Journal of Advanced Zoology
ISSN: 0253-7214
*Volume **44** Issue **-3 Year 2023** Page **1449:1456***

INTEGRATED DIAGNOSING OF SKIN DISEASE DETECTION USING KNN

[1]D.V.N.Bharathi, [2]M. Vamsi Priya, [3]Yarraguntla Karuna Manjusha, [4]N.M.Hema Devi
[1]Assistant Professor, Department of E.C.E, Sagi Rama Krishnam Raju Engineering College(A), Bhimavaram, A.P, India
[2]Assistant Professor, Department of C.S.E(data science), CMR technical campus, kandlakoya, medchel, Hyderabad, Telangana, India
[3]Assistant Professor, Department of C.S.E, Andhra Loyola Institute of Engineering and Technology, Vijayawada, A.P, India
[4]Assistant Professor, Department of C.S.E, Andhra Loyola Institute of Engineering and Technology, vijayawada.
Mail id:dvnbharathi@srkrec.edu.in

Article History
Received: 12 Aug 2023
Revised: 10 Sept 2023
Accepted: 29 Sept 2023

ABSTRACT: Today, a wide range of illnesses affect people of all ages. Skin cancer is shown to be one of the most common problems and it has a serious impact on human life and health. An allergy, a fungal infection, a bacterium, harmful UV rays from sunburn, etc. could be the cause of a number of skin diseases. It is possible to recover if the disease can be diagnosed earlier and more accurately. Currently, Artificial Intelligence (AI) has a significant impact on the medical industry. Skin diseases, also known as Cutaneous diseases, affect nearly two out of every three people. One of the most common medical environments is skin disease, and when compared to other diseases, the visual representation of skin disease is especially important. Dermatological diseases are the most common diseases in the world. Despite its prevalence, its diagnosis is highly complex and requires extensive practical experience. An efficient automated technique for identifying people with skin diseases is critically needed. In this approach, the K-NN model is recommended for detecting various skin diseases at an early stage. The recommended procedure will provide the highest level of accuracy for detecting skin diseases. Finally, the recommended model works more efficiently than other existing models.

KEYWORDS: Skin disease, KNN (K-Nearest Neighbor), Artificial Intelligence, conventional model.

I. INTRODUCTION

The skin serves a number of functions and is the most significant organ in the human body. The skin is made up of three layers: the epidermis, the dermis and the subcutaneous layer. However, in order to maintain skin barrier function, the skin must minimize lipid and water leakage between the epidermis and dermis.

It serves to protect other delicate parts of our body, however it is not unbreakable part. Skin disease develops as a result of both genetic and environmental causes. Because skin is exposed to the outside world, illnesses and infections affect the skin more frequently. Lesion areas are infected areas of the skin. The earliest clinical symptoms indications of many skin diseases are skin lesions. Today, computer-aided diagnosis is more prevalent in the medical industry.

Early identification of skin disease is more challenging for unskilled dermatologists due to the vast range of relevant symptoms. It is possible to diagnose skin

conditions without having to touch the skin by incorporating digital image processing for the purpose of skin disease detection. These elements have contributed to the importance of studying the advancement of image recognition technology. Even in the hands of inexperienced dermatologists, dermoscopy has been shown to improve the diagnostic accuracy of skin conditions. The remaining sections of this approach examine common skin conditions with noticeable manifestations that can be used as image recognition objects.

Cancer was an unsolvable problem for the scientific community up until now because there is currently no therapy for the issues caused by this terrible disease. Despite being discontinued, research into it has continued for 50 years. The advancement of scientific innovation aims to develop new treatment methods. Melanoma disease develops in an unregulated manner, spreading to the surrounding Injuries and reproducing within bodily cells.

Despite the fact that skin cancer occurs with less consistency than many other cancer forms, it has a high death rate and is therefore quite important. Skin cancer has different characteristics such as squamous cell carcinoma, malignant melanoma, and basal cell carcinoma. With the exception of lesion melanoma, melanoma typically progresses in two stages. One melanoma serves as a sign for the vertical progression that forms with the outer layer of the skin in the initial evaluation.

