Eliminating Racial Bias at the Time of Detection Melanoma using Convolution Neural Network (CNN)

Md. Abdullah Al Noman Majumder[1], Eimon Hossain Taief[2], Md. Nurul Amin Bhuiyan[3], M. F. Mridha[4], Aloke Kumar Saha[5]

University of Asia Pacific, Bangladesh[1,2,3,5]

Bangladesh University of Business and Technology, Bangladesh[4]

Corresponding author: Md. Abdullah Al Noman Majumder, Email: abdullahalnoman642@gmail.com

Melanoma considers deadly cancer that can cause the death of a person if not distinguished at an initial stage. Although Melanoma is most common in white skin people and can be detected at an early stage using AI, white skin people have a much lower death rate from this cancer. But when black skin people have Melanoma, AI can't detect it at an early stage because most of the time, the machines are trained with Dermoscopic pictures of white people, which leads to a higher mortality rate for black skin people. As a result, people don't want to trust the AI system at the time of Melanoma detection. In this paper, we proposed a model with whatever black or white skin it can easily detect using machine learning. In this case, we will use the Convolution Neural Network (CNN) of machine learning to detect Melanoma at an early stage so that the death rate caused by Melanoma cancer can be reduced. The proposed method can detect Melanoma with an accuracy of 88.9% for both skin people which may significantly decrease the mortality rate.

Keywords: Melanoma, Machine Learning, CNN, Dermoscopic picture, Artificial Intelligence.

1 Introduction

Melanoma is the cancer of the skin that has the most elevated mortality rate of all dermatological cancer. Melanoma is particular from other skin cancers in

2021. In Raju Pal & Praveen K. Shukla (eds.), *SCRS Conference Proceedings on Intelligent Systems*, 271–285. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-93-91842-08-6-27

that it more often than not an expansion of pigmented cells, even though not continuously; Melanoma can moreover originate in numerous parts of the mucosal track, counting the head and neck and the GI tract and the bottoms of the palms and the soles, which is usually doesn't happen in other sorts of skin cancers. Melanoma is the preeminent perilous frame of skin cancer since it is most likely to proliferate to other organs. Different types of skin cancers rarely spread from the exterior of the skin's surface, and Melanoma's ability to metastasize makes it the foremost dangerous. In stage one, Melanoma contains a cure rate above 95% on the off chance that you'll be able to identify it in an early stage. Melanoma is one of the significant common cancers in young and adults. It is observed that Ultra-violate (UV) rays are contemplated as one of the most common bearers of skin cancer [1]. Statistics show that no other cancer get treated as skin cancer per year [2]. Usually, melanoma can be cured if it is detected at an early stage, it increases the chances of cure, and the longer it takes to identify, the more likely it is that the melanoma will spread throughout the body from which it is not possible to recover later. Statistics show that there are 99 % chances of getting a cure if melanoma can be detected at an initial stage, increasing the survival rate of more than five years [22]. In statistics, it has shown that in a calendar year, almost 250,000 new cases of Melanoma skin cancer were found across all over the world [6]. In 2019, Melanoma skin cancer ranked 5th in the United States of America (USA) [7]. Skin people are usually affected by melanoma; thus, when AI is applied to recognize melanoma cancer, skin people get trained by AI. Therefore, when it comes to detecting melanoma, cancer in the skin people worked very well, but in black skin people, it often fails to detect melanoma cancer [3]. As a result, the death rate of black skin people in melanoma is high. You can see that there is a racial bias in the time of training machine; thus, people don't want to trust AI. Trust in AI is considered a two-dimensional build comprising trust and doubt, where trust is related to sentiments of calmness and security, and doubt includes fear [4]. Unarguably, the trust could be a complex social preparation with an assortment of components deciding the degree to which people trust AI agents. But when the term detection comes to AI, which depends on only the machine's accuracy, the more accuracy can predict melanoma precisely. Dermoscopy picture is the main ingredients to detect the Melanoma or other types of skin cancer [5].

