International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 02, ISSN No. 2455-2143, Pages 186-191
Published Online June 2023 in IJEAST (http://www.ijeast.com)

SKIN DISEASE DETECTION USING PYTHON AND DEEP LEARNING

Usha Dhankar, Seema Jain
Assistant Professor,
HMR Institute of Technology and Management, New Delhi, India

Saba Zaidi
Assistant Professor,
Panipat Institute of Engineering and Technology, Panipat, Haryana, India

Shubham Chauhan
Student,
HMR Institute of Technology and Management, New Delhi, India

Abstract: **Skin diseases are dangerous and highly contagious especially melanoma, basal cell carcinoma and toxic epidermal necrolysis (TEN). These skin diseases can be cured if detected early. The fundamental problem with this is that only experienced dermatologists can recognize and classify such conditions. Doctors may misclassify the disorder and prescribe the wrong drug for the patient. The paper proposed a skin disease detection tool based on image processing, machine learning and deep learning techniques. The proposed tools are non-invasive, easy-to-use and accurate to identify appropriate skin conditions. The patient must provide images of the infected skin area as input. Also, with the help of image model training and high-precision algorithms, skin diseases can be detected with a very low percentage error. In this review, based on the accuracy, we used two different real-time skin detection methods & compared convolutional neural networks (CNN) and random forests (RF). Real-time test results are also displayed.**

Key Words: **Skin disease, Convolutional Neural Network, Random Forest, Feature Extraction, Deep Learning.**

I. INTRODUCTION

Skin diseases, also known as dermatological conditions, are a group of different diseases that affect the skin, hair and nails. Common examples include acne, eczema [1], psoriasis, rosacea and skin cancer. Acne is a common skin condition that occurs when hair follicles become clogged with sebum and dead skin cells [2]. Occurring most commonly on the face, neck, shoulders and chest, it is characterized by the presence of pimples. Acne can be treated with a combination of over-the-counter and prescription medications [3], as well as maintaining a healthy diet and avoid touching the face. It can be treated with lifestyle changes such as Eczema, also known as atopic dermatitis, is a chronic inflammatory skin disease that causes dry, itchy and red skin. It is most common in children but can occur in adults as well. Eczema is usually treated with moisturizers, over-the-counter or prescription [4]-[6] topical corticosteroids and sometimes oral medications such as antihistamines. It is a chronic autoimmune disease. It occurs most commonly on the scalp, elbows, and knees, but other areas of the body can also be affected. It is typically treated with topical agents such as creams, ointments and gels as well as phototherapy, oral drugs and in severe cases, biologics. Rosacea is a chronic skin disease that [7] causes flushing, hot flashes and visible blood vessels. It is most common in fair-skinned people and usually affects the cheeks, nose, forehead and chin. Rosacea can be treated with topical and oral medications [8], as well as lifestyle changes such as avoiding triggers such as spicy foods, alcohol and the sun. Skin cancer is the most common type of cancer in the United States [9], [10]. It is caused by abnormal growth of skin cells and can take many forms such as basal cell carcinoma, squamous cell carcinoma and melanoma. Other symptoms, which may vary, include persistent pain, red or scaly patches on the skin and lumps or bumps on the skin. It is important to note that most skin conditions are not serious but can lead to complications if left untreated. If you suspect a skin condition, seek medical attention and treatment. is important. Regular self and professional examination are also recommended for early detection and prognosis of skin cancer.

As per study, there are currently about 8 million citizens in the UK suffering with skin diseases. Skin diseases don't just hurt the skin [11], it has an impact on people's daily lives and can destroy a person's self-confidence. Skin diseases have a serious impact on the mental health of patients. It leads to loss of trust [12] and can even drive patients into depression. Therefore, skin diseases can be fatal. So, it is necessary to detect skin diseases early [13] and prevent their spread. Human skin is an unpredictable and most challenging area [14], [15] due to the complexity of its irregularities, damaged structures, bruises, tone, dense hair present and other mitigating features. Early detection of skin diseases has proven to be inexpensive and available in remote areas. Identifying the infected areas of the skin and recognizing the nature [16] of the disease can help in early detection. Disease detection depends on many factors, including the parameters considered in disease detection. We first choose a picture, apply a filter to remove noise, slice the image to extract key characteristics [17], execute attribute extraction based on the input parameters [18] and finally apply the proper classifier used to categorize illnesses. A picture must be converted into digital form in order to be processed digitally so as to produce an improved image or extract valuable information from it. An image is used as the input and an image with the same characteristics is used as the output [19]. The majority of the input sampled two-dimensional signals used in image processing models correspond to a fixed signal processing strategy. It is a technology that is commonly utilized nowadays and has various uses in the commercial world [20]. Moreover, it is a developing field of study in engineering and computer science.

