Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 08, No. 01, March 2022.

MELANOMA SKIN LESION CLASSIFICATION USING IMPROVED EFFICIENTNETB3

Saumya R. Salian and Sudhir D. Sawarkar

(Received: 4-Nov.-2021, Revised: 26-Dec.-2021, Accepted: 12-Jan.-2022)

ABSTRACT

Malignant skin cancer is one the most common and lethal type of skin cancer. Early detection of cancerous skin lesions will increase the possibility of patient survival. In recent years, implementation of models built on deep neural networks in building medical diagnostic imaging systems is quite beneficial to medical experts. In this study, we present an improved and fine-tuned EfficientNetB3 model to classify malignant skin lesions using the concept of fine-tuning transfer learning. We have performed a comparative analysis of different deep learning pre-trained models, like ResNet50, InceptionV3, InceptionResNetV2 and EfficientNet B0-B2 models. The analysis findings signify the ability of utilizing fine-tuned EfficientNetB3 in the mission of melanoma detection and development of a computer-aided diagnostic system. All experimental procedures were carried out on ISBI-ISIC 2017 dataset. To check the efficiency of the proposed model, we compare the proposed model with EfficientNetB3 baseline model and present state-of-art pre-trained methods and approaches. The proposed EfficientNetB3 model obtained an accuracy of 87.12%, a recall of 87.00%, a precision of 87.00% and an F1 score of 85.00%. The proposed model achieved good computational results and efficaciously addressed the problem of model over-fitting and abated false negative labels.

KEYWORDS

Melanoma, Deep learning, Classification, Skin lesions, Computer vision.

1. INTRODUCTION

Cancer disease is minacious to human life. At present, over 100 different cancers are known that affect humans. One of the most fast-expanding and lethal cancers is skin cancer. Malignant melanoma is a class of skin lesion cancer that is considered as most precarious and prevalent type of cancer. Melanoma skin cancer occurs due to abnormal reproduction of pigmented melanocytes cells. The key factor that risks to the expansion of melanoma skin cancer cases is the increased exposure to natural or artificial ultra violet rays, tanning beds and sunburns. According to Skin Cancer Foundation statistics, it is estimated that new cases of melanoma cancer would intensify close to 5.8% and count of fatality rate is likely to intensify by 4.8% [1]. If skin cancer is not diagnosed at an early stage, 80% of cases may result in death. Therefore, it becomes important that identification of melanoma at a very preliminary stage will remarkably escalate the survival chance of a patient. Medical experts or dermatologists usually examine skin surface using dermotoscopy to identify the skin lesions [2]. Visual inspection using dermatoscopy is a subjective technique and its diagnosis often depends on the medical practitioner's experience. For unbiased diagnosis of melanoma skin cancer, an automated and computerized image analysis system is a prime requirement. Computer-aided diagnostic systems will aid medical experts to utilize technological advances, but also have second opinion. To support dermatologists to achieve faster results in diagnosis of skin cancer with lesser computational time, it is necessary to develop computer-aided image processing systems which automatically classify malignant skin lesions. In the last few years, there have been a notable increase in the quest to build computer-aided diagnostic solutions to diagnose malignant skin lesions. Before 2015, research was typically based on conventional machine learning computer vision algorithms to detect skin cancer. Deep-learning models have accomplished outstanding results in building diagnostic computer-aided systems.

Convolutional Neural Network (CNN) has achieved remarkable results on tasks, like melanoma classification, object recognition, image recognition, …etc. In recent years, deep neural networks along with transfer learning models have been used together to work on large datasets [3]. Transfer learning is a technique of reusing knowledge gained by a model trained on a specific domain to build the solutions for related problems [12]. Transfer learning assists in lowering the time required to train the network

S. R. Salian and S. D. Sawarkar are with Department of Computer Engineering, Datta Meghe College of Engineering, Navi Mumbai, India. Emails: srs.cm.dmce@gmail.com and sudhir_sawarkar@yahoo.com

and reducing out-of-sample error. ImageNet is a huge dataset having more than 15 million labelled images. The diverseness and ordered organization of ImageNet provide exceptional opportunities to explore and further research in the field of computer vision systems. All the deep learning models that we are analyzing are pre-trained on ImageNet [13] dataset consisting of 1000 classes and containing features with respect to their weights and biases. In order to use these models to adapt to our two class skin lesion classification problem, fine-tuning of the network is important. Pre-trained model is fine-tuned when the target dataset is different from the original dataset on which the model was trained. Performance of the pre-trained model can be remarkably enhanced by fine-tuning the model on target dataset instead of training from the beginning. In fine tuning transfer learning, we take underlying weights of a pre-trained model and adjust (fine tune) them to fit to our dataset.

In this work, we report a deep learning-based image classification system that classifies the dermoscopic images into two classes of malignant and benign melanoma. We propose an improved and fine-tuned EfficientNetB3 model for melanoma lesion classification. EfficientNetB3 model belongs to the family of EfficientNet B0-B7 [15] models that are previously trained on huge datasets such as ImageNet. In most of the baseline Convolution Neural Networks (CNNs), initial layers of the model extract much less features and the last layers of the model bring out comprehensive features. Instead of building the CNN architecture from scratch, we have utilized network architecture of EfficientNetB3 model by freezing the initial layers of the model and then fine tuning their weights in order to classify them accurately to our two-class skin lesion classification task. We use various data augmentation techniques and add customized layers of global average pooling, drop out layers to minimize the model over-fitting problem and change the last classification layer with softmax layer that classifies the benign and malignant skin lesions.

Most of the pre-trained CNN architectures suffer from model over-fitting problem. With the purpose to obliterate this drawback, fine tuning of pre-trained network plays an essential role in building an efficient computer-aided diagnostic model for classification objective. In this study, we examine the efficiency of our proposed model in detection of malignant melanoma skin lesion. We compare our fine-tuned EfficientNetB3 model with baseline EfficientNetB3 model without additional layers of data augmentation, global average pooling and dropout. A comparative analysis of the proposed model with other state-of-the-art advanced neural networks, like ResNet50, InceptionV3, InceptionResNetV2 and EfficientNetB0-B2 is carried out in this paper. We also compare our proposed model with deep-learning networks proposed by other researchers. We carry out a comparative analysis of model performance evaluated on metrics like accuracy, recall, precision and F1-score. We have carried out our experiments on dataset available from ISBI-ISIC 2017 Challenge with the purpose to improve the diagnosis of melanoma [4]. Experiment analysis indicates that our proposed EfficientNetB3 model performs better than all other pre-trained models and provides a promising result that tackles the problem of model over-fitting in deep neural pre-trained models. Our model can be used as a decision support solution to help dermatologists in the diagnosis process. The paper is outlined in following manner. Section 2 provides a detailed quintessence of recent works and Section 3 furnishes a detailed description of the proposed methodology. Section 4 contains an analysis and a comparison of the results obtained. We deduce the paper along with future scope of research in Section 5.

