Grenze International Journal of Engineering and Technology, June Issue

Skin Disease Identification using online and Offline Data Prediction using CNN Classification

Minakshi M. Sonawane[1], Ali Albkhrani [2], Bharti W. Gawali[3], Ramesh R. Manza[4] and Sudhir Mendhekar[5]
[1-4]Department of Computer science and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS), India
Email: minakshi919@gmail.com, albakhrani2017@gmail.com, drbhartirokade@gmail.com, india.manzaramesh@gmail.com
[5]Department of Dermatology, neurology, and leprosy Government Medical College, Aurangabad (MS), India.
Email: sudhir.mendhekar@gmail.com

***Abstract*—In this study, a convolution neural network (CNN) is used to classify images for the detection of skin illness. We collected in a database from the government medical hospital in Aurangabad and the HAM10000 online data. The seven classes are in the skin diseases dataset such as Basel Cell Carcinoma, Psoriasis, Ringworm, Impetigo, Leprosy, and Eczema. The seven additional categories of skin disease are in the database. We have used pre-processing techniques to improve the model accuracy such as resizing images, and normalization of a dataset. We have used a deep learning algorithm for the classification of skin diseases in the database. We have used a deep learning algorithm for the classification of skin disease. It is given an 80.2% percent accuracy rate and its overall accuracy is 78%. Acne disease identification is got 100 accuracies while testing for it is 97.6% accurate. For the classification of skin diseases, we used a deep learning system. Its total accuracy is 78% and it has an accuracy rate of 82.2%. Identification of the acne disease has a 100 accuracy rating, while testing for it has a 97.6% accuracy rating.**

***Index Terms*— CNN, Skin Disease Dataset, Deep Learning, Convolution Neural Network, Image Processing.**

I. INTRODUCTION

The most common disease nowadays that affects people of all ages is skin disease and lesions, however young children and the elderly are immunity powerless compared to other. The investigation of patient's medical history and symptoms, skin scrapping, dermoscopic examination, and skin biopsy is a common method for skin disease diagnosis. But these methods of diagnosis are exhausting, time-consuming, and prone to error. Most of them request an expert dermatologist with superb vision. Medical imaging technologies are sophisticated and trustworthy in diagnosing skin diseases. However, people in low-resource contexts are healthcare institutions. In the healthcare sector, digital imaging cameras and sensing platforms have recently emerged as an alternate method of disease diagnosis. The most recent generation of the camera allows for high-resolution digital image capture images in high resolution to its high-definition camera, enormous storage capacity, and high-resolution digital image capture to its high-definition camera, enormous to its probability, affordability, and connectivity [1].
Computer-aided diagnosis is important and required because it may analysis of different types of skin diseases. The bulk of regularly used algorithms for forecasting skin diseases involve deep learning. This approach will

Grenze ID: 01.GIJET.9.2.806

enable finding the inspected bits in the discovered data pattern, which will considerably improve the performance of even the simplest computational models [2]. The majority of chronic skin conditions, such as impetigo, ringworm, eczema, basal cell carcinoma, and psoriasis, are categorized as serious fitness problems that have an adverse impact on one's physical, mental, and financial health. This dataset contains 2000 dermatoscopic images. A digital camera is employed in many different settings due to its likelihood, affordability, and connectivity [3]. The majority of the initiatives, which focused on skin disease images and aimed to identify specific body parts, were dependent on the availability of an online public dataset. The rest of this article is structured as follows: More information on related studies work for the diagnosis of skin conditions is provided in sections 3 and 4. Section 3 describes the finding and the discussion of the result followed by Section 5's conclusion.

II. LITERATURE REVIEW

Since the past ten years, several studies have been published on Skin disease. Andre Esteva and Brett Kuprel investigated clinical screening and histological testing to categorise skin cancers at the dermatologist level that have substantial neurological systems. They initially showed how to classify a skin illness using a single CNN, and then they became ready to classify images utilising two essential binary inputs, where the disease is only represented by a single pixel [4].

Convolution and Artificial Neural Networks (ANN) Neural networks are commonly used techniques in radiological imaging and diagnosis. The ANN-based model for early detection of breast cancer through image processing or either neural network approach method requires enormous training and testing models considering performance, which requires a lot of computational effort. Furthermore, in ANN, as image resolution increases, so does the number of trainable parameters, resulting in massive training efforts. Furthermore, for the validation set, the classifier achieved an accuracy of 89.90%. The K-Nearest Neighbours (KNN) classification model is widely used for casting and prediction models. This model is also divided into training and testing phases. Furthermore, the accuracy of the KNN model is quite high [5]. KNN models are not suitable for use with large-scale data models because performing prediction models can take a long time. Poor performance when working with high-dimensional datasets with inappropriate feature information may impact the model's accuracy and prediction performance.

