iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35299

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Skin Cancer Detection and Classification using KNN Technique

Apeksha R Swamy[1], Dr. Prasanna Kumar S C[2]
[1, 2]R V College Of Engineering, India

Abstract: Skin cancer is a major health issue worldwide. Skin cancer detection at an early stage is key for an efficient treatment. Lately, it is popular that, deadly form of skin cancer among the other types of skin cancer is melanoma because it's much more likely to spread to other parts of the body if not identified and treated early. The advanced medical computer vision or medical image processing take part in increasingly significant role in clinical detection of different diseases. Such method provides an automatic image analysis device for an accurate and fast evaluation of the sore. The steps involved in this project are collecting skin cancer images from PH2 database, preprocessing, segmentation using thresholding, feature extraction and then classification using K-Nearest Neighbor technique (KNN). The results show that the achieved classification accuracy is 92.7%, Sensitivity 100% and 84.44% Specificity.
Keywords: Melanoma skin cancer, Image processing, Features, Thresholding, K-Nearest Neighbour technique.

I. INTRODUCTION

The skin is the largest part of the human body. Skin cancer is a disease that affects our skin cells, which are particularly susceptible to ultraviolet radiation. It starts when normal cells change and grow out of control and form a mass known as a tumor. This can become carcinogenic or harm to the body. Skin cancer affects people of all skin types, including those with dark skin. But white people are more on to have skin cancer. It is caused by the formation of abnormal cells that are able to move to different lump of the body. There are four forms of skin cancer namely Basal cell carcinoma, Squamous cell carcinoma, Merkel’s cancer cell, Melanoma skin cancer. We Indian’s, commonly get Melanoma skin cancer.

A. Melanoma Skin Cancer

Melanoma is a form of skin damage that begins in melanocytes. Skin color is made of melanocytes. It is identified as melanin, which gives the skin its dark color and ensures deeper layers of the skin from the destructive effects of the sun. When melanocytes develop malignant mutations, melanoma forms a disease, growing out of control and aggressively attacking the nearby tissues. Melanoma can affect the skin or spread to other organs and bones through blood or lymph system. There is no specific definition of all melanomas, but the brightness of UV rays may increase the chance of getting melanoma. Reducing your exposure to UV rays will reduce your chances of developing melanoma. On all accounts, especially in people under the age of 40, the incidence of melanoma increases. According to the statistics made by skin cancer collaboration, globally, it is estimated that the number of new melanoma cases diagnosed in 2021 will increase by 5.8 percent. The number of melanoma deaths is expected to increase by 4.8 percent in 2021.[9]

B. Signs And Symptoms Of Melanoma Skin Cancer

1) Melanomas will grow on your body anywhere i.e., on the back, knees, arms and face.
2) Melanomas may also appear in areas that do not get much sunlight such as below your knees, palms, and nails.
3) Changes in existing mole.
4) Improving the formation of other colored or unusual pigments on your skin.
5) Usually, melanoma does not start as a mole. In any case, it is also possible for normal skin manifestations.

Figure [1]: Mole and the Melanoma

II. LITERATURE SURVEY

Noel B. Linsangan et al, [1] in their work, the centre of attention was on the geometrical characteristics of the lesion of the information in order to detect and classify skin cancer. Geometric features were characterized by the asymmetry boundaries and diametric parameters of the ABCD Rule Dermoscopy. This study explores the three classes of skin lesions as a malignant melanoma, benign form of melanoma and the unknown. The classification of the injury, the skin is carried out pictures by means of the k-nearest neighbor algorithm (k-NN), and displays an accuracy of 90.0% for the test.

Vidya M et al, [2] proposed an automatic lesion detection method to find the accuracy of the system, efficiency and overall performance. In this paper, the authors used hybrid feature extraction method to classify skin cancer as cancerous or non-cancerous and automatic detection of skin sore is done with ABCD rule, feature extraction using HOG and GLCM technique and classification using machine learning algorithms. Here, for classification, different machine learning techniques such as SVM, KNN and Naïve bayes was proposed by collecting the images from ISIC datasets. On comparing with all the classification methods, SVM has highest accuracy of 97.8%, Sensitivity was 86.2% and data obtained was 85% using KNN.

