Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Malignant and Benign Skin Cancer Classification using Modified EfficientNetB3 Deep Learning Model and LIME

Rambabu Pasumarthy, Lakshmeeswari Gondi

Research Scholar, Dept of CSE, GITAM School of Technology, GITAM Deemed to be University, Visakhapatnam, India.
Email: rpasumar2@gitam.in

Separating out the many types of skin cancer well is key to catching them early and devising effective treatment strategies. Diagnosing skin cancer early, be it melanoma or other forms is so that medical professionals can create the best treatment plan possible and give diagnosed patients an accurate prognosis. Skin cancer has several subtypes, all with unique features and treatment options. This work modifies the EfficientNetB3 model for classifying types of skin cancer (benign or malignant). The Local interpretable Model-agnostic Explanations (LIME) explainable Artificial Intelligence (AI) model is used to interpret the EfficientNetB3 model outputs. Through this study, the EfficientNetB3 with LSTM model is described which will have to strike a trade off between both accuracy of the model and complexity. Moreover, it employs compound scaling strategy for the first time to make training and testing faster with larger computational resources. LIME is a good tool to understand and interpret predictions of deep learning models especially for image classification. LIME provides easy-to-understand rationales behind each prediction, bringing out the salient features in images that influence them. LIME can be used to detect biases, particularly if the model was trained on one dataset yet is being applied it to a different test set. Regarding the classification of benign and malignant skin tumors from these images, proposed method achieved 90% accuracy rate. This is 3% more accurate than current models.

Keywords: Skin cancer, classification, benign, malignant, EfficientNetB3, LIME.

1. Introduction

Melanoma manifests as uncontrolled proliferation of skin cells and eventually becomes a mass like lump or tumor commonly seen in cases of cancer. Melanomas tend to occur in people who have a history of sun exposure. Melanoma is a prevalent kind of cancer [1]. The main cause of this type of cancer is the exposure to ultraviolet (UV) radiation, which can come from the sun or tanning beds. Approximately one-third of all cancer cases worldwide manifest as skin

cancer [2]. Individuals with fair complexion, blond or red hair, and blue or green eyes have a higher susceptibility to skin cancer. However, it is important to note that everybody, regardless of their skin color or ethnicity, is at risk [3]. Because untreated skin cancer may spread and become potentially fatal, early identification and fast treatment are very necessary for optimal treatment results [4].

There are three basic varieties of skin cancer, each of which gets its name from the kind of skin cell from which it originates. Most often, it appears as a pink or red patch of skin, a pearly, glittering lump, or an open wound that does not heal [5]. Even if it expands slowly, it almost never spreads to other areas of the body. Squamous Cell Carcinoma (SCC) manifests itself clinically as a scaly, red area, a hard, elevated lump, or an open wound that does not heal [6]. If treatment is not received, it may develop more quickly than Basal Cell Carcinoma (BCC) does and can spread to other areas of the body. Melanoma is the worst kind of skin cancer, and it begins in the melanocytes, which are the cells that are responsible for producing color in the skin. It could show up as a new mole or an existing one that changes in size, shape, or color [7].

An estimated 5 million Americans receive a skin cancer diagnosis each year, making it the most common kind of cancer diagnosed there. By the time they are 70 years old, one in five Americans will have skin cancer [8–9]. Men work outside more frequently than women do, therefore their likelihood of developing skin cancer and dying from it is probably higher. They also tend to wear sunscreen less frequently. Less than 1% of all malignancies in India are skin cancers, making them a very rare type of the illness. Even if skin cancer prevalence is rising in India, it is still decreasing there compared to western countries. The incidence of skin cancer varies greatly by area in India [10], with greater rates observed in the south along with lower rates in the north.

For a successful course of therapy and better results, early identification of skin cancer is essential [11]. For skin cancer, there are many diagnostic options. Regular skin self-examination is seen to be a crucial component of a skin cancer early detection approach, and any changes in the moles' size, shape, color, or texture should be checked out. A dermatologist thoroughly examines the patient's skin to look for any skin cancer warning signs. Every area of the patient's skin is examined throughout the examination, including areas that are seldom exposed to sunlight. A biopsy may be performed to ascertain whether any suspected lesions are malignant. A series of digital photographs of the whole skin surface are taken during the mole mapping procedure to capture both the appearance and location of moles and other skin abnormalities. These images may be used to spot any new or developing lesions as well as track how changes change over time.

Deep learning models have the capacity to analyze medical pictures in a highly accurate and effective manner, as shown [12,13]. To recognize characteristics that are indicative of skin cancer, these models analyze photographs of skin lesions. Convolutional neural network (CNNs) are built with the purpose of recognizing patterns in pictures via the examination of several layers of the image and the extraction of characteristics that are pertinent to the classification process. CNNs may be taught to identify skin cancer by being shown enormous datasets of photos of benign and malignant skin lesions. CNNs with this ability are able to understand the little differences between the two types of skin lesions. It is important to

remember that even while deep learning models have shown promise in detecting skin cancer, a dermatologist's clinical examination should always come first. Alternatively, by aiding in the early detection of skin cancer, they might be utilized as an additional resource to improve patient outcomes. In this paper, a novel deep learning-based model for skin cancer diagnosis is proposed. An LSTM and an EfficientnetB3 model have been developed in order to accurately forecast skin cancer. The organization of the paper as follows: The section-II describes the literature survey, and the proposed model is explained in section-III. The simulation results are discussed in section-IV.

2. Literature

The study in [14] aims to create a CNN framework that can classify different types of skin cancer and help detect them early. The scripting programming languages will be used to build the segmentation model, with different network designs tested using various types of layers, including Convolutional, Dropout, Pooling, and Dense layers. Transfer Learning was also employed to achieve early agreement.

Vijayalakshmi et al [15] proposed a system for the identification of dermatological diseases using photographs of lesions. This machine intervention contrasted with the traditional method of medical diagnosis, which is dependent on the examination of a patient by a medical professional. The development of the suggested framework is broken down into three stages: the first is data gathering and augmentation; the second is model design; and the third and final stage is prediction. Authors formed a better structure by combining several AI algorithms, with image processing tools, which resulted in a greater accuracy of 85%. This was accomplished via the application of different AI algorithms.