1.1 Contribution

- We present an improved EfficientNetB3 model that aids in diagnosis of malignant melanoma with better accuracy.
- We perform a comparative analysis of existing pre-trained models and investigate the model performance on malignant melanoma classification task.
- We employ the concept of fine tuning by adjusting the learned weights of EfficientNetB3 model on our malignant melanoma skin lesion classification task and add custom layers of data augmentation, global average pooling (GAP) layer, dropout layer and dense softmax classification layer during the training and testing phases to achieve better accuracy.
- The proposed model provides promising results to handle the problem of model over-fitting which is usually observed in pre-trained deep neural models.

2. RELATED WORKS

Zabir et al. [5] presented an automatic segmentation and classification model built on deep learning with transfer learning. To scale up training data, different augmentation methods are applied on the training dataset. The authors carry out semantic segmentation using U-Net model on augmented images and further apply various deep convolution neural networks to extract features, such as VGG16, VGG19, ResNet50, InceptionResNetV2, DensetNet201, InceptionV3 …etc. and compare classification results with several classifiers, such as SVM, Random Forest, Decision Tree and AdaBoost. The proposed algorithm achieved an accuracy of 0.92 with DenseNet201 used as feature extraction model and SVM as classification model.

Nils et al. [6] submitted an ensemble-based deep neural network model for lesion classification. They applied cropping techniques to address the problem of images with different resolutions and loss balancing approach to handle the class imbalance dataset. Multi-resolution EfficientNet model coupled with comprehensive data augmentation were researched and an accuracy of 0.928 was achieved for melanoma lesion classification.

Hasib et al. [7] proposed a two-stage unified deep learning framework for automatic skin cancer image classification using adversarial training fused with transfer learning to comprehensively handle inter-class diversity and imbalance class dataset. The classifier is trained by continually decreasing the focal loss function that supports the model in learning. The MelaNet model achieved an F1 score value of 94%, which is better compared to other baseline models, like VGG-Gap and VGG-Gap+Augment.

Jason et al. [8] proposed a melanoma classification deep learning-based approach combined with conventional image processing method to achieve superior results in terms of accuracy. Fusion model with cross-validation achieved a classification accuracy of 0.94, which is better than ResNet-50 classifier and traditional classifiers for image processing. AUC values of 0.87 and 0.90 were observed when ResNet-50 and a traditional image processing-based classifier were applied, respectively.

Arthur et al. [9] presented a computer based diagnostic system employing CNN framework. The proposed diagnostic system classified the lesion images into classes, such as seborrheic keratosis, nevi and melanoma. The authors applied an ensemble and aggregation method along with directed acyclic graph technique on three-class CNN in the context of improving the model performance. The proposed network accomplished an accuracy of 76.6%.

Maria et al. [10] presented a comparative analysis of various deep learning models on ISIC dataset for melanoma classification. The authors investigated the influence of pre-processing stage on varied neural network approaches, like 2D-CNN, self-organizing neural networks and ResNet. Canny edge detector along with morphological techniques, like Ostu thresholding and dilation operation, were used to identify hair artefacts. Further, image pre-processing was carried out using inpainting method to reconstruct the original image without hair artefacts. Skin lesion segmentation to highlight the lesion region was performed using bitwise AND binary operator. ResNet model and achieved an accuracy of 81.5% on ISIC dataset for melanoma classification.

Jasil et al. [11] reported a comparative study for identification of skin lesion type employing transfer learning technique, utilizing deep-learning networks. All experimental analyses were examined on ISIC dataset and pre-prcoessing techniques, like image normalization, image resizing and augmentation, were performed on datasets to raise the preciseness of neural network models. In particular, InceptionV3, VGG16 and VGG19 models were reviewed for skin lesion classification task. From the experimental evaluation, VGG16 furnished superior results with an accuracy of 77%.

3. PROPOSED METHODOLOGY

We provide a comprehensive description of our advanced method for malignant skin lesion classification in the current section. The proposed methodology consists of the following sub-sections: (1) Dataset and data preparation, (2) Data augmentation, (3) EficientNetB3 network, (4) Fine-tuning with global average pooling, (5) Fine-tuning with dropout. (6) Fine-tuning fully connected classification layer.

3.1 Dataset and Data Preparation

We have carried out our research on ISBI-ISIC 2017 dataset [4] accessible from "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC)". The dataset consists of 3297 lesion images classified into 1800 benign class skin lesion images and 1497 malignant class skin lesion images. Training and testing of the pre-trained model are carried out ISBI-ISIC 2017 dataset. The dataset was spilt in the ratio 80:20 training and testing set comprising 2637 training data images and 660 testing data images. Figure 1 depicts sample images of ISBI-ISIC 2017 dataset of malignant class. Images of skin lesion were of varied pixel sizes in the RGB color space. For the model training and testing, we resized the image into 224×224 pixels. The images of skin lesion were resized to 224×224 pixel resolution to make images compatible with EfficientNetB3 model.

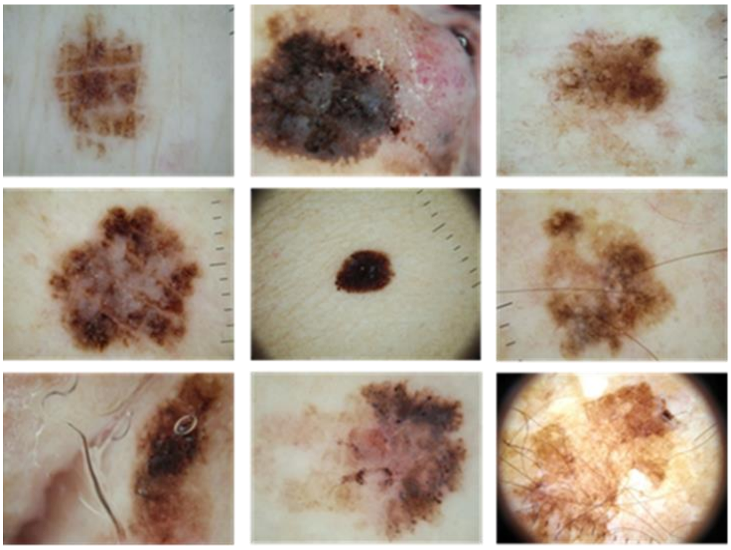


Figure 1. Skin lesion images from ISBI-ISIC dataset.