Minakshi Waghulde et al, [3]in their project, they take the help of image-processing techniques to find the melanoma in a picture. First of all, they applied preprocessing method to make the image sound free. Median filters are used to filter the image and the filtered image is transformed into HSI color image form. At last, they used PNN classifier to classify the image as normal or abnormal. Total 20 images were used for the proposed algorithm.

Sreena S et al, [4] this paper talks about the techniques and algorithms used in order to perform the analysis of the data. The solution uses the combination of K-Means clustering technique and thresholding for skin sore segmentation. For preprocessing, median filter and sobel operator is used. The system achieves an accuracy of 90% efficient.

Deva Ratnasari et al, [5] in this paper, the authors aim and apply to find basal cell carcinoma skin cancer on android gadget. Here, the authors discuss ways to get better of this problem by using the release of the ABCD and KNN feature as a category. Built with a framework that can be used by all science circles, this app has an easy-to-use android application template, so users do not need trained lT skills. Based on the results of the tests performed, the results of the release of the feature get an accuracy of 91.6%.

Prashant bhati et al, [6] this paper suggests and explains the process, of classifying skin lesions as bad or dangerous. Their research main is to develop, use, evaluate the app, by taking clinical images using high-speed cameras, to find melanoma early. The system is based on the calculation of TDS (Dermoscopic Score) with pre-processing and OSTU classification of skin lesion classification as non-cancerous and cancerous. The project has been tested in a database of 30 lesions. The test result shows that the forward steps and separation of OTSU and ABCD algorithm give good results.

Pratik Dubal et al, [7] in this paper, they have suggested how to find and identify how dangerous or serious skin lesions are based on photographs taken with standard cameras. Images are classified, features extracted using ABCD law and the Neural network is trained to distinguish lesions with a high degree of accuracy. A trained neural network obtained 76.9% complete separation of 463 image databases, divided into six different categories.

A. Scope Of Work

Skin cancer is on the rise. According to the WHO, 132,000 melanoma skin cancers come throughout the world every year. If the disease spreads the prognosis is worse, so it is good to be treated early. [8]

III. METHODOLOGY

The main purpose of the project is to differentiate and identify images of skin cancer. Given a picture of skin cancer, the first step improves the image, the second step is processing, the third stage separates the skin cancer image using morphological function, the fourth step forms a descriptive element that accurately describes a large set of data and the final step is classification using KNN algorithm.

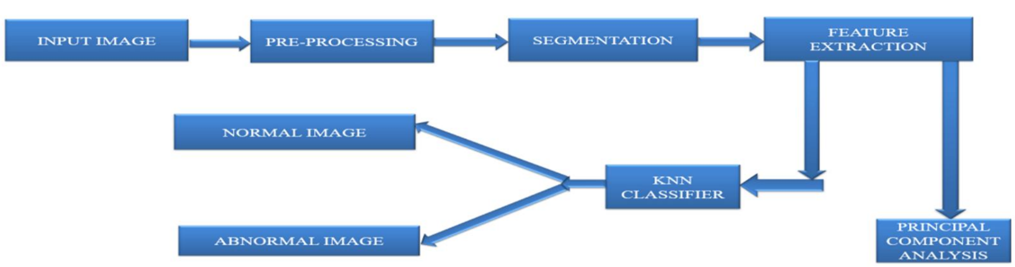


Figure [2]: Block diagram of skin cancer detection and classification using KNN technique.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Figure [2] shows a block diagram for skin cancer detection and classification using the KNN procedure. Image processing is a way of doing certain tasks in an image, to get an improved image or to obtain some useful information from it. Here, we will do the step-by-step process:

A. Input Image

The input color image is 3-Dimensional array. For the purpose of the project, I have collected a PH2 database of skin cancer images as an input image. The color input skin cancer image is converted into a gray scale image in order to enhance the image quality. Figure [3] shows the input color image.

