emphasizes minimizing false negatives, precision focuses on minimizing false positives. However, these metrics alone may not provide a comprehensive assessment of the model's performance. Therefore, the F1 score (13), which harmonizes precision and recall, is often utilized. The F1 score (13) is the harmonic mean of precision and recall and is given by the formula:

$$F1 = \frac{2 \times Precison \times Recall}{Precison+Recall} \quad (13)$$

In skin cancer classification, where both false positives and false negatives can have serious consequences, achieving a balance between precision and recall is crucial. Thus, the F1 score (13) serves as a consolidated measure, incorporating both precision and recall, providing a more holistic evaluation of efficacy in classifying skin lesions accurately.

B. *Scenario 1: The Result of Classifying Skin Diseases Into Nine Classes: Actinic Keratosis, Basal Cell Carcinoma, Dermatofibroma, Melanoma, Nevus, Pigmented Benign Keratosis, Seborrheic Keratosis, Squamous Cell Carcinoma, and Vascular Lesion.*

TABLE I. THE ACCURACY OF CLASSIFYING SKIN DISEASES INTO NINE CLASSES IN TRANSFER LEARNING AND FINE TUNING, FOR EACH DEEP LEARNING MODEL

Model	Transfer learning				
	Val acc	Test acc	Precision	Recall	F1
Our Model	**63.71%**	**65.59%**	**65.73%**	**65.59%**	**64.84%**
ResNet50V2	53.39%	55.27%	55.07%	55.27%	54.10%
MobileNetV2	45.17%	44.06%	44.80%	44.06%	43.77%
EffecientNetB3	63.26%	61.15%	61.18%	61.15%	60.42%
ResNet50	72.03%	69.37%	69.25%	69.37%	68.55%
MobileNet	51.83%	50.72%	50.44%	50.72%	50.06%
-	**Fine tuning**				
Our Model	**85.13%**	**84.91%**	**85.62%**	**84.91%**	**84.68%**
ResNet50V2	60.49%	57.49%	59.21%	57.49%	57.35%
MobileNetV2	56.60%	53.50%	54.78%	53.50%	53.39%
EffecientNetB3	84.91%	83.24%	83.69%	83.24%	82.86%
ResNet50	80.69%	80.58%	80.74%	80.58%	80.08%
MobileNet	58.82%	58.49%	58.52%	58.49%	57.19%

Following Table II, our model achieved an accuracy of 65.59% in transfer learning for classifying nine classes of skin diseases. To clarify, the custom model attained the second position after ResNet50, which reached 69.25%. However, our model significantly improved and reached 84.91% in test accuracy of fine-tuning phase, marking a 19.32% increase and placing it first among test models. Additionally, ResNet50 and EfficientNetB3 showed moderate improvements at 80.58% and 83.24%, respectively. Thus, this significant increase in performance led to the successful classification of skin diseases by our model.

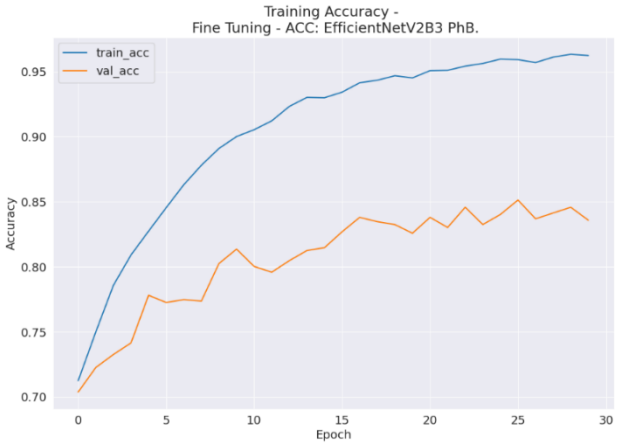

Fig. 8. Training accuracy and validation accuracy by fine tuning of our model by classifying nine classes.

Additionally, Fig. 8 displays the accuracy of a model during both the training and validation phases. The validation accuracy demonstrates how well the model generalizes to unseen data, helping to identify overfitting or underfitting issues. Ideally, both training and validation accuracies should increase as training progresses.