III. DATASET DESCRIPTION

TABLE I. LOCAL DATABASE FOR SKIN DISEASE

Sr. No.	Name of Institute/Organization	Database Size	*Name of Disease*	*Resolution*	*Year*
1	Govt. Hospital (GHATI), Aurangabad	612	Acne, psoriasis, Eczema, Wart, Ringworm, Vitiligo, Skin Cold.	6016*3384	2020-2021
2.	https://www.kaggle.com/datasets	10015	Melanocytic –Nevi, BCC, Benign Keratosis-lesions, Melanoma, Dermatofibroma, Vascular-lesson, Akira	478 x 600	2014 A dataset ingested by data. The world may have a maximum size of 1GB and up to 250 individual files

We have captured images from Government Medical Hospital Aurangabad, under the observation of the dermatology lab. We used a direct-current light source to avoid the flickering effect of alternating current (AC). Additionally, information is withheld unless physicians completely and honestly disclose the objectives behind the collection of their data. Both dermatologists and patients are aware that we are only gathering this information for the study. This data was gathered from every patient, and our study was authorized by the institute's ethics committee.

In Source _1 (Sony HD Camera) we have data from seven different skin datasets such as Acne, psoriasis, Eczema, Wart, Ringworm, Vitiligo, and Skin Cold., which are mentioned in table 2. The data indicates a 6016*3384 resolution.

Another is the Source 2 we have taken the pictures from Kaggle [6]. We have considered skin infection pictures with the natural parts. It has been seen that the proposed framework yield exactness differs as for skin illnesses.

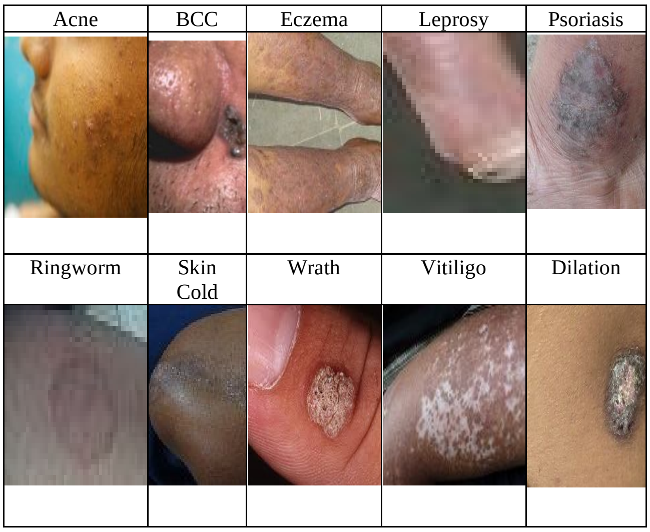


Figure 1. Skin Disease Images Dataset

TABLE II. LOCAL DATABASE FOR SKIN DISEASE

Skin Disease	***Train Data (80%)***	***Test Data (20%)***
Acne	180	25
Basal Cell Carcinoma- BCC	45	10
Psoriasis	275	120
Ringworm	50	21
Impetigo	45	35
Leprosy	122	79
Eczema	118	23
Total	825	313

We have additionally gathered pictures from the web. More than 1012 images with a resolution of 678*600 have been downloaded on seven different infections: melanocytic, bcc, benign keratosis lesion, melanoma, dermatofibroma, vascular lesion, and Akira. In the underlying preparation stage, trademark properties of ordinary picture highlights are confined, and, in light of these, a one-of-a-kind portrayal of every characterization classification is made for seven distinct classes.

The classes are skin inflammation, acne infection, leg infection, hand infection, dermatitis subcutaneous, lichen simplex, stasis dermatitis, and ulcers [7]. In the testing stage, these component space allotments are utilised to create group picture highlights.