3.2 Data Augmentation

One of the significant techniques to reinforce the efficiency of neural network method, it is important to train the network with enormous wealth of data. Most of the computer vision tasks consist of lesser dataset, which results in building poor classification models. Pre-trained deep-learning models are trained on a huge dataset such as ImageNet and perform substantially well on huge datasets. To increase the size of our training data and to build an efficient melanoma skin lesion classification model, we use data augmentation technique. Data augmentation is an approach of increasing the number of copies of training data artificially by slight modification on original data without actually gathering new data. Image augmentation is one the most common techniques used to improve the performance of deep-learning models and reduce over-fitting problem.

To increase the efficacy of our, we add an augmentation layer on top of EfficientNetB3 neural network layer. To enhance the size of our training data, we create altered variants of training images that belong to the same class. Training dataset is expanded using different augmentation techniques, like horizontal random flip, rotation range, zoom range, width range and height. In our data augmentation layer, we use various data augmentation transformation techniques with specific range values, as shown in Table 1. To use data augmentation right within our model, we'll create a Keras sequential model consisting of only data pre-processing layers. We then use this sequential model within another functional model. Figure 2 indicates the data augmentation layer created using various techniques.

Table 1. Data augmentation techniques.

Data Augmentation Technique	**Description**
Random Flip	Images are flipped horizontally.
Random Rotation	Images are rotated randomly by 0.2 value.
Random Zoom	Images are Zoomed randomly by 0.2 value.
Random Height	Image height is shifted randomly by 0.2 value.
Random Width	Image width is shifted randomly by 0.2 value.

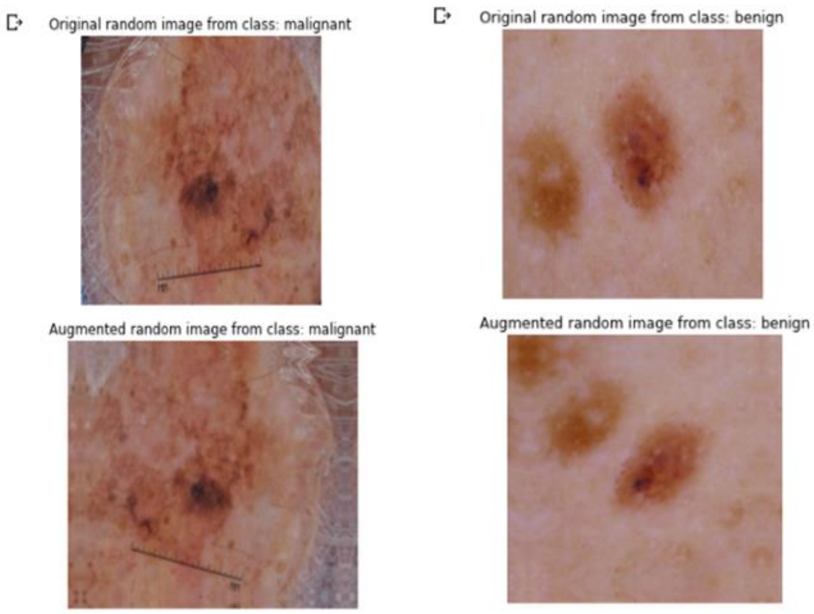


Figure 2. Random augmented images.

3.3 EfficientNetB3 Network Architecture

EfficientNet [15] consists of a family of models from B0 to B7 and is considered as one of the most computationally efficient deep-learning models trained over ImageNet. EfficientNet models are based on compound scaling method which deftly expands a baseline convolution network model size to target model size in an efficient manner, attaining top model accuracy gain. Compound scaling method allows the network to be uniformly scaled across width, depth and resolution. Figure 3 shows the compound scaling strategy in EfficientNet compared to baseline model. The EfficientNet model is comprised of different types of mobile inverted bottleneck convolution blocks MBConv with varied kernel size of 3x3 and 5x5. We expand the network depth, width and resolution uniformly using compound scaling coefficient ∅ in the following manner:

$$d=\alpha^{\emptyset},\ w=\beta^{\emptyset},\ r=\gamma^{\emptyset}\ \text{ such that } \alpha\beta^2\gamma^2 \approx 2\ ,\ \alpha\geq 1, \beta\geq 1,\ \gamma\geq 1 \qquad (1)$$

where d is depth, w is width, r is resolution and α, β, γ are constants.

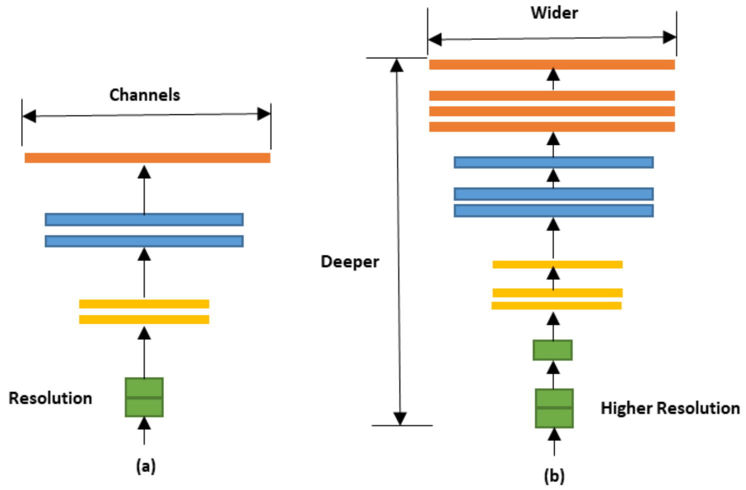


Figure 3. Scaling of EfficientNet model: (a) Baseline model and (b) Compound scaling model.

The value of $\emptyset$ is user-specified and helps scale up the network and identify computation resources available for the model. EfficientNetB0 model is built up by using values of $\emptyset$= 0, w=1, d=1, r=1 representing the baseline model. EfficientNetB0 is comprised of MBconv1 and MBcon6 layers. Likewise, EfficientNetB3 model is constructed by using values of $\emptyset$= 3, w= α^3, d= β^3, r= γ^3, indicating that more resources are available to acquire superlative performance. EfficientNetB3 model consists of more layers of MBCon6 with inverted residual connection. EfficientNetB3 model consists of deeper network compared to baseline model, which apprehends intricate and richer features and generalizes well on new missions. EfficientNetB3 consists of a wider network that can extract optimal features and patterns that are beneficial for classification task. Along with improved accuracy, EfficientNetB3 model also ameliorates efficiency by decreasing FLOPS (Floating Point Operations per Second) and parameters. Figure 4 shows EfficientNetB3 network architecture.

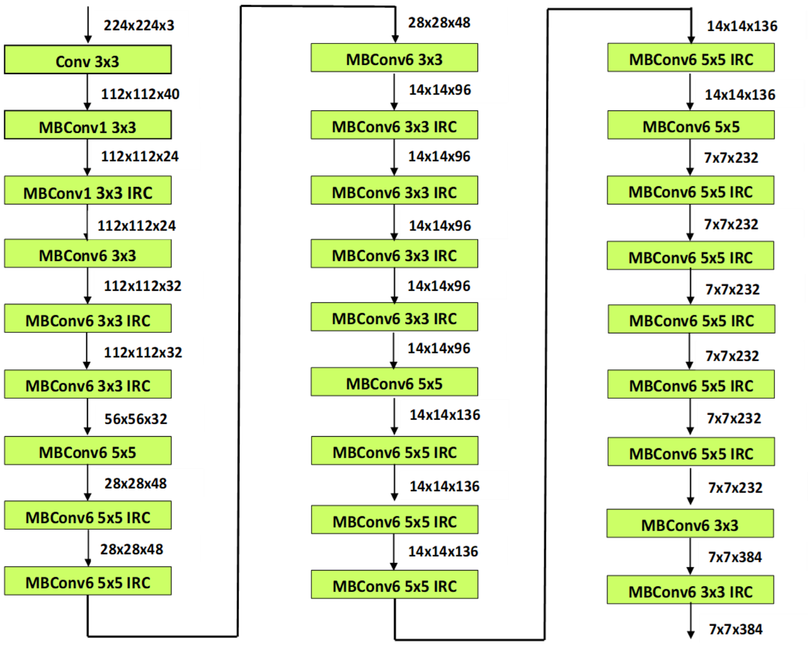


Figure 4. EfficientNetB3 network architecture.

3.4 Fine Tuning with Global Average Pooling

EfficientNetB3 model outputs a list of feature vectors or feature maps which are used to extract patterns. The motive behind using global average pooling is to generate a single feature map for each benign and malignant lesion category of the classification task. Rather than appending fully connected layers over feature maps, we take average of each feature map to generate a feature vector belonging to each class. Feature maps generated through global average pooling layer represent confidence maps of each benign and malignant category. Global average pooling layer diminishes spatial dimensionality of the feature maps, thus shrinking the count of feature parameters in the model and improving model performance by reducing the problem of model over-fitting [19]. We use tf.keras.layers.GlobalAveragePooling2D() layer to transform 4-dimensional tensor into 2-dimensional tensor by condensing the input tensor shape of (1,4,4,3) to (1,3).

3.5 Fine Tuning with Dropout Layer

To train the EfficientNetB3 model faster and to avert the neural network from over-fitting, we add a dropout layer. Due to its large architecture, EfficientNetB3 model produces a large number of parameters which are often slow to use during training time. With dropout technique, one can overcome this problem by randomly dropping units from the neural network during training, thus thinning the neural network. By using dropout layer, it becomes easier for the model during the testing phase to estimate the averaging results of all these thinned network predictions by utilizing unthinned network having smaller weights [18]. We choose 0.2 value of probability p of dropout in our proposed model to achieve a higher accuracy and reduce the over-fitting problem from the proposed neural network. In this study, we show that dropout layer improves the performance of EfficientNetB3 neural network for melanoma skin lesion classification task.

3.6 Fine-tuning Fully Connected Classification Layer

In this stage, we merge together all the feature vectors that we have received from the previous stages. We pass the feature vectors received from the dropout layer further into the network for classification as an output layer. The top layer of baseline EfficientNetB3 pre-trained model outputs 1000 classes, because it was pre-trained on ImageNet. To adjust to our two-class skin lesion classification problem, we fine-tune the pre-trained network to adapt to only two classes of malignant and benign skin lesions. In order to calculate the class probabilities of each input, we have used softmax layer as an output layer for EfficientNetB3 model classifier that works on our target dataset of two classes of skin lesions. The real and predicted value difference is calculated using loss function. We build an output activation layer using tf.keras.layers.Dense() and meld the inputs with the output using the model tf.keras.Model(). Figure 5 shows the architecture of our proposed methodology.

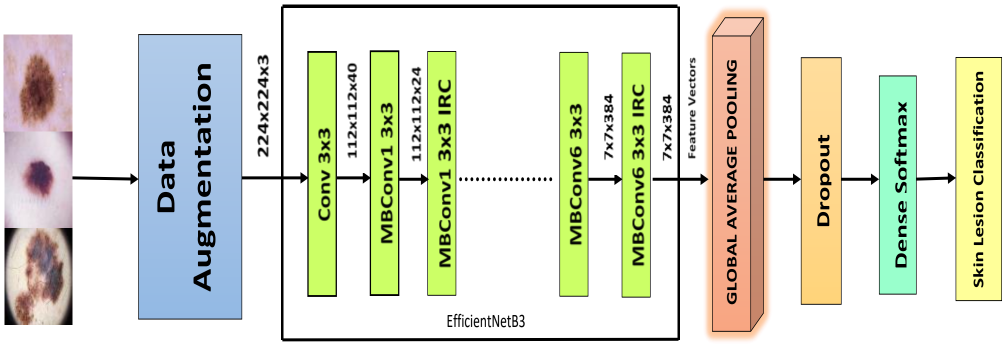


Figure 5. Proposed methodology.