Teck Yan Tan et.al [16] proposed an intelligent skin cancer diagnosis decision support system. Lesion classification relies on constructing an effective lesion representation, which exploits the discriminative potential of distinct characteristics. Two modified PSO models for feature optimization are then shown. The first model solves stagnation through re-initialization, many distant leaders, customizable acceleration coefficients, and depth sub-dimension feature search. The second model enhances both variety and intensity by using non-adaptive random acceleration coefficients rather than adaptive ones. These coefficients are derived from non-linear circle, sine, and helix functions. The development of ensemble classifiers involves training each base model with a distinct optimum feature subset in order to construct the final product. PSO models are used in order to make adjustments to the hyper-parameters of a deep convolutional neural network. Experimental experiments employing ALL-IDB2 image data, dermoscopic skin lesion data, medical data from the UCI machine learning repository, and data on the efficacy of models are carried out in detail.

The paper by Pradhumn Agrahari et.al [17] suggested a very accurate method for identifying skin lesions that is based on a pre-trained MobileNet framework for the diagnosis of skin cancer. This strategy may enhance clinical advancement in the medical field. According to the research, a multiclass skin cancer detection algorithm that is as effective as a dermatologist can be created. The model has top-3 accuracy of 96.26%, top-2 accuracy of 91.25%, and a categorical accuracy of up to 80.81% in detecting skin lesions. This is a fast and adaptable

approach that could result in flawless clinical development, expanding the reach and breadth of primary healthcare practice, among other things.

Hatice Catal Reis et.al [18] presented a convolutional neural network based on deep learning to identify benign and malignant lesions. The International Skin Imaging Collaboration HAM10000 pictures from ISIC 2018, 2019, and 2020 are employed to measure the method's performance. Computational time and accuracy of the offered method were compared to those of the currently used machine learning techniques. In the relevant ISIC 2018, 2019, and 2020 datasets, the outcomes show that the developed model surpasses the other methodologies, with high accuracy levels. In addition to conventional approaches, deep learning algorithms may provide solid findings since they remove the human aspect from diagnosis.

In research study [19], Chaturvedi et.al suggest an automated computer-aided diagnostic approach for the highly accurate categorization of several classes of skin cancer (MCS). For MCS cancer classification, the proposed strategy achieved better than both experienced dermatologists and existing deep learning methods. To compare the efficacy of five pre-trained CNNs and four ensemble models, we fine-tuned the HAM10000 dataset across seven classes. In this research, the greatest accuracy of 92.83% for the ensemble model and 93.20% for the individual model among the set of models are presented. Because of ResNeXt101's optimized design and capacity for greater accuracy, we recommend using it for the MCS cancer classification.

Ali et al [20] presented a deep CNN (DCNN) model to correctly identify benign and malignant skin lesions. Before preprocessing it is necessary to filter out noise and artifacts with a kernel. The data augmentation step preprocess the input images and extract features to help in accelerating classification are the steps second stage. The third step is using data augmentation, which helps the model increase images dataset and possibly improving classification accuracy. The comparison of the proposed model to other transfer learning techniques, is carried out. Assessment were accomplished with the HAM10000 dataset, which led to commendable training and testing accuracy scores of 93.16%,and 91.93%, correspondingly. The models were unequivocally found to have both more dependability and even greater longevity not using the classical transfer learning methods. This proved very well the utility and long-term survivability of the DCNN model with this dataset.

3. Proposed Model

In this section, the LIME architecture for prediction and the EfficientNetB3 cancer detection framework are presented. Both ideas are related to cancer detection. Here is an example of the blocks that make up the EfficientNetB3 model, which are all composed of activation functions, batch normalization, convolutional layers and pooling layer. It is the building block of the EfficientNetB3 model. Its special trait is the compound scaling method it utilizes for better performance. At the same time, increasing both depth and breadth of a network at fixed resolution is what this technology achieves by balancing efficiency with accuracy. Several modifications to the original EfficientNetB3 model are made for tuning up the updated EfficientNETB3 model. A large number of changes has been done with network architecture like adding layers. This update may increase the number of convolutional layers or filters in

each layer so that it can enhance feature extraction as well capture more complex patterns. Another modification is the skip connections or residual connections between some levels or blocks. These connections allow for the easier flow of information and help alleviate vanishing gradient problem by means such as easier path to backpropagation. By adding these links the other way around, this model will benefit more from reinforcement learning and perform better overall than one without them. Techniques such as regularization are used in trained EfficientNetB3 models for better generalization and prevent overfitting. Examples of these are weight decay which adds a penalty term to the loss function and dropout where some random fraction of neurons is turned off during training. These regularization methods are useful for the model to learn better representations and avoid being overly specifically over-fitting. The modified EfficientNetB3 has the following characteristics:

- Stem convolutional layer: The model starts with a stem convolutional layer that procedures the input image.
- Repeated blocks: The suggested approach consists of multiple blocks that are repeated throughout the network.
- Depth and width scaling: The quantity of layers and the quantity of filters in each layer are scaled proportionally to the network's depth and width.
- Resolution scaling: The input image's resolution is scaled according to the network's depth and width.
- Squeeze-and-excitation (SE) blocks: The model uses SE blocks to improve the feature representation of the network.
- Classification layer: The final layer of the model is a fully connected layer that performs the classification task.

3.1 Proposed EfficientNetB3 model

EfficientNetB3 architecture was designed using a wide range of layers: From an activation to batch normalization, convolutional and pooling layer. The convolutional layers of the model make use a combination of pointwise and depth wise convolutions in order to balance accuracy with base usage. By normalising the inputs to each layer using batch normalization layers thus aiding in stabilizing the training process. While at the same time, those nonlinear activation layers transform it into a non-linear model. The noteworthy structure of the EfficientNetB3 approach for skin cancer images diagnosis is shown in Figure 1. The model is a stack of reusable blocks, each having layers mentioned above as parts. EfficientNetB3 also uses compound scaling to improve the model performance by simultaneously scale up its width, depth, and resolution. This method of scaling balances the trade-off between accuracy and computational efficiency.