IV. METHOD AND TECHNIEQUES

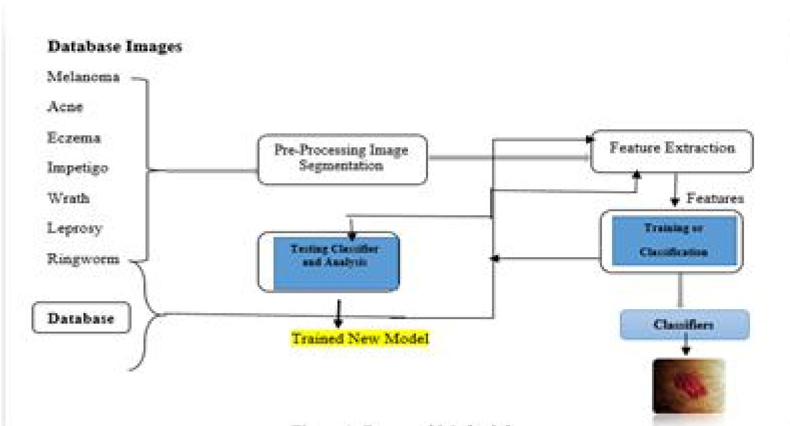


Figure 2. Proposed Methodology

A. Image Acquisition

With the use of two sources, including dermatologist photographs from a digital camera and the Kaggle website as well as online and offline disease images.

Data Preparation

At the point when we gathered our pictures, all the pictures were in an alternate measurement. Our informational index is diverse in its height, width, and size. In any case, the profound neural classifier needs a comparable informational index for preparing and testing the informational index. So we set the pixels to 100 x 100. At that point, we'll prepare our model. Our all-out picture number post-growth is 3000. We used 2400 images for preparation and 600 for testing [8].

B. Convolution neural Network

We developed our algorithm based on pre-trained 825 images we adjust the final layer and used on, Our dataset as inputs, CNN have similar functions, where the calculated feature are combined with each other. The simplified framework of the entire process is shown in figure 2. CNN is often used in real life for image recognition and natural language processing. Each pixel in the input images was transformed into element in a matrix. If there are 100 images input images, the input matrix would be 825 images. 256*100 dimensional images this also called input layers.

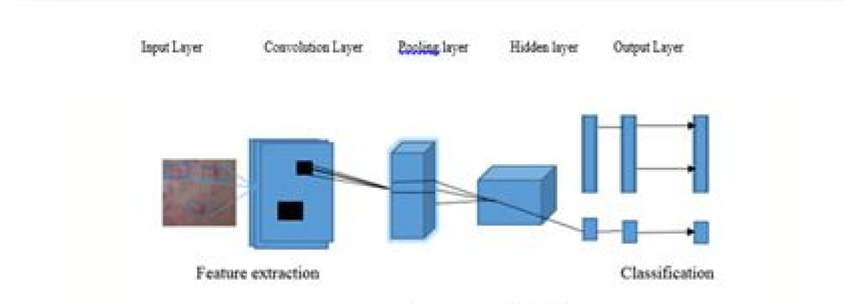


Figure 3. Convolutio Model

Image data generator generates argumentation of images in real time while the model is still training.
One can apply random transformation on each training images as it is passed to model.
Data are distributed in following layers.

1. Convolutional
2. Pooling
3. Dropout
4. Flatten
5. Dense

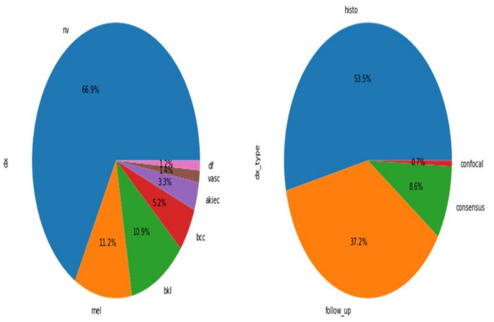
Our idea is to build up a new CNN model, in our model, we have 13 layers. We also have 5 convolutional layers:
• The first layer has 32-3 × 3 filters and ‘linear’ as an activation function.
• The second layer has 64-3 × 3 filters and ‘linear’ as an activation function.
• The third layer has 128-3 × 3 filters and ‘linear’ as an activation function.
• The fourth layer has 256-3 × 3 filters and ‘linear’ as an activation function.
Additionally, we may state that the average size of the five max-pooling layers is 2 X 2. The parameters of our two dropout layers are 0.3 and 0.4, respectively. In our model, a level layer exists. Finally, there are two thick layer capabilities known as "softmax" and "straight". Both skills are used in the beginning process. But to determine the likelihood of our five classes, "softmax" is used [9].