A dermatologist is a specialist who diagnoses skin conditions. An accurate diagnosis at an early stage is essential for providing patients with effective care and reducing the financial burden caused by skin conditions. A dermatologist uses a combination of two procedures to identify skin diseases. The first method uses patterns found through visual study of the lesion. Most of this typically occurs automatically. As one gains more practise with the visual analysis of skin problems, the diagnosis advances since humans get better at automatically recognising patterns with experience. The patient's medical history, current symptoms, and the results of any diagnostic tests are carefully examined in the second approach.

This disorder, which can cause social isolation, depression, and even self-harm, can also be categorised as a mental disease. Skin disease has consequently evolved as one of the most important topics in medicine. To effectively treat skin diseases, reduce their effects, and increase survival rates, early diagnosis and treatment are essential. Deaths are caused by malignant melanoma (the deadliest kind). Melanoma is a deadly disease that is not treatable. If irregular development of skin melanocytes is recognised early, the survival rate has increased; if it is discovered later, the survival rate is only lowered. Therefore, early detection and treatment are necessary to reduce the undesirable threats brought on by skin

diseases. Furthermore, the accuracy of diagnosing skin diseases isn't high enough due to the similarity of many skin conditions and the lack of physicians with the necessary training. Detecting skin diseases has become a major scientific challenge. The skin cancer that occurs second most frequently is squamous cell carcinoma.

Carcinoma is a type of cancer that affects the squamous cells in the epidermis. Small, flat cells called squamous cells have the appearance of scales. They can be found in the epidermis of the skin, the lining of the body's hollow organs, and the lining of the digestive and respiratory systems. Skin that is frequently exposed to sunlight or other ultraviolet radiation is frequently the site of this cancer. Bowen disease, which has not yet spread to other injuries, is a form of squamous cell

carcinoma which first presents. Keratosis, a type of carcinoma lesion, can lead to squamous cell carcinoma in sun-exposed areas.

To find a solution to the problem of skin diseases diagnosis and treatment, people have adopted computer diagnosis to automatically recognize skin diseases based on skin disease images.As Artificial Intelligence technologies developed, KNN increased advanced.

The remaining analysis is represented as follows: The literature review on skin disease detection has been summarized in section II. In section III, the proposed skin disease detection model is discussed. Results analysis is covered in detail in section IV. The analysis is finished in Section V.

II. LITERATURE SURVEY

Anurag Kumar Verma, Saurabh Pal, B. B. Tiwari, et al. [1] The model has been developed to predict skin diseases. A hybrid feature selection technique was used to enhance the results. For a satisfactory conclusion, the composite approach was also used.

Li-sheng Wei , Quan Gan, and Tao Jiet. al. [4] Skin conditions were identified using a model. They attempt to identify skin diseases using the color and texture of skin images. The SVM classification technique is used as a classifier in this case to detect skin diseases.

Adria Romero Lopez, Jack Burdick, Xavier Giro-i-Nieto, and Oge Marques et al. [5] This method provides a remedy for dermatologists' analysis of skin lesions. The method is developed to predict skin disorders and uses the metaphor of a skin lesion (deeply trained using CNN) to plan and execute two classes of lexemes, determining whether the input is malicious (a threat) or completely harmless. The recommended method achieves high accuracy to detect these diseases.

D. Seth, K. Cheldize, D. Brown, E.F et al. [6] The Dermascope is a special device that is used to diagnose and treat various kinds of skin diseases. To analyse the skin layers covering the skin and through the epidermis, the image is enhanced 10 times with polarised light. Millions of individuals have access to portable dermoscopy, but mobile dermoscopy apps continue to be useful. Images taken by smartphones are still being investigated for medical diagnosis.

Candra, Y. Kurniawan, and K. Het al.[7] Based on the fact that "hues of human skin occupy a substantial region in the color space," a visual RGB algorithm is used. Image registration, which includes aligning the thermal picture and visual image, which are of different alternate modalities, follows the segmentation process.

The feature method is used to create a single image after image registration is finished. For the thermal image pre-processing, patients must enter a lab under carefully controlled conditions, maintaining this instrument is an expensive undertaking.