Figure [3]: Color input image

B. Preprocessing

Preprocessing is done to reduce friction and improve the damaged image of the skin for next process. The reason for these methods is primarily to improve the image so that we can get more certainty and understanding in the diagnosis of skin disease.

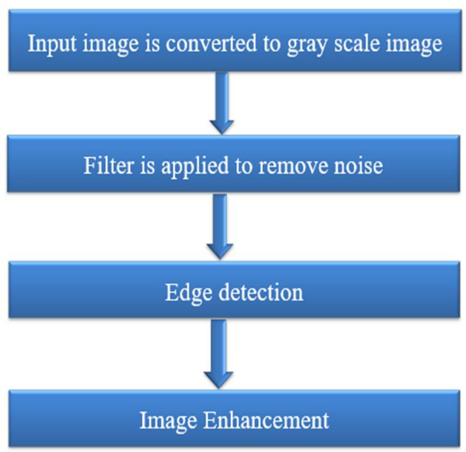


Figure [4]: Preprocessing steps

The steps in figure [4] are proposed here since some of the obtained images are not homogenous due to inappropriate illumination during image acquisition. Median filter is done to remove noise in an image in order to reduce the effect of hair cover on the skin in the final image which is used for classification.

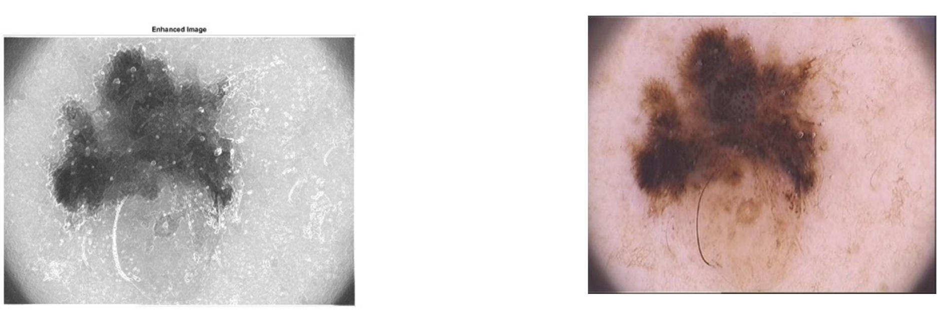


Figure [5]: Image before and after preprocessing process

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

C. Segmentation

The second step is to detect and segment the region of interest (ROI), which represents the affected area. The segmentation step includes: image thresholding, image filling, image opening and then converting extracted region to binary form. Here, we make use of Otsu thresholding technique since the ROI is homogenous. Now, morphological operations such as image opening and filling is done. Image filling is used to remove background pixels from binary image to make the ROI clear. Image opening is used to remove small objects from the binary image. Finally, the obtained binary image and the enhanced image is multiplied to get the segmentation image.

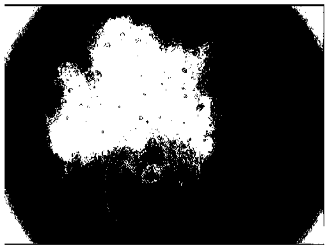


Figure [6]: Threshold image

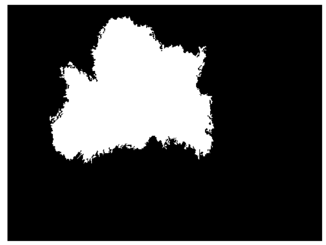


Figure [7]: Binary image

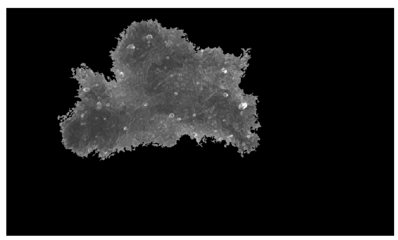


Figure [8]: Segmented image

D. Feature Extraction

Feature extracting is a method of minimizing size when the first set of raw data is reduced to control groups for processing. After extracting ROI in segmentation part, the structural parameters, colour parameters and textural parameters are extracted from ROI for further classification. Total 24 features are extracted here to get a good classification output to distinguish the cancerous or non-cancerous.