Training Model
Adam optimizer is used for the compilation of our model. For training purposes, we use 80% of our training dataset, and then the rest of the 20% dataset is used for testing purposes. our training dataset consists of 2400 images. So we can say that the number of training sets consists of 1920 images and validating set consist of 670 images. Our classifier’s batch size 78. 50 epochs were used by us to train the model.

V. Performance Evaluation

Preparing precision is regarded as the model's accuracy when applied to the data we prepare. The model's accuracy when applied to a small sample of data from any class is referred to as the approval precision. The diagram in the illustration depicts the creation and acceptance of exactness. to evaluate the performance of the proposed model, we conduct a set of experiments by comparing the proposed model to several state-of-art in diagnosis models. A convolutional neural network (CNN) system is used for Deep Learning [10]. In image processing, such a method is commonly used to classify the object as well as to perform the ROI detection and segmentation process. There are a number of layers in CNN to detect various features of the input layer learning.

At different resolutions, filters are applied to every trained image and the outcome of the convolution layer. The CNN algorithm used for disease detection is based on layer. Our aim behind using CNN for skin disease detection is to improve the recognition results compared to other classifiers.



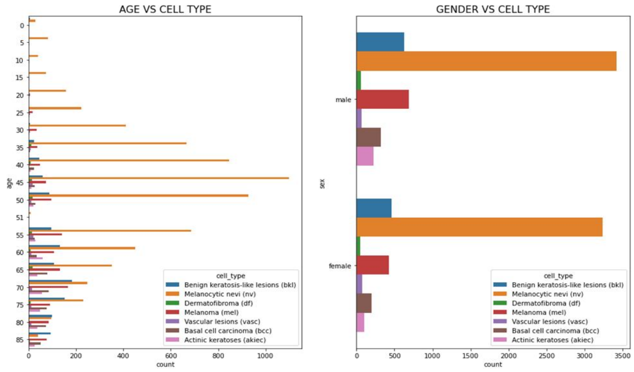


Figure 4. Compare Skin disease Ratio

Figure 5. The Ratio of DiseasePrediction

we have shown in figure 6. The ratio of disease prediction result. In the graph is affected among people and more prominent in men and infection on the lower extremity of the body is more visible in women. and some unknown regions also show infection and its visible in men and women, that acral surface show at least cases that too in men.only gender groups don't show this kind of infection.

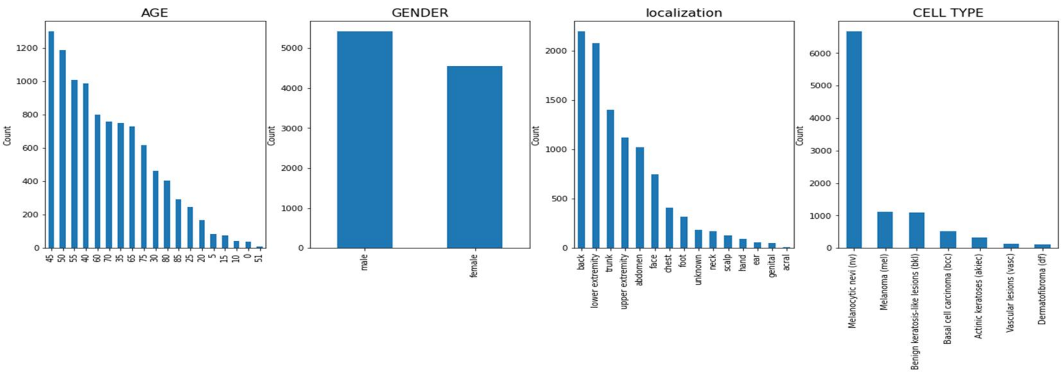


Figure 6. Gender Wise Disease Prediction Ratio

In the online dataset, skin diseases were found maximum 45 year old patient and below 10 year old. Maximum of 10 below. We observed that the probability of having a disease ratio is increased compared to men and women who are facing skin cancerous disease; we have found most of the melanocytic lesions in dermatofibroma disease. The age group between 0-75 years is infected the most by melanocytic nevi. On the other hand, people aged 80-90 are affected more by benign keratosis-like lesions. All the gender groups are affected the most by melanocytic nevi.