There are several model build on detecting melanoma using AI and almost all of them trained by using white skin people Dermoscopic picture which create a racial bias at the time of detecting melanoma. Our model will help to remove this racial bias using CNN where machine will be trained by both skin people.

The rest of the paper sketched out as follows. Section II discussed about the existing systems and related works of Melanoma. In section III, we proposed our model to eliminate the racial bias in the time of Melanoma detection. Section IV contains the experimental results that are performed to evaluate CNN models. Section V discusses future works on this topic. In Section VI, we concludes the paper.

2 Literature Review

Skin cancer like Melanoma, BSC are expanding at a tremendous rate due to the changes in climate in the present world. The statistic shows that more than 4500 cases of Melanoma can be found as the depletion of the ozone layer by 10%, which may be considered as a causes significant rate of Melanoma due to changes of climate [13]. The connection between daylight exposure and melanoma is less clear and questionable [8][9]. Melanoma, the original deadly shape of human skin cancer, appears at an emotional rate of increment around the world and causes a few millinery deaths per year within the United States of America (USA)[10][11]. According to epidemiological evidence discontinuous presentation to solid daylight, especially amid childhood and youthful youth, may advance the arrangement of melanoma [12][13].

Numerous analysts have been working on the Computer vision approach for skin cancer detection. For segmentation of skin lesions within the input picture, existing frameworks either utilize manual, semi-automatic, or fully automatic border detection strategies. There are some research group who was endeavored how to classify Melanoma utilizing numerous data-sets where they used some popular machine learning algorithm such as SVM, CNN, ANN etc. In [13] the authors have utilized the method of partitioning the input picture into various clinically critical districts using the Euclidean distance transform for the extraction of color and texture features. Numerous developed Computer-based diagnosing models analyze skin cancer; these depend on future extraction and classification. For illustration, ABCD rules, pattern generation, seven-point checklist, and utilizing well-developed classifiers like SVM are confined to dermoscopy or histopathology picture information [14][15]. There are some researcher detect melanoma using machine learning algorithm [17] where some discuss about how to detect melanoma using machine learning [16]. Adekanmi Adegun and Serestina Viriri [23] discussed traditional-based classification techniques and the recently developed state-of-the-art techniques for Melanoma detection. They also discussed some methods for the classification of Melanoma, such as Basal Cell

Carcinoma, Actinic Keratosis, Squamous Cell Carcinoma, Seborrheic Keratosis, Solar Lentigo, Dermatofibroma, etc.Wang Haoxiang et al. [24] proposed an innovative frame work termed as the MC-SVM (Multi Class Support Vector Machine) to devise the scheduling of the workflow in the cloud paradigm.

3 Proposed System

We make a proposal of a modern and proficient model to detect whether a person has Melanoma or not without doing any medical methodology. Our model is based on CNN, which can be more accurate and precise than the physician, where racial bias can be removed. The proposed model is used to detection of Melanoma in any type of skin person. Our algorithm approach to Detect Melanoma utilizing image processing where the input of the models will be the skin lesion image, which is supposed to be a Melanoma lesion. CNN's use image recognition and classification in arrange to distinguish objects, recognize faces, etc. They are consisting of neurons with learn-able weights and biases. As we are used CNN to detect Melanoma to eliminate racial bias, we try to make a structure of how it should work. Its look like Fig 1.

Basically, the computer doesn't see the image like as a human, so that they reads the image as three channels R, G, B (Red, Green, Blue) image by reading pixel value if the image size is M X N X 3 then it has M rows, N columns and three channels of RGB. To make this image processing happen, the Convolution Neural Network algorithm is mainly used recently. CNNs are deep learning algorithms that take input images and convoluted them with filters or kernels to extricate features.