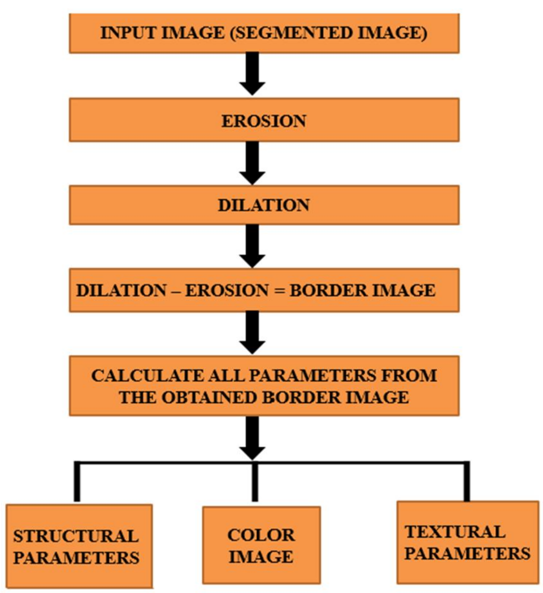


Figure [9]: Flowchart of feature extraction process

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Erosion reduces or degrades objects in a binary image and Dilation is the operation of enlarging or consolidating objects with a binary image.

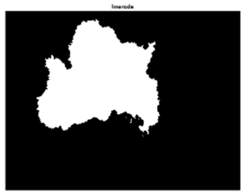


Figure [10]: Eroded image

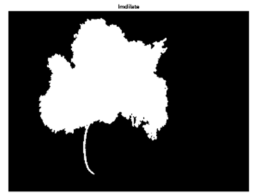


Figure [11]: Dilated image

1) *Border*
   a) Now, to get border, dilated image is subtracted from eroded image.
   b) Dilation – Erosion = Border
   c) From the obtained border image, structural, textural parameters are calculated.

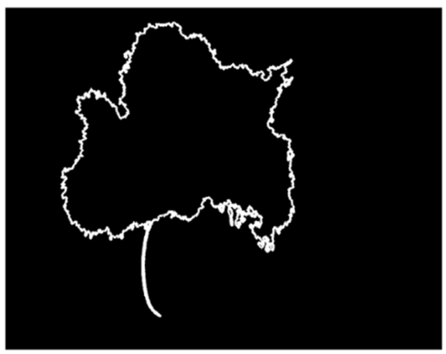
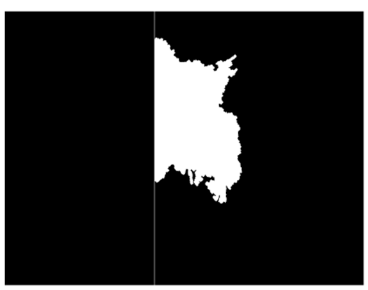


Figure [12]: Border image



BORDER RIGHT (BR)

BORDER LEFT (BL)

BORDER UPPER(BU)

BORDER BOTTOM(BB)

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

a) Phase 1: In this phase, we deal with binary, where the structural parameters are calculated. Total 11 structural parameters are calculated and it is shown in table [1].
b) Phase 2: Here, we go with the original image in RGB format, which contains three channels of colors i.e Red, Green, Blue. Color features are taken out by finding the standard deviation, mean value and variance.
c) Phase 3: In this phase, we deal with textural parameters such as entropy, energy, homogeneity and contrast by applying gray level co-occurance matrix (GLCM).