II. LITERATURE REVIEW

Deep CNN has lately been the method of choice for features extraction and image classification. With a high-performance GPU, a model can learn on a big dataset to provide precise outcomes. The classic approaches employing deep CNN got higher efficacy than people in object categorization, according to a variety of studies [21]-[23] utilizing ImageNet [24]. Recently, Esteva et al. [25] developed a strategy for classifying all types of skin diseases by modifying the VGG16 and VGG19 models used to facilitate network training. Their network considerably outperformed social skills in their studies, achieving Top-1 and Top-3 classification performance of 60.0% and 80.3%, accordingly. They also urged others to employ a similar strategy in order to get greater results.In light of the absence of labeled training dataset, Tajbakhsh et al. [26] described that it is preferable to use a pre-trained infrastructure instead of training a deep CNN from the bottom up. They used photographs from other medical disciplines to rectify the skin disorder difficulties using a pretrained network rather than training a deep CNN from the start. Deep CNN was used by Giotis et al. [27] to create a decision support system that made use of a new collection of characteristics, including colour, morphological and lesion topography. Haenssle [28] developed a technique for classifying dermoscopy pictures with discrete diagnostic categories employing deep CNN. ECOC SVM, constructed using deep CNN, was created by Dorj et al. [29] to arrange skin cancer pictures into four disease criteria. A deep CNN image classifier was put out by Han et al. [30] to categorize the diagnostic pictures of 12 different types of skin diseases. The database of four different types of skin conditions included in all trials, carcinoma, was classified using an imaging classification approach by Mohamed et al. [31]. Almansour and Jaffar [32] developed a method to categorize carcinoma using k-means clustering and Support Vector Machine (SVM); they also presented the comparison outcomes. By applying AdaBoost MC independently, German et al. [33] described the skin cancer diagnosis technique. For melanoma identification, an information from a separate class of skin lesions is employed. For employing the many sets of characteristics including visual diagnosing traits, hue, lesion appearance, damaged area and damage extent for carcinoma, Ioannis et al. [27] built a support network using the image processing technique and the deep CNN.

III. METHODOLOGY

Our proposed procedure includes 2 processes:

A. PREPROCESSING

Image preprocessing is step one to identify the affected regions. Several steps are performed during the preprocessing stage to make the image suitable for the attribute extraction process. Anomalies in the input image are detected and preprocessed for the following purposes:

- To avoid uneven illumination
- To enhance the contrast between image background pixels and exudate
- To eliminate the sound in the image

In this research work, the techniques used for the preprocessing phase are:

Image resizing: There are numerous methods for resizing images. The nearest neighbor interpolation technique is a simple way to lengthen the image. With this, the nearest pixel in the output is used to replace each pixel. This indicates that there are several pixels of the same color for scaling purposes. Although remapping can happen when fixing lens distortion or rotating the image, image scaling is required if the overall pixel count needs to be raised or lowered. Increasing the area of pixels in an image makes it appear as though finer details are visible when magnified.

Color transformation: Images are captured in RGB (red, green, blue) format. Grayscale is the spectrum of gray tones without any discernible color. Black is the darkest color. This indicates that neither light is reflected nor transmitted. White, which perfectly transmits or reflects light at all visible wavelengths, is the lightest color that is possible. Equal brightness levels in transmitted and reflected light for

International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 02, ISSN No. 2455-2143, Pages 186-191
Published Online June 2023 in IJEAST (http://www.ijeast.com)

the three basic colors—red, green, and blue—represent intermediate shades of gray. The brightness levels of the red (R), green (G), and blue (B) components of transmitted light (such as graphics on a computer screen) are each represented by a decimal or binary value ranging from 0 to 255.00000000 to 11111111. Red, Green, and Blue (RGB) grayscale images have R = G = B for each pixel. The number indicating the primary color's level of lightness is exactly related to the lightness of gray. White is represented by R=G=B=255 or R=G=B=11111111, whereas black is represented by R=G=B=0 or R=G=B=00000000Because the binary representation of grayscale uses 8 bits, this imaging technique is known as 8-bit grayscale. In rare circumstances, three extra parameters are defined for grayscale instead of the RGB or CMY color models. Color, saturation, and lightness are these. Each pixel in a grayscale image has an apparent hue and saturation of 0, respectively. The only aspect of a pixel that may change is brightness (apparent brightness). Brightness is at least 0 (black) and at most 100. (white).

B. FEATURE EXTRACTION

Applying a number of filters to the image allows a CNN to extract features from it. These filters are intended to find particular motifs and characteristics in photos. B. Edge or Texture. As the filter moves over the image, it produces a feature map that highlights the areas of the image where the detected features reside. In order to extract increasingly complicated features, CNNs then apply many layers of these filters, each layer building on the feature map created by the previous layer. A CNN's final output is a collection of feature maps that reflect the essential elements of an image and may be applied to tasks like object recognition and picture categorization.

To extract features for this study, we used the histogram of oriented gradients. A feature descriptor called HOG or histogram of directional gradients, is used in computer vision to identify objects in pictures and videos. In CNN, his HOG features can be used as input to the network instead of raw pixels. The first step in extracting HOG features is to split the image into small cells, after which each cell's gradient orientations are computed as a histogram. In order to create a feature vector for each cell, these histograms are then concatenated. The CNN can receive this feature vector as input. The CNN then applies a series of filters to the feature vector to produce a feature map that emphasizes the regions of the image where the detected features reside. CNNs extract increasingly complex features through multiple layers of filters. The final output of a CNN is a set of feature maps representing key features of an image, which can be utilized for tasks like object detection and image categorization.

IV. CONVOLUTIONALNEURALNETWORK

An example of a deep learning neural network is the convolutional neural network (CNN), which is frequently employed in applications involving computer vision like object and picture recognition. CNNs are made to analyze data with a structure that resembles a matrix, like a picture, which is made up of a pixel grid. An input layer, many hidden units and an output layer are only a few of the layers that make up CNNs. Convolutional layerscapture characteristics from the source picture and pooling layers that are employed to down-sample the extracted features created by the convolutional layers, are two examples of the hidden layers in a CNN.

In order to identify certain structures or characteristics in the source image, like borders or textures, the convolutional layers employ a series of filters. These filters create image features when combined with the source image and they are sent via a non-linear activation function like ReLU. The characteristic maps' spatial dimensions are reduced whilst also containing the most crucial data thanks to the pooling layers. The ultimate product of the CNN is a collection of extracted features that reflect the most crucial aspects of the picture and may be applied to applications like object recognition or image categorization.

V. RANDOM FOREST

For classification and regression tasks, Random Forest is a form of ensemble learning that builds various decision trees during training and produces classes that represent the means of individual classes (classification) or mean forecasts (regression). In a decision tree, data is continuously split into a series of branches and leaves based on certain decision criteria. Each leaf represents a class or value of the target variable. With the help of random selections of features and training data, random forest creates several decision trees. Each decision tree in the random forest produces a prediction as input is fed into it throughout the prediction process. Random Forest then obtains the mode or average of all predictions made by Decision Tree. Due to the randomness of the data and feature subsets used in each tree, individual trees are likely to make different predictions, and combining those predictions often outperforms a single tree. You get Random forests are widely used in various machine learning applications such as image classification, speech recognition, and bioinformatics.

VI. ACCURACY IMPROVISATION

In order to anticipate skin diseases, precision improvisation is crucial.