C. L. Aruta , J. K. Gameng, M. V. Prudentino, C. R. Calaguas, A Anthony and C.J. Lubaton.et. al. [9]The mechanism has been put in place for diagnosing skin conditions. The expertise framework utilizes the picture segmentation method with classification after focusing on the affected area of human skin. Mean filtering and Gaussian smoothing are used to improve an image's region of focus by reducing noise. To classify the segmented diseased region, various classification algorithms such as Ada Boost and Naive Bayes are utilised. A questionnaire was utilised as part of the system to collect information on the patient's age, gender,

and duration of infections. This technique is helpful in detecting eczema, impetigo, and melanoma with high accuracy. Access to this model is restricted since this product is only offered as a Windows application.

R. Sumithra, M. Suhilb, and D. S. Gurucet. al.[10]To effectively segment and categorise skin lesions, a model using KNN and Support Vector Machine (SVM) classifiers was recommended, and it produced satisfactory results.

R. Yasir, M. S. I. Nibir, and N. Ahmedet. al. [11]For the purpose of diagnosing skin conditions, A computer vision-based system has been developed. It is compatible with both mobile phones and computers.

L. Ballerini, R. B. Fisher, B. Aldridge, and J. Rees et al.[12] For the purpose of treating melanoma skin disease, the

hierarchical K-NN classification approach, which recommends using three classifiers for hierarchical merging and feature selection to modify the feature set of each classifier in accordance with its purpose. The accuracy of detection was increased using feature extraction.

Teresa Mendonça, Jorge S. Marques, Andre R. S. Marcal, Pedro M. Ferreira and Jorge Rozeiraet al. [13] An automated system for classifying skin diseases has been developed. A skin colour image has been used to detect skin diseases as well as to explaining independent components. Skin disease is classified based on a brief examination of the skin's texture.

III. CLASSIFICATION OF SKIN DISEASE

This section focuses into the presented integrated diagnosis of skin disease detection using KNN. Figure.1 represents the skin disease detection model's block diagram.

Dermatology dataset was used in this investigation, as well as different methodologies for modeling datasets. 200 dermoscopy photos with a resolution of 768x560 were used, and the clinical data from those images was utilized. Lesion segmentation, clinical and histological diagnosis, and appraisal of a number of dermoscopic criteria (blue-white veil, colors, pigment network, dots/spheres, streaks, regression zones, etc.) were all part of the clinical annotation of the images in the database. There are 200 cases and 13 different features in the data set. There are three types of class values in this data set. The first is a common nevus, a risk factor for skin cancer apart from melanoma. The second type of lesion is an unusual, which indicates a high risk of melanoma skin cancer, and melanoma is the third type of nevus, which indicates cancer.