A. Performance Measures

Experiments have been carried out to validate the efficiency of the proposed model. The experiment was carried out with a core i5, 2.3 GHz processor with 8 GB RAM using python. Comparisons with other models conducted to measure the performance of the classification are evaluated in terms of classification sensitivity, specificity, and accuracy from the confusion matrix. The measures are computed by using the equations described below with the following convolutions. In this study, the confusion matrix was used to calculate several metrics. This matrix forms four indices which are true positive (TP), false positive (FP), false negative (FN), and true negative (TN). TP and TN match the number of correctly predicted hypoxic and normal samples, whereas the FP and FN match the number of incorrectly predicted hypoxic and normal samples, respectively.

Accuracy, Recall, and F1-score have been determined from our test dataset which contains 600 pictures. So we can say that out Precession normal is 0.76, Recall normal is .78 and F1-score normally us 0.78. Finally, we can say that our classifier is quite acceptable. Characterization table is given underneath. The total accuracy we got 78 %.

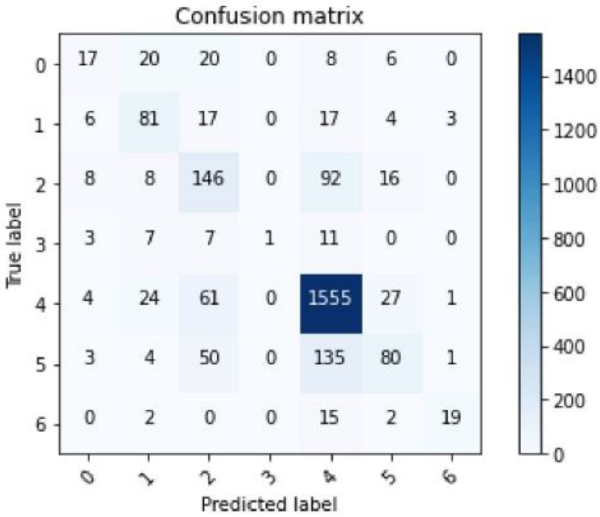


Figure 7 Confusion Matrix

TABLE IV. CNN MODEL-1, PERFORMANCE MEASURES DETECTION OF OFFLINE SKIN DISEASE

Diagnosis class in the dataset	Precision	Recall	F-score
Acne	100%	100%	100%
Basal Cell Carcinoma-BCC	96.01%	98.6%	98%
Psoriasis	95%	97.9%	97.5%
Ringworm	92.6%	96%	96%
Impetigo	92.4%	97.3%	96%
Leprosy	74.4%	94.23%	91%
Eczema	68.6%	85.5%	80.2%

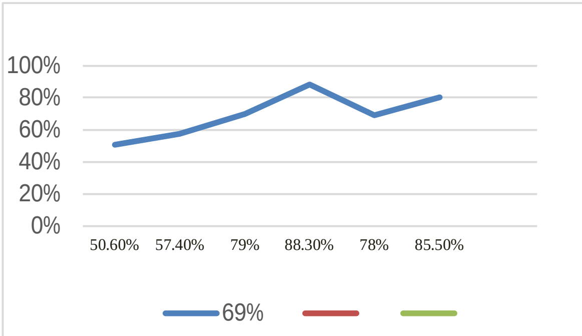


Figure 8. Disease Prediction

Table V. Show that the proposed system can produce high accuracy when we apply it to multiclass skin disease. The result shows that the proposed system correct identifiers all Acne, BCC (Basel cell carcinoma), and psoriasis patients with a diagnosis. Finally, Acne skin diagnosis has the highest accuracy compared with other diseases.

TABLE V. CNN MODEL-1, PERFORMANCE MEASURES OF DETECTION OF ONLINE SKIN DISEASE

Diagnosis class in the dataset	Precision	Recall	F-score
Melanocytic – Nevi	69%	69%	69%
BCC	50.01%	50.6%	51%
Benign Keratosis-lesions (BKL)	55%	57.4%	57.6%
Melanoma –MEL	79.6%	79%	70%
Dermatofibroma – DF	88.4%	88.3%	88.3%
Vascular Lesions –VASC	74.4%	78%	69%
Akira	68.6%	85.5%	80.2%

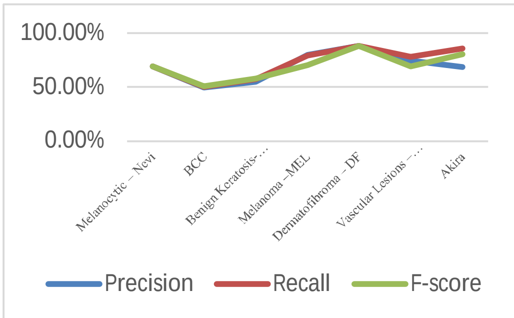


Figure 9. Performance Accuracy with Offline Dataset

Table VI. Show that the proposed system has higher performance in terms of accuracy specification and sensitivity and f-score than the system proposed system by using higher accuracy compared with other diseases using online data using skin disease. The result shows that the proposed system nearby correctly identifies all patients with DF, MEL, and AKIRA disease. Finally, dermatofibroma skin disease has also can be diagnosed with the nearby accuracy state with other state of art skin diagnosis systems.