Following actions were done to increase the algorithm's accuracy:

International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 02, ISSN No. 2455-2143, Pages 186-191
Published Online June 2023 in IJEAST (http://www.ijeast.com)

- **Increase the training dataset size:** Using a larger CNN may learn from a greater variety of examples thanks to the training dataset, potentially enhancing generalization. performance.
- **Data enrichment:** Data augmentation techniques such as mirroring, rotation, cropping and scaling can be used to artificially increase the size of the training dataset and introduce more variation to the training data.
- **Hyper parameter tuning**: CNN accuracy can be improved by tuning hyper parameters such as learning rate, stack size and number of filters. A technique called grid search or random search can be used to find the best set of hyper parameters for a given data set.

Figure 1: Comparison of Loss and Accuracy with Training and Validation Loss

VII. RESULT AND DISCUSSIONS

Since the algorithm is now accurate and precise the disease detection is more stable and correct.

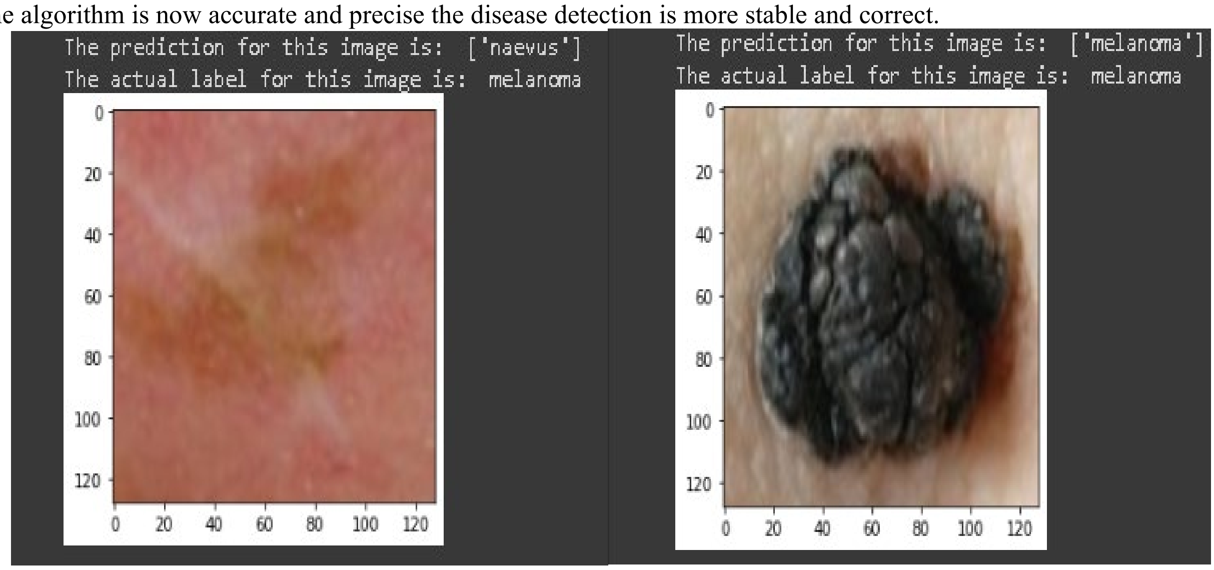

Figure 2: Sample images taken for prediction

Eventually after trying this algorithm on around 1230 images. Sometimes the prediction matches with the actual label of the image, which proves that it identifies the disease correctly (If and only if the actual label of the disease is for the correct skin disease). The data set used here is bin rushed and consists of four classes of disease. The total number of images is 280. It was tested with a 4-class test and showed an accuracy of 90.02%. To get more disease classes, we split the images into 8 classes. The convolutional neural network algorithm showed an overall accuracy of 88.5%, while other multiclass classification algorithms failed to exceed 50%. We also tested different scenarios for login pages and different types of images. Our algorithm gave better results in most cases. This may not be the exact replacement of the current but with some improvements it can potentially become one's reliable detection tool.

