

Deep Learning for Automatic Diagnosis of Skin Cancer Using Dermoscopic Images

S. Rama Krishna,¹ Anand Gudur,² Siddharth Jain,³ Shanmugavel Deivasigamani,⁴ Mohit Tiwari,⁵
K. G. S. Venkatesan,⁶ and Karthik Raj V⁷

¹Department of Computer Science & Engineering GSoT, GITAM (Deemed-to-be-University), Rudraram, Telangana, India

²Department of Onchology, Krishna Institute of Medical Sciences, Krishna Vishwa Vidya Peeth, Karad, MS, India

³Inueron, Flat 401, Natraj Tower, Waltair Uplands, Visakhapatnam, India

⁴Computer Science and Engineering, Sri Muthukumaran Institute of Technology, Anna University, India

⁵Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, A-4,
Rohtak Road, Paschim Vihar, Delhi, India

⁶Department of Computer Science & Engineering, MEGHA Institute of Engineering & Technology for Women,
Edulabad, Hyderabad, Telangana, India

⁷Department of Accreditations and Trainings, Andhra Pradesh MedTech Zone Limited (AMTZ), AMTZ Campus,
Pragati Maidan, VM steel Project SO, Visakhapatnam, India

(Received 24 September 2023; Revised 28 November 2023; Accepted 31 December 2023; Published online 23 February 2024)

Abstract: Over the past few years, the healthcare industry has seen a dramatic increase in the use of intelligent automation enabled by artificial intelligence technology. These developments are made to better the standard of medical decision-making and the standard of treatment given to patients. Fuzzy boundaries, shifting sizes, and aberrations like hair or ruler lines all provide difficulties for automatic detection of skin lesions in dermoscopic images, slowing down the otherwise efficient process of diagnosing skin cancer. However, these difficulties may be conquered by employing image processing software. To address these issues, the authors of this paper provide a novel intelligent multilevel thresholding with deep learning (IMLT-DL) model for intelligent dermoscopic image processing. Multilevel thresholding and DL are brought together in this model. Top hat filtering and inpainting have been included into IMLT-DL for use in image processing. In addition, mayfly optimization has been used in tandem with multilayer Kapur's thresholding to identify specific trouble spots. For further investigation, it uses an Inception v3-based feature extractor, and for data classification, it makes use of gradient boosting trees (GBTs). On the International Skin Imaging Collaboration (ISIC) dataset, this model was shown to outperform state-of-the-art alternatives by a margin of 0.992% over the duration of trial iterations. These advances are a major step forward in the quest for faster and more accurate skin lesion detection.

Keywords: deep learning; dermoscopic images; skin cancer

I. INTRODUCTION

The most prevalent form of skin cancer, melanoma, is also one of the most frequent forms of cancer overall. Melanoma, squamous cell carcinoma, basal cell carcinoma, and intraepithelial carcinoma are the four most frequent forms of skin cancer, although there are many more. Intraepithelial carcinoma and basal cell carcinoma are two other kinds. Human skin consists of three layers: the hypodermis, the epidermis, and the dermis. Melanocytes are a kind of cell that may show up in the dermis, or the top layer of skin. Rapid melanin production is a unique property of these cells [1]. Exposure to greater levels of UV radiation from light over a longer period of time, for instance, may initiate melanin synthesis. The most lethal kind of skin cancer, known as malignant melanoma, has been linked to an increase in melanocytes. The American Cancer Society

has released estimates predicting that there would be 96,480 new cases of melanoma in 2019 with an associated death toll of 7,230. A proper treatment plan can only be implemented if an accurate diagnosis of melanoma has been made. There is a 92% likelihood of survival for at least five years if melanoma is caught early [1].

Despite this, the similarity between malignant and noncancerous skin lesions is one of the main hurdles to proper melanoma detection argued by [2]. This means that even for doctors with extensive experience, detecting melanoma might be difficult. Visual inspection of a lesion may not always provide clear results in terms of the lesion's classification. In recent years, several imaging techniques have become more popular for capturing skin. Dermoscopy, which uses an immersion fluid and light magnification equipment to provide a visual image of the skin's surface, is a noninvasive imaging approach. However, as it relies on the expert's prior knowledge, the simpler visualization approach for recognizing melanoma in skin lesions may be prone to subjectivity, failure to repeat, or inaccuracy [1].

Corresponding author: S. Rama Krishna (e-mail: rsankara@gitam.edu).

According to Nawaz *et al.* [3], the success percentage of dermoscopic images collected by nonspecialists for melanoma prognosis evaluation is between 75% and 84%. Computer-aided diagnostic (CAD) techniques are crucial for melanoma diagnosis because they offer medical professionals with an analytical system. We cannot expect to solve these problems unless we do this. The melanoma CAD model incorporates the preprocessing, segmentation, feature extraction, and classification phases of the diagnostic procedure [3]. Accurate melanoma detection relies heavily on the CAD system's lesion segmentation method. Many different types of skin lesions may be seen during dermoscopy, and they may be identified by their size, color, location, and texture. This is an important stage, but it might be difficult to do because of the variety argued by [4]. Hair, dark frames, air bubbles, colored illumination, ruler lines, and blood vessels are just a few of the many distractions that may make lesion segmentation more challenging. Differentiating skin lesions was a major topic of discussion throughout this session. A convolutional neural network (CNN) is a recent deep learning (DL) method that has been used to improve the CAD model [5]. Alex Net, Mobile Net, and ResNet are three well-known examples of popular DL architectures. Numerous factors, including the widespread interest in the Inception film series, led to the trilogies being chosen as the theoretical basis for this investigation. Because of its low processing efficiency, the model utilized in Inception only partially and inconsistently realizes all of its parameters. It boosts efficiency, makes efficient use of computer resources, and contributes very little to the network's total computational load, all of which contribute to Inception's impressively powerful performance [3].

We use intelligent multilevel thresholding in conjunction with DL to build a model for the segmentation and classification of skin lesions in dermoscopic images (IMLT-DL). The relevant intelligent multilevel thresholding with deep learning (IMLT-DL) model employs top hat filtering and in painting to perform preliminary image processing on dermoscopic images. To pinpoint the damaged regions even further, we use a technique called mayfly optimization (MFO), which is based on multilayer Kapur's thresholding-based segmentation [6]. The segmented image is then sent into an Inception v3-based feature extractor, which draws out relevant information to form a set of feature vectors [7]. The last phase involves classifying the dermoscopic images using a technique adapted from the gradient boosting tree (GBT) model. A simulation is run using data from the International Skin Imaging Collaboration (ISIC) to test the IMLT-DL model proposed, and the results are analyzed using many criteria. Following this plan will give you a general idea of how to structure the paper you are

writing. In the next section, we'll have a look at the state-of-the-art methods for segmenting skin lesions [7]. The submitted IMLT-DL model will be analyzed in Section III, and Section IV will check the simulation results. At this moment, the IMLT-DL model has reached its final stage [6].

To illustrate the findings of a study on the detection of skin lesions, we provide the following sample comparison analysis table.

The purpose of this table is to provide a synopsis of the methods employed and the results obtained in studies of skin lesion detection. The purpose of this table is to provide an overview of the research methods and results. New information or research may need a reorganization of the document's content and appearance.

II. LITERATURE REVIEW

Daghrir *et al.* [8] argue that a variety of approaches for identifying and categorizing skin lesions are described. An overview of the Grab cut and K-means clustering algorithm-based semi-supervised approach to skin lesion segmentation is presented. After the melanoma has been segmented using graph cuts, K-means clustering is used to refine the lesion's margins. Prior to being read in, the picture undergoes preprocessing steps including noise reduction and image normalization [8]. The image is then prepared for pixel classification. Features were extracted using a scale-invariant feature transform approach. In order to classify medical images, CNNs have been used. The artificial bee colony (ABC) approach is used in a similar way when segmenting lesions [9]. Before a digital picture can be analyzed using the swarm-based method, it must go through certain preliminary processing. Otsu thresholding is used to determine the best threshold value after the segmentation of a melanoma lesion. Delaunay triangulation is used to create a final picture, and this technique may be used to divide up the final image (DTM) [8].

Shorfuzzaman [9] argue that this technique uses a parallel segmentation strategy to generate two images that are then fused into a single lesion mask. The next step is to apply this design as a mask on the image. After the artifacts have been removed, the skin is filtered out of the image using a single model to provide a binary mask of the lesions. DTM offers the ability to reduce time compared to other approaches due to its fully automated nature and lack of need on a trained model [10]. In this work, we provide a new automated method for image segmentation by combining image-wise supervised learning (ISL) with multi-scale super pixel-based cellular automata (MSCA). The researchers eliminated the need for the user to submit a description of the seed selection

Study	Methodology	Key findings
Wei <i>et al.</i> [1]	Visual inspection, dermoscopy.	Improved prognosis is associated with earlier diagnosis of melanoma. Inaccuracy and subjectivity in less complex visualization approaches.
Tahir <i>et al.</i> [2]	Comparative study of carcinogenic and benign tumors.	Identifying subtypes is challenging even for seasoned medical professionals.
Nawaz <i>et al.</i> [3]	Analysis of dermoscopic images by inexperienced observers, computer-aided design.	The average success percentage for those who are not experts is 75% to 84%. The value of CAD in early diagnosis.
S M <i>et al.</i> [4]	Lesion segmentation issues in dermoscopy.	Segmentation is complicated by visual noise such as hair, bubbles, and lines.
Fraivan and Faouri [5]	CAD models based on CNN.	Enhancing CAD using Convolution Neural Network (CNN) Architectures (AlexNet, MobileNet, and ResNet).
Waheed <i>et al.</i> [6]	Multi-threshold mayfly optimization.	Lesion segmentation using multi-field optimization and multilayer thresholding.

process by using probability maps. The MSCA method was then used to segment the skin lesions [11]. A fully convolutional networks (FCNs)-based method is employed to accurately segment the skin lesion. Accurate segmentation of skin lesions is achieved (compared to prior challenges) without the customary preprocessing procedures since several phases of the FCN are included within the picture features to be learned (such as improving contrast, removing hair, and so on) [9].

Ameri [12] argue that convolutional and DE convolutional neural networks (CDNNs) are used for the purpose of mechanically segmenting skin lesions. This technique has an advantage over others that need several pre- and postprocessing steps since it largely depends on training methodologies. This approach encodes the likelihood that a group of pixels will acquire melanoma as a function of their location on the map. A 10-fold cross-validation method is used to train the algorithm, and a U-Net framework is then used to segment skin lesions based on a probabilistic mapping of the image's dimensions [12].

Reshma *et al.* [13] argue that artificial intelligence (AI) has gone a long way in the preceding decade, notably in the domains of DL and CNNs, enabling for the creation of trustworthy image-based medical systems for screening and diagnosis. Recently, automated classification using DL as the major technique of research has supplanted image segmentation (the separation of relevant areas in an image) and feature extraction [14]. Recently, this tendency has become clear. Similar patterns have been seen in the study of skin cancer detection and screening, with the use of sophisticated DL AI replacing the more traditional approach of image processing to filter out noise (such as hair). This direction was adopted by the majority of studies. These more contemporary algorithms are typically invulnerable to picture noise and do not require the explicit extraction of characteristics. However, they are typically computationally time-consuming [13].

Arif *et al.* [15] argue that using digital hair removal (DHS), hair was eliminated from the skin lesion picture, and the impact of the hair removal was analyzed using intra-structural similarities (Intra-SSIM). An original technique for the segmentation of lesion pictures according to areas of interest was created utilizing DL (ROI). With the aid of trained neural networks, we developed a novel representation of intermediate features to extract information from the ROI (such as ResNet and DenseNet). CNNs were employed to differentiate dermoscopic characteristics and lesions [8], instead of utilizing a pertained model or excessive augmentation [15]. A unique approach for data separation leverages high-resolution convolutional neural networks. The suggested technique beats the other DL segmentation algorithms without the need for any further image processing, and the results were statistically significant. A new technique for segmentation is based on a deep fully convolutional network with 29 total layers [16]. CNN segmentation of dermoscopy pictures using attention approach maintains edge features. An innovative DL model is based on Gabor wavelets for distinguishing between melanoma and seborrheic keratosis. Seven separate CNN models that made use of Gabor wavelets were integrated to form the ensemble used to construct this model. More specifically, their technique combines a Gabor wavelets-based model with a CNN model for images [8]. Performance assessment findings showed that a collection of image and Gabor wavelet-based models performed better than any of the individual models. This ensemble performed better than a subset of CNN models constructed just using Gabor wavelets [8].

Kaur *et al.* [17] argue that recently it has been shown that deep transfer learning may be used for robust automated internal feature

extraction in the field of medical imaging. There were three phases to find melanoma. The first thing to do is to make sure the dataset is error-free and resize the photographs to the right dimensions. Subsequently, feature extraction is improved with the help of deep transfer learning [18]. It compared the performance of six different transfer learning networks in classifying multiclass lesions using these characteristics to train a set of classic classification methods, including support vector machine (SVM), logistic label propagation (LLP), and k-nearest neighbors (KNN). However, their findings were contingent on a larger dataset [19]. The primary objective of picture augmentation is to enhance existing photographs without producing duplicates. Creating many variants of the same picture in the dataset with varying degrees of improvement confuses findings and does not accurately reflect the effectiveness of the system [20].

III. METHODOLOGY

The technique through which images of skin lesions were used in the process of establishing a classification system for skin cancer is shown in Fig. 1. The approaches that were employed in this inquiry did not require the extraction of features or the segmentation of data [21].

Rather, both of these tasks were performed automatically (i.e., separation of lesions from the rest of the image). The several layers and procedures that are included inside the DL model are responsible for automatically resolving each and every one of these issues. An in-depth description of each component is provided in the following passages [21].

A. DATASET

For this paper, the authors used the dataset HAM1000 (Human against Machine), which contains dermoscopic pictures of the most prevalent skin malignancies [22]. Seven classes may be used to classify each set of pictures: there were 1113 melanomas, 6705 nevi, and 142 vascular lesions altogether. Random x-y scaling in the range and random x-y translation in the pixel range were utilized as enhancing techniques [7].

B. DEEP LEARNING MODELS

Numerous applications, especially those involving medical imaging, have shown the efficacy of transfer learning. This is accomplished by swapping out poorly constructed and tested ad hoc deep CNN architectures with more robust pretrained models [23]. These models' bottom layers are taught to recognize and respond to common visual cues like color and brightness [24]. The layers that follow the input must be retrained and customized to account for the unique aspects of the job they perform. This method has been proved to be useful in a wide range of studies. Thirteen DL models were tweaked, retrained, tested, and compared in this study for their ability to classify skin lesions into the aforementioned seven categories from the HAM1000 dataset. This was done to find out which model was the most efficient at solving the problem [7].

C. METRICS AND FRAMEWORK FOR PROGRAM EVALUATION

According to [10], a performance's F1 rating is based on how well it performs across four criteria: accuracy, precision, recall, and

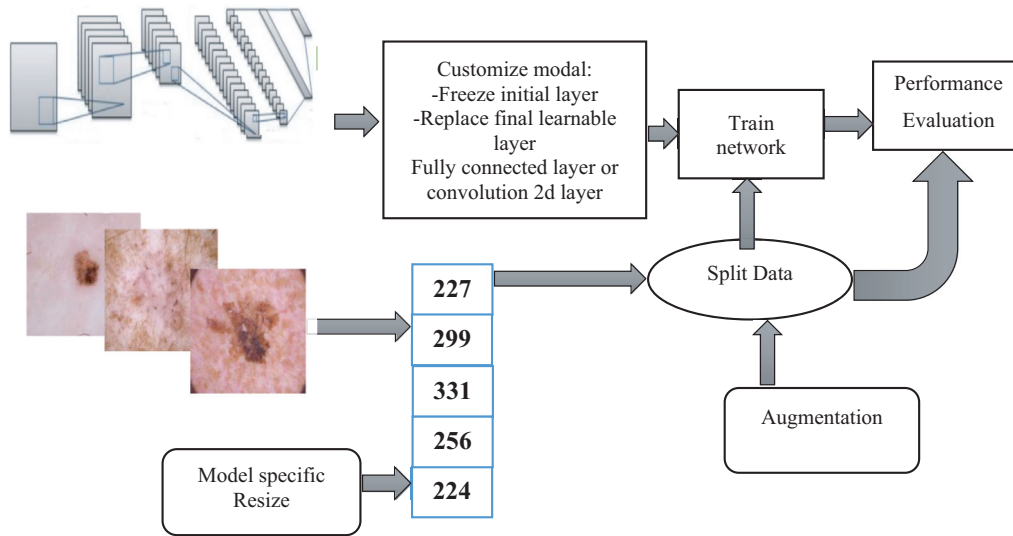


Fig. 1. A diagrammatic summary of the overarching procedures discussed in this article.

specificity. Accuracy is ranked by how many photos out of a total set were correctly labeled [23]. The accuracy of a test is defined as the percentage of true positives. In consideration of all important factors, we calculate the proportion of correct answers (as assessed by recall, also called sensitivity). A photograph’s specificity, sometimes called its selectivity, is the proportion of negative photos relative to the total number of images. Classification performance on imbalanced datasets is measured using the F1 score, the harmonic mean of the recall, and accuracy scores. The ratio of confirmed to suspected negative pictures is often accepted as the most accurate measure of precision. Definitions are provided for the following five types of quantitative research: if all goes according to plan, the reported scores for each course will be the class mean [25]. The data were successfully divided into separate datasets for research and development using three different methods. The models’ ability to generalize without becoming too particular to the input photographs will be evaluated using these. When we talk about “random axis translation,” we were referring to the process of arbitrarily shifting an image along the x and y axes [7]:

$$Accuracy = \frac{TP + TN}{P + N} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (5)$$

IV. EXPERIMENTAL RESULTS

From the results of the experiments, it was determined that the IMLT-DL model accurately detected and classified photos of skin lesions. For instance, while trying to classify instances into the “Angioma” category, the IMLT-DL model has a sensitivity of 0.952, specification of 1, accuracy of 0.997, precision of 1, and G-measure of 0.976. The IMLT-DL method successfully classified the “Nevus” subset with a sensitivity of 0.957, specificity of 0.996, accuracy of 0.991, precision of 0.978, and a G-measure of 0.967. The IMLT-DL method successfully classified the “Melanoma” category with a sensitivity of 0.980, specificity of 1, acc. of 0.997, accuracy of 1, and G-measure of 0.990 [26].

In the end, the DLN model has shown even better results than the CDNN model, with a sensitivity of 0.732, a specificity of 0.754, and an accuracy of 0.743. More recently, the CDNN model has demonstrated a sensitivity of 0.825, a specificity of 0.975, and an accuracy of 0.934. The ResNets model has improved sensitivity by 0.02, specificity by 0.985, and accuracy by 0.934%. The findings of the DCCN-GC model are similarly rather good: sensitivity = 0.908, specificity = 0.927, and acc. = 0.934. As an added bonus, the DL-ANFC model has been used to provide almost optimal results, with a sensitivity of 0.934, specificity of 0.987, and accuracy of 0.979. Comparable approaches are only 0.97 sensitive, 0.99 specific, and 0.99 trustworthy, according to the IMLT-DL model [27]. Figure 2 shows the comparative result analysis of accuracy, specificity, and sensitivity.

V. RESULTS

Correctly classifying photos of skin lesions into one of two categories (benign vs. melanoma) or three groups has been shown in the past (benign, melanoma, and nonmelanoma). One such instance is as follows: this has been shown by previous research to be correct. The purpose of this research was to test the ability of deep convolutional network models to generalize their knowledge of skin lesions from one dataset to another [28]. Ten training iterations were run, and the average results were provided, to allow

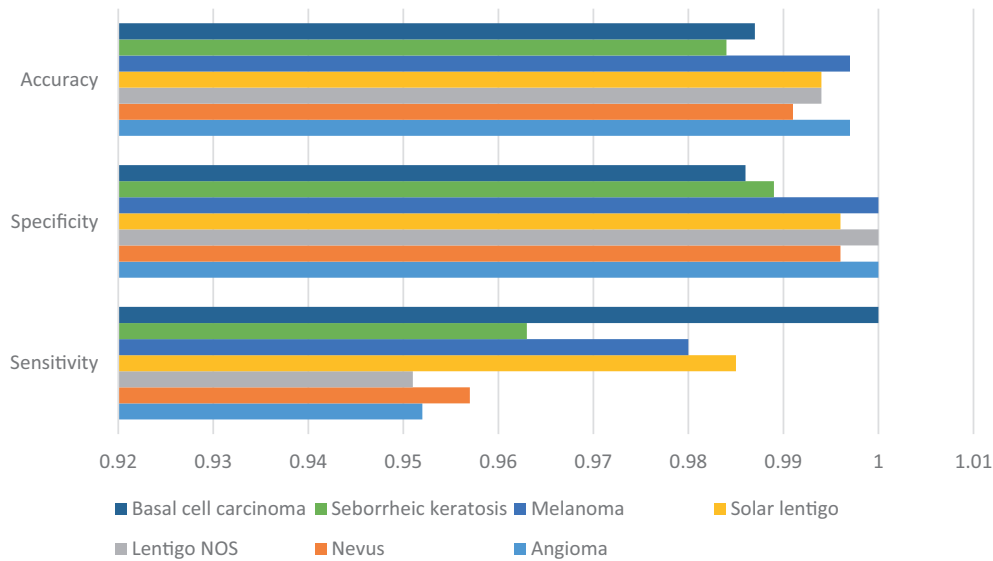


Fig. 2. Results analysis.

Table I. Graphs illustration for the texture model

76	7	19	2	54	66%	33%		
45	47	76	4	32	7	68%	38%	
65	13	76	24	3	58%	41%		
3	633	32	15	2	52%	47%		
6	58	12	14	237	5	71%	29%	
19	8	7	76	34	1733	4	82%	17%
15	11	100	43	34	43	22	58%	41%

for any variations caused by the random distribution of data. This action was taken to guarantee more uniform outcomes. The training and validation times were taken into account in addition to the high computing cost of DL models [10].

For ease of comparison, Table I summarizes the average performance metrics from 10 iterations of each of the 13 DL models, using 70% of the data for training. All of the models performed rather well in terms of accuracy; however, Resnet101 performed best (76.7%). Sample rows and columns from a confusion matrix are summarized in Fig. 3 for in-depth analysis. The imbalance between the number of images in each category and the accuracy values causes the F1 score to be lower than the accuracy values. There is also evidence that shows smaller class sizes lead to less precision overall [29]. The bulk of the pictures were assigned to the NV class because they had the highest levels of accuracy (92.5%) and recall (see NV summary for details on the NV columns) (82.5 percent; see NV summary for more information about the NV rows). However, only 43.1% of melanoma patients were properly identified according to their specific subtype, despite the fact that 71% of patients were correctly categorized overall (recall). The gaps in accuracy and recall between the other groups are less pronounced, but they still exist [30] (Table II).

Figure 3 depicts a progress curve for training and validation that was created for Resnet101 by dividing the data into 70/30 halves. This graph implies two things:

- (a) The model is not able to consistently minimize loss and provide high testing accuracy, even when the number of epochs is raised [29].
- (b) The disparity between validation and testing performance is noticeable since there are less photos available for most classes (DL requires large datasets).

Compare and contrast the IMLT-DL model’s performance metrics with those of other available techniques. Emphasize the model’s usefulness and the areas in which it advances upon prior methods. The improvement amounted to 0.992%. In the case of skin cancer, when early identification is critical for improving patient outcomes, please outline the significance that even tiny percentage gains in diagnostic accuracy may have on patient outcomes and the efficacy of therapy. Please elaborate on the implications of incremental improvements in diagnosis accuracy in the context of skin cancer. Analyze and report on the causes of the IMLT-DL model’s excellent algorithmic performance. Describe how the model’s efficiency and accuracy were improved by including multilevel thresholding, DL, and optimization techniques. Challenges and Restriction: take note of the difficulties that arose throughout the process of evaluating the IMLT-DL model. Discuss how to cope with the computational complexity of the model, as well as any potential biases it may have and any application circumstances in which it may still be unsuccessful. What Comes Next? Suggest some avenues that may be explored in the future of this field of study. There have been considerable developments in AI technology, as well as improvements in datasets and clinical validation. The opportunities presented by these variables for further improvement of the IMLT-DL model should be underlined. Talk about how the IMLT-DL model may be used in clinical settings and whether or not it is ready for widespread adoption. There are a number of considerations that need to be made, including scalability, clarity, and interoperability with current healthcare infrastructures. In addition to an analysis of the data, the discussion of the study’s findings may include an examination of the implications of the study’s findings and predictions for the future of skin lesion

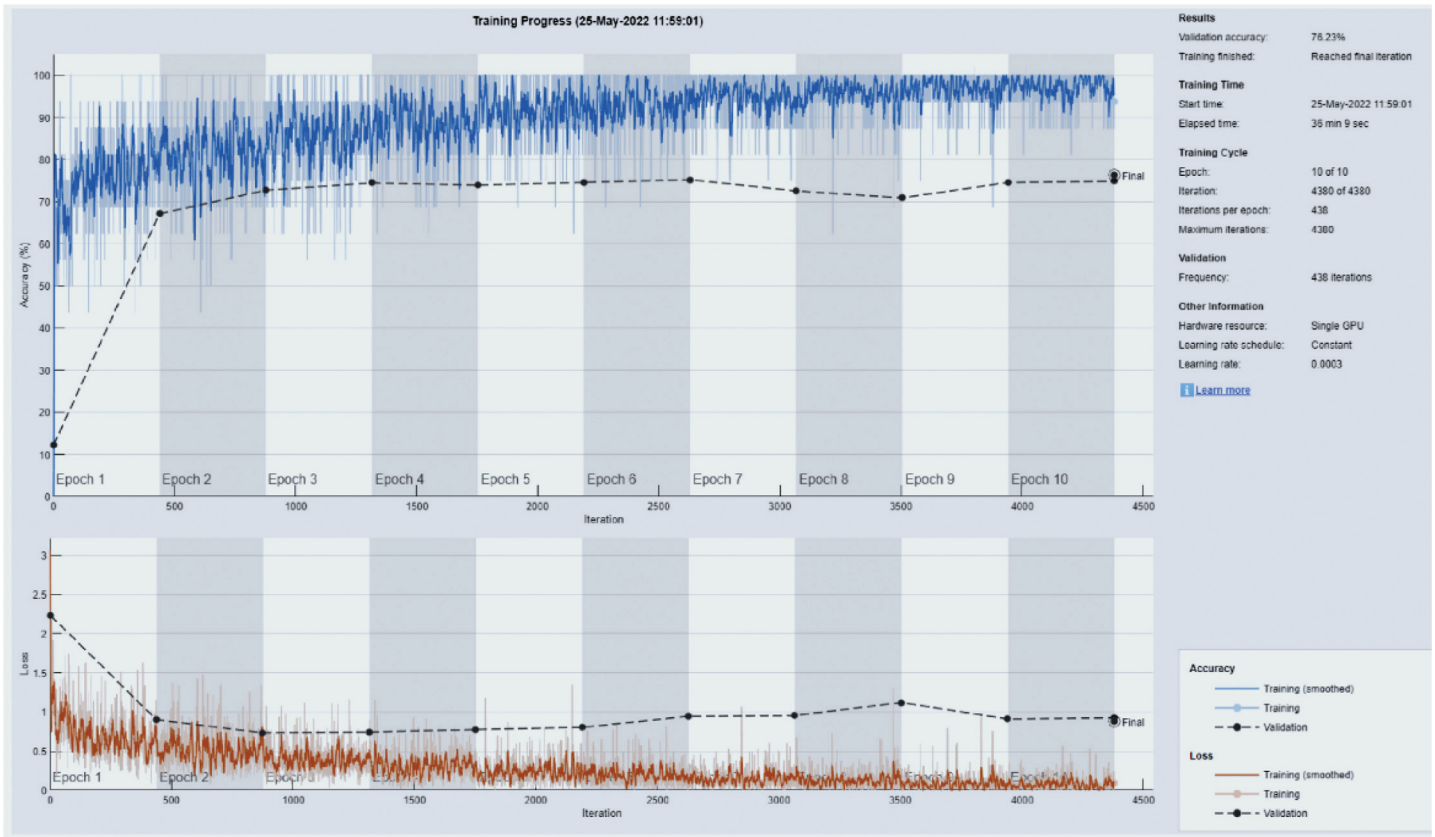


Fig. 3. A strength and conditioning curve illustration using Resnet101 and a 70/30 split of the data.

Table II. The mean values for each deep learning model’s overall accuracy, *F*-score, precision, recall, and specificity with 70/30 data split

Model	F1 Score (%)	Precision (%)	Recall (%)	Specificity (%)	Accuracy (%)
SqueezeNet	49.2	60.1	50.6	89.8	72.7
GoogLeNet	45.4	60.2	53.4	93.3	72.0
Inceptionv3	59.5	62.5	59.7	93.5	73.2
DenseNet201	60	69.9	59.7	93.7	76.8
MobileNetv2	61.0	70.2	60.4	93.1	72.6
Resnet101	62.3	70.6	59.8	94.0	71.7
Resnet50	59.4	69.5	59.4	93.7	76.4
Resnet18	61.3	65.7	60.8	93.6	72.3
Xception	57.9	65.5	60.2	93.7	71.4
Inception-ResNet-v2	58.4	66.3	59.8	93.4	74.5
ShuffleNet	59.6	63.9	60.7	92.5	72.6
DarkNet-53	62.9	62.8	59.9	93.5	75.6
EfficientNetb0	60.6	69.3	55.7	92.1	76.8

diagnosis and healthcare AI. One way to do this would be to elaborate on the themes brought up in the results section.

6. Discussions

The mean overall performance metrics for each of the 13 DL models are provided in Table IV; these metrics are based on 10

separate runs, with 80% of the available data being used for training purposes [31]. Because the highest possible F1 score that could be achieved was 66.1%, this indicates that increasing the size of the training set by 10% had no impact on the performance measures. The confusion matrix that may be seen in Table III demonstrates that a substantial source of mistakes was the incorrect categorization of NV pictures as melanoma [16]. The vast majority of courses

Table III. The following is an example of a confusion matrix for the DenseNet201 model and the 80/20 data split

35		20		7	5	5	56%	43%
34	59	21	1	3	3	34	58%	40%
23	34	119		29	1	24	48%	41%
13	21	21		42		53	52%	41%
12	1	10	11	188		15	71%	30%
1	2	3		12	22	52	78%	15%
8	3	4		31		21	56%	31%

were effective in attaining a high level of accuracy but only managed to get a low level of recall. Figure 4 demonstrates, in addition, the same propensity toward training and overfitting that was shown previously in the paper [31].

According to Jojoa *et al.* [32], the ratio of test pictures to train images ended up being 90% after an additional 10% was added to both the testing dataset and the training dataset. The frequency with which each of the 13 DL models was executed is summarized in Table III, along with the models’ average overall performance metrics. When compared to ResNet101 and DarkNet53, both of these networks scored more than 70% in F1. The table indicates that the majority of models get more accurate as additional training data is used, with the exception of the fragile SqueezeNet model. DL models, as opposed to more conventional machine learning models, could gain more from larger datasets [32]. This might be the reason why DL models seem to have superior performance. The

conventional machine learning models do not gain any advantage from using more extensive datasets. Figure 5 is an illustration of an example of the confusion matrix that may be constructed using DarkNet-53. In comparison to past iterations, there has been a considerable increase in the number of rows and columns in this matrix that include at least one inaccurate classification. This trend can be seen both horizontally and vertically. A training/validation progress curve that demonstrates overfitting is depicted even if it has been reduced [32].

According to Naeem *et al.* [7], even though a larger size of the training dataset showed some signals of promise, there is still more work to be done before a trustworthy diagnostic system that goes beyond the criterion of screening can be established. On the other hand, some of the findings were skewed since there were so few photographs available to represent each category. Even a little inaccuracy may have a significant impact on how well a DL model performs overall, despite the fact that the baseline prerequisites for a successful DL model are fairly simple. You can get more information by looking at Table IV, but in a nutshell, it demonstrates how much time each model required to train, divided down according to how it separates the data. In most cases, the amount of time necessary to complete a task rises in a linear fashion in proportion to the size of the training dataset, with increments of less than 10% for each order of magnitude [33].

The major contributions of this research are as follows:

- A novel approach for evaluating dermoscopy pictures is discussed in the study titled “Model Improvement Using Multi-Level Thresholding and Deep Learning” (IMLT-DL).

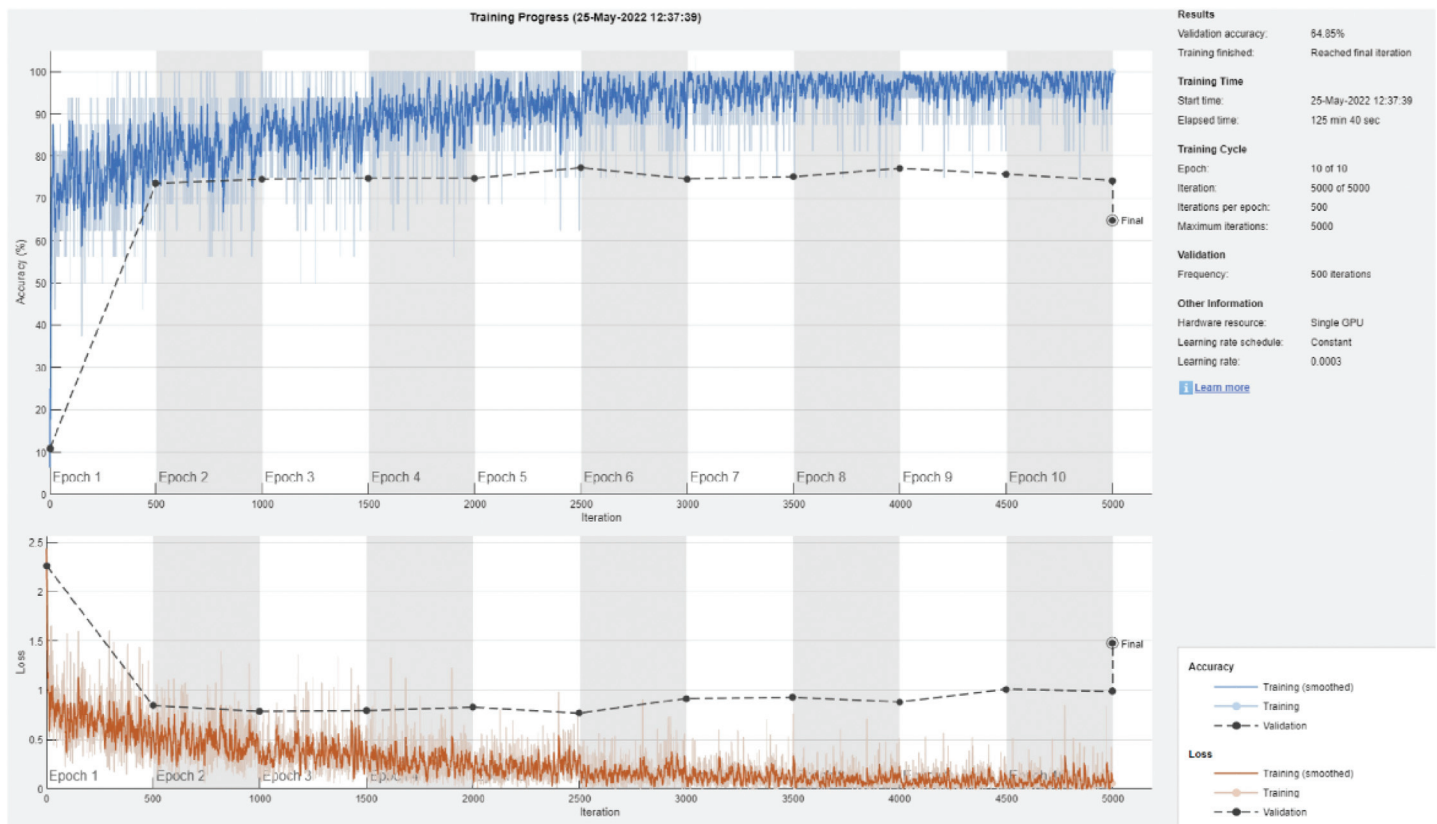


Fig. 4. This is an example of the strength and conditioning curve for DenseNet201 with an 80/20 split of the data.

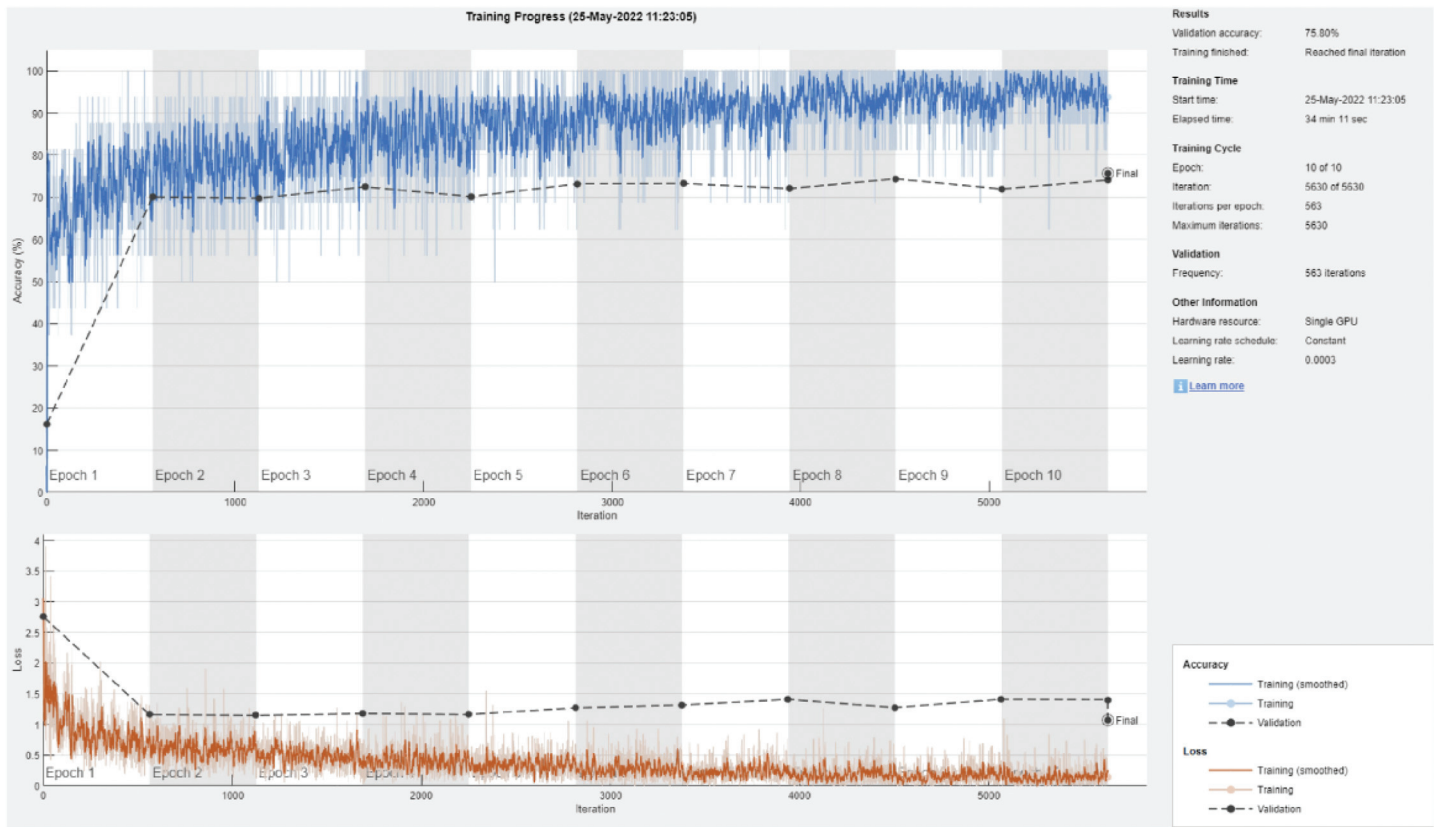


Fig. 5. Training and validation curve example for DarkNet-53 with a 90/10 data split.

Table IV. Mean overall performance metrics for 13 deep learning models

Model	F1 Score (%)	Precision (%)	Recall (%)	Specificity (%)	Accuracy (%)
SqueezeNet	49.6	60.0	49.8	89.4	58.0
GoogLeNet	53.2	69.0	52.8	83.4	58.5
Inceptionv3	59.1	63.2	59.6	84.0	58.8
DenseNet201	59.1	69.7	59.3	84.3	53.5
MobileNetv2	59.5	55.9	58.2	83.9	53.0
Resnet101	52.3	59.0	58.2	84.2	60.2
Resnet50	53.2	60.7	58.8	83.9	57.7
Resnet18	52.2	62.7	53.2	83.8	59.6
Xception	46.1	59.3	49.9	84.0	60.2
Inception-ResNet-v2	48.5	56.9	53.7	83.8	57.4
ShuffleNet	51.2	69.2	47.8	83.3	60.0
DarkNet-53	51.4	69.7	48.5	83.6	60.2
EfficientNetb0	49.0	59.8	42.6	83.6	62.2

- Complex image processing methods, such as top hat filtering, inpainting, MFO, and multilayer Kapur’s thresholding, are used to precisely locate problem areas in skin diseases. MFO and multilayer Kapur’s thresholding are two further methods.
- Dermoscopy image classification using a feature extractor built on top of Inception v3 and GBTs to generate fast, high-quality feature vectors.
- In a comprehensive test conducted on the ISIC dataset, the suggested IMLT-DL model displayed improved performance when compared to state-of-the-art approaches for the segmentation and classification of skin lesions by a margin of 0.992%.
- The research provides a clear and organized foundation for comprehending the methods and results through its structured approach to paper organization. The framework includes a

discussion of the state-of-the-art methods, a presentation of the IMLT-DL model, and an analysis of the simulation results.

This study's main contribution is the development and testing of the IMLT-DL model for dermoscopy image processing in skin lesion identification. This model was used to identify skin lesions. Based on an evaluation with data from the ISIC, our model outperformed state-of-the-art techniques by a margin of 0.992%. Multi-level thresholding, DL methods, and optimization algorithms like MFO and top hat filtering are all a part of the IMLT-DL model, allowing it to significantly improve the speed and accuracy of skin lesion detection for the purpose of identifying skin cancer. In the search for more accurate, efficient, and trustworthy medical diagnostics, the development of this conclusion represents a major breakthrough.

VII. CONCLUSION

Building a diagnostic system that is more accurate than the screening approaches that are now in use will need a significant amount of effort. It has been shown that encouraging results may be achieved by using a larger training sample, but there are still a number of obstacles to overcome. On the other hand, some of the assumptions were incorrect since there were not enough images to accurately portray each category. Even if there are not many variables in a good DL model, even a little piece of inaccurate information could have a substantial impact on how well the model works. Table IV presents a summary of the required training time for each model, divided down according to the manner in which it separates the data, as well as potential connections to more material on this subject. Across orders of magnitude, task completion durations have a tendency to rise in a linear fashion in proportion to the size of the training dataset, often by a margin of less than 10%. This is accurate in almost every circumstance. DarkNet-53 is the most effective model in terms of both the amount of time needed for training and the accuracy of categorization. Resnet101 is now in the position of runner-up. DarkNet-53 may have an advantage during training, but SqueezeNet is the superior model in terms of both speed and throughput.

VIII. FUTURE SCOPE

"Smart labs" and "smart imaging systems" have emerged as a result of the increasing prevalence of digitalization, roboticization, and automation in recent years. These companies are able to quickly and economically react to the expanding expectations that the healthcare sector places on them. The results of our research reveal that some new technologies are already making diagnostics better, while other technologies have the potential to completely revolutionize the area in the not too distant future. The widespread use of biosensors and the expanded use of companion diagnostics are two major developments that will define the future of diagnostics. Other major trends that will have an impact on the future of diagnostics include the increased use of liquid biopsies, direct to consumer testing, automation, and the transformation of pathology and radiology via the application of AI and sophisticated analytics. Partnerships with companies that specialize in consumer technology will be of considerable assistance in the creation of new treatments and the modification of existing regimens. Because the new value-based payment models will reward all participants for successful management of preventive care, early detection, and wellness, the only way they can be constructed is via partnerships.

CONFLICT OF INTEREST STATEMENT

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

- [1] L. Wei, K. Ding, and H. Hu, "Automatic skin cancer detection in dermoscopy images based on ensemble lightweight deep learning network," *IEEE Access*, vol. 8, pp. 99633–99647, 2020. DOI: [10.1109/ACCESS.2020.2997710](https://doi.org/10.1109/ACCESS.2020.2997710).
- [2] M. Tahir, A. Naeem, H. Malik, J. Tanveer, R. A. Naqvi, and S.-W. Lee, "DSCC_Net: multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images" *Cancers*, vol. 15, p. 2179, 2023. DOI: [10.3390/cancers15072179](https://doi.org/10.3390/cancers15072179).
- [3] M. Nawaz, Z. Mehmood, T. Nazir, R. A. Naqvi, A. Rehman, M. Iqbal, and T. Saba, "Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering," *Microsc. Res. Tech.*, vol. 85, pp. 339–351, 2022. DOI: [10.1002/jemt.23908](https://doi.org/10.1002/jemt.23908).
- [4] S. M. Jaisakthi, P. Mirunalini, C. Aravindan, and R. Appavu, "Classification of skin cancer from dermoscopic images using deep neural network architectures," *Multimed. Tools Appl.*, vol. 82, pp. 15763–15778, 2023. DOI: [10.1007/s11042-022-13847-3](https://doi.org/10.1007/s11042-022-13847-3).
- [5] M. Fraiwan and E. Faouri, "On the automatic detection and classification of skin cancer using deep transfer learning," *Sensors*, vol. 22, p. 4963, 2022. DOI: [10.3390/s22134963](https://doi.org/10.3390/s22134963).
- [6] Z. Waheed, A. Waheed, M. Zafar, and F. Riaz, 2017. "An efficient machine learning approach for the detection of melanoma using dermoscopic images," in *2017 Int. Conf. Commun., Comput. Digital Syst. (C-CODE)*. Presented at the 2017 International Conference on Communication, Computing and Digital Systems (C-CODE), pp. 316–319. DOI: [10.1109/C-CODE.2017.7918949](https://doi.org/10.1109/C-CODE.2017.7918949).
- [7] A. Naeem, T. Anees, M. Fiza, R. A. Naqvi, and S.-W. Lee, "SCDNet: a deep learning-based framework for the multiclassification of skin cancer using dermoscopy images," *Sensors*, vol. 22, p. 5652, 2022. DOI: [10.3390/s22155652](https://doi.org/10.3390/s22155652).
- [8] J. Dagher, L. Tlig, M. Bouhouicha, and M. Sayadi, "Melanoma skin cancer detection using deep learning and classical machine learning techniques: a hybrid approach," in *2020 5th Int. Conf. Adv. Techn. Signal Image Process. (ATSIP)*. Presented at the 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 2020, pp. 1–5. DOI: [10.1109/ATSIP49331.2020.9231544](https://doi.org/10.1109/ATSIP49331.2020.9231544).
- [9] M. Shorfuzzaman, "An explainable stacked ensemble of deep learning models for improved melanoma skin cancer detection," *Multimed. Syst.*, vol. 28, pp. 1309–1323, 2022. DOI: [10.1007/s00530-021-00787-5](https://doi.org/10.1007/s00530-021-00787-5).
- [10] D. Adla, G. V. R. Reddy, P. Nayak, and G. Karuna, "Deep learning-based computer aided diagnosis model for skin cancer detection and classification," *Distrib Parallel Databases*, vol. 40, pp. 717–736, 2022. DOI: [10.1007/s10619-021-07360-z](https://doi.org/10.1007/s10619-021-07360-z).
- [11] S. Lafraxo, M. E. Ansari, and S. Charfi, "MelaNet: an effective deep learning framework for melanoma detection using dermoscopic images," *Multimed Tools Appl.*, vol. 81, pp. 16021–16045, 2022. DOI: [10.1007/s11042-022-12521-y](https://doi.org/10.1007/s11042-022-12521-y).
- [12] A. Ameri, "A deep learning approach to skin cancer detection in dermoscopy images," *J. Biomed. Phys. Eng.*, vol. 10, pp. 801–806, 2020. DOI: [10.31661/jbpe.v0i0.2004-1107](https://doi.org/10.31661/jbpe.v0i0.2004-1107).
- [13] G. Reshma, C. Al-Atroshi, V. K. Nassa, B. T. Geetha, G. Sunitha, M. Gouse, and N. Subramani, "Deep learning-based skin lesion

- diagnosis model using dermoscopic images,” *Intell. Automat. Soft Comput.*, vol. 31, 2021. DOI: [10.32604/iasc.2022.019117](https://doi.org/10.32604/iasc.2022.019117).
- [14] T. Mazhar et al., “The role of machine learning and deep learning approaches for the detection of skin cancer,” *Healthcare*, vol. 11, p. 415, 2023. DOI: [10.3390/healthcare11030415](https://doi.org/10.3390/healthcare11030415).
- [15] M. Arif, F. M. Philip, F. Ajesh, D. Izdrui, M. D. Craciun, and O. Geman, “Automated detection of Nonmelanoma skin cancer based on deep convolutional neural network,” *J. Healthcare Eng.*, vol. 2022, p. e6952304, 2022. DOI: [10.1155/2022/6952304](https://doi.org/10.1155/2022/6952304).
- [16] S. Bechelli and J. Delhommelle, “Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images,” *Bioengineering*, vol. 9, p. 97, 2022. DOI: [10.3390/bioengineering9030097](https://doi.org/10.3390/bioengineering9030097).
- [17] R. Kaur, H. GholamHosseini, R. Sinha, and M. Lindén, “Automatic lesion segmentation using atrous convolutional deep neural networks in dermoscopic skin cancer images,” *BMC Med. Imaging*, vol. 22, p. 103, 2022a. DOI: [10.1186/s12880-022-00829-y](https://doi.org/10.1186/s12880-022-00829-y).
- [18] Y. Wu, B. Chen, A. Zeng, D. Pan, R. Wang, and S. Zhao, “Skin cancer classification with deep learning: a systematic review,” *Front. Oncol.*, vol. 12, p. 893972, 2022.
- [19] A. A. Malibari, J. S. Alzahrani, M. M. Eltahir, V. Malik, M. Obayya, M. A. Duhayyim, A. V. Lira Neto, and V. H. C. de Albuquerque, “Optimal deep neural network-driven computer aided diagnosis model for skin cancer,” *Comput. Electr. Eng.*, vol. 103, p. 108318, 2022. DOI: [10.1016/j.compeleceng.2022.108318](https://doi.org/10.1016/j.compeleceng.2022.108318).
- [20] R. Patil and S. Bellary, “Machine learning approach in melanoma cancer stage detection,” *J. King Saud Univ. – Comput. Inf. Sci.*, vol. 34, pp. 3285–3293, 2022. DOI: [10.1016/j.jksuci.2020.09.002](https://doi.org/10.1016/j.jksuci.2020.09.002).
- [21] S. H. Kassani and P. H. Kassani, “A comparative study of deep learning architectures on melanoma detection,” *Tissue Cell*, vol. 58, pp. 76–83, 2019. DOI: [10.1016/j.tice.2019.04.009](https://doi.org/10.1016/j.tice.2019.04.009).
- [22] F. Grignaffini, F. Barbuto, L. Piazza, M. Troiano, P. Simeoni, F. Mangini, G. Pellacani, C. Cantisani, and F. Frezza, “Machine learning approaches for skin cancer classification from dermoscopic images: a systematic review,” *Algorithms*, vol. 15, p. 438, 2022. DOI: [10.3390/a15110438](https://doi.org/10.3390/a15110438).
- [23] A. Bassel, A. B. Abdulkareem, Z. A. A. Alyasser, N. S. Sani, and H. J. Mohammed, “Automatic Malignant and Benign skin cancer classification using a hybrid deep learning approach,” *Diagnostics*, vol. 12, p. 2472, 2022. DOI: [10.3390/diagnostics12102472](https://doi.org/10.3390/diagnostics12102472).
- [24] S. Malik and V. V. Dixit, “Skin cancer detection: state of art methods and challenges,” in A. Kumar and S. Mozar (eds.), *ICCCE 2021, Lecture Notes in Electrical Engineering*. Singapore: Springer Nature, 2022, pp. 729–736. DOI: [10.1007/978-981-16-7985-8_76](https://doi.org/10.1007/978-981-16-7985-8_76).
- [25] R. Kaur, H. GholamHosseini, R. Sinha, and M. Lindén, “Melanoma classification using a novel deep convolutional neural network with dermoscopic images,” *Sensors*, vol. 22, p. 1134, 2022b. DOI: [10.3390/s22031134](https://doi.org/10.3390/s22031134).
- [26] G. Alwakid, W. Gouda, M. Humayun, and N. U. Sama, “Melanoma detection using deep learning-based classifications,” *Healthcare*, vol. 10, p. 2481, 2022. DOI: [10.3390/healthcare10122481](https://doi.org/10.3390/healthcare10122481).
- [27] R. L. Araújo, F. H. D. de Araújo, and R. R. V. e. Silva, “Automatic segmentation of melanoma skin cancer using transfer learning and fine-tuning,” *Multimed. Syst.*, vol. 28, pp. 1239–1250, 2022. DOI: [10.1007/s00530-021-00840-3](https://doi.org/10.1007/s00530-021-00840-3).
- [28] H. Tabrizchi, S. Parvizpour, and J. Razmara, “An improved VGG model for skin cancer detection,” *Neural Process Lett.*, 2022. DOI: [10.1007/s11063-022-10927-1](https://doi.org/10.1007/s11063-022-10927-1).
- [29] H. Chang, “Skin cancer reorganization and classification with deep neural network,” 2017. DOI: [10.48550/arXiv.1703.00534](https://doi.org/10.48550/arXiv.1703.00534).
- [30] I. A. Alfi, M. M. Rahman, M. Shorfuzzaman, and A. Nazir, “A non-invasive interpretable diagnosis of Melanoma skin cancer using deep learning and ensemble stacking of machine learning models,” *Diagnostics*, vol. 12, p. 726, 2022. DOI: [10.3390/diagnostics12030726](https://doi.org/10.3390/diagnostics12030726).
- [31] A. Murugan, S. A. H. Nair, A. A. P. Preethi, and K. P. S. Kumar, “Diagnosis of skin cancer using machine learning techniques,” *Microprocess. Microsyst.*, vol. 81, p. 103727, 2021. DOI: [10.1016/j.micpro.2020.103727](https://doi.org/10.1016/j.micpro.2020.103727).
- [32] M. F. Jojoa Acosta, L. Y. Caballero Tovar, M. B. Garcia-Zapirain, and W. S. Percybrooks, “Melanoma diagnosis using deep learning techniques on dermatoscopic images,” *BMC Med. Imaging*, vol. 21, p. 6, 2021. DOI: [10.1186/s12880-020-00534-8](https://doi.org/10.1186/s12880-020-00534-8).
- [33] L. Talavera-Martínez, P. Biliboni, and M. González-Hidalgo, “Hair segmentation and removal in dermoscopic images using deep learning,” *IEEE Access*, vol. 9, pp. 2694–2704, 2021. DOI: [10.1109/ACCESS.2020.3047258](https://doi.org/10.1109/ACCESS.2020.3047258).