International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

SKIN CANCER DETECTION AND DIAGNOSIS USING DEEP CONVOLUTION NEURAL NETWORK

ALI ABDULLAH SALEH
Electrical and Computer Engineering,
Altinbas University, Istanbul, Turkey.
engalibayatie@gmail.com

NAMARIQ AYAD SAEED
Department of Computer Science,
University of Diyala
Diyala, Iraq.
gemy91gemy@gmail.com

OMAR ABDULLAH SALEH
Information Technology,
Altinbas University,
Istanbul, Turkey.
Omar.saleh@ogr.altinbas.edu.tr

Abstract

Cancer ranks as the world's second leading cause of mortality. Cancer is the fast formation of aberrant cells that expand past their usual borders and subsequently infiltrate neighboring body sections, spreading to other organs. Skin cancer, particularly melanoma, is regarded as one of the most severe kinds of cancer, with a significant increase in death rates due to a lack of understanding about the signs and prevention. Thus, early identification at an early stage is required to avoid the spread of cancer. Even professional doctors struggle to predict the early stages of skin cancer. This research aims to develop deep-learning models that can categorize dermal cell pictures and diagnose skin cancer at an early stage. In this study, a Deep Convolution Network (DSC-DCNN) system is used to identify and diagnose skin cancer at an early stage. In the preprocessing stage, dermoscopic images are used as input, and Gaussian Blur is used to remove noise and improve the image. The experimental analysis uses the PH2 dataset of 200 dermoscopic images. DCNN was used for classification, with an accuracy of around 92.45%.

Index Term: Skin Cancer, DCNN, CNN, Gaussian Filter, PH2, Neural Network.

I. Introduction

The term "cancer" refers to a wide range of illnesses that can impact any part of the body [1]. Skin cancer is the sixth most common type of cancer that is increasing

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
SKRGC Since 2012
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

internationally. In general, tissues are made up of cells, which make up the skin. Therefore, aberrant or unchecked cell development in the related tissues or other nearby tissues is the cause of cancer. Cancer may arise due to several factors, such as a weakened immune system, family history, exposure to ultraviolet light, etc. It is possible to classify this kind of erratic cell growth pattern as benign or cancerous. Benign tumours are a kind of malignancy that are typically regarded as harmless moles. On the other hand, malignant tumours are treated as life-threatening cancer. They can also cause damage to other tissues in the body. If cancer is not diagnosed at an early stage, it can spread through other organs of the body, making it extremely difficult to cure[2]. Skin cancer is classified into three types: melanoma, basal cell carcinoma, and squamous cell carcinoma [2]. Melanoma is the most unexpected. Skin cancer is diagnosed through clinical screening, thermoscopic analysis, histological testing, and biopsy. This approach is uncomfortable, intrusive, and requires laboratory testing, so it takes time. Furthermore, because of the similarities, doctors face the most significant obstacle in detecting and diagnosing cancer sickness [3]. Information technology plays a significant role in disease diagnosis.

Advances in information technology have created numerous prospects for improving healthcare delivery in disease diagnosis, management, and support [4].

Several strategies for detecting and diagnosing skin cancer have been proposed utilizing information technology. Some of them used machine learning techniques like SVM [5], Naive Bayes [6], and neural networks. The best outcomes in skin cancer identification and diagnosis have been obtained with deep learning approaches such as Fuzzy Neural Network [7] and CNN [8].

Skin cancer diagnosis using Deep Learning (DL) classifiers is an interesting avenue to pursue. Recent research has demonstrated that CNNs' layers' potential for abstraction allows them to automatically extract both low-level (edges and forms) and high-level (texture and semantics) data [9]. Consequently, Deep CNN will be used in this effort to detect and diagnose skin cancer.

In this work, we propose a system for early detection and diagnosis skin cancer using Deep CNN, which is one of the most effective deep learning algorithm for classifying skin cancer. Gauss Blur was used for noise removal and image smoothing as pre-processing step. Experimental results demonstrate the efficiency of the model over various melanoma skin cancer detection techniques.