	FEATURE NAME	FEATURE VALUES	
f1	Solidity	0.0793	
f2	Extent	0.0591	
f3	equidiameter	124.4546	
f4	Compactness	0.1634	
f5	Aspect ratio	0.8293	Structural parameters
f6	Center of gravity	0.8297	
f7	Circularity	0.0268	
f8	BL – BR/ Filled Area	0.0970	
f9	BB-BU/ Filled Area	- 0.1738	
f10	BL / BR	1.2142	
f11	BB / BU	0.7050	
f12	Red – Mean value	23.4195	
f13	Red – variance	15.7913	
f14	Red – standard deviation	14.7476	
f15	Green – Mean value	7.5757	
f16	Green – variance	7.5756	Color image check
f17	Green – standard deviation	7.5748	
f18	Blue – Mean value	3.8573	
f19	Blue – variance	3.8573	
f20	Blue – standard deviation	3.8572	
f21	Entropy	0.7856	
f22	Energy	0.6233	Textural parameters
f23	Homogeneity	0.9955	
f24	Contrast	0.0194	

Table [1]: Feature parameters and their values for single image

Inference

- In table [1], 24 feature values of single image are tabulated.
- Feature f1 to f11 is structural parameters, f12 to f20 is color parameters and f21 to f24 is textural parameters.
- We can observe that, red colour feature parameter has higher values because the cancer images are reddish in color.

E. Classification Using KNN

KNN is most popular machine learning technique that can be used for classification of data. K-Nearest Neighbor classifier is easy to use, works powerful with large data set and it is quick.

The training and testing data are given to KNN classifier where each class is calculated by using nearest distance. A confusion matrix is a matrix which represents number of images consists of both normal and abnormal images. Here, total 173 images are taken, where 96 images are taken as testing images and 77 images are taken as training images.

To evaluate the system performance, different parameters are calculated:

1) Ture Positives (tp) is the number of abnormalities detected in patients;
2) Ture Negative (tn) is the number of normal people detected as normal;
3) False Positives (fp) refer to the number of healthy persons misdiagnosed as abnormal;
4) False Negative (fn) is equivalent to the number of persons missed.
5) Accuracy is the ratio of correctly predicted observation to the total observations.

$$\text{Accuracy} = \frac{\text{true positive} + \text{true negative}}{\text{total population}}$$

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

- Sensitivity – finds the proportion of positives that have been correctly found.
- Specificity – finds the proportion of negatives that have been correctly identified.

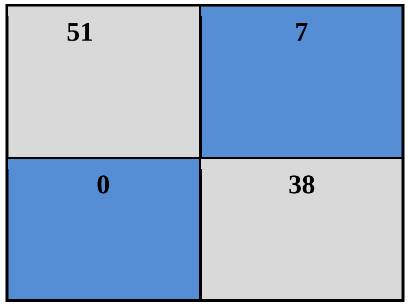


Figure [13]: Confusion matrix of KNN classifier

From the figure [13], we can say that number of true positives is 51, true negative is 38, false positive is 7, false negative is 0.

IV. RESULTS

The proposed method is applied to skin cancer images collected from PH2 database. The dataset consists of around 200 images in which total 173 images are taken, where 96 images are taken as testing images and 77 images are taken as training images. The system achieves 92.70% accuracy, 100%, Sensitivity and 84.44% Specificity. Figure [14] shows overall system performance values.