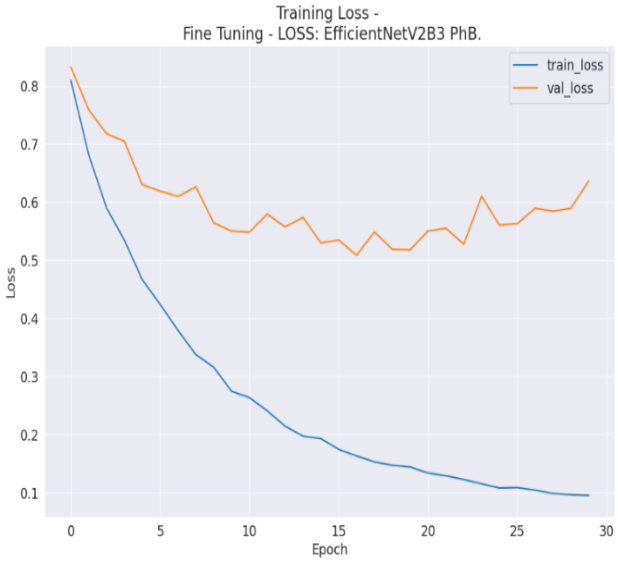

Fig. 9. Training loss and validation loss by fine tuning of our model by classifying nine classes.

Besides, Fig. 9 shows the training and validation loss over epochs. The loss represents a measure of how well the model is performing: lower loss indicates better performance. The training loss depicts how well the model fits the training data, while the validation loss indicates how well the model generalizes to unseen data.

A confusion matrix in Fig. 10 is a tabular representation of predicted classes versus true classes. For classifying skin diseases, the confusion matrix can help evaluate the performance of a classification model by providing insights into the types of errors it makes. From this information, adjustments to the model or further data collection efforts can be made to improve classification accuracy.

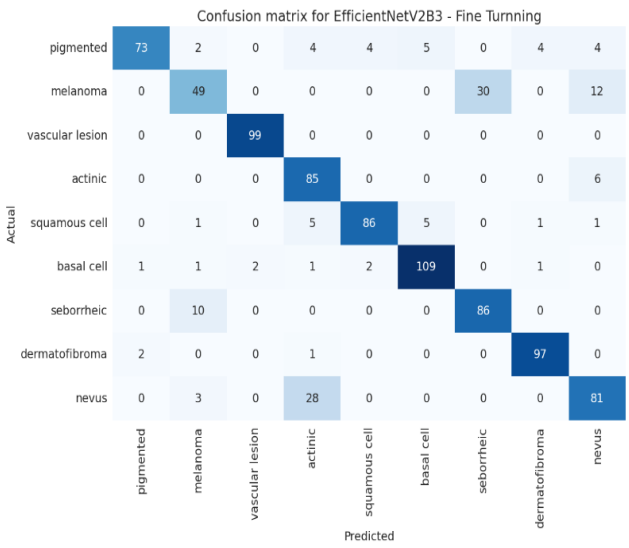

Fig. 10. Confusion matrix in fine tuning for our model by classifying nine classes.

C. *Scenario 2: The Result of Classifying Skin Diseases Into Two Classes: Benign and Malignant*

Using both transfer learning and fine-tuning strategies, Table II presents a comparative examination of different deep-learning models in the classification of skin disease photos into two classes: benign and malignant. During the transfer learning phase, the ResNet50 model hit the top performance with an accuracy of 90.00%. Next to that, ResNet50V2, EffecientNetB3, and our model reached with test accuracy of 88.28%, 86.60%, and 86.20%, respectively. However, During the fine-tuning phase, the custom model successfully hit the highest top of the test models with a 94.00% accuracy score (i.e., increasing 7.8% when compared with transfer learning). As a result, the EffecientNetV2B3 model with our extra layer works effectively with both nine and two classes classifying when compared with other models although it ran unsmooth in transfer learning.

The utilization of training and validation line graphs in Fig. 11 and Fig. 12 aid in monitoring the performance of machine learning models over training epochs, ensuring optimal accuracy and minimal loss. Meanwhile, Fig. 13 facilitates a comprehensive assessment of model performance, enabling targeted improvements and insights into misclassifications for enhanced diagnostic accuracy.

TABLE II. THE ACCURACY OF CLASSIFYING SKIN DISEASES INTO TWO CLASSES IN TRANSFER LEARNING AND FINE TUNING, FOR EACH DEEP LEARNING MODEL

Model	Transfer learning				
	Val acc	Test acc	Precision	Recall	F1
Our Model	**85.80%**	**86.20%**	**86.23%**	**86.20%**	**86.20%**
ResNet50V2	87.03%	88.28%	88.33%	88.28%	88.28%
MobileNetV2	85.29%	85.54%	85.54%	85.54%	85.54%
EffecientNetB3	87.80%	86.60%	87.10%	86.60%	86.55%
ResNet50	89.40%	90.00%	90.06%	90.00%	90.00%
MobileNet	83.29%	81.30%	81.79%	81.30%	81.22%
-	**Fine tuning**				
Our Model	**95.20%**	**94.00%**	**94.01%**	**94.00%**	**94.00%**
ResNet50V2	88.53%	91.52%	91.53%	91.52%	91.52%
MobileNetV2	89.03%	88.78%	88.78%	88.78%	88.78%
EffecientNetB3	95.40%	92.00%	92.04%	92.00%	92.00%
ResNet50	94.00%	92.80%	91.53%	91.52%	91.52%
MobileNet	91.02%	92.02%	92.17%	92.02%	92.01%

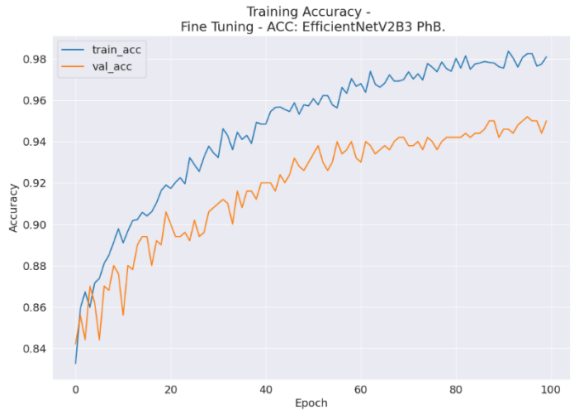

Fig. 11. Training accuracy and validation accuracy by fine tuning of our model by classifying two classes.

Fig. 12. Training loss and validation loss by fine tuning of our model by classifying two classes

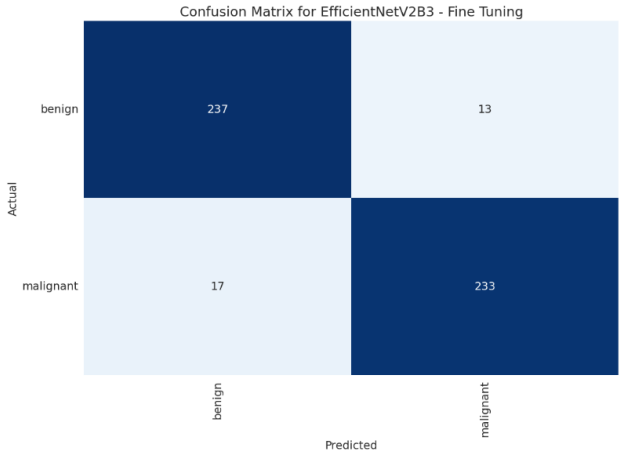

Fig. 13. Confusion matrix in fine tuning for our model by classifying two classes.

D. *Scenario 3: The Result of Classifying Skin Diseases Into Six Benign Classes: Actinic Keratosis, Dermatofibroma, Nevus, Pigmented Benign Keratosis, Seborrheic Keratosis, and Vascular Lesion.*

In classifying six benign classes, Table III indicates our model working extremely well at the fine-tuning phase which reached an accuracy of 89.56%. But ResNet50 and EfficientNetB3 show a little better in performance with an accuracy of 90.40% and 91.58%. Thus, this scenario demonstrates our model still has a limit and needs to improve in the future.

TABLE III. THE ACCURACY OF CLASSIFYING SKIN DISEASES INTO SIX BENIGN CLASSES IN TRANSFER LEARNING AND FINE TUNING, FOR EACH DEEP LEARNING MODEL

Model	Transfer learning				
	Val acc	Test acc	Precision	Recall	F1
Our Model	**89.06%**	**86.70%**	**87.17%**	**86.70%**	**86.53%**
ResNet50V2	79.12%	76.43%	77.56%	76.43%	76.41%
MobileNetV2	67.68%	61.62%	62.86%	61.62%	61.86%
EffecientNetB3	87.04%	84.01%	84.74%	84.01%	83.46%
ResNet50	88.22%	86.36%	86.71%	86.36%	86.32%
MobileNet	73.06%	69.02%	69.44%	69.02%	68.71%
-	**Fine tuning**				
Our Model	**92.59%**	**89.56%**	**90.08%**	**89.56%**	**89.55%**
ResNet50V2	79.97%	75.59%	76.31%	75.59%	75.60%
MobileNetV2	71.38%	64.31%	64.89%	64.31%	64.09%
EffecientNetB3	93.43%	91.58%	92.08%	91.58%	91.48%
ResNet50	91.41%	90.40%	90.87%	90.40%	90.39%
MobileNet	78.28%	76.77%	76.83%	76.77%	76.27%

Furthermore, Training and validation on both accuracy and loss scores are presented in Fig. 14 and Fig. 15. Following the figures, the evaluation performance of our model presents the balance with validation accuracy achieved of 92,59% and validation loss gained of 0.26 when the dataset is changed. Moreover, Fig. 16 is provided for evaluating, optimizing, and understanding the performance of deep learning models.