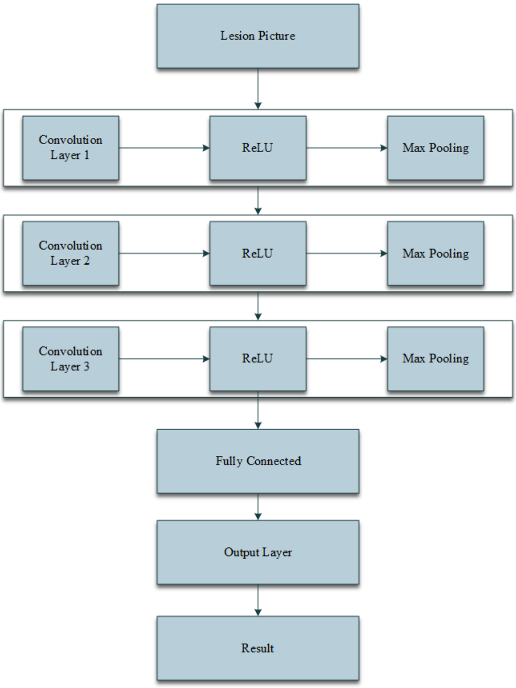

Figure 1: The figure demonstrate how CNN works in our proposed model. There are multiple hidden layer which perform nonlinear transformations of the inputs entered into the network.

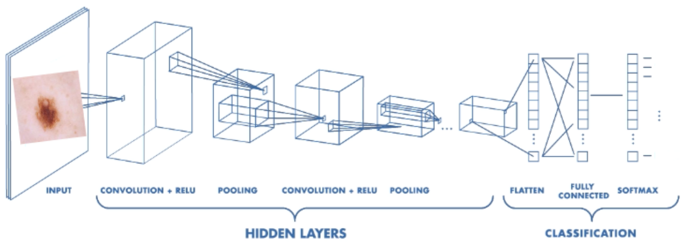

Figure 2: This figure shows the architectural model of CNN model in Machine Learning. Here full working process of CNN which used in our model are shown.

In Fig. 2 the full working process of CNN which used in our model are shown. Here you can see Convolution layer and Activation layer are combined together in hidden layers where Pooling function works after Convolution and Activation function works. In classification, our model predict whether the given lesion has Melanoma or not.

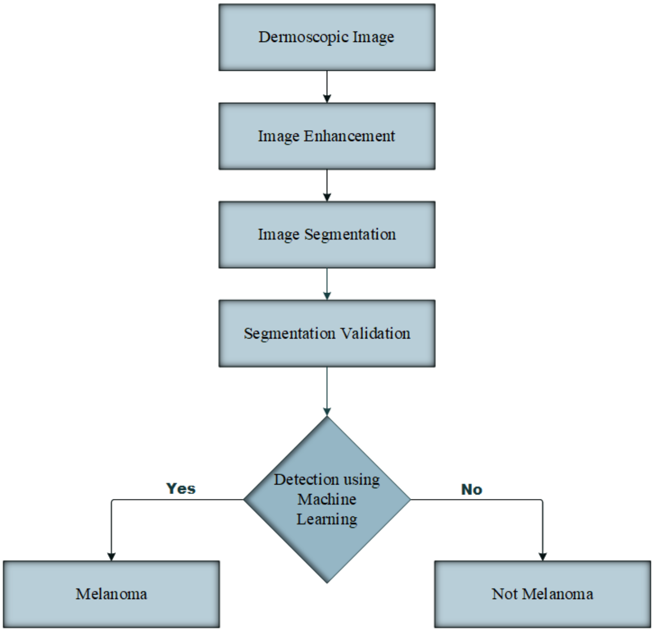

Figure 3: This figure illustrate the flowchart of our model how to detect Melanoma using Machine Learning (CNN).

Before going to detection, some pre-processing are conducted on the Dermoscopic picture. They are briefly discussed below: From Fig. 3 we can see that some contents are there to discuss for better understanding. They are:

i. Image Enhancement

ii. Image Segmentation

3.1 Image Enhancement

The noise and artifacts were evacuated using a 2D anisotropic diffusion filter at the initial stage [19]. The anisotropic diffusion filtering process is utilized to diminish picture noise without influencing the imperative components of the picture details (Fig. 4), more often than not edges, lines, or elective points fundamental for the perception of the picture. After sharpening the picture, color steadiness using gray world normalization was applied on three channels together [20]. Hairs in Dermoscopic photographs act as equivocalness. The hairs were identified using gray world normalization and impainting. The recognized hairs were expelled using impainting strategy.

In our model, we utilize the concept of Image Enhancement as there are so much noise in our Dermoscopic picture, so that we can remove the noise (like hair) from the picture and get a proper image of the lesion to make the algorithm efficient.

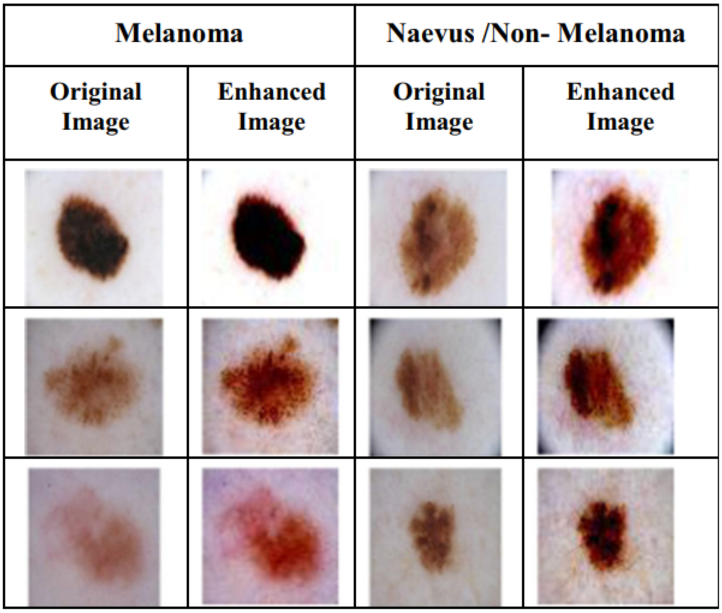

Figure 4: This figure shows, noise (like hair) of the image removed from picture the after Image Enhancement.

From Fig. 4 we can see that after the image enhancement of white skin people noises are removed from the skin lesion. But when we try to enhance the Dermoscopic picture of black skin people noise can't reduce after one or two enhancement. So we have to enhance the picture so many times to remove the noise from the picture.

In Fig. 5 we can see the Dermoscopic picture of black skin people need to enhance several times to reduce the noise so that our model can train with a higher accuracy.

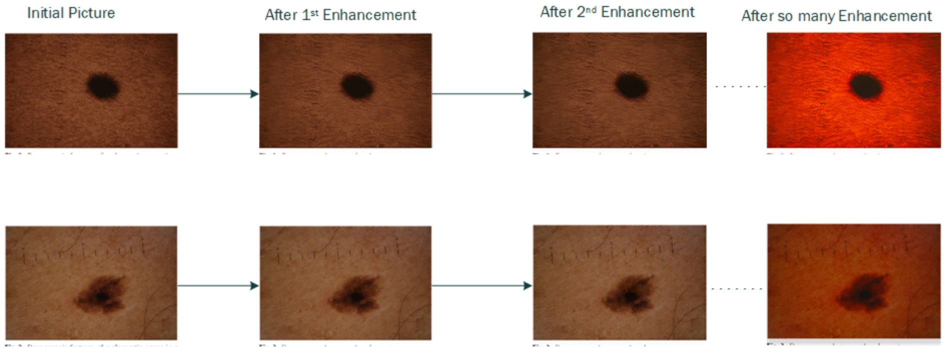

Figure 5: This figure shows that Image enhancement of Black skin people processing is going through many steps as we can not find the noise free image after 1st iteration.

After doing these image enhancements specially for black skin people (Fig. 5), image segmentation will help the divide the image into many segments and can achieve the important part of the picture rather than full image.

3.2 Image Segmentation

Image Segmentation is the method by which a digital image is divided into different subgroups (of pixels) called Image Objects, which can diminish the complexity of the image, and in this way analyzing the picture becomes simpler. In other word, Image segmentation is the method of break down a digital image into different particular districts containing each pixel(sets of pixels, moreover known as super-pixels) with similar attributes. By the partition of an image into segments, we can process only the image's important segments instead of processing the entire image. Image segmentation's objective is to alter the representation of an image into something more significant and easier to analyze.

After the image enhancement, we will divide the picture into different parts called segments. It's not a great thought to process the whole image simultaneously as there will be districts or regions within the image that don't contain any information. By partitioning the picture into segments, we can utilize the vital segments for processing the image (Fig: 6), so that the efficiency of image processing can improve. From Fig. 6(a) we can see there are so many object are there in the picture which are not needed, so after the segmentation of picture from Fig. 6(b) we can able to reduce the unwanted segments from the image.

After the image segmentation, segmentation validation will take the place where our model will check whether the segmented picture is valid or not. After these we will apply the CNN model so that the machine can learn about the lesion picture and further it's can predict for an input image whether the picture whatever the skin color (White skin or Black skin) has deadly Melanoma or other skin cancer.

Figure 6: This figure illustrate image segmentation using only vital segments which are needed for further procedure.

After all process applied our model will look like this:

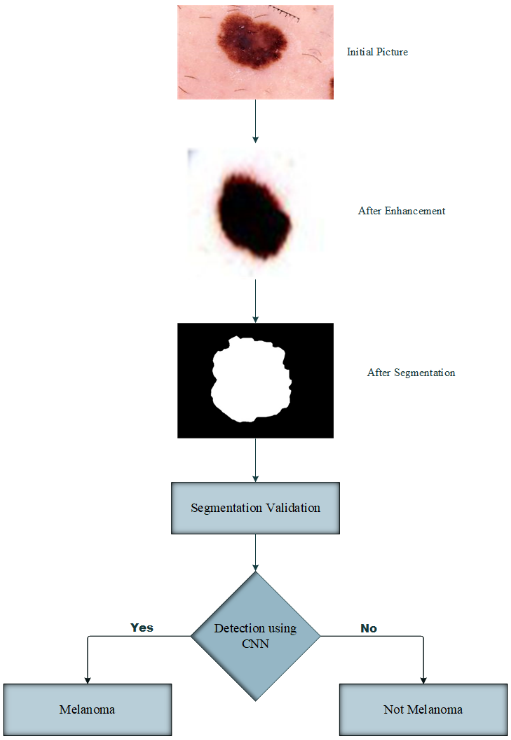

Figure 7: The figure demonstrate the full working process of our model.

Data-set

To discover the precise result for a given skin tissue picture, we took the dataset of affected skin pictures from ISIC. The accepted set is isolated to create a training dataset and testing dataset fundamental within the developing model. We had developed this model utilizing the Convolution neural networks algorithm in deep learning by utilizing python language.

Table 1: The table shows the Dataset Division of training, validation and testing for the detection of Melanoma.

Skin Lesion	Melanoma	Not Melanoma	Total
Training	5341	5102	10443
Training (%)	51.1%	48.9%	100%
Validation	1781	1502	3283
Validation (%)	54.2%	45.8%	100%
Testing	1781	1680	3461
Testing (%)	51.5%	48.5%	100%

For this model we used 10443 pictures to train our machine where 5341 pictures contain melanoma and 5102 pictures contain not melanoma. To make our model more accurate to reduce racial bias we trained the machine with a huge number of Dermoscopic pictures.

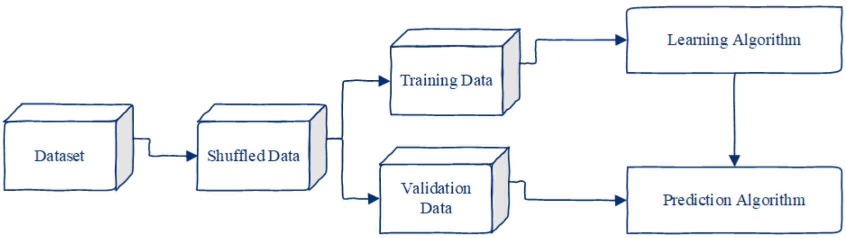

Figure 8: This figure illustrates how data-set are used for train the model.