The rest of the paper formed as follows. Section 2 illustrates the related works on skin cancer detection whereas section 3 demonstrates research methods. Section 4 describes the proposed DSC-DCNN system. Section 5 illustrates a discussion of experimental results obtained. Conclusions are provided in Section 6.

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

II. Related Work

The study of skin cancer diagnosis using image analysis has evolved greatly. Several strategies have been tested.

L. Bi et al. [10] employed an automatic melanoma identification technique for dermoscopic images based on multi-scale lesion-biased representation (MLR) and joint reverse classification (JRC). Closely connected histograms represent skin lesions. The JRC model provides unique new information for melanoma identification. The PH2 public database is used to evaluate and test the suggested approach. The results indicate an accuracy of 92%, sensitivity of 87.50%, and specificity of 93.13%.

Many strategies proposed in the literature for preparing melanoma images focused on eliminating hairs and enhancing contrast. Lee et al. [11] presented one such method, Dullrazor, for removing hair and picture artifacts. It is one of the most well-known apps for dermoscopic images. Soumya et al. [12] used 84 directional filters to eliminate hairs from images before segmenting them using an active contour-based technique. Finally, color correlogram and texture features are retrieved and fed into the Bayesian classifier. When tested on the PH2 database, the approach achieved a detection accuracy of 91.5%. Ajed et al. [13] used the Dullrazor approach to remove hairs and improve image quality. They then retrieved structural and textural data for melanoma and non-melanoma classification using an SVM classifier. The approach had an accuracy of 86.07%.

III. Materials and methods

A. *Dataset*

The dataset utilized in this study is PH2. A skin lesion dataset known as PH2[14] has been created for benchmarking and research purposes, allowing for a comparative examination of both classification and segmentation methods for dermoscopic pictures. Tuebinger Mole Analyzer equipment was used at Hospital Pedro Hispano's Dermatology Service to analyze similar cases at a magnification of 20 ×. The collection consists of 200 photos and 160 benign lesions, while the remaining 40 are 8-bit RGB color photographs with a resolution of 768x560 pixels, all available in jpeg format.

B. *Gaussian Blur*

The results and discussion may be presented separately, or in one combined section, and may optionally be divided into headed. The images obtained from the dermoscope could include undesired elements including hair, gel, and air bubbles. Image processing methods, including contrast enhancement and noise reduction, are used to prepare the dermoscopy images. Unwanted sounds are eliminated during the pre-processing stage called noise removal. The technique of contrast enhancement helps the image's features stand out more. Contrast manipulations entail adjusting an image's value range to boost contrast for improved quality [15]. Gaussian blur, sometimes referred to as Gaussian smoothing, results from applying a Gaussian function to blur an

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

image It is a commonly used effect in graphics software, usually intended to lessen detail and noise in images.

The "Gaussian Blur" results from convolutioning the image's pixels with a kernel characterized by a Gaussian function. In the discrete situation, the convolution is provided by [16], as seen in the following equation:

$$f * g[n] = \sum_{m=-\infty}^{\infty} f[m].g[n-m] \qquad (1)$$

The two-dimensional Gaussian function is the function that generates the kernel. The following is the definition of this function [17]

$$g(x,y) = \frac{1}{M}\sum f(x,y) e^{\frac{-(x-i)^2+(y-j)^2}{2\sigma^2}} \quad (i,j) \in S \qquad (2)$$

Where S is every pixel set in the neighborhood. And,

$$M = \sum e^{\frac{-(x-i)^2+(y-j)^2}{2\sigma^2}} \qquad (3)$$

C. Convolution Neural Network

One particular kind of neural network (NN) designed to extract visual features from images is the convolutional neural network (CNN). As of right now, this is the deep learning technique that performs the best at classifying images. Three layers make up the typical CNN: convolutional, pooling, and fully connected layers [18]. Most of the computing is done by the convolutional layer, which is the most significant. It is made up of kernels, which are weighted combinations that use the input images' visual qualities to learn new ones. This layer's output, a feature map, is created by convolving each kernel across the entire image. More formally, the kth output feature mapYk can be computed as follow [19].

$$y[i] = \sum_{k=1}^{k} x[i + r.k]w[k] \qquad (4)$$

Where x is the input image and w[k] represents the convolutional filter associated with the kth feature map.