VIII. CONCLUSION AND FUTURE SCOPE

A crucial first step in lowering mortality, disease transmission and the emergence of skin diseases is skin disease detections. Clinical procedures for detecting skin diseases are very expensive and time consuming. Imaging method aids in the early development of automated dermatological screening systems. The characterization of skin diseases relies heavily on feature extraction. In this paper, we created a detection technique employing the techniques Random Forests (RF) and Convolutional Neural Networks (CNN). Iteratively training the model's pictures on the disease dataset improves the algorithm's accuracy. This paper will help detect the skin diseases that people are suffering from at a very early stage. It saves you money and if you don't use it as your primary tool, it's also useful as a second opinion.

Future scopes of improvement in present methodologies are:

- All sorts of skin diseases should be identified using a similar paradigm.
- Support for the development of multilingualism to improve usability.
- To increase the multi-platform introducing iOS compatibility and demonstrating capability.

IX. REFERENCE

[1]. Hay, Roderick, Sandra E. Bendeck, Suephy Chen, Roberto Estrada, Anne Haddix, Tonya McLeod, and AntoneMahé. "Skin diseases." Disease Control Priorities in Developing Countries. 2nd edition (2006).

[2]. Ayo, Femi Emmanuel, Roseline OluwaseunOgundokun, Joseph Bamidele Awotunde, Marion Olubunmi Adebiyi, and Abidemi Emmanuel Adeniyi. "Severe acne skin disease: A fuzzy-based method for diagnosis." In Computational Science and Its Applications– ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part VI 20, pp. 320-334. Springer International Publishing, 2020.

[3]. Leyden, James J., James Q. Del Rosso, and Guy F. Webster. "Clinical considerations in the treatment of acne vulgaris and other inflammatory skin disorders: focus on antibiotic resistance." Cutis 79, no. 6 Suppl (2007): 9-25.

[4]. Saunders, Charles W., Annika Scheynius, and Joseph Heitman. "Malassezia fungi are specialized to live on skin and associated with dandruff, eczema, and other skin diseases." PLoS pathogens 8, no. 6 (2012): e1002701.

[5]. Magin, Parker J., Jon Adams, Gaynor S. Heading, and C. Dimity Pond. "Patients with skin disease and their relationships with their doctors: a qualitative study of patients with acne, psoriasis and eczema." Medical Journal of Australia 190, no. 2 (2009): 62-64.

[6]. Hamnerius, Nils, Ann Pontén, Ola Bergendorff, Magnus Bruze, Jonas Björk, and Cecilia Svedman. "Skin exposures, hand eczema and facial skin disease in healthcare workers during the COVID-19 pandemic: a cross-sectional study." Acta dermato-venereologica 101, no. 9 (2021): adv00543-adv00543.

[7]. Picardo, Mauro, and Monica Ottaviani. "Skin microbiome and skin disease: the example of rosacea." Journal of clinical gastroenterology 48 (2014): S85-S86.

[8]. Millikan, Larry. "Recognizing rosacea: could you be misdiagnosing this common skin disorder?" Postgraduate medicine 105, no. 2 (1999): 149-158.

[9]. Linares, Miguel A., Alan Zakaria, and Parminder Nizran. "Skin cancer." Primary care: Clinics in office practice 42, no. 4 (2015): 645-659.

[10]. Gloster Jr, Hugh M., and Kenneth Neal. "Skin cancer in skin of color." Journal of the American Academy of Dermatology 55, no. 5 (2006): 741-760.

[11]. Barankin, Benjamin, and Joel DeKoven. "Psychosocial effect of common skin diseases." Canadian Family Physician 48, no. 4 (2002): 712-716.

[12]. Basra, Mohammad KA, and Muhammad Shahrukh. "Burden of skin diseases." Expert Review of Pharmacoeconomics & Outcomes Research 9, no. 3 (2009): 271-283.

[13]. Jowett, Sandra, and Terence Ryan. "Skin disease and handicap: an analysis of the impact of skin conditions." Social science & medicine 20, no. 4 (1985): 425-429.

[14]. Wu, Z. H. E., Shuang Zhao, Yonghong Peng, Xiaoyu He, Xinyu Zhao, Kai Huang, Xian Wu et

al. "Studies on different CNN algorithms for face skin disease classification based on clinical images." IEEE Access 7 (2019): 66505-66511.