From Table-1, we had given a point-by-point depiction of the number of Melanoma and non-melanoma training, validation, and testing data collections.

Programming Language

To create this model, we utilized python, where different libraries or modules like Keras, Tensorflow, Numpy are being used to develop this proficient model. This Python code is executed in Jupyter Notebook[21] as it is an open-source web application.

4 Experimental Results

Our model had tested by utilizing different irregular input pictures from our taken dataset by altering the number of epochs, and the number of move tests for epochs at that point diverse comes about as organized underneath in Table 2. From Table 2 when the number of epochs is the same, but the sample of each epoch is different, then the difference in accuracy can be noticed. The reason for this difference is that whenever the number of epochs is the same, but the sample of epochs is more, then the machine gets more samples per epochs to train and increase the accuracy.

$$Precision = \frac{TP}{TP+FP}$$

$$Accuracy = \frac{TP+TN}{TOTAL}$$

Where, TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative. Accuracy is calculated later by using these equation.

Table 2: The table shows the accuracy for sample per dataset and Number of epochs needed.

Number of Epochs	Sample Per Epochs	Accuracy
8	100	78.2%
10	80	79.7%
12	68	81.3%
5	160	76.2%
10	200	83.5%
10	1044	86.4%
30	348	88.9%

The model has trained the image sample with 30 epochs and 348 transitions per each epoch it predicted as Melanoma with 88.9% accuracy shown in Table 2 .

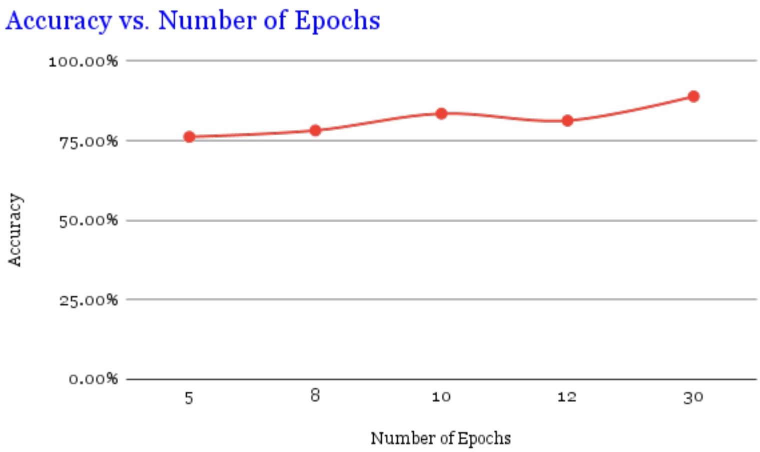

Figure 9: This figure illustrates how accuracy changed when the number of epochs are different.

5 Conclusion

After a few research, we came to know Melanoma is the foremost dangerous type of skin cancer. Melanoma may show up as a new spot or alter within an existing mole or spot's appearance. Melanoma treatment depends on whether cancer has spread to other regions of the body. But if it is possible to detect Melanoma cancer at an early stage, it might be cured. Though Melanoma can be detected in an early stage for white skin people using AI, our model which based on CNN will help to detect Melanoma for both white skin people and black skin people with an accuracy of 88.9%, which is sometimes better than the physician. Due to the lack of equipment assets, we used fewer epochs; we can move forward its efficiency by expanding the number of epochs and the number of transitions per epoch.

6 Future Work

There are numerous analysts have been working on the Computer vision approach for skin cancer detection. Though our work eliminate racial bias in time of detection Melanoma but the accuracy may increase using Deep Neural Network (DNN) which can be a phenomenal work in this sector.