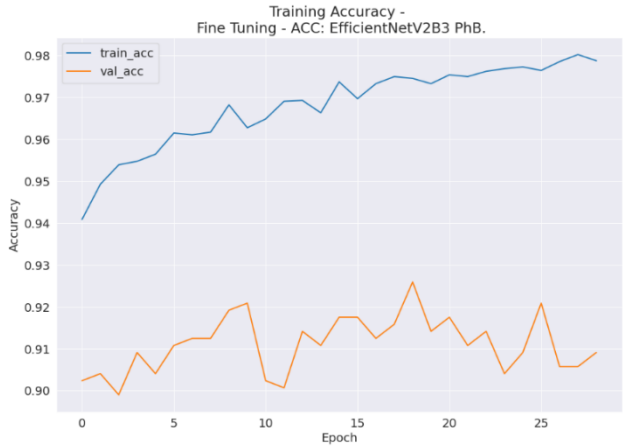

Fig. 14. Training accuracy and validation accuracy by fine tuning of our model by classifying six classes.

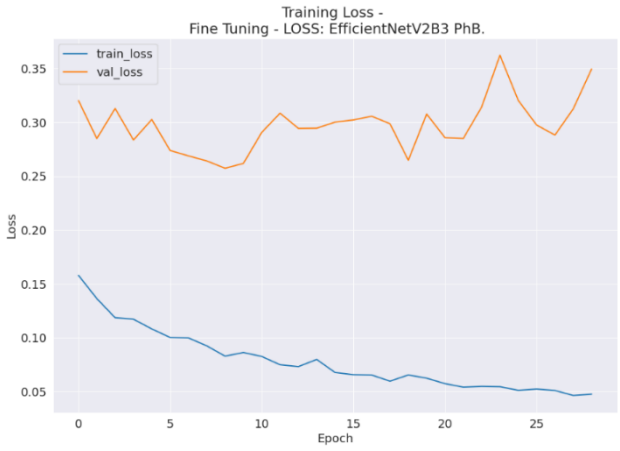

Fig. 15. Training loss and validation loss by fine tuning of our model by classifying six classes.

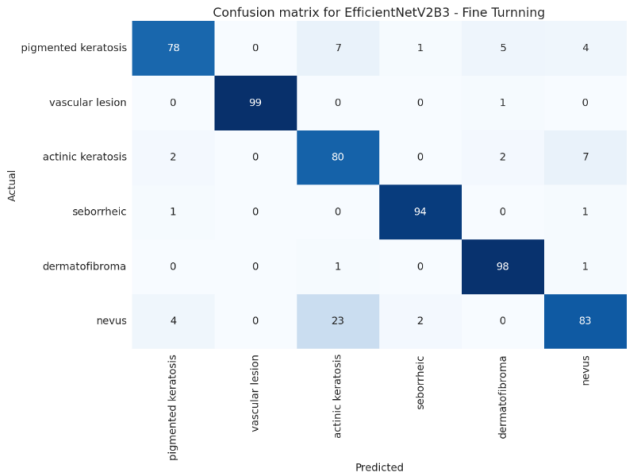

Fig. 16. Confusion matrix in fine tuning for our model by classifying six classes.

E. Scenario 4: The Result of Classifying Skin Diseases Into Three Malignant Classes: Basal Cell Carcinoma, Melanoma, and Squamous Cell Carcinoma

A successful classification is shown in Table IV when our model achieves a significant increase in test accuracy of 96.74% of fine-tuning or a growth of 13.03%. In addition, some scores including f1, recall, and prediction hit a peak. As a result, our model effectively proved that it performs better than previous models at classifying images into three classes of malignant. However, EfficientNetB3 and ResNet50 present a dramatic growth in performance (i.e., with an accuracy of 94.79% and 93.49%, respectively) when working with three classes. Besides ResNet50V2, MobileNetV2, and MobileNet rise marginally.

TABLE IV. THE ACCURACY OF CLASSIFYING SKIN DISEASES INTO THREE MALIGNANT CLASSES IN TRANSFER LEARNING AND FINE TUNING, FOR EACH DEEP LEARNING MODEL

Model	Transfer learning				
	Val acc	Test acc	Precision	Recall	F1
Our Model	**87.95%**	**83.71%**	**83.74%**	**83.71%**	**83.71%**
ResNet50V2	79.48%	73.94%	76.16%	73.94%	74.23%
MobileNetV2	73.29%	70.68%	71.11%	70.68%	70.85%
EffecientNetB3	89.90%	79.48%	79.67%	79.48%	79.55%
ResNet50	87.62%	85.02%	85.14%	85.02%	85.04%
MobileNet	71.66%	68.08%	69.15%	68.08%	68.30%
-	**Fine tuning**				
Our Model	**98.70%**	**96.74%**	**96.74%**	**96.74%**	**96.74%**
ResNet50V2	78.18%	71.34%	70.61%	71.34%	70.67%
MobileNetV2	73.94%	72.96%	72.99%	72.96%	72.81%
EffecientNetB3	97.07%	94.79%	94.84%	94.79%	94.79%
ResNet50	95.11%	93.49%	93.53%	93.49%	93.50%
MobileNet	77.52%	75.57%	76.51%	75.57%	75.82%

Furthermore, Fig. 17 and Fig. 18 illustrate our model's development, almost reaching the pinnacle with a surprising validation accuracy of 98.70%. Additionally, training and validation loss obtained a substantial decrease, reaching 0.06. For further information, Fig. 19 provides an overall confusion matrix for the research result.

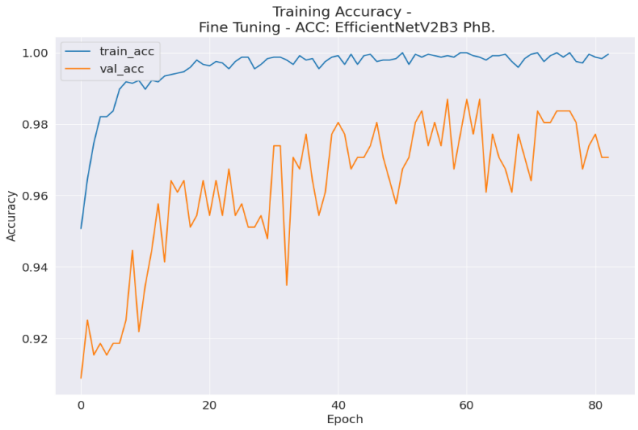

Fig. 17. Training accuracy and validation accuracy by fine tuning of our model by classifying three classes.

Fig. 18. Training loss and validation loss by fine tuning of our model by classifying three classes.

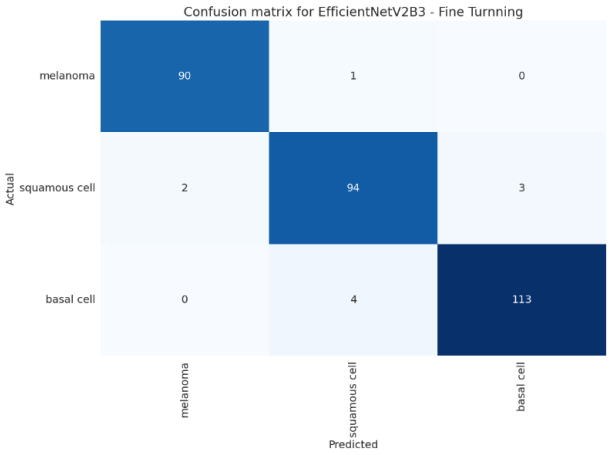

Fig. 19. Confusion matrix in fine tuning for our model by classifying three classes.

V. Results and Comparison

A. Results Explaination

A quick look at all experiments, all of the results in our model reached an impressive performance although it still has a limit that needs to be enhanced in the future. Specifically, Scenario 1: The Result of Classifying Skin Diseases Into Nine Classes: Actinic Keratosis, Basal Cell Carcinoma, Dermatofibroma, Melanoma, Nevus, Pigmented Benign Keratosis, Seborrheic Keratosis, Squamous Cell Carcinoma, and Vascular Lesion. shows that the proposed pipeline doing a good job in classifying nine classes with an accuracy of 84.91%. However, our ambitions are higher when future research should reach a test accuracy larger than 90%. In addition, the model illustrates the limit on the performance in Scenario 3: The Result of Classifying Skin Diseases Into Six Benign Classes: Actinic Keratosis, Dermatofibroma, Nevus, Pigmented Benign Keratosis, Seborrheic Keratosis, and Vascular Lesion. when classifies six classes. This led to the way for our research to fix it in the future. Besides, the performance of our model in Scenario 2: The Result of Classifying Skin Diseases Into Two Classes: Benign and Malignant and Scenario 4: The Result of Classifying Skin Diseases Into Three Malignant Classes: Basal Cell Carcinoma, Melanoma, and Squamous Cell Carcinoma achieved the highest test accuracy and other scores when compared with test models and other state-of-the-art methods. This demonstrates the customized EfficientNetV2B3 model achieved a success in classifying skin diseases. The summary of the outcome of these scenarios is shown in Fig. 20.

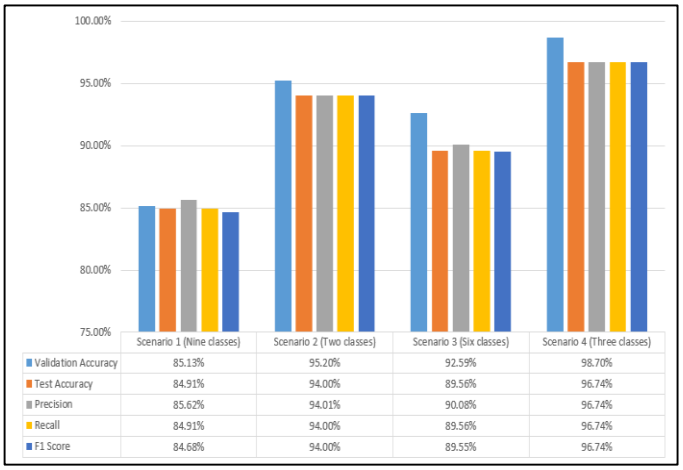

Fig. 20. The result of fine-tuning of our model through four scenarios.

Furthermore, the Grad-Cam was added to visualize areas of focus in skin cancer images in Fig. 3. It highlights regions contributing to predictions and enhances model transparency for medical professionals. Additionally, K-means clustering enables detailed analysis of different regions in Fig. 4. Integrating Grad-CAM for visualization and K-means clustering for feature extraction enhances the interpretability and effectiveness of skin cancer classification models, facilitating more accurate diagnoses and treatment decisions.

B. Comparison with others State-of-the-art Methods

Comparing skin disease classification models with state-of-the-art methods in CNN is crucial for assessing performance, identifying areas for improvement, and validating innovations. By benchmarking against existing approaches, researchers gain insights into their model's effectiveness and efficiency. Such comparisons help highlight strengths and weaknesses, guiding further optimizations. Ensuring CNN models perform competitively against state-of-the-art methods is essential for their practical utility and reliability in real-world scenarios. This process fosters the development of robust diagnostic tools, potentially enhancing healthcare outcomes. Thus, the result of this comparison is presented in Table V.

TABLE V. Comparison with Others State-of-the-Art Methods in ISIC Dataset

Ref.	Proposed	Year	Classes	Accuracy
Ahmed Abdelhafeez et al [19]	SVNSs, DarkNet, and GoogleNet	2023	8 classes	85.74%
Maad M. Mijwil et al [21]	InceptionV3	2023	2 classes	86.90%
Solene Bechelli et al [23]	VGG16	2023	2 classes	88%
Ayesha Atta et al [27]	Customized CNN	2022	2 classes	86.23%
Vatsala Anand et al [28]	VGG16	2022	2 classes	89.09%
Dipu Chandra Malo et al [29]	VGG16	2023	2 classes	87.60%
Mohammed Rakeibul Hasan et al [31]	VGG16	2021	2 classes	93.18%
Abdurrahim Yilmaz et al [32]	NASNetMobile	2022	3 classes	82%
Chandran Kaushik Viknesh et al [33]	AlexNet, LeNet, and VGG-16	2021	2 classes	91%
Proposed model			**9 classes**	**84.91%**
			2 classes	**94.00%**
			6 classes	**89.56%**
			3 classes	**96.74%**

C. Limit and Future Work

While the research has reached promising results, it also shows certain limitations. Despite achieving high accuracy rates, there may still be instances of misclassification. Additionally, the dataset may not encompass all possible variations of skin diseases, necessitating ongoing expansion and diversification. Looking ahead, the study sets the stage for future endeavors aimed at refining the model and methodologies. Plans include increasing data preprocessing and incorporating advanced

visualization techniques to gain deeper insights into model performance and image characteristics. Moreover, expanding the dataset to encompass a broader spectrum of skin diseases will be a priority ensuring greater robustness and generalization of the model.

VI. Conclusion

The article developed a specialized model tailored to classify skin disease images for medical applications. Our custom model showcased remarkable accuracy, achieving 84.91% in classifying nine different classes of skin cancer. Notably, it also demonstrated an impressive 94.00% accuracy in discerning between malignant and benign cases. Further experiments revealed its proficiency in distinguishing between various types of benign and malignant skin diseases, with accuracies of 89.56% for six benign classes and 96.74% for three malignant classes.

One of the key techniques employed to boost the performance was transfer learning and fine-tuning. In this case, the EfficientSkinCaSV2B3 framework was proposed by adding dense and dropout layers into the EfficientNetV2B3 model while fine-tuning its parameters. As a result, this process significantly improved accuracy. In addition, Grad-Cam were used to provide insights into the model's decision-making process. Furthermore, k-means clustering was employed to segment images.

In conclusion, the research contributes to the intersection of medicine and computer science by advancing the classification and segmentation of skin disease images. Through the judicious application of transfer learning, visualization techniques like Grad-Cam, and clustering methods such as k-means, the aim is to continue improving diagnostic accuracy and ultimately enhance patient care in dermatology.

Acknowledgment

We express sincere appreciation to Huong Hoang Luong, Hao Van Tran, and Phuc Tan Huynh for their invaluable contributions. Their dedication and expertise have been instrumental in the success of the project. We are deeply grateful for their hard work and collaboration.

References

[1] S. A. Umar, S. A. Tasduq, "Ozone layer depletion and emerging public health concerns-an update on epidemiological perspective of the ambivalent effects of ultraviolet radiation exposure," Frontiers in Oncology, vol. 12, p. 866733, 2022.

[2] J. Ferlay, M. Colombet, I. Soerjomataram, D. M. Parkin, M. Piñeros, A. Znaor, F. Bray, "Cancer statistics for the year 2020: An overview," International Journal of Cancer, vol. 149, no. 4, pp. 778–789, 2021.

[3] C. Xia, X. Dong, H. Li, M. Cao, D. Sun, S. He, F. Yang, X. Yan, S. Zhang, N. Li, "Cancer statistics in China and United States, 2022: profiles, trends, and determinants," Chinese Medical Journal, vol. 135, no. 05, pp. 584–590, 2022.

[4] S. N. Hồng, H. Le Thanh, S. N. Huu, "The epidemiology of skin cancer at the national hospital of dermatology and venereology from 2017 - 2021," Tạp chí Da liễu học Việt Nam, 2023.

[5] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, J. Barata, "Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook," IEEE Access, vol. 8, pp. 220121–220139, 2020.

[6] A. Jamwal, R. Agrawal, M. Sharma, A. Giallanza, "Industry 4.0 technologies for manufacturing sustainability: A systematic review and future research directions," Applied Sciences, vol. 11, no. 12, p. 5725, 2021.

[7] D. Mourtzis, J. Angelopoulos, N. Panopoulos, "A Literature Review of the Challenges and Opportunities of the Transition from Industry 4.0 to Society 5.0," Energies, vol. 15, no. 17, p. 6276, 2022.

[8] H.L. Duc, T.T. Minh, K.V. Hong, and H.L. Hoang, "84 birds classification using transfer learning and efficientnetb2," in International Conference on Future Data and Security Engineering. Springer, 2022, pp. 698–705.

[9] K.V. Hong, T.T. Minh, H.L. Duc, N.T. Nhat, and H.L. Hoang, "104 fruits classification using transfer learning and densenet201 fine-tuning," in Computational Intelligence in Security for Information Systems Conference. Springer, 2022, pp. 160–170.

[10] K.D.D. Le, H.H. Luong, and H.T. Nguyen, "Patient classification based on symptoms using machine learning algorithms supporting hospital admission," in Nature of Computation and Communication: 7th EAI International Conference, ICTCC 2021, Virtual Event, October 28–29, 2021, Proceedings 7. Springer, 2021, pp. 40–50.

[11] H.T. Nguyen, N.K.T. Nguyen, C.L.H. Tran, and H.H. Luong, "Effects evaluation of data augmentation techniques on common seafood types classification tasks," in Biomedical and Other Applications of Soft Computing. Springer, 2022, pp. 213–223.