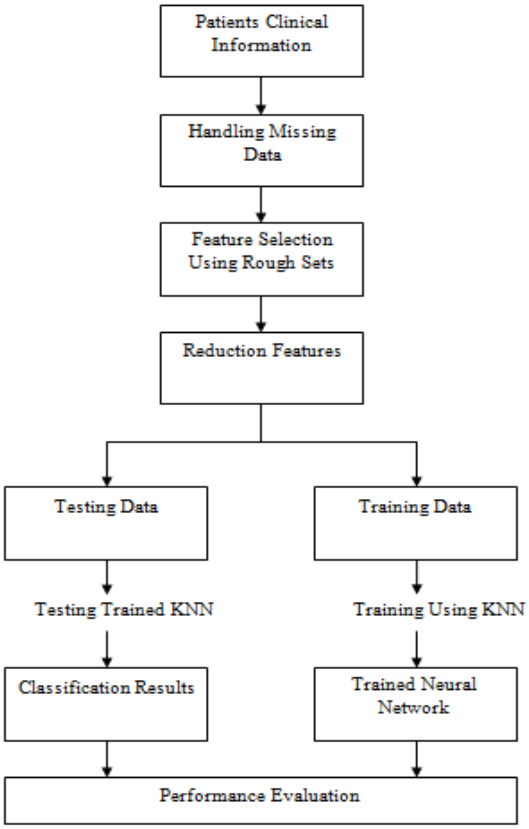

Fig 1: THE BLOCKDIAGRAM OF PRESENTED SKIN DISEASE DETECTIONMODEL

To achieve outstanding performance in the classification of skin diseases, some difficulties that occurred during data loading, such as color contrast and image size, needed to be overcome. This application includes a module that addresses this issue. The images are all enlarged by the Python image resizing technique before being uploaded to the server for processing. As a result, the main objective of this step is to eliminate numerous noises from the skin disease image, including fundamental variations and growth and air bubbles. The techniques have been applied image is processed using the centering, mean, average, and histogram approaches to remove noise and produce a smoother image. The shape and boundaries of the skin disease image that was defined in the dataset are currently undergoing post-processing. To deal with missing data and identify relevant features from a dataset, a processing block will be used to process the obtained skin cancer dataset during the preprocessing stage.

The highest estimate of the class mark component in a tuple is used to account for all missing values in order to enable robust dataset management of missing data. The rough sets method has been widely used in feature selection data analysis for the rough set theory. Equivalent relationship between the features of the dataset is considered as the feature reduction and selection is done as a subset of features. It consists of similar elements of the whole properties and the Core (C) is said to be the crossing part of the reduction and the Reduction (red (R)) is the minimal subdivision of the properties. The total number of minimal subsets of features after reduction in all temporal features and empty sets is reduced in progress by the following equation.

$$S=2^{n}-2 \qquad (1)$$

Where n denotes the total number of attribute digits. When a digit of insignificant subsection of characteristics (S_i) is produced in a single number digit of characteristics, the Combination (C) of rules can record that digit. It progresses to become more

$$S_i=C_{ni}^{N} \qquad (2)$$

Where ni is the quantity of characteristics in the small subset. To clean the dataset, remove all noise. Mapping data sets are used to create training and test sets. The model is developed on the training data set, then splits and tests on the cleaned data set.

KNN is used as the main classification technique in this instance. The K-Nearest Neighbour algorithm is a simple method for supervised machine learning that is primarily utilized in regression and classification tasks. The KNN algorithm accurately classifies these samples while allowing unusual sampling to a small group of n samples. The number of nearby records in KNN is determined by the value of the parameter k. In the testing phase, KNN directly applied the training data also referred to as lazy learning. Later, the performance of the model was improved, and they were located using RF (Random Forest) and LDA classifiers. The effectiveness of all classifiers is evaluated after fitting and testing these approaches in order to choose the best way for the quick and accurate detection of skin diseases.

IV. RESULT ANALYSIS

This section presents an integrated diagnosis of skin disease detection using KNN result analysis. The presented approach is implemented using Python.

In order to accurately identify the types of skin diseases, these methods are applied to three different ML (Machine Learning) classifiers. To find the best classifier among the three, the accuracy, recall, and confusion matrices of the three classifiers

KNN, RF, and LDA (Linear Discriminant Analysis) are evaluated. By assessing the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), which are defined as follows, the performance of these classifiers is evaluated:

True Positive (TP): If a particular instance is actually positive despite being appropriately classified as positive

True Negative (TN): When a negative example is presented and it is in fact negative.

False Positive (FP): If an occurrence is described as positive but is truly negative.

False Negative (FN): a situation that is presented as negative is however actually positive

Accuracy: It is estimated as the proportion of accurate occurrences to all instances.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

Recall: The capacity to locate all relevant instances in a dataset is shown by recall. As a result, categorization models define all instances in a recall that are pertinent as

$$Recall = \frac{TP}{TP+FN} \times 100 \quad (4)$$

Confusion Matrix: Commonly the confusion matrix, is a crucial idea in performance because it uses tabular visualization of model predictions rather than ground-truth labeling. The confusion matrix's rows and columns each represent instances from a predicted class and an actual class, respectively. The performance of a classification model on a test data set is frequently described by a table of confusion matrices called real values.

Table 1: PERFORMANCE ANALYSIS

Performance Metrics	KNN	RF	LDA
Accuracy (%)	96.7	92.3	94.1
Recall (%)	89.6	73.2	76.5
Confusion Matrix	79.3	75.7	70.9

Among these four algorithms, KNN has better performance follow by LDA and RF. The KNN classifiers effectively detects the different types of skin diseases. The performance comparison between these models is shown in Fig. 2.