EfficientNetB3, a model architecuture that involves several layers to learn features from images input and make prediction. The following is a comprehensive view of the layers used in EfficientNetB3 model:

- Input Layer: The EfficientNetB3 model's input layer accepts a three-channel RGB picture that is 224x224 pixels in size.

- Convolutional Layers: The model starts with a first layer of 3×3 convolutional stride-2 which is the cradle for the family of convolutional layers. A series of seven blocks will be introduced as a result, with each block containing both pointwise and depth-wise separated convolutional layers. As lambda is unknown, this will increase the number of channels with pointwise convolutions. Thus, at any given time step this has created overall fewer parameters in the framework but as good performance as it should.

- Batch Normalization: Following every convolutional layer is a batch normalization layer. This layer helps to speed up the training process while also normalizing the input to the layer.

- Activation Layers: The activation layers serve to enhance the model's performance by introducing non-linearity into the equations that it uses. The old ReLU activation function is replaced in the EfficientNetB3 model with the Swish activation function, which provides a more seamless experience than its predecessor.

- Pooling Layers: In order to reduce computational complexity along with downsample the feature maps, the EfficientNetB3 approach combines max as well as average pooling. Whereas the max pooling layer determines the value that is maximum over all feature maps, the average pooling layer computes the value that is averaged over all feature maps.

- Fully Connected Layers: The fully connected layers are responsible for making the final predictions.

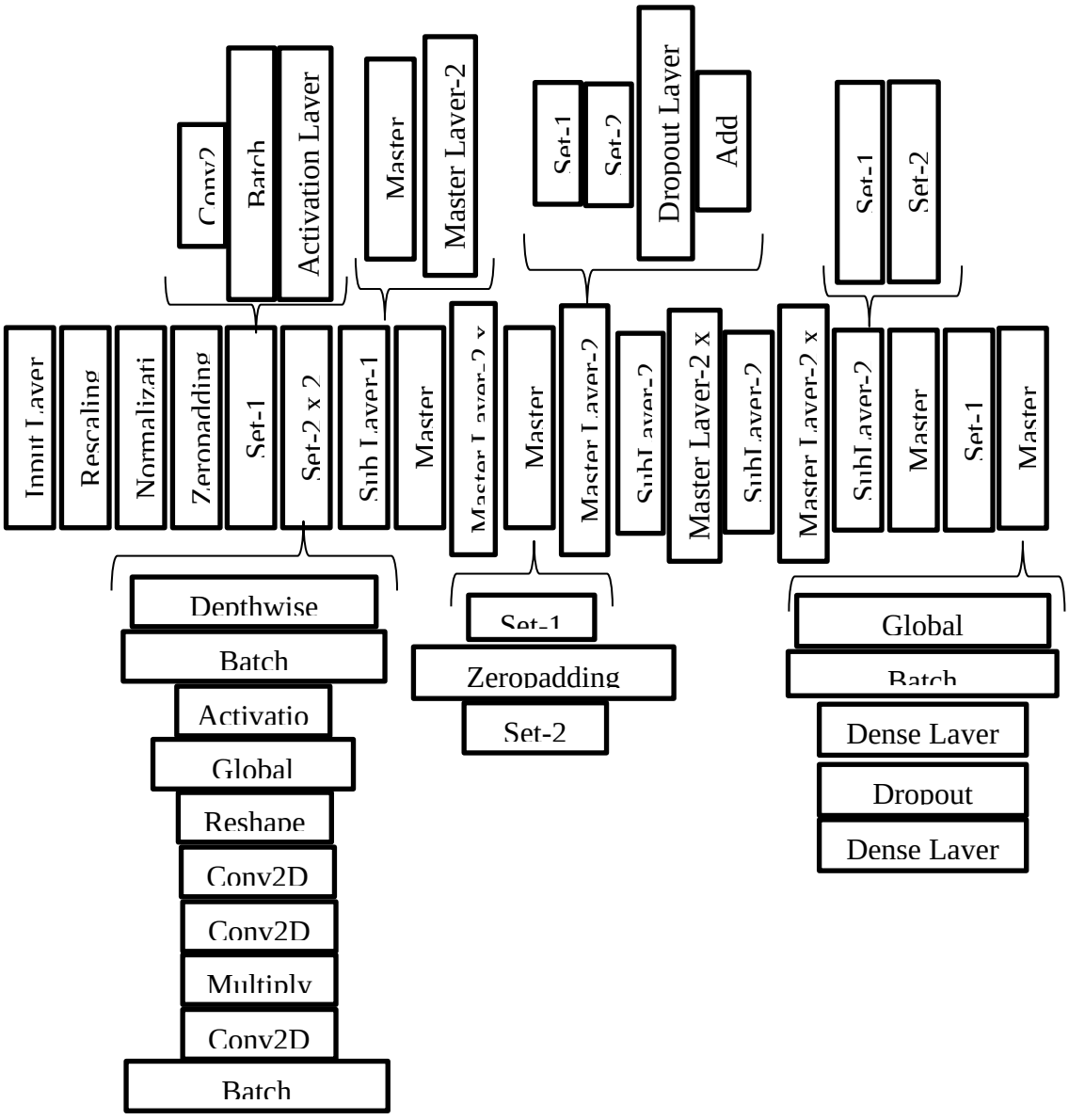


Figure 1: Proposed EfficientNetB3 model Architecture

Overall, the EfficientNetB3 model architecture is designed to be efficient and effective, using a combination of layers which depicted in Fig.1, to process input images and make accurate predictions.