[12] H.T. Nguyen, Q.T. Quach, C.L.H. Tran, and H.H. Luong, "Deep learning architectures extended from transfer learning for classification of rice leaf diseases," in International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Springer, 2022, pp. 785–796.

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

[14] Z. Zhu, K. Lin, A. K. Jain, J. Zhou, "Transfer learning in deep reinforcement learning: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.

[15] S. Yao, Q. Kang, M. Zhou, M. J. Rawa, A. Abusorrah, "A survey of transfer learning for machinery diagnostics and prognostics," Artificial Intelligence Review, vol. 56, no. 4, pp. 2871–2922, 2023.

[16] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, K. Aberman, "Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 22500–22510.

[17] W. Chen, Y. Liu, W. Wang, E. M. Bakker, T. Georgiou, P. Fieguth, L. Liu, M. S. Lew, "Deep learning for instance retrieval: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.

[18] H. Rasheed, M. U. Khattak, M. Maaz, S. Khan, F. S. Khan, "Fine-tuned clip models are efficient video learners," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 6545–6554.

[19] A. Abdelhafeez, H. K. Mohamed, A. Maher, N. A. Khalil, "A novel approach toward skin cancer classification through fused deep features and neutrosophic environment," Frontiers in Public Health, vol. 11, 2023.

[20] P. Nadiger, R. Pavasker, S. C., S. Hangal, V. S. Bhat, "Skin Cancer Detection and Classification Using Deep Learning," International Journal For Science Technology And Engineering, vol. 11, no. 5, pp. 7182–7185, 2023.

[21] M. M. Mijwil, "Skin cancer disease images classification using deep learning solutions," Multimedia Tools and Applications, vol. 80, no. 17, pp. 26255–26271, 2021.

[22] K. Ali, Z. A. Shaikh, A. A. Khan, A. A. Laghari, "Multiclass skin cancer classification using EfficientNets--a first step towards preventing skin cancer," Neuroscience Informatics, vol. 2, no. 4, p. 100034, 2022.

[23] S. Bechelli, J. Delhommelle, "Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images," Bioengineering, vol. 9, no. 3, p. 97, 2022.

[24] K. Mridha, M. M. Uddin, J. Shin, S. Khadka, M. F. Mridha, "An interpretable skin cancer classification using optimized convolutional neural network for a smart healthcare system," IEEE Access, 2023.

[25] D. Keerthana, V. Venugopal, M. K. Nath, M. Mishra, "Hybrid convolutional neural networks with SVM classifier for classification of skin cancer," Biomedical Engineering Advances, vol. 5, p. 100069, 2023.

[26] S. Jain, U. Singhania, B. Tripathy, E. A. Nasr, M. K. Aboudaif, A. K. Kamrani, "Deep learning-based transfer learning for classification of skin cancer," Sensors, vol. 21, no. 23, p. 8142, 2021.

[27] A. Atta, M. A. Khan, M. Asif, G. F. Issa, R. A. Said, T. Faiz, "Classification of Skin Cancer empowered with convolutional neural network," in 2022 International Conference on Cyber Resilience (ICCR), 2022, pp. 01–06.

[28] V. Anand, S. Gupta, A. Altameem, S. R. Nayak, R. C. Poonia, A. K. J. Saudagar, "An enhanced transfer learning based classification for diagnosis of skin cancer," Diagnostics, vol. 12, no. 7, p. 1628, 2022.

[29] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, 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, p. 100036, 2021.

[30] D. C. Malo, M. M. Rahman, J. Mahbub, M. M. Khan, "Skin Cancer Detection using Convolutional Neural Network," in 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC), 2022, pp. 0169-0176.

[31] M. R. Hasan, M. I. Fatemi, M. M. Khan, M. Kaur, A. Zaguia, "Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks," Journal of Healthcare Engineering, vol. 2021, 2021.

[32] A. Yilmaz, M. Kalebasi, Y. Samoylenko, M. E. Guvenilir, H. Uvet, "Benchmarking of lightweight deep learning architectures for skin cancer classification using ISIC 2017 dataset," arXiv preprint arXiv:2110.12270, 2021.

[33] C. K. Viknesh, P. N. Kumar, R. Seetharaman, D. Anitha, "Detection and Classification of Melanoma Skin Cancer Using Image Processing Technique," Diagnostics, vol. 13, no. 21, p. 3313, 2023.