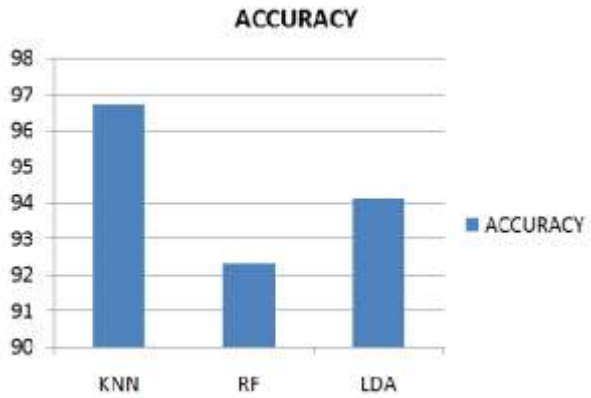

Fig. 2: ACCURACY PERFORMANCE COMPARISON BETWEEN DIFFERENT METHODS

As a result, the KNN is more accurate than the other methods in the above figure.

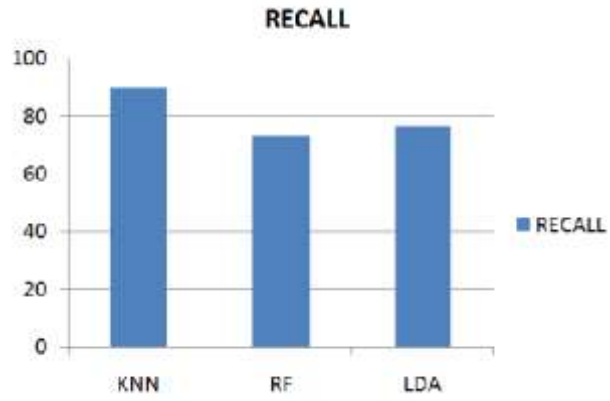

Fig.3 : RECALL PERFORMANCE COMPARISON BETWEEN DIFFERENT METHODS

Therefore the KNN has better Recall than the RF and LDA.

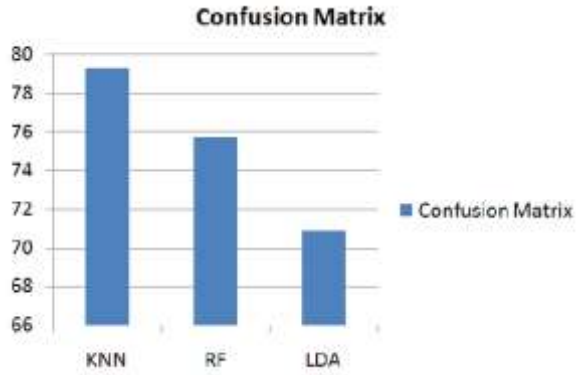

Fig. 4: CONFUSION MATRIX COMPARISON BETWEEN DIFFERENT METHODS

Hence the presented KNN for skin disease detection has high Confusion Matrix than other algorithms.

V. CONCLUSION

This approach use KNN to provide a comprehensive and detailed diagnostic for skin disease detection.In the medical field, identifying and classifying skin diseases is a very difficult process. Due to the characteristics of human skin and the visual similarity of the diseases, determining the particular type of skin disease can be difficult. Consequently, identifying and classifying the skin condition early on is critical. The investigation of the patient's symptoms and pattern identification on the skin lesion are the two main methods used to diagnose skin diseases. The data inside this analysis is first pre-processed to optimize performance. The performance of three different classifier types LDA, KNN, and RF was subsequently evaluated in terms of precision, recall, and confusion matrix to determine the most effective technique. KNN performs better than other algorithms. As a result, the recommended model is effective at diagnosing skin conditions. The investigation concludes that, when compared to Random Forest (RF) and LDA, KNN is the basic learner algorithm that obtains the maximum precision, recall, and confusion matrix.

VI. REFERENCES

[1] Anurag Kumar Verma, Saurabh Pal, B. B. Tiwari, “Skin disease prediction using ensemblemethods and a new hybrid feature selection technique “,Iran Journal of Computer Science, Published Online: 04 Aprill 2020.