4. Experimental Results and Discussion

To check the efficiency of our model, we carried out our experimental analysis in 3 approaches. We first compare the fine-tuned proposed EfficeinetNetB3 model with the pre-trained base EfficientNetB3 model on ISBI-ISIC 2017 dataset. In the second approach, we perform comparative analysis of the proposed model with present state-of-the-art pre-trained classification methods, such as ResNet50, InceptionV3, InceptionResNetV2 and EfficientNetB0-B2 on ISBI-ISIC 2017 dataset. In the third approach, we carry out comparative analysis of our improved model with models proposed by other authors.

4.1 Experimental Specification

We have performed all our experiments on Google Colab Python notebook which provided access to free 12Gb Tesla K80 NVIDIA GPU of SMI 470.74. We have used Adam as our optimization algorithm with a rate of learning of 0.001 to compile the models. All the methods are trained with a count of 35 epochs with 32 batch size. We investigated the efficiency of our proposed model by computing recall, F1 score, precision and accuracy. The metrics are computed on True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) cases.

True Positive (TP)- correct label is positive and predicted label is positive.
False Positive (FP)- correct label is negative and predicted label is positive.
True Negative (TN)- correct label is negative and predicted label is negative.
False Negative (FN)- correct label is positive and predicted label is negative.

Accuracy: It is an evaluation metric that finds the model performance across all classes. It is the fraction of the sum of correct predictions to the sum of total predictions.

$$\text{Accuracy} = \frac{((TN+TP))}{((TN+TP+FN+FP))} \tag{2}$$

Precision: It is an evaluation metric that calculates the fraction of the total positive samples over the sum of total positive samples either classified precisely or imprecisely.

$$\text{Precision} = \frac{TP}{((TP+FP))} \tag{3}$$

Recall: It is an evaluation metric that calculates the fraction of the total positive samples over the sum of total positive input samples classified correctly.

$$\text{Recall} = \frac{TP}{((TP+FN))} \tag{4}$$

F1-score: It is an evaluation metric that finds the accuracy of the mode by combining both precision and recall by giving more weightage to false positives and false negatives.

$$\text{F1} - \text{score} = \frac{2TP}{((2TP+FP+FN))} \tag{5}$$

4.2 Comparison of Proposed Model with Base Pre-trained EfficientNetB3

In our proposed model, we incorporated data augmentation technique and additional layers of global average pooling (GAP), dropout and dense softmax classification to increase the exactness of skin lesion classification results. We compare our proposed model with baseline EfficientNetB3 model without data augmentation, global average pooling layer and dropout layer. The final classification layer of base pre-trained EfficientNetB3 model is replaced by dense softmax layer with the correct number of output classes (two class) to adjust the model to our classification task. From Table 2, it is observed that baseline EfficientNetB3 model performed much less with an accuracy of 56.97%, precision, F1-score and recall of 58.00%. Our proposed model gave an accuracy of 87.12% and precision, F1-score and recall of 87.00%.

Table 2. Comparative result of proposed model with baseline EfficientNetB3.

Approach	Dataset	Accuracy	Precision	Recall	F1-score
Baseline EfficientNetB3	ISBI-ISIC 2017	56.97	58.00	58.00	58.00
Proposed Model	**ISBI-ISIC 2017**	**87.12**	**87.00**	**87.00**	**87.00**

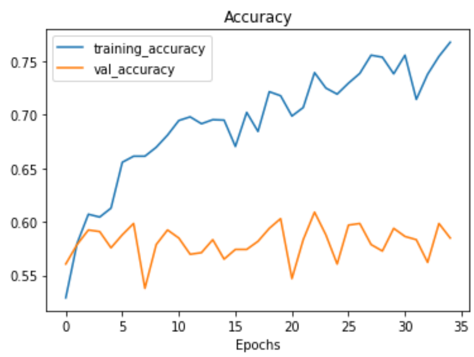


Figure 6. Accuracy curve of baseline EfficientNetB3.

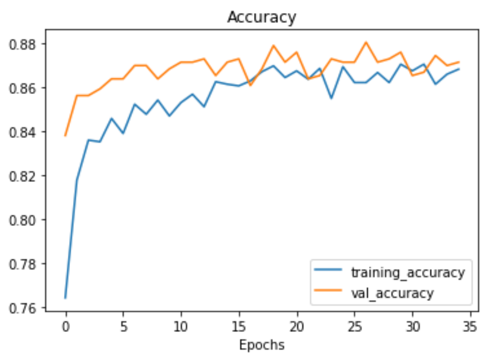


Figure 7. Accuracy curve of the proposed model.

Figure 6 shows the accuracy curve of baseline EfficientNetB3 and Figure 7 shows the accuracy curve of the proposed fine-tuned model. In Figures 6 and 7, x-axis represents the epochs and y-axis represent the accuracy value. From Figure 6, it was observed that baseline EfficientNetB3 gave inferior results and showed a huge gap between training and validation accuracy curves resulting into model over-fitting. From Figure 7, it was noticed that the proposed model provided superior results, as the gap among accuracy curves of training and validation is fairly reduced. In some cases, the training and testing accuracy curves are overlapping, which indicates the proposed model avoids over fitting problem.

4.3 Comparison of Proposed Model with Other Pre-trained Models

We further review the potency of the proposed model by comparing it with present state-of-the-art baseline pre-trained methods. We present comparative analysis of our proposed model with pre-trained models, such as ResNet50, InceptionV3, InceptionResNetV2, EfficientNetB0, EfficientNetB1 and EfficientNetB2. ResNet-50 is a variant of residual neural network with 50 layers pre-trained on ImageNet. ResNet-50 model includes five stages consisting of convolution block and identity block. At each block of convolution and identity, 3 convolutional layers are present, respectively [14]. Applying the concept of skip connections is the strength of ResNet model. Skip connection assists in reducing the vanishing gradient problem in deep neural network by skipping through some of the layers during the training. InceptionV3 is a deep-learning model pre-trained on ImageNet dataset comprising of recurrent inception modules. Inception module consists of convolution layer of 1×1, 3×3, 5×5, max pooling layer and concatenation layer [16]. InceptionResNetV2 [17] is another deep-learning model trained on ImageNet dataset. It is a hybridized model elicited from residual connection and inception modules. It consists of varied-sized convolutional filters mapped with residual connections with 164 deep layers. Concatenation layer of the inception module is missing in the InceptionResNetV2 model. Combination of inception architecture with residual connection accelerates the training speed. EfficientNetB0-B2 [15] belongs to the family of EfficientNet models based on compound scaling method. These models are considered to attain top accuracy gain and are also computationally efficient deep-learning models pre-trained on ImageNet dataset.