Data	Values
Accuracy	92.70%
Sensitivity	100%
Specificity	84.44%

Figure [14]: System performance values

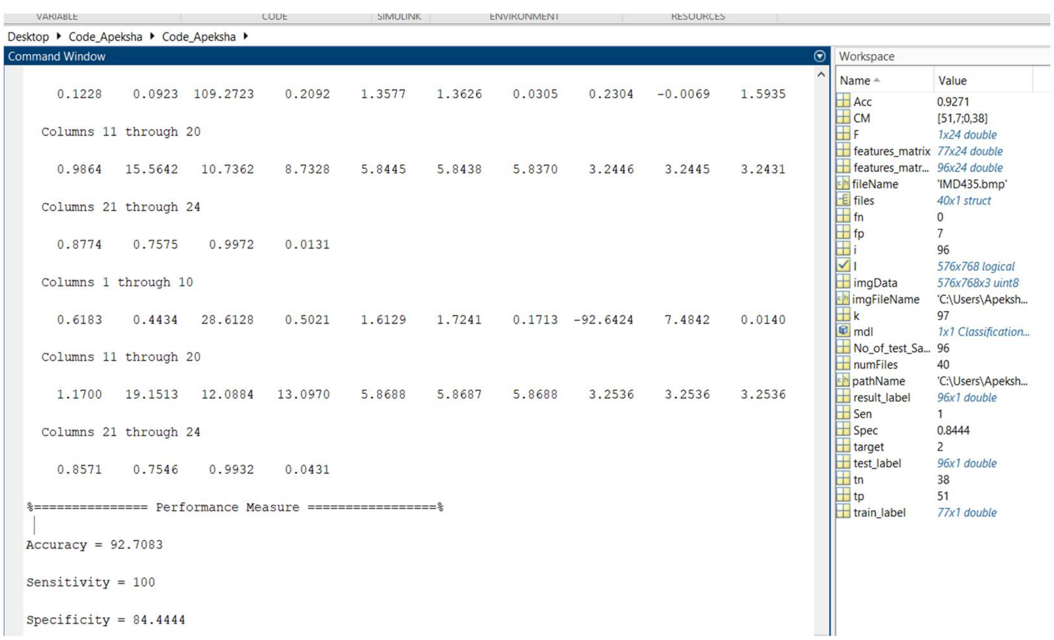


Figure [15]: Screenshot of final output showing in command window

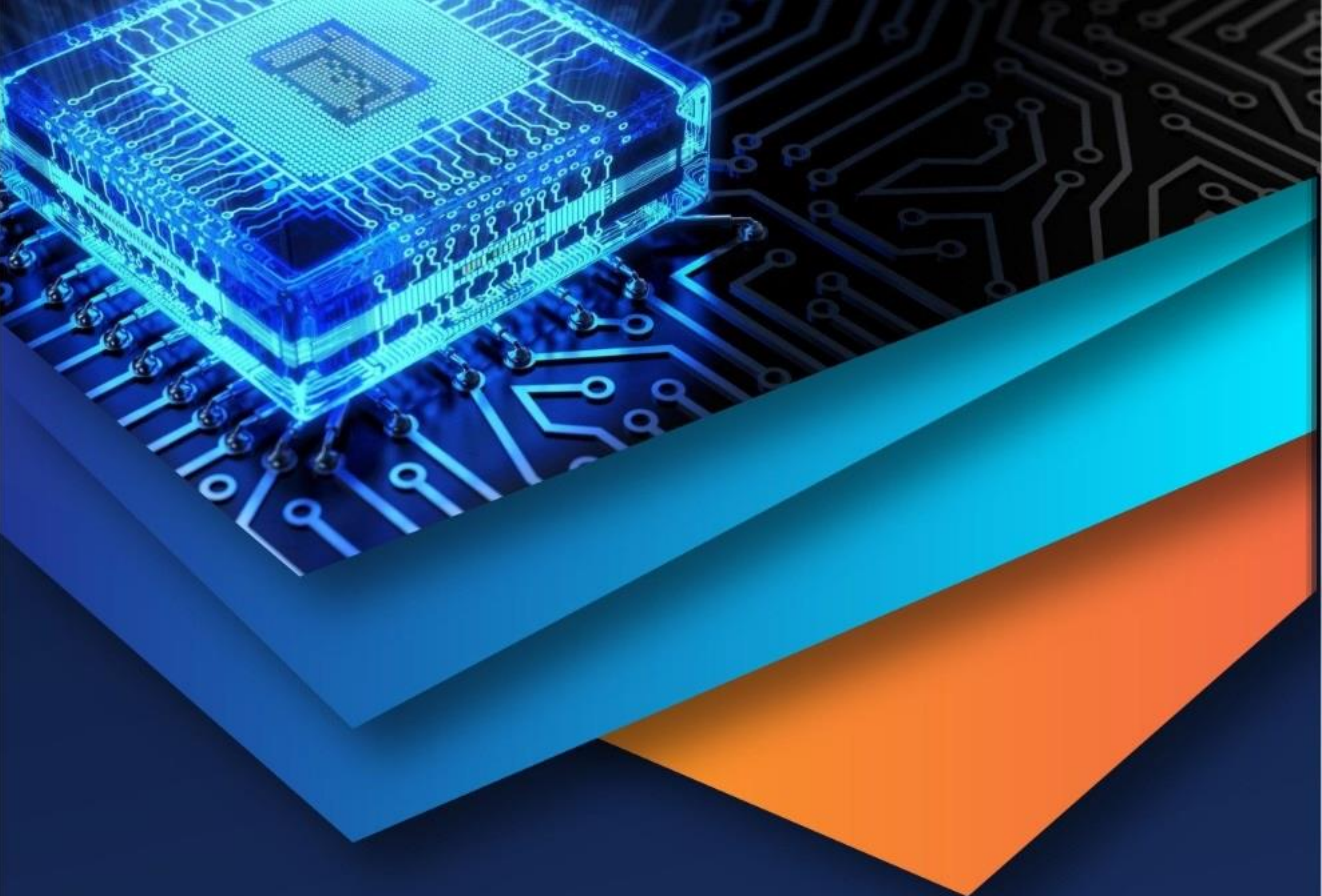
V. CONCLUSION AND FUTURE WORK

In this study, we introduced a robust tool for detection, extraction and classification of skin sore using KNN technique. The system achieves the accuracy of 92.70% with KNN, owning higher performance than other methods.

The future work on the skin cancer detection approach can be more precise and well organized where the system can be executed in the stand-alone mobile application and therefore make the system more authentic and real.

REFERENCES

[1] Noel B. Linsangan; Jetron J. Adtoon; Jumelyn L. Torres, "Geometric Analysis of Skin Lesion for Skin Cancer Using Image Processing", 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM).
[2] M. Vidya; Maya V. Karki, "Skin Cancer Detection using Machine Learning Techniques", 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT).
[3] Minakshi Waghulde; Shirish Kulkarni; Gargi Phadke, "Detection of Skin Cancer Lesions from Digital Images with Image Processing Techniques", 2019 IEEE Pune Section International Conference (PuneCon).
[4] S Sreena; A Lijiya, "Skin Lesion Analysis Towards Melanoma Detection", 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT).