[2] Cerwall, P. Report, E. M. Ericssons mobility report at www.ericsson.com/49da93/assets/local/mobilityreport/documents/2020/june2020-ericsson-mobility-report.pdf. 2020.

[3] J. Vasquez and B. E. Comendador, ‘‘Competency discovery system: Integrating the enhanced ID3 decision tree algorithm to predict the assessment competency of senior high school students,’’ Int. J. Adv. Sci., Eng. Inf. Technol., vol. 9, no. 1, p. 60, Jan. 2019, doi: 10.18517/ijaseit.9.1.7763.

[4] Li-sheng Wei , Quan Gan, and Tao Ji, “Skin Disease Recognition Method Based on Image Color and Texture Features ”, Hindawi, Computational and Mathematical Methods in Medicine Volume 2018, Article ID 8145713, 10 pages.

[5] Adria Romero Lopez, Xavier Giro-i-Nieto, Jack Burdick and Oge Marques, "Skin Lesion Classification from Dermoscopic Images Using Deep Learning Techniques", 2017 13th IASTED International Conference on Biomedical Engineering (BioMed), pp. 49-54, 2017.

[6] D. Seth, K. Cheldize, D. Brown, E.F. Freeman. Global Burden of Skin Disease: Inequities and Innovations. Curr Dermatol Rep. 2017;6(3):204- 210. DOI:10.1007/s13671-017-0192-7

[7] Candra, Y. Kurniawan , and K. H. Rhee. Security analysis testing for secure instant messaging in android with study case: Telegram. In 2016 6th International Conference on System Engineering and Technology (ICSET), pages 92-96. IEEE, 2016 .

[8] B. He, C. Huang, G. Sharp, S. Zhou, Q. Hu, C. Fang, Y. Fan, and F. Jia, ‘‘Fast automatic 3D liver segmentation based on a three-level AdaBoostguided active shape

model,'' Med. Phys., vol. 43, no. 5, pp. 2421–2434, Apr. 2016
[9] C. L. Aruta , C. R. Calaguas, J. K. Gameng, M. V. Prudentino, A Anthony & C.J. Lubaton. Mobile-based medical assistance for diagnosing different types ofskin diseases using case-based reasoning with image processing . IJ CI, 2015 :3, 2015 .
[10] R. Sumithra, M. Suhilb, and D. S. Guruc, "Segmentation and classification of skin lesions for disease diagnosis," Procedia Computer Science, vol. 45, pp. 76–85, 2015.
[11] R. Yasir, M. S. I. Nibir, and N. Ahmed, "A skin disease detection system for financially unstable people in developing countries," Global Science and Technology Journal, vol. 3, no. 1, pp. 77–93, 2015.
[12] L. Ballerini, R. B. Fisher, B. Aldridge, and J. Rees, ''A color and texture based hierarchical K-NN approach to the classification of non-melanoma skin lesions,'' in Color Medical Image Analysis (Lecture Notes in Computational Vision and Biomechanics), The Netherlands: Springer, 2013, pp. 63–86.
[13] Teresa Mendonça, Pedro M. Ferreira, Jorge S. Marques, Andre R. S. Marcal and Jorge Rozeira, "PH2- A dermoscopic image database for research and benchmarking", 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5437-5440, 2013.
[14] R. Garnavi, M. Aldeen, and J. Bailey, ''Computer-aided diagnosis of melanoma using Border- and wavelet-based texture analysis,'' IEEE Trans. Inf. Technol. Biomed., vol. 16, no. 6, pp. 1239–1252, Nov. 2012, doi: 10.1109/TITB.2012.2212282.
[15] S. Arivazhagan, R. N. Shebiah, K. Divya, and M. P. Subadevi, "Skin disease classification by extracting independent components," Journal of Emerging Trends in Computing and InformationSciences, vol. 3, no. 10, pp. 1379–1382, 2012
[16] P. L. Gutierrez and S. Siva, Classification and Regression Tree (CART). Amsterdam, The Netherlands: Springer, 2008.