The main purpose of the pooling layer is to minimize the size of the feature map. As such, it lowers the number of training parameters required for the subsequent layers and aids in managing overfit. It introduces non-linearity into the network in conjunction with a non-linear activation filter. Max pooling, in formal terms, chooses the largest element in each receptive field so that [19]

$$Y_{kij} = \max_{(p,q)\in R_{ij}} X_{kpq} \qquad (5)$$

Where the output of the pooling operation, associated with the kth feature map, is denoted by Ykij, xkpq denotes the element at location (p,q)contained by the pooling region Rij, which embodies a receptive field around the position (i, j).

The final feature map supplied by the preceding layer is connected to the fully connected layer, a conventional NN. In conclusion, the feature extractor comprises convolutional and pooling layers, and the

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
Since 2012
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

classifier is the fully-connected layer [18]. Increasing the depth and width of a deep neural network (DNN) is the most straightforward method of improving it.

D. *Evaluation Methods*

The most popular criterion for assessing the effectiveness of the suggested method is accuracy. It gauges how well the classifier can generate a diagnosis at a certain degree of accuracy [20]. We may see the accuracy formula in equation (6).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

IV. Diagnosis Skin Cancer based Deep Neural Network DSC-CNN System

The proposed system consists of four main stages, namely: load image, pre-processing, dataset split and shuffle, and classification. The model proposed is illustrated in Figure 1 utilizing a block diagram and every block is demonstrated in detail below

A. *Load Image*

Images of melanoma are obtained from the PH2 database [14]. These datasets are employed in research. The skin image from the dermoscope is provided as the input image.

B. *Pre-Processing*

Medical images are frequently prone to noise, primarily due to poor lighting, hair, and air bubbles. Artifacts are produced due to the noise being included in the photographs. These artifacts can skew the segmentation results, leading to erroneous detection outcomes. Eliminating noise is an important first step toward a precise diagnosis. Gaussian filtering is strongly advised to smooth the image since it eliminates speckle noise introduced during the acquisition procedure.

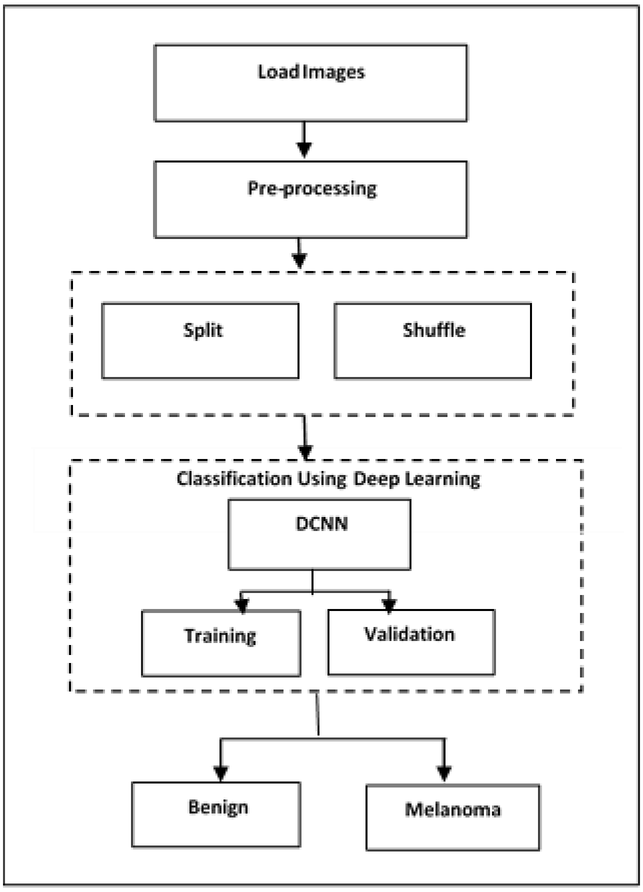


Fig.1 Proposed DSC-DCNN System

C. *Dataset Split and Shuffle*

The dataset is divided into two sections, the first containing 70% of the data needed to train the suggested system. In contrast, the remaining thirty percent has been allocated for testing the proposed system. The

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
Since 2012
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

"random. Shuffled" function in Python has partitioned and shuffled the dataset.

D. Classification

The task of classifying involves identifying related classes from the available data. This study classified benign lesions and melanoma using a deep convolution neural network. The weight-sharing feature of CNN, which lowers the number of trainable parameters in the network, aids in better generalization and prevents overfitting, is one of the primary arguments in favour of DCNN.

V. Results and Discussion

This section covers the model's performance over a range of parameters and includes a comparison analysis demonstrating the suggested methodology's superiority over the current melanoma diagnosis methods.

A. Testing Dataset

This paper shows that a deep learning approach to automatic image recognition-based skin cancer diagnosis is technically feasible. Making use of PH2 databases. There are 200 dermoscopy pictures of melanocytic lesions in this image set, 160 of which are non-melanoma (benign) and 40 melanoma (malignant). Of these, 60 photos were used for model validation and 140 images for training. All research uses 8-bit RGB color images at 768 x 560 pixels. The data was split up into training and validation ratios of 70:30.

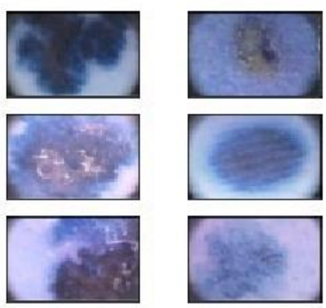
B. Simulation Specification

Because of Python's increasing renown as a scientific programming language and the abundance of free state-of-the-art image processing tools inside its ecosystem, the system is constructed utilizing the Keras framework on the Python 3 kernel. Our platform consisted of an Intel (R) Core (TM) i7-5500U CPU running at 2.40GHz with 16 GB of RAM.

C. Pre-Processing Result

Images are pre-processed to remove undesirable things that could lead to inaccurate results. The primary objective of noise removal methods is to mitigate the impact of body hair, which can be found in a significant portion of the dataset's photos. In certain instances, tiny lesions and textures surrounding the affected skin area may also lead to incorrect classification task assessment. For this reason, blurring is necessary to lessen the impact of such areas. Images are sharpened at this stage to prepare them for the next actions. Gaussian blur is used for noise reduction and picture improvement, as section 3.2 explains. Fig. 2 lists the pre-processing processes that were completed.

(a)



International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

(b)

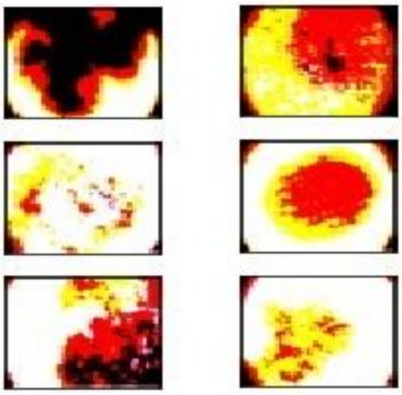


Fig 2. Pre-processing step. a original images b. filtered image

D. *Classification Result*

In this study, DCNNs function as iterative automatic feature extractors from dermal pictures, promoting the most discriminative characteristics in the image. Fully connected layers come next, which use the retrieved features to complete the classification task after convolutional layers. Sixty sample photos from the validation dataset were used to validate the model. To assess the effectiveness of our model, we have created a loss and accuracy curve. As the number of iterations increases, these curves demonstrate that accuracy rises and loss falls. Figures 4 and 5 display the accuracy and loss curves, respectively. We have also contrasted the performance of our model with previous research like L. Bi et al.[7], Soumya et al.[12],and Ajed et al. [13] shown in table 1. In these study have been worked on two classes, L. Bi et al.[10] reported 92% classification accuracy, Soumya. et al[12] reported 91.5% classification accuracy, Ajed et al.[13] reported 86.07% accuracy. We have achieved the highest classification accuracy of 92.45%.