All these models are pre-trained on ImageNet dataset and contain features with respect to 1000 classes. To adjust these models to our ISBI-ISIC 2017 dataset, we use softmax layer as the last classification network layer with two classes. Only the top layer of these pre-trained models is adjusted and the rest of layers remain frozen. All these pre-trained models are trained with a count of 35 epochs with 32 batch size. Adam optimization algorithm with a rate of learning rate of 0.001 is used to compile all the models. We evaluate their performances on the ISBI-ISIC 2017 dataset for melanoma skin lesion classification task. From Table 3, it is observed that ResNet50 achieved the second best result and provided an accuracy of 84.85%, precision, recall and F1-score of 85.00%. Figure 8 indicates the accuracy curve of ResNet50. InceptionV3 and InceptionResNetV2 provided accuracy values of 82.73% and 83.33%, respectively. Figure 9 shows the accuracy curve of InceptionResNetV2. In Figures 8 and 9, x-axis represents the epochs and y-axis represents the accuracy value. EfficientNetB0, EfficientNetB1 and EfficientNetB2 gave accuracy values of only 60.15%, 55.75% and 58.48%, respectively. The proposed model outperformed present state-of-the-art deep-learning methods on ISBI-ISIC 2017 dataset with highest F1-score of 87.00%. To investigate the efficacy of the proposed model, we analyze the confusion matrix of our model with ResNet50, InceptionV3 and InceptionResNetV2 models. Figure 10 shows the confusion matrix of the proposed model having lesser number of false negative and false positive labels as compared to other state-of-the-art advanced neural network methods ResNet50, InceptionV3 and InceptionResNetV2.

Table 3. Comparative results of the proposed model with state-of-the-art pre-trained methods.

Approach	Dataset	Accuracy	Precision	Recall	F1-score
ResNet50	ISBI-ISIC 2017	84.85	85.00	85.00	85.00
InceptionV3	ISBI-ISIC 2017	82.73	83.00	83.00	83.00
InceptionResNetV2	ISBI-ISIC 2017	83.33	83.00	83.00	83.00
EfficientNetB0	ISBI-ISIC 2017	60.15	65.00	60.00	59.00
EfficientNetB1	ISBI-ISIC 2017	55.75	55.00	56.00	55.00
EfficientNetB2	ISBI-ISIC 2017	58.48	66.00	66.00	66.00
Proposed Model	**ISBI-ISIC 2017**	**87.12**	**87.00**	**87.00**	**87.00**

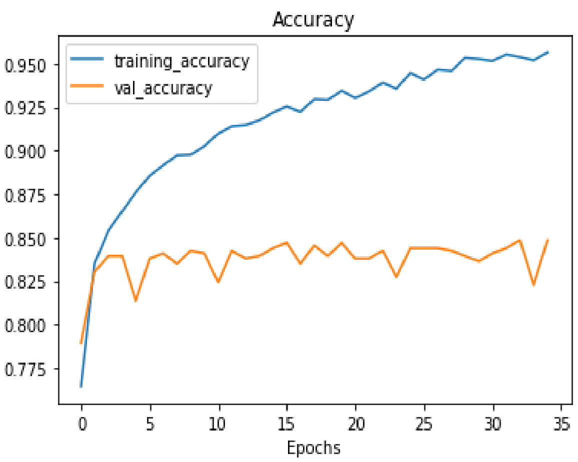


Figure 8. Accuracy curve of ResNet50.

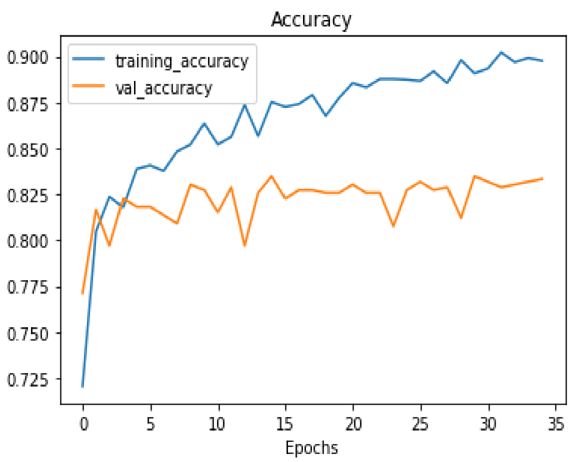
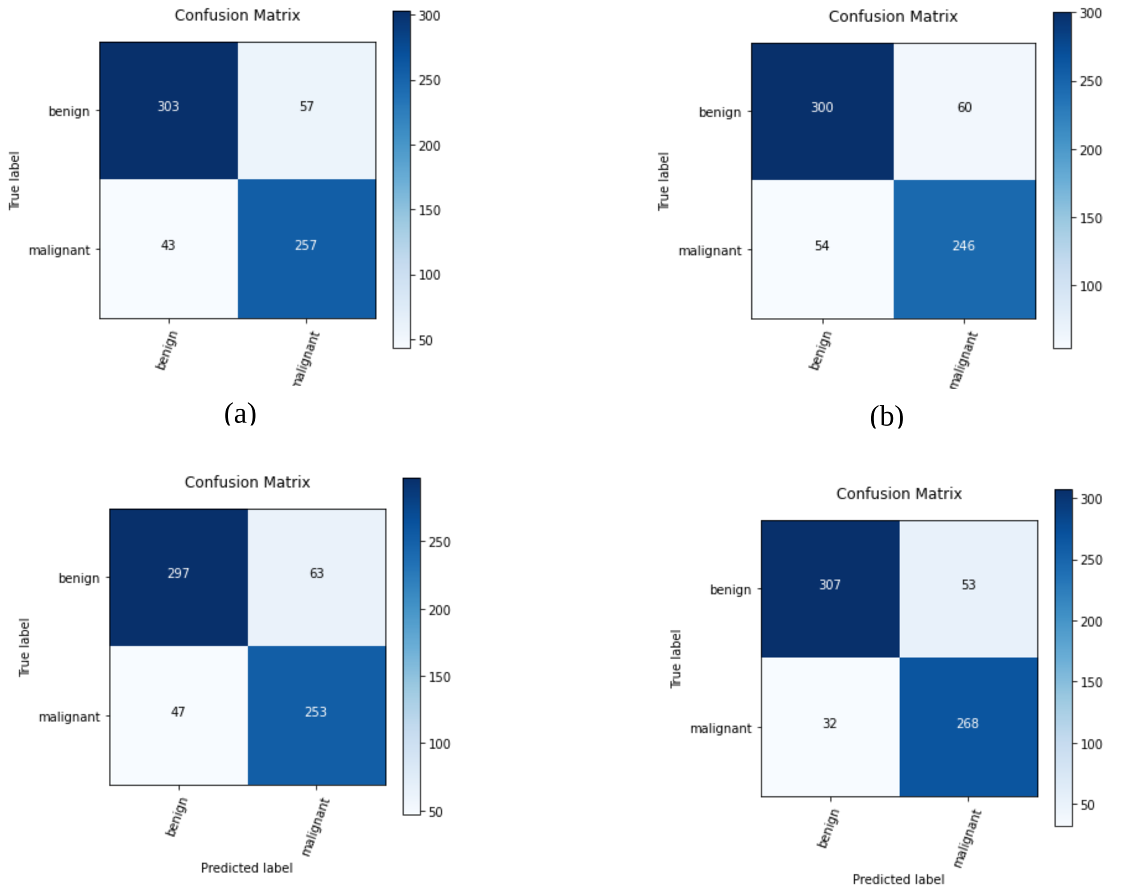


Figure 9. Accuracy curve of InceptionResNetV2.



(a) (b) (c) (d)

Figure 10. (a) Confusion matrix of ResNet50, (b) Confusion matrix of InceptionV3, (c) Confusion matrix of InceptionResNetV2 and (d) Confusion matrix of the proposed model.

4.4 Comparison of the Proposed Model with Other Approaches

In addition, we analyze the findings of the proposed model with previous works carried out on ISIC dataset. In Table 4, we summarize the evaluation outcomes of the proposed model with those of other approaches. The proposed model accomplished an accuracy of 87.12% on ISBI-ISIC 2017 dataset, whereas Hasib et al. [7] achieved an accuracy of 81.18% using MelaNet deep-learning model with adversial training and transfer learning. Arthur et al. [8] used a combination of decision-directed acyclic graph with VGG_19 CNN deep-learning model and obtained an accuracy of 76.6%. Maria et al. [10] presented a deep neural network-based melanoma classification approach along with pre-processing technique for removal of hair artefacts and lesion segmentation technique. The authors achieved an accuracy of 81.5% on ResNet model and an accuracy of 74.1% on 2D CNN model on ISBI-ISIC 2017 dataset. Jasil et al. [11] inspected the performance of deep-learning networks, like InceptionV3, VGG16 and VGG19 using pre-processing techniques, like image normalization, image resizing and augmentation on ISIC dataset. According to [11], VGG19 achieved an accuracy of 76%. It can be noted that the proposed approach furnished more accurate results compared to other approaches from Table 4.

Table 4. Comparative results of the proposed model with other approaches.

Reference	Dataset	Approach	Accuracy %
Proposed Model	**ISBI-ISIC 2017**	**Fine-tuned EfficientNetB3**	**87.12**
Hasib et al. [7]	ISBI-ISIC 2016	MelaNet	81.18
Arthur et al. [9]	ISIC-2018	DDAG VGG_19_2	76.6
Maria et al. [10]	ISBI-ISIC 2017	ResNet	81.5
		2D CNN	74.1
Jasil et al. [11]	ISIC-2018	VGG16	77
		VGG19	76
		InceptionV3	74

5. CONCLUSIONS

Melanoma skin cancer is fatal for human being and timely diagnosis of skin cancer is needed. Building a computer-aided diagnostics system for classification of melanoma skin lesion is important. In this paper, we propose an improved, comprehensive and fine-tuned EfficientNetB3 deep neural network model to categorize lesions into malignant and benign. We employ fine-tuning transfer-learning concept by utilizing network architecture of EfficientNetB3 model by freezing the initial layers of the model and then fine-tuning their weights in order to classify accurately to our two-class skin lesion classification task. We employ a range of data augmentation approaches to upgrade the correctness of the model in the training phase. We fine-tune EfficientNetB3 model, by adding customized layers of global average pooling (GAP), dropout and dense softmax classification node. In this paper, we provided comparative analysis of pre-trained deep-learning models for malignant melanoma classification. We reviewed the performance of ResNet50, InceptionV3, InceptionResNetV2 and EfficientNet B0-B2 pre-trained models. From the results, it was observed that these pre-trained models suffered model over-fitting problem. We tried to tackle the problem of model over-fitting by utilizing augmentation techniques and adding layers of global average pooling, dropout and dense softmax, which achieved promising results on EfficientNetB3 model. In this paper, we furnished an improved EfficientNetB3 model to categorize lesions into malignant and benign with promising results. The investigational results reflected the capability of employment of our proposed model in the task of malignant melanoma classification. Future scope of our research would be training and testing on more datasets and checking the efficiency of our proposed model aiding to build a reliable computer-aided diagnostic system for melanoma diagnosis. The study can be extended for future work by inspecting larger architectures and building deeper fine-tuned models.

REFERENCES

[1] Cancer.org, "Cancer Facts and Figures 2021," [Online], Available: https://www.cancer.org/ content/ dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2021/cancer-facts-and-figures-2021.pdf, 2021.

[2] S. Sonthalia, S. Yumeen and F. Kaliyadan, "Dermoscopy Overview and Extradiagnostic Applications," StatPearls [Internet], Treasure Island (FL): StatPearls Publishing; PMID: 30725816, [Online], Available: https://www.ncbi.nlm.nih.gov/books/NBK537131/, Jan. 2022.

[3] K. Munir, H. Elahi, A. Ayub, F. Frezza and A. Rizzi, "Cancer Diagnosis Using Deep Learning: A Bibliographic Review," Cancers (Basel), vol. 11, no. 9, p. 1235, DOI: 10.3390/cancers11091235, 2019.

[4] Isic-archive, "ISIC Challenge Datasets," [Online], Available: https://challenge.isic-archive.com/data/.