3.2 LIME Architecture

The LIME is a model explanation approach that is used for the purpose of providing human-interpretable explanations for the predictions that are generated by machine learning models. The objective of LIME is to offer an elucidation for the outcomes anticipated by any "black box" model, which refers to a model with obscure or unknown internal mechanisms. The

structure of LIME comprises the subsequent stages.:

- Sampling: To explain the prediction made by the black box model, LIME first generates a set of perturbed data points by sampling around the instance to be explained. This is done by randomly perturbing the feature values of the instance while keeping the label fixed. The perturbed instances are then employed to train a new model.

- Model Training: It trains a local model through LIME which basically approximates the black box model behavior in region of instance to be explained. This is done by training a linear model with the modified examples that are generated in step 1. The linear model is trained to narrow the gap between predictions of a black box and local model.

- Feature Importance: Once the local model is trained, LIME calculates how import each feature was for that particular prediction by Black Box Model. This is achieved by calculating weights of the trained linear model for each feature.

- Explanation: Finally, LIME provides an explanation for the prediction of black-box model where it highlights features that are extremely important to get a prediction. To do this, the user can be presented with a plot or table that gives satiable weights on each feature by the local model.

The LIME method is designed to generate explanations for any model predictions, even from so-called "black-box" models. Given an input instance x, the algorithm proceeds to generate a set of similar samples and then constructing training local dataset around x. A technique used by this algorithm, LIME, approximates how complex black-box models behave in nearby that specific point represented as x. This is done by training a smaller, understandable model on a local dataset with similar k samples that are x. These are obtained by taking random perturbations of x based on the individual feature distributions within the respective set. This step produces examples which are very similar to x except for differences in feature values. It calculates the similarity between each sample and x based on distance metric Euclidean distance. The k most similar ones are then selected and put together as the local dataset D. with their feature values, along with black-box model f outputs in labels. An interpretable model g is then trained on this local dataset D, where the outputs for f are used as true labels during training. Then the interpretable model coefficients are used to explain the black-box models prediction for just this instance x without of sample test data. The aim here is letting an accurate approximate explanation learn so that to generate a new low-degree model capable enough mimicking nearby at this time referencing x.

The algorithm steps involved in LIME architecture is as follows:

Input:
- x: instance to explain (e.g. a row of a dataset)
- f: black-box model to explain
- g: interpretable model (e.g. linear regression) to explain f
- k: number of samples to generate
- r: radius of the local neighborhood to explain

Output:
- Explanation of the prediction for the input instance x

Step 1. Generate k samples by perturbing the input instance x:
- Sample a perturbation from a normal distribution centered at 0 with a standard deviation of 1 for each feature.
- Scale each perturbation by the range of the corresponding feature in the dataset.
- Add the perturbations to the input instance x to obtain the k samples.

Step 2. Compute the similarity between each sample and the input instance x using a distance metric (e.g. euclidean distance).

Step 3. Select the k-nearest samples to the input instance x and construct a local dataset D.

Step 4. Compute the outputs of the black-box model f for each sample in D.

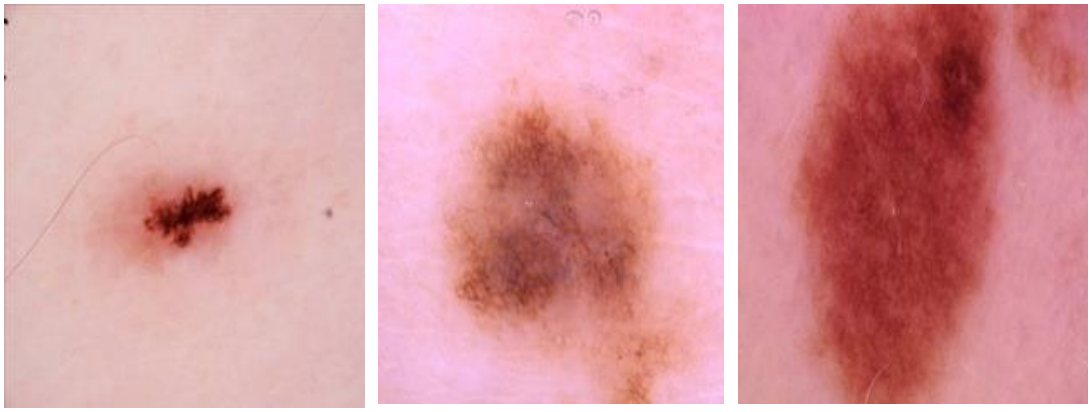
Step 5. Train the interpretable model g on D, using outputs of f as labels.

Step 6. Calculate the coefficients of the interpretable model g to interpret a local prediction of the black-box model f over the input instance x, compute corresponding explanations from the explained models .

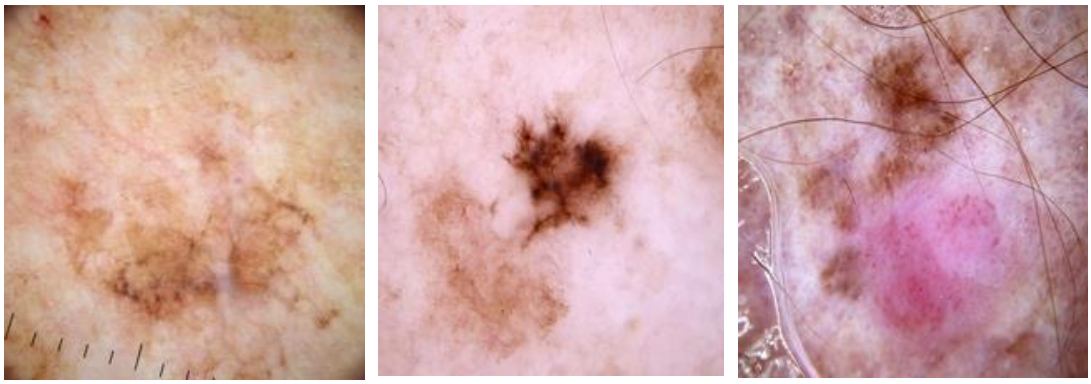
The coefficients of g are calculated and utilized to interpret f prediction for x, shedding light on factors which are heavily-used in the forecast justification as how these features effectively tilt a model output.