References

[1] National Toxicology Program (2002) Ultraviolet radiation related exposures: broad-spectrum ultraviolet (UV) radiation, UVA, UVB, UVC, solar radiation, and exposure to sunlamps and sunbeds. Rep Carcinog Carcinog Profiles 10:250–254

[2] Siegel, R. L., Miller, K. D. and Jemal, A. (2018). Cancer statistics 2018. CA Cancer J Clin 68(1):7–30.

[3] Noor, P. (2020). Can we trust AI not to further embed racial bias and prejudice?. *BMJ*, 2020;368:m363.

[4] Sethumadhavan, Arathi. (2018). Trust in Artificial Intelligence. *Ergonomics in Design: The Quarterly of Human Factors Applications*, 27. 106480461881859.

[5] Bafounta, M. L., Beauchet, A., Aegerter, P. and Saiag, P. (2001). Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests. *Arch Dermatol*,137(13):43–50.

[6] Bray, F. et al. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394-424.

[7] Siegel, R. L., Miller, K. D. and Jemal, A. (2019). *CA: a cancer journal for clinicians*, 69(1):7–34.

[8] Tucker, M. A. (2008). Is sunlight important to melanoma causation?. *Cancer Epidemiol Biomarkers Prev.*,17:467–468.

[9] Jhappan, C., Noonan, F. P. and Merlino, G. (2003). Ultraviolet radiation and cutaneous malignant melanoma. *Oncogene*,22(20):3099–3112.

[10] Gilchrest, B. A. et al. (1999). The pathogenesis of melanoma induced by ultraviolet radiation. *The New England Journal of Medicine*, 340:1341–1348.

2021. In Raju Pal & Praveen K. Shukla (eds.), *SCRS Conference Proceedings on Intelligent Systems*, 271–285. Computing & Intelligent Systems, SCRS, India.
https://doi.org/10.52458/978-93-91842-08-6-27

[11] Leiter, U. and Garbe, C. (2008). Epidemiology of melanoma and non-melanoma skin cancer–the role of sunlight. *Advances in Experimental Medicine and Biology*, 624:89–103.

[12] Tatalovich, Z. et al. (2006). The objective assessment of lifetime cumulative ultraviolet exposure for determining melanoma risk. *Journal of Photochemistry and Photobiology B: Biology*, 85:198–204.

[13] Celebi, M. et al. (2007). A methodological approach to the classification of dermoscopy images. *Computerized Medical Imaging and Graphics*, 31:362–373.

[14] Zhang, N. et al. (2020). Skin cancer diagnosis based on optimized convolutional neural network. *Artificial Intelligence in Medicine*, 102:101756.

[15] Karthiga, M., Priyadarshini, R. K. and Bazila B. A. (2020). Malevolent Melanoma diagnosis using Deep Convolution Neural Network. *Research Journal of Pharmacy and Technology*, 13(3):1248-1252.

[16] Vocaturo. (2019). Machine Learning Techniques for Automated Melanoma Detection. in *IEEE International Conference on Bioinformatics and Biomedicine.*

[17] Sanketh, R. S. et al. (2020). Melanoma Disease Detection Using Convolutional Neural Networks. in *4th International Conference on Intelligent Computing and Control Systems.*

[18] Roslin, S. E. (2020). Classification of melanoma from Dermoscopic data using machine learning techniques. *Multimedia Tools and Applications*, 79(5):3713-3728.

[19] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639.

[20] Van, D. W. J. and Schmid, C. (2006). Coloring local feature extraction. in *9th European Conference on Computer Vision*, 334–48.

[21] Jupyter[Online].http://jupyter.org/

[22] Skin Cancer Organization (https://www.skincancer.org/)

[23] Adegun, A. and Viriri, S. (2021). Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art. *Artificial Intelligence Review*, 54(2):811-841.

[24] Haoxiang, W. and Smys, S. (2020). MC-SVM Based Work Flow Preparation in Cloud with Named Entity Identification. *Journal of Soft Computing Paradigm*, 2(2):130-139.