[15]. Sanz-Bernardo, Beatriz, Rey Suckoo, Ismar R. Haga, NajithWijesiriwardana, Alice Harvey, Sanjay Basu, Will Larner et al. "The acquisition and retention of lumpy skin disease virus by blood-feeding insects is influenced by the source of virus, the insect body part, and the time since feeding." Journal of Virology 96, no. 15 (2022): e00751-22.

[16]. Anand, Vatsala, Sheifali Gupta, and Deepika Koundal. "Skin disease diagnosis: challenges and opportunities." In Proceedings of Second Doctoral Symposium on Computational Intelligence: DoSCI 2021, pp. 449-459. Springer Singapore, 2022.

[17]. Li, Ling-Fang, Xu Wang, Wei-Jian Hu, Neal N. Xiong, Yong-Xing Du, and Bao-Shan Li. "Deep learning in skin disease image recognition: A review." IEEE Access 8 (2020): 208264-208280.

[18]. Abunadi, Ibrahim, and Ebrahim Mohammed Senan. "Deep learning and machine learning techniques of diagnosis dermoscopy images for early detection of skin diseases." Electronics 10, no. 24 (2021): 3158.

[19]. ALEnezi, Nawal Soliman ALKolifi. "A method of skin disease detection using image processing and machine learning." Procedia Computer Science 163 (2019): 85-92.

[20]. Ajith, Archana, Vrinda Goel, Priyanka Vazirani, and M. Mani Roja. "Digital dermatology: Skin disease detection model using image processing." In 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 168-173. IEEE, 2017.

[21]. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, ‘‘ImageNet: A large-scale hierarchical image database,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 248–255.

[22]. S. Ioffe and C. Szegedy, ‘‘Batch normalization: Accelerating deep network training by reducing internal covariate shift,’’ 2015, arXiv:1502.03167. [Online]. Available: https://arxiv.org/abs/1502.03167

[23]. K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ 2014, arXiv:1409.1556. [Online]. Avail- able: https://arxiv.org/abs/1409.1556

[24]. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, ‘‘Ima- geNet large scale visual recognition challenge,’’ Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Apr. 2015.

[25]. A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, ‘‘Dermatologist-level classification of skin cancer with deep neural networks,’’ Nature, vol. 542, no. 7639, pp. 115–118, Jan. 2017.

[26]. N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, ‘‘Convolutional neural networks for medical image analysis: Full training or fine tuning?’’ IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1299–1312, May 2016.

[27]. I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman, and N. Petkov, ‘‘MED-NODE: A computer-assisted melanoma diagnosis sys- tem using non-dermoscopic images,’’ Expert Syst. Appl., vol. 42, no. 19, pp. 6578–6585, Nov. 2015.

[28]. L. Oakden-Rayner, ‘‘Reply to ‘Man against machine: Diagnostic perfor- mance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists’ by haenssle et al.’’ Ann. Oncol., vol. 30, no. 5, pp. 1836–1842, May 2019.

[29]. U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, ‘‘The skin cancer classifi- cation using deep convolutional neural network,’’ Multimedia Tools Appl., vol. 77, pp. 9909–9924, Feb. 2018.

[30]. S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, ‘‘Clas- sification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm,’’ J. Investigative Dermatol., vol. 138, no. 7, pp. 1529–1538, Jul. 2018.

[31]. M. Anas, K. Gupta, and S. Ahmad, ‘‘Skin cancer classification using K-means clustering,’’ Int. J. Tech. Res. Appl., vol. 5, no. 1, pp. 62–65, 2017.

[32]. E. Almansour and M. A. Jaffar, ‘‘Classification of Dermoscopic skin cancer images using color and hybrid texture features,’’ Int. J. Comput. Sci. Netw. Secur., vol. 16, no. 4, pp. 135–139, 2016.

[33]. G. Capdehourat, A. Corez, A. Bazzano, R. Alonso, and P. Musé, ‘‘Toward a combined tool to assist dermatologists in melanoma detection from dermoscopic images of pigmented skin lesions,’’ Pattern Recognit. Lett., vol. 32, no. 16, pp. 2187–2196, Dec. 2011.