4. Simulation Results

In this section, the simulation results of the proposed model is discussed. The dataset includes photos of skin moles that are both benign and malignant, providing a balanced representation of skin mole types. The dataset balancing is done by using Synthetic Minority Oversampling Technique (SMOTE). The dataset [21] is organized into two folders, each of which has 1800 photos (224x244) depicting one of the two classes of moles. Fig.2 illustrates some of the sample images that are included in the dataset.



(a) benigin



(b) malignant

Figure 2: Input images

Malignant pictures are represented in Fig.2.(b), whereas benign images are depicted in Fig.2.(a) of the same figure. To understand patterns and characteristics that are linked with malignant tumors, the EfficientNetB3 architecture is often trained on a large dataset consisting of photographs of skin cancer. Based on these previously learnt patterns, the model may then be used to categorize newly acquired pictures as either benign or malignant.

4.1 EfficientNetB3

To utilize the skin images for training the EfficientNetB3 model for skin cancer detection, several preprocessing steps are required. The EfficientNetB3 architecture is then selected as the model for this task, which is fine-tuned using transfer learning. This involves freezing the early layers' weights that have learned general image features and training only the later layers on the skin cancer dataset to learn the specific features essential for skin cancer detection. The model is trained on the skin cancer dataset using Adamax optimizer and categorical crossentropy loss function. Figure 3 shows validation loss and accuracy metrics.

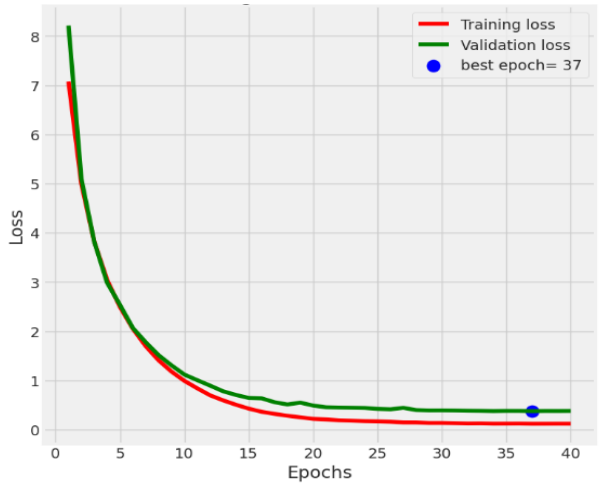


Figure 3: Training and Validation loss graphs

The proposed EfficientNetB3 model's training and validation loss is depicted in Fig.3, which is a standard way of visualizing a model's performance during training. The figure's y-axis depicts the loss, while the x-axis shows the number of epochs that have passed. To train the model, the loss function, which evaluates the amount by which the model's predictions deviate from the actual labels of the skin pictures, is optimized in order to reach the highest potential level of performance. Fig. 4 shows the Accuracy of training along with validation plot.

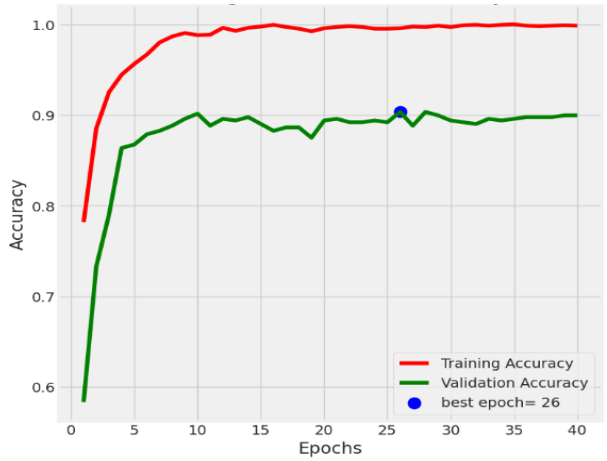


Figure 4: The Training and Validation Accuracy graphs

The accuracy plots for the EfficientNetB3 model represent how well the model can classify images during the training and validation phases. Confusion matrix is presented in Fig.5.

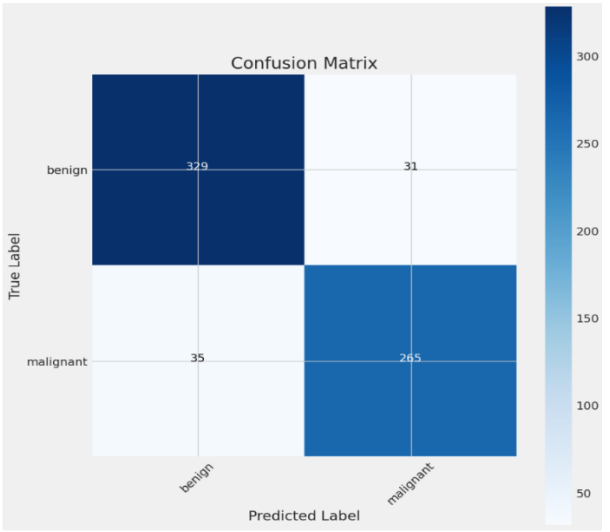


Figure 5: Confusion matrix of the proposed model

The efficiency of a framework for classification is assessed using a table called a confusion matrix, as shown in Figure 5. The study provides a thorough examination of the model's predictions, as well as the actual class labels of the data. The table is organized such that the rows represent the actual classes, while the columns represent the classes that were anticipated. In this matrix, each column reflects the number of samples that were assigned a certain combination of actual and predicted class labels.

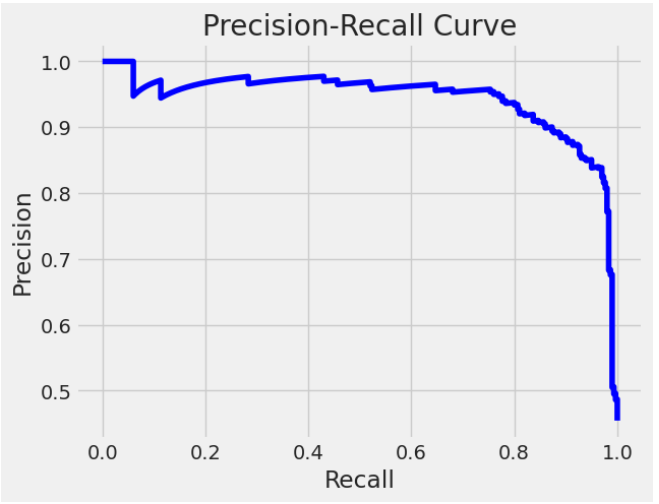


Figure 6: Precision-Recall Curve of the Proposed Model

Figure 6 shows how the precision changes for different threshold levels given recall. A Precision Recall Curve for the model demonstrates several performance aspects which includes precision, i.e. a fraction of true positive predictions to all predicted positives and recall where it is a from amongst the True Positive predictions against actual total Positives.

The curve consistently holds a very high precision value throughout most of the recall values, showing that proposed model has quite successfully identified true positive without degrading its precision.

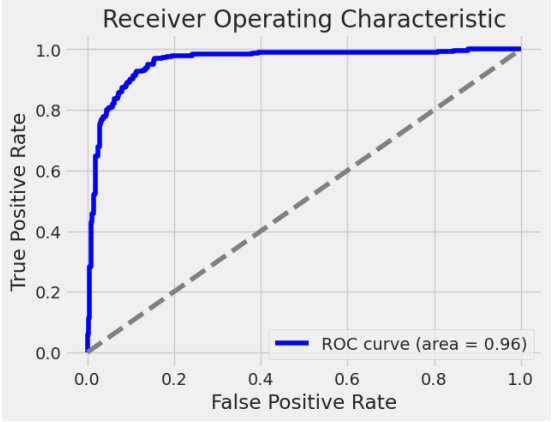


Figure 7: RoC Curve of the Proposed Model

Figure 7 shows the ROC Curve of the proposed model showing how well the models distinguish between positive and negative cases in different thresholds. ROC curve represents the trade-off between true positive rate and false positive rate for different possible cut-offs of a diagnostic test. The curve is very near to the top left corner, which means a high true positive rate and low false positives. This performance is reported using the Area Under the Curve (AUC), which in this case 0.96. The 0.96 AUC means the model has very high prediction, i.e it predicts true call vs false positives extremely well. In figure 8, the Calibration Curve of proposed model indicates how accurate to prediction probabilities based on a plot between true probabilities and predicted ones. In an effective model, this curve will hug the diagonal line meaning that proposed predicted probabilities are close to equal with true event possibilities. The blue line in the curve above is calibration of proposed model.

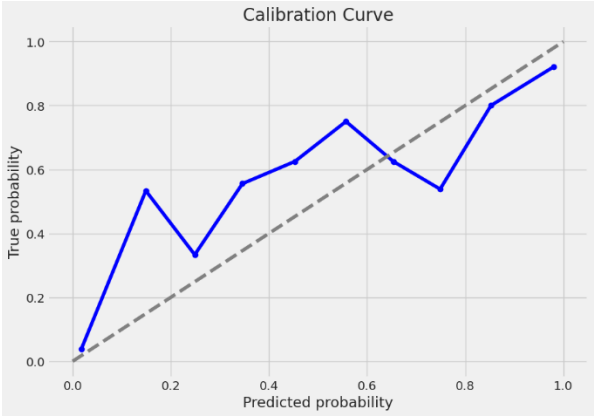


Figure 8: Calibration Curve of the Proposed Model

On the calibration curve, since blue line remains close to the diagonal, the assessment appears to be quite calibrated and predicted probabilities appear to correlate with its true values. This

evidence on the calibration curve also implies that the model provides adequate assessment on variance of probabilities, but they could be further enhanced to achieve more precise probability estimation across all range of predicted probabilities. Table 1 presents the metric results of the suggested approach.

Table1: Performance metrics of the Proposed model

	Precision	Recall	F1-score
benign	90%	91%	91%
malignant	90%	88%	89%
Accuracy	90%		
Macro average	90%	90%	90%
Weighted average	90%	90%	90%

Precision is the measure of the proportion of successfully predicted positive cases, known as "true positives," to the overall number of positive examples predicted by the model, which includes both "true positives" along with "false positives." A model with a high degree of accuracy will result in a reduced number of "false positive" predictions. Recall, on the other hand, is a technique that computes the proportion of correctly identified positive instances out of the total number of actual positive cases (i.e., true positives + false negatives). The formula for measuring it is 2 × (precision × recall) / (precision + recall). When there is a disparity in the dataset, with one class having a greater number of samples than the other, the F1-score can be a useful indication to employ.

4.2 LIME model

In the LIME model, initially select the skin cancer image to explain the prediction. Then a set of perturbations of the image is generated by adding or removing pixels, changing colors, or applying other transformations. Each perturbation will create a new instance of the data. After that the proposed EfficientNetB3 model is used to make a prediction for each perturbation instance. For each perturbation instance, create an interpretable model that explains the prediction made by the proposed EfficientNetB3 model. This can be a linear model, a decision tree, or another simple model that is easy to understand. Then after feature importance scores are calculated for each feature of the original image, based on the weights of the interpretable models. The feature significance scores are utilized to emphasize the most significant regions of the picture that influenced the prediction provided by the EfficientNetB3 approach. The input image is used for LIME is shown in Fig.9.

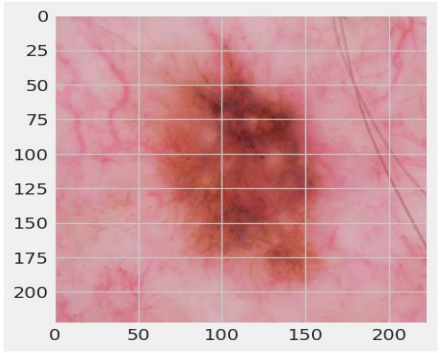


Figure 9: Input Image used in LIME model

Prediction confidence is utilized to assess the accuracy of the explanation in LIME. The prediction confidences assigned by the original model to the perturbed instances are often used to assess how well the interpretable model approximates the original model. If the interpretable model accurately captures the original model's behavior, then it should assign similar prediction confidences to the perturbed instances as the original model did. The prediction confidence of the LIME is depicted in Fig.10.