VI. CONCLUSION

In this study, we have presented a CNN method for the diagnosis of dermatological disease in brief. We gather information from internet and original datasets with various photos, including those of skin conditions including vitiligo, psoriasis, warts, eczema, and skin cold. Another source is online, where 1012 pictures of seven distinct infections melanocytic, bcc, benign keratosis lesion, melanoma, dermatofibroma, vascular lesion, and Akira have been obtained. In terms of identifying skin lesions, we have seen some quite encouraging findings. We identified seven different skin conditions in certain hairy photos. Finally, we ran a statistical analysis to compare performance with the results of our objective investigation. The results of the statistical tests conducted on both datasets' photos to assess performance led to the conclusion that our technique is the statistically best algorithm. As a consequence, when each class was examined independently, our accuracy rate values in multiple classifications rose, and using CNN classification resulted in findings with varied degrees of accuracy. In terms of identifying skin lesions, we have seen some quite encouraging findings. As a consequence, when each class was examined independently, our accuracy rate values in multiple classifications rose and using CNN classification resulted in findings with varied degrees of accuracy. The method displays disease skin detection accuracy in an online database. It has an accuracy rating of 82.2% and a 78% total accuracy. Acne disease detection accuracy is 100 percent, and test results are 97.6 percent accurate.

REFERENCE

[1] Minakshi M. Sonawane, Bharti W. Gawali, Ramesh R.Manza, Sudhir Mendhekar, "analysis of skin disease technique using smartphone and digital camera identification of skin disease", *Research Article*, vol. 4,issue. 3, pp. 529–551, July - September 2022.

[2] Minakshi M. Sonawane, Ramdas D. Gore, Bharti W. Gawali, Ramesh R.Manza, Sudhir Mendhekar, "Computed Aided Diagnosis System for Skin Disease Identification", *International Conference on IoT-based Control Networks and Intelligent Systems (ICICNIS 2020*), pp. 656-667, 2020.

[3] R. J. Hay, N. E. Johns, H. C. Williams, I.W. Bolliger, R. P. Dellavalle, and D. J. Margolis, "The global burden of Skin disease in 2010: An Analysis of the prevalence and impact of skin conditions", *55 J, Investigative Dermatology*, vol. 134, no. 6, PP NO.1527_1534, 2014.

[4] Sardana K, Mahajan S, Sarkar R. Spectrum of skin diseases among Indianchildren. Pediatr Dermatol, 26 (1) pp no:6–13,2009.

[5] Palak Mehta, Bhumika Shah, "Review on Techniques and Steps of Computer Aided Skin Cancer Diagnosis" *International Conference on Computational Modeling and Security (CMS2016)*. https://doi.org/10.1016/j.procs.2016.05.28.

[6] Online available: https://www.medicalnewstoday.com/articles/154322.

[7] Abraham Getachew Kelbore, Philip Owiti ,Anthony J. Reid , Efa Ambaw Bogino, Lantesil Wondewosen and Blen KassahunDessu, "Pattern of skin disease in childeren attending a dermatology clinic in a referral hospital in wolaita soda , southern Ethiopia ",BMC Dermatology, https://doi.org/10.1186/s12895-019-0085-5, pp no- 3-8, 2019.

[8] Housman TS, Feldman SR, Willi ford PM, Fleischer AB Jr., Goldman ND, et al., "Skin cancer is among the most Costly of all Cancers to treat for the Medicare population", J Am Acad Dermatol 48: pp. 425_429, 2003.

[9] Z. Hu and C. S. Yu, "Functional research and development of skin barrier", *Chinese Journal of Clinicians*, vol. 7, No. 7, PP NO 3101–3103, 2013.

[10] A.F. Jerant, J. T. Johnson, C. D. Sheridan, and T. J. Caffrey. "Early detection and treatment of skin cancer", Am Fam Physician 62 (2): 357–68, 375–6,381–2. PMID 10929700, July 2000.