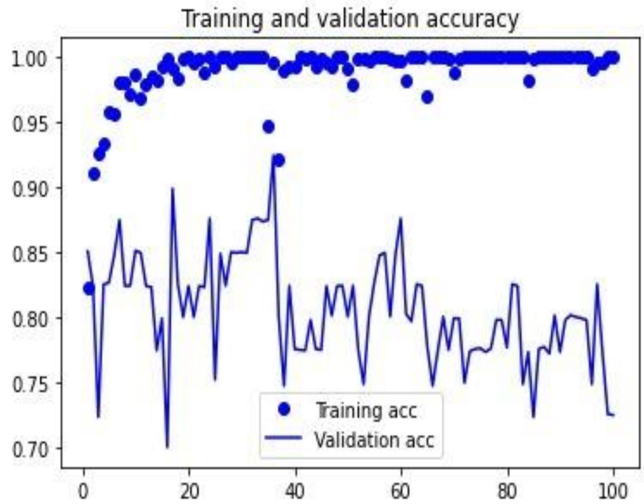


Fig.3 Categorical Accuracy of the Model

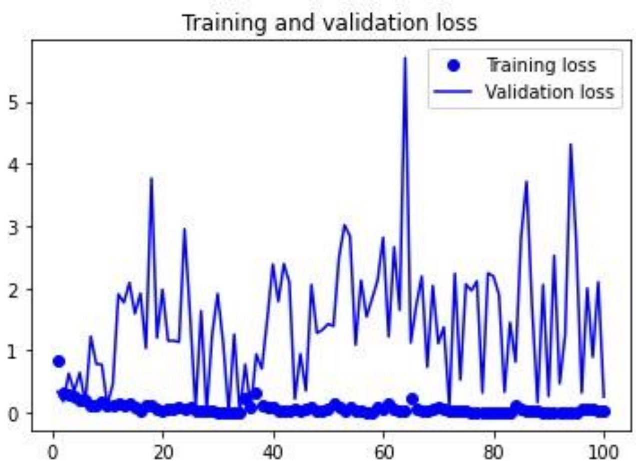


Fig.4 Training and Validation Loss

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

Table 1. Performance Comparison between the Suggested and Current Techniques

Authors	Method	Accuracy
L. Bi et al.[10]	MLR, JRC	92%
Soumya et al[12]	Active contour based method. and Bayesian classifier	91.5%
Ajed et al.[13]	Dullrazor method and Svm classifier	86.07%
Proposed method	Gaussian Blur filter and DCNN	92.45%

VI. Conclusions

The version of this template is V2. Most of the formatting instructions in this document have been compiled by Causal Productions from the IEEE LaTeX style files. Causal Productions offers both A4 templates and US Letter templates for LaTeX and Microsoft Word. The LaTeX templates depend on the official IEEEtran.cls and IEEEtran.bst files, whereas the Microsoft Word templates are self-contained. Causal Productions has used its best efforts to ensure that the templates have the same appearance.

Causal Productions permits the distribution and revision of these templates on the condition that Causal Productions is credited in the revised template as follows: "original version of this template was provided by courtesy of Causal Productions (www.causalproductions.com)".

References

[1] American Cancer Society. Skin Cancer Facts & Figures 2020. Atlanta. American Cancer Society; 2020. Accessed at https://www.cancer.org/content/dam/cancer-org/cancer-control/en/booklets-flyers/skin-cancer-fact-sheet.pdf on December 3, 2020.

[2] Guiling Sun, Rock CL, Thomson C, Gansler T, et al. American Cancer Society guideline for diet and physical activity for cancer prevention. CA Cancer J Clin. 2020;70(4). doi:10.3322/caac.21591.Accessed at https://onlinelibrary.wiley.com/doi/full/10.3322/caac.21591 on June 9, 2020.

[3] Haseeb Younis, Muhammad Hamza Bhatti, Muhammad Azeem, " Classification of Skin Cancer Dermoscopy Images using Transfer Learnin ", IEEE, 2019.

[4] L. Hudson, M. E. Cohen, "THE ROLE OF INFORMATION TECHNOLOGY IN DISEASE MANAGEMENT", Pmc of the 4th Annual IEEE Conf on Information Technology Applications in Biomedicine, UK IEEE, 2003.