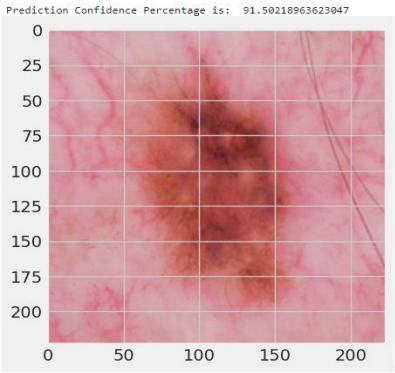


Figure 10: The prediction confidence of LIME

Prediction confidence % is a quantitative measure employed in the field of machine learning to express the degree of certainty that a model possesses on its prediction. The term "estimated probability" refers to the likelihood that the anticipated outcome is accurate, taking into account the input data as well as the model's training. The prediction confidence percentage in this study is 91.50%, indicating that the model has a 91.50% level of certainty that the picture exhibits a malignant condition. The segmented and final outcome of the LIME model is shown in Fig.11 and 12.

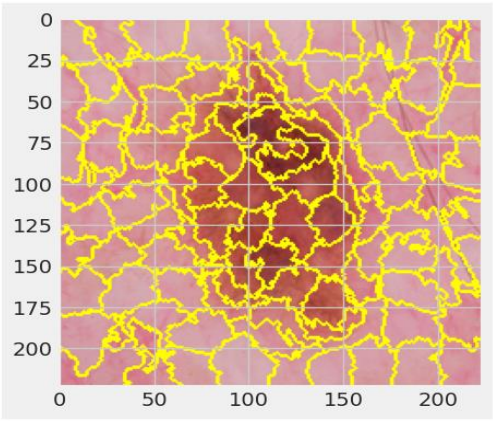


Figure 11: Segmented representation in LIME model

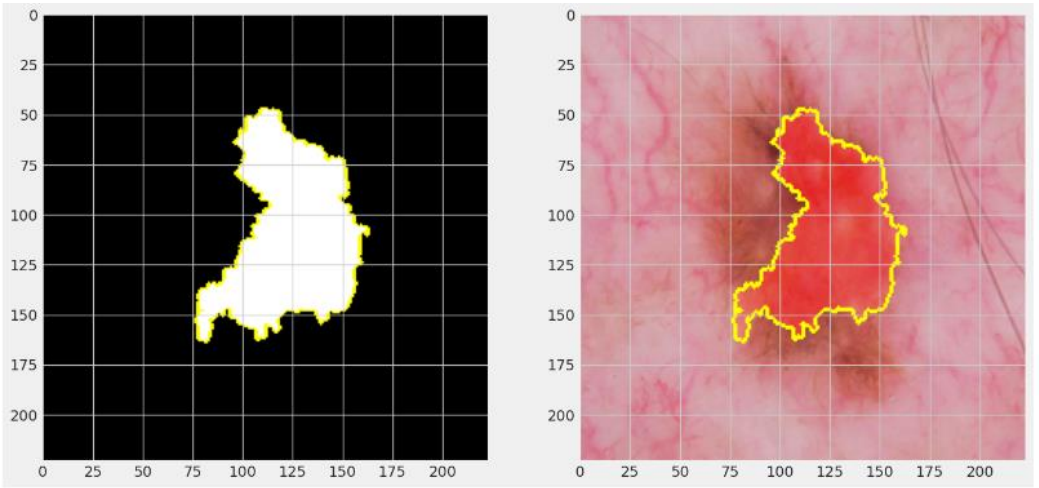


Figure 12: Malignant portion detection using LIME model

LIME use heatmaps to visually display the segmentation representation, with the color of each pixel indicating its level of relevance to the prediction. The heatmaps displayed in Figures 11 and 12 highlight the areas of the image that exerted the greatest impact on the model's prediction. This offers perceptive data on the conduct of the model. The initial step in generating the segmentation representation in LIME involves employing a segmentation algorithm to partition the picture into distinct segments. The super pixel method, which groups pixels based on similar qualities such as color or texture, is the most commonly employed approach. The LIME algorithm then delivers an explanation for each segment separately, treating each segment as an independent input instance. LIME determines the feature significance scores for each segment and then maps these values onto the relevant pixels in the picture to provide a heatmap depiction of the segmentation explanation. Brighter colors indicate sections of the image that were more significant for the prediction in the generated heatmap. To validate the suggested model, a comparison is made with models that are thought to be state-of-the-art; the results of this comparison are shown in Table.2.

Table 2: Comparison results of the proposed model

Model	Precession	Recall	F1-score	Accuracy	Computational Time (second)
Resnet50	0.51	0.53	0.84	85%	3.7
Xception [22]	0.90	0.88	0.89	89%	3.87
InceptionV3 [23]	0.92	0.68	0.83	83%	3.0
DenseNet121 [24]	0.87	0.90	0.89	89%	5.80
Proposed EfficientB3	0.90	0.91	0.91	90%	5.55

The ResNet50, Xception, InceptionV3, and DenseNet121 state-of-the-art models are compared with the proposed EfficientNetB3 model. The InceptionV3 model achieved least accuracy with 83%. The Resnet50 model is getting 85% accuracy. The Xception and DenseNet121 models achieved 89% accuracy. When compared to all other models, the proposed model's accuracy of 90% was higher. Overall, an EfficientNetB3 model for skin cancer detection uses LIME to understand its predictions. LIME may support to increase the model's reliability, accuracy, and trust in its decision-making by identifying the most crucial

aspects in the input image that led to the prediction.