[5] A. Nazi, Zabir and Tasnim Azad Abir, "Automatic Skin Lesion Segmentation and Melanoma Detection: Transfer Learning Approach with U-Net and DCNN-SVM," Proc. Int'l conf. on Computational Intelligence, pp. 371-381. Springer, DOI: 10.1007/978-981-13-7564-4_32, Singapore, 2020.

[6] N. Gessert, M. Nielsen, M. Shaikh, R. Werner and A. Schlaefer, "Skin Lesion Classification Using Ensembles of Multi-resolution EfficientNets with Meta Data," MethodsX, vol. 7, no. 100864, p. 100864, DOI:10.1016/j.mex.2020.100864, 2020.

[7] H. Zunair and A. B. Hamza, "Melanoma Detection Using Adversarial Training and Deep Transfer Learning," Phys. Med. Biol., vol. 65, no. 13, p. 135005, DOI: 10.1088/1361-6560/ab86d3, 2020.

[8] J. R. Hagerty, R. J. Stanley, H. A. Almubarak et al., "Deep Learning and Handcrafted Method Fusion: Higher Diagnostic Accuracy for Melanoma Dermoscopy Images," IEEE Journal of Biomedical and Health Informatics, vol. 23, pp. 1385-1391, DOI: 10.1109/JBHI.2019.2891049, 2019.

[9] A. C. Foahom Gouabou, J.-L. Damoiseaux, J. Monnier, R. Iguernaissi, A. Moudafi and D. Merad, "Ensemble Method of Convolutional Neural Networks with Directed Acyclic Graph Using Dermoscopic Images: Melanoma Detection Application," Sensors (Basel), vol. 21, no. 12, p. 3999, DOI: 10.3390/s21123999, 2021.

[10] M. Frasca, M. Nappi, M. Risi, G. Tortora and A. A. Citarella, "A Comparison of Neural Network Approaches for Melanoma Classification," Proc. of the 25th IEEE International Conference on Pattern Recognition (ICPR), pp. 2110–2117, DOI:10.1109/ICPR48806.2021.9412893, Milan, Italy, 2021.

[11] S. P. G. Jasil and V. Ulagamuthalvi, "Deep Learning Architecture Using Transfer Learning for Classification of Skin Lesions," Journal of Ambient Intelligence and Humanized Computing, vol. 2021, DOI: 10.1007/s12652-021-03062-7, 2021.

[12] F. Zhuang, Q. Zhiyuan, D. Keyu, X. Dongbo, Z. Yongchun, Z. Hengshu, X. Hui and H. Qing, "A Comprehensive Survey on Transfer Learning," Proceedings of the IEEE, vol. 109, no. 1, pp. 43–76, DOI: 10.1109/JPROC.2020.3004555, 2021.

[13] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural

Networks," Commun. of the ACM, vol. 60, no. 6, pp. 84–90, DOI: 10.1145/3065386, 2017.

[14] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, DOI: 10.1109/CVPR.2016.90, 2016.

[15] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Proc. of the Int. Conf. on Machine Learning, ArXiv, vol. abs/1905.11946, pp. 6105-6114, 2019.

[16] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818–2826, DOI: 10.1109/CVPR.2016.308, Las Vegas, NV, USA, 2016.

[17] C. Szegedy, S. Ioffe, V. Vanhoucke and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," Proc. of the 31^{st} AAAI Conference on Artificial Intelligence (AAAI'17), vol. 31, no. 1, pp. 4278–4284, 2017.

[18] H. Lim, "A Study on Dropout Techniques to Reduce Overfitting in Deep Neural Networks," Proc. of Advanced Multimedia and Ubiquitous Engineering, Part of the Lecture Notes in Electrical Engineering Book Series, vol. 716, pp. 133-139, DOI: 10.1007/978-981-15-9309-3_20, 2021.

[19] T.-Y. Hsiao, Y.-C. Chang, H.-H. Chou and C.-T. Chiu, "Filter-based Deep-compression with Global Average Pooling for Convolutional Networks," Journal of Systems Architecture, vol. 95, pp. 9–18, DOI: 10.1016/j.sysarc.2019.02.008, 2019.

[20] A. El-Halees and M. Tafish, "Breast Cancer Severity Predication Using Deep Learning Techniques," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 06, no. 01, pp. 94-102, 2020.

[21] S. F. Abuowaida and H. Y. Chan, "Improved Deep Learning Architecture for Depth Estimation from Single Image," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 06, no. 04, pp. 434-445, 2020.

ملخص البحث:

يُعدّ سرطان الجلد الخبيث أحد أخطر أنواع السرطان وأوسعها إنتشاراً. وإنّ الاكتشاف المبكّر لتقرُّحات سرطان الجلد يَزيدُ من احتمال نجاة المريض. وفي السّنوات الأخيرة، فإنّ تطبيق نماذج تقوم على الشّبكات العصبية العميقة في أنظمة التّصوير التّشخيصية يُعدّ أمراً ذا فائدة للخبراء في الحقل الطبي.

في هذه الدّراسة، نقدّم نموذجاً محسَّناً دقيق الضّبط باستخدام (EfficientNetB3) لتصنيف تقرُّحات الجلد الضّارّة باستخدام مفهوم تعليم النّقل ذي الضّبط الدّقيق. كذلك نجري مقارنةً تحليليةً لنماذج مختلفة من نماذج التعلّم العميق المدرّبة مسبقاً، مثل: (ResNet50) و(InceptionV3) و(InceptionResNetV2) و(EfficientNetB0-B2). وقد أظهرت نتائج التّحليل انّ نموذج (EfficientNetB3) دقيق الضّبط من الممكن الاستفادة منه في الكشف عن سرطان الجلد وتطوير أنظمة حاسوبية للتّشخيص. وقد جرى تطبيق كلّ الإجراءات على مجموعة البيانات (ISBI-ISIC2017). ولاختبار فاعلية النّموذج المقترح، تمّت مقارنته بنموذج (EfficientNetB3) الأساسي القائم على الطّرق القائمة المستندة الى التّدريب المسبق. وقد حقّق النّموذج المقترح دقّة بلغت 87.12%، بينما كانت النّسبة 87% لكلٍّ من الاستعادة والضّبط، و 85% لدرجة F1. كذلك حقّق النّموذج المقترح في هذه الدّراسة نتائج حسابية جيدة، وعالج مشكلة الملاءمة الفائضة، ووضع حدّاً للأوسام السّالبة الخاطئة (FN).