[5] M. K. Monika, N. A. Vignesh, C. Usha, M. N. V. S. S. Kumar, and E. L. Lydia, "Skin cancer detection and classification using machine learning," *Mater. Today Proc.*, no. xxxx, 2020.

[6] Grzesiak-Kopeć K., Nowak L., Ogorzałek M. (2015) Automatic Diagnosis of Melanoid Skin Lesions

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 02, 2024. ISSN: 2348-6600
Reference ID: IJCS-496 PAGE NO: 3423-3432

Using Machine Learning Methods. In: Rutkowski L., Korytkowski M., Scherer R., Tadeusiewicz R., Zadeh L., Zurada J. (eds) Artificial Intelligence and Soft Computing. ICAISC 2015. Lecture Notes in Computer Science, vol 9119. Springer, Cham. https://doi.org/10.1007/978-3-319-19324-3_51

[7] F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, and A. C. Bovik, "Melanoma Classification on Dermoscopy Images using a Neural Network Ensemble Model," *IEEE*, pp. 1–11, 2016.

[8] A. G. C. Pacheco and R. A. Krohling, "The impact of patient clinical information on automated skin cancer detection," *Comput. Biol. Med.*, vol. 116, 2020.

[9] S.Ramya Silpa1, Chidvila V," A REVIEW ON SKIN CANCER". INTERNATIONAL RESEARCH JOURNAL OFPHARMACY, 4(8),pp.83-88, 2013.

[10] L. Bi *et al.*, "Automatic Melanoma Detection via Multi-scale Lesion-biased Representation and Joint Reverse Classification," *IEEE*, pp. 1055–1058, 2016.

[11] A. C. D. M. TIM LEE, VINCENT NC, RICHARD GALLAGHER, "A SOFTWARE APPROACH REMOVAL FROM IMAGES TO HAIR," *Pergamon*, vol. 21, no. 6, pp. 533–543, 1997.

[12] R. S. Soumya, S. Neethu, T. S. Niju, A. Renjini, and R. P. Aneesh, "Advanced Earlier Melanoma Detection Algorithm Using Colour Correlogram," in *International Conference on Communication Systems and Networks (ComNet)* |, 2016, no. July, pp. 190–194.

[13] F. Adjed, S. Jamal, S. Gardezi, F. Ababsa, I. Faye, and S. C. Dass, "Fusion of structural and textural features for melanoma recognition," *IET Comput. Vis.*, vol. 12, no. 2, pp. 185-195r, 2016.

[14] Mendonça T, Ferreira PM, Marques J, Marcal ARS, Rozeira J. PH2 - a dermoscopic image database for research and benchmarking. 35th International Conference of the IEEE Engineering in Medicine and Biology Society, July 3-7, Osaka, Japan. 2013https://www.fc.up.pt/addi/ph2%20database.html.

[15] S. Saravanan, B. Heshma, A. V. A. Shanofer, and R. Vanithamani, "Materials Today : Proceedings Skin cancer detection using dermoscope images," *Mater. Today Proc.*, 2020.

[16] Tiwari, S., Shukla, V. P., Singh, A. K., & Biradar, S. R. (2013). Review of motion blur estimation techniques. Journal of Image and Graphics, 1(4), 176-184.

[17] A. Noori, A. Al-jumaily, and A. Noori, "Comparing the Performance of Various Filters on Skin Cancer Images," *Procedia - Procedia Comput. Sci.*, vol. 42, no. 02, pp. 32–37, 2014.

[18] W. Rawat and Z. Wang, "Deep Convolutional Neural Networks for Image Classification :A Comprehensive Review," *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, 2017.

[19] B. Harangi, “Skin lesion classi fi cation with ensembles of deep convolutional neural networks,” *J. Biomed. Inform.*, vol. 86, no. June, pp. 25–32, 2018.

[20] Saeed N.A., Al-Ta’i Z.T.M. (2020) Heart Disease Prediction System Using Optimization Techniques. In: Al-Bakry A. et al. (eds) New Trends in Information and Communications Technology Applications. NTICT 2020. Communications in Computer and Information Science, vol 1183. Springer, Cham. https://doi.org/10.1007/978-3-030-55340-1_12