5. Conclusion

In this study, an improved EfficientNetB3-based deep learning model is employed to classify skin cancer into benign and malignant subtypes. The adjustments allowed for improved computing performance by employing a compound scaling strategy, which attempted to balance model size and accuracy. The LIME explainable AI model is employed to determine how well out enhanced EfficientNetB3 model is performing. Additionally, LIME can prove to be a useful technique for image classification since it provides human-readable and likewise are easily understood explanations of each prediction, hence in turn also helping us understand more about the decision making process of the model. According to the results, the detection technique identified skin cancers as malignant or benign with 91% accuracy. The upgraded version EfficientNetB3 and LIME technique have the potential to come with the possibility of great flexibility to show high reliability for classification models. This development has great promise in early skin cancer detection and treatment.

References

1. Godic, Aleksandar, Borut Poljšak, Metka Adamic, and Raja Dahmane. "The role of antioxidants in skin cancer prevention and treatment." Oxidative medicine and cellular longevity 2014 (2014).
2. Soundharaj, Sowmiya, M. Ramachandran, and Chinnasami Sivaji. "The Role of Ultraviolet Radiation in Human Race." Environmental Science and Engineering 1, no. 2 (2022): 48-56.
3. Li, Yueyao, Wen-Qing Li, Tricia Li, Abrar A. Qureshi, and Eunyoung Cho. "Eye color and the risk of skin cancer." Cancer Causes & Control (2022): 1-8.
4. LeQuang, Jo Ann. "Using Gene Expression Profiling to Personalize Skin Cancer Management." The Journal of Clinical and Aesthetic Dermatology 15, no. 11 Suppl 1 (2022): S3-S15.
5. Chmiel, Paulina, Martyna Kłosińska, Alicja Forma, Zuzanna Pelc, Katarzyna Gęca, and Magdalena Skórzewska. "Novel Approaches in Non-Melanoma Skin Cancers—A Focus on Hedgehog Pathway in Basal Cell Carcinoma (BCC)." Cells 11, no. 20 (2022): 3210.
6. Wako, Beshatu Debela, Kokeb Dese, Roba Elala Ulfata, Tilahun Alemayehu Nigatu, Solomon Kebede Turunbedu, and Timothy Kwa. "Squamous Cell Carcinoma of Skin Cancer Margin Classification From Digital Histopathology Images Using Deep Learning." Cancer Control 29 (2022): 10732748221132528.
7. Aljohani, Khalil, and Turki Turki. "Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks." Ai 3, no. 2 (2022): 512-525.
8. Ukharov, A. O., I. L. Shlivko, I. A. Klemenova, O. E. Garanina, K. A. Uskova, A. M. Mironycheva, and Y. L. Stepanova. "Skin cancer risk self-assessment using AI as a mass screening tool." Informatics in Medicine Unlocked 38 (2023): 101223.
9. Kaur, Ranpreet, Hamid GholamHosseini, Roopak Sinha, and Maria Lindén. "Melanoma classification using a novel deep convolutional neural network with dermoscopic images." Sensors 22, no. 3 (2022): 1134.
10. Zakhem, George A., Akshay N. Pulavarty, Jenna C. Lester, and Mary L. Stevenson. "Skin cancer in people of color: a systematic review." American Journal of Clinical Dermatology (2022): 1-15.
11. Arora, Ginni, Ashwani Kumar Dubey, Zainul Abdin Jaffery, and Alvaro Rocha. "Bag of feature

and support vector machine based early diagnosis of skin cancer." Neural Computing and Applications (2022): 1-8.
12. Silva, Carina V., Caitlin Horsham, and Monika Janda. "Review of educational tools for skin self-examination: A qualitative analysis of laypeople's preferences." Health Promotion Journal of Australia 33, no. 2 (2022): 386-394.
13. Narasimhulu, C. Venkata. "An automatic feature selection and classification framework for analyzing ultrasound kidney images using dragonfly algorithm and random forest classifier." IET Image Processing 15, no. 9 (2021): 2080-2096.
14. Saturi, Rajesh, and Prem Chand Parvataneni. "Histopathology breast cancer detection and classification using optimized superpixel clustering algorithm and support vector machine." Journal of The Institution of Engineers (India): Series B 103, no. 5 (2022): 1589-1603.
15. Vijayalakshmi, M. M. "Melanoma skin cancer detection using image processing and machine learning." International Journal of Trend in Scientific Research and Development (IJTSRD) 3, no. 4 (2019): 780-784.
16. Tan, Teck Yan, Li Zhang, and Chee Peng Lim. "Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models." Applied Soft Computing 84 (2019): 105725.
17. Agrahari, Pradhumn, Archit Agrawal, and N. Subhashini. "Skin cancer detection using deep learning." In Futuristic Communication and Network Technologies: Select Proceedings of VICFCNT 2020, pp. 179-190. Springer Singapore, 2022.
18. Reis, Hatice Catal, Veysel Turk, Kourosh Khoshelham, and Serhat Kaya. "InSiNet: a deep convolutional approach to skin cancer detection and segmentation." Medical & Biological Engineering & Computing (2022): 1-20.
19. Chaturvedi, Saket S., Jitendra V. Tembhurne, and Tausif Diwan. "A multi-class skin Cancer classification using deep convolutional neural networks." Multimedia Tools and Applications 79, no. 39-40 (2020): 28477-28498.
20. Ali, Md Shahin, Md Sipon Miah, Jahurul Haque, Md Mahbubur Rahman, and Md Khairul Islam. "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models." Machine Learning with Applications 5 (2021): 100036.
21. https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign.
22. Lu, Xinrong, and Y. A. Firoozeh Abolhasani Zadeh. "Deep learning-based classification for melanoma detection using XceptionNet." Journal of Healthcare Engineering 2022 (2022).
23. Alam, Talha Mahboob, Kamran Shaukat, Waseem Ahmad Khan, Ibrahim A. Hameed, Latifah Abd Almuqren, Muhammad Ahsan Raza, Memoona Aslam, and Suhuai Luo. "An Efficient Deep Learning-Based Skin Cancer Classifier for an Imbalanced Dataset." Diagnostics 12, no. 9 (2022): 2115.
24. Gajera, Himanshu K., Deepak Ranjan Nayak, and Mukesh A. Zaveri. "A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features." Biomedical Signal Processing and Control 79 (2023): 104186.