[5] Deva Ratnasari; Budhi Irawan; Casi Setianingsih, "Early Detection of Superficial Basal-Cell Carcinoma Skin Cancer with Extraction Method ABCD Feature Based on Android", 2019 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA).
[6] Prashant bhati et al, "Pre-Phase Stage Detection and Melanoma Separation", International Conference on Communication, Control and Intelligence Systems (CCIS), IEEE 2015.
[7] Pratik Dubal et al, "Skin Cancer Detection and Separation", IEEE 2017.
[8] https://www.who.int/news-room/q-a-detail/radiation-ultraviolet-(uv)-radiationand skincancer#:~:text=Currently%2C%20between%202%20and%203,skin%20cancer%20in%20their%20lifetime
[9] "Skin Cancer Facts and Statistics" https://www.skin cancer.org/skin cancer information/skin cancer facts/
[10] Landis, Sarah H., et al. "Cancer statistics, 1999." CA: A cancer Journal for Clinicians 49.1 (1999): 8-31.
[11] Wiltgen, Marco, A. Gerger, and Josef Smolle. "Tissue counter analysis of benign common nevi and malignant melanoma." International journal of medical informatics 69.1 (2003): 17-28.
[12] Hiam Alquran., et al. "The Melanoma Skin Cancer Detection and Classification using Support Vector Machine", 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT).



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

TOGETHER WE REACH THE GOAL

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY