International Journal Of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704124 |

Early Diagnosis of Multiple Age Skin Disease Using SVM and CNN

J.RAMYA, MADALA PAVAN KALYAN, PAYAVULLA JAYANTH, SANGU NAVEEN KUMAR REDDY

Assistant Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: Skin lesions affect millions of people worldwide. They can be easily recognized based on their typically abnormal texture and color but are difficult to diagnose due to similar symptoms among certain types of lesions. The motivation for this study is to collate and analyze machine learning (ML) applications in skin lesion research, with the goal of encouraging the development of automated systems for skin disease diagnosis. To assist dermatologists in their clinical diagnosis, several skin image datasets have been developed and published online. Such efforts have motivated researchers and medical staff to develop automatic skin diagnosis systems using image segmentation and classification processes. This paper summarizes the fundamental steps in skin lesion diagnosis based on papers mainly published since 2013. The applications of ML methods (including traditional ML and deep learning (DL)) in skin disease recognition are reviewed based on their contributions, methods, and achieved results. Such technical analysis is beneficial to the continuing development of reliable and effective computer-aided skin disease diagnosis systems. We believe that more research efforts will lead to the current automatic skin diagnosis studies being used in real clinical settings in the near future.

KEYWORDS: skin image segmentation; skin lesion classification; machine learning; deep learning

I.INTRODUCTION

Skin Disease are occurring almost on all groups of ages among people. The rate of skin disease has been increased due to lifestyle and changing environments. In the USA country, it is observed that every one out of five people are infected with any kind of skin disease. They are usually caused by factors like different organism's cells, a different diet, and internal and external factors, such as the hierarchical genetic group of cells, hormones, and immune system of conditions. These factors may act together or in a sequence of skin disease. There are chronic and incurable diseases, like eczema and psoriasis, and malignant diseases like malignant melanoma. Recent researchers have found the availability of cures for these diseases if they are detected in the early stage. From the literature survey, authors of this paper found that, the creation of an expert application of skin disease detection using methods like Naive Bayes, CNN, SVM methods was felt to be very necessary to help all people who want to know about skin diseases that are being experienced or need information about skin diseases.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704124 |

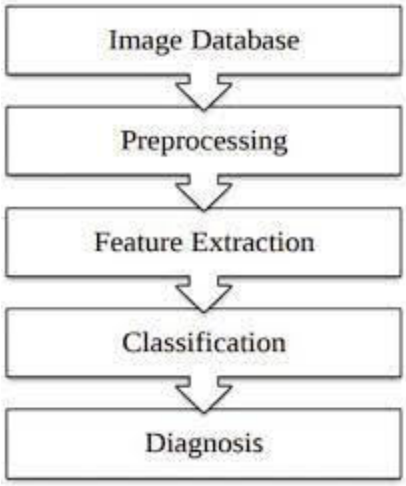

Fig 1: Convolutional Neural Networks for Classifying

To detect these diseases using the image processing method many research papers has been published and many researchers has contributed a lot which paved a way for our application and gave us a right direction. Without the previous works of these fellow researchers my work on developing any application would never have been easier [5][6]. Skin diseases are often quite hard to detect at an early stage and it is even harder to classify them separately. Recently, it is well known that, the most dangerous form of skin cancer among the other types of skin cancer is melanoma because it is much more likely to spread to other parts of the body if not diagnosed and treated early. In order to classify these skin diseases, “Support Vector Machine (SVM)” a Machine Learning Algorithm can be used Image classification is one of classical problems of concern in image processing. Support Vector Machine are classified under supervised learning models and is a part of machine learning algorithm which used to analyze structured and unstructured data such as text and images. As an input SVM always requires clean data. In skin disease detection, classifying the images into different types of skin diseases is the problem. This paper gives us the complete overview on existing machine learning and image processing algorithms for detection of skin disease through android application development.

II.RELATED WORK

To develop skin diagnosis models, various datasets of different sizes and skin lesion types have been created by educational institutions and medical organizations. These datasets can serve as platforms to educate the general public and as tools to test newly developed diagnosis algorithms. The first dermoscopic image dataset, PH2, had 200 carefully selected images with segmented lesions, as well as clinical and histological diagnosis records. The Med-Node dataset contains a total of 170 macroscopic images, including 70 melanoma and 100 nevus images, collected by the Department of Dermatology, University Medical Center Groningen. Due to the limited image quantity and classes, these labeled datasets are often used together with others in the development of diagnosis models.

The most popular skin disease detection dataset, ISIC Archive, was created through the collaboration of 30 academic centers and companies worldwide with the aim of improving early melanoma diagnosis and reducing related deaths and unnecessary biopsies. This dataset contained 71,066 images with 24 skin disease classes [3], but only 11 of the 24 classes had over 100 images. The images were organized by diagnostic attributes, such as benign, malignant, and disease classes, and clinical attributes, such as patient and skin lesion information, and image type. Due to the diversity and quantity of images contained, this dataset was used to implement the ISIC challenges from 2016–2020 and dramatically contributed to the development of automatic lesion segmentation, lesion attribute detection, and disease classification [3].

The first ISIC 2016 dataset had 900 images for training and 350 images for testing under 2 classes: melanoma and benign. The dataset gradually increased to cover more disease classes from 2017 to 2020. Images in ISIC 2016–2017 were fully paired with the disease annotation from experts, as well as the ground truth of the skin lesion in the form of binary masks. The unique information included in the ISIC dataset is the diameter of each skin lesion, which can help to clarify the stage of melanoma. In the ISIC 2018 challenge, the labeled HAM 10,000 dataset served as the training set, which included 10,015 training images with uneven distribution from 7 classes. The BCN 20,000 dataset was utilized as

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704124 |

labeled samples in the ISIC 2019 and ISIC 2020 challenge, which consisted of 19,424 dermoscopic images from 9 classes. It is worth noting that not all ISIC datasets were completely labeled in the ISIC 2018–2020 datasets.

To prompt research into diverse skin diseases, images from more disease categories have been (and continue to be) collected. Dermofit has 1300 macroscopic skin lesion images with corresponding segmented masks over 10 classes of diseases, which were captured by a camera under standardized conditions for quality control. DermNet covers 23 classes of skin disease with 21,844 clinical images. For precise diagnosis, melanoma and nevus can be further refined into several subtypes. For example, the EDRA dataset contains only 1011 dermoscopic images from 20 specific categories, including 8 categories of nevus, 6 categories of melanoma, and 6 other skin diseases. The images under melanoma were further divided into melanoma, melanoma (in situ), melanoma (less than 0.76 mm), melanoma (0.76 to 1.5 mm), melanoma (more than 1.5 mm), and melanoma metastasis. summarizes these datasets according to image number, disease categories, number of labeled images, and binary segmentation mask inclusion. It can be seen that the ISIC Archive is the largest public repository with expert annotation from 25 types of diseases, and PH2 and Med-Node are smaller datasets with a focus on distinguishing between nevus and melanoma.

III.METHODS

Traditional image segmentation methods use pixel, region, and edge-based approaches to extract skin lesions from images. Pixel-based segmentation methods, such as binary or Otsu thresholding, can outline each pixel into two categories (i.e., healthy skin or skin lesion). However, these methods can generate discontinuous results, particularly from dermoscopic images, mainly due to low contrast and smooth transitions between lesions and healthy skin. Region-based segmentation can identify and combine adjacent pixels to form skin lesion regions by merging or growing regions. Merging combines adjacent pixels with similar intensity together, while growing starts from a point and checks nearby pixels to expand region coverage. The difficulties faced in implementing these methods lie in the variety of colors and textures of individual skin lesions. Edge-based segmentation methods, such as the watershed algorithm, utilize intensity changes between adjacent pixels to outline the boundary of a skin lesion. These methods are susceptible to noise, such as hair, skin texture, and air bubbles, which can lead to convergence, especially around noisy points, and produce erroneous segmentation results. In general, traditional segmentation methods often struggle to achieve accurate results when segmenting images with noise, low contrast, and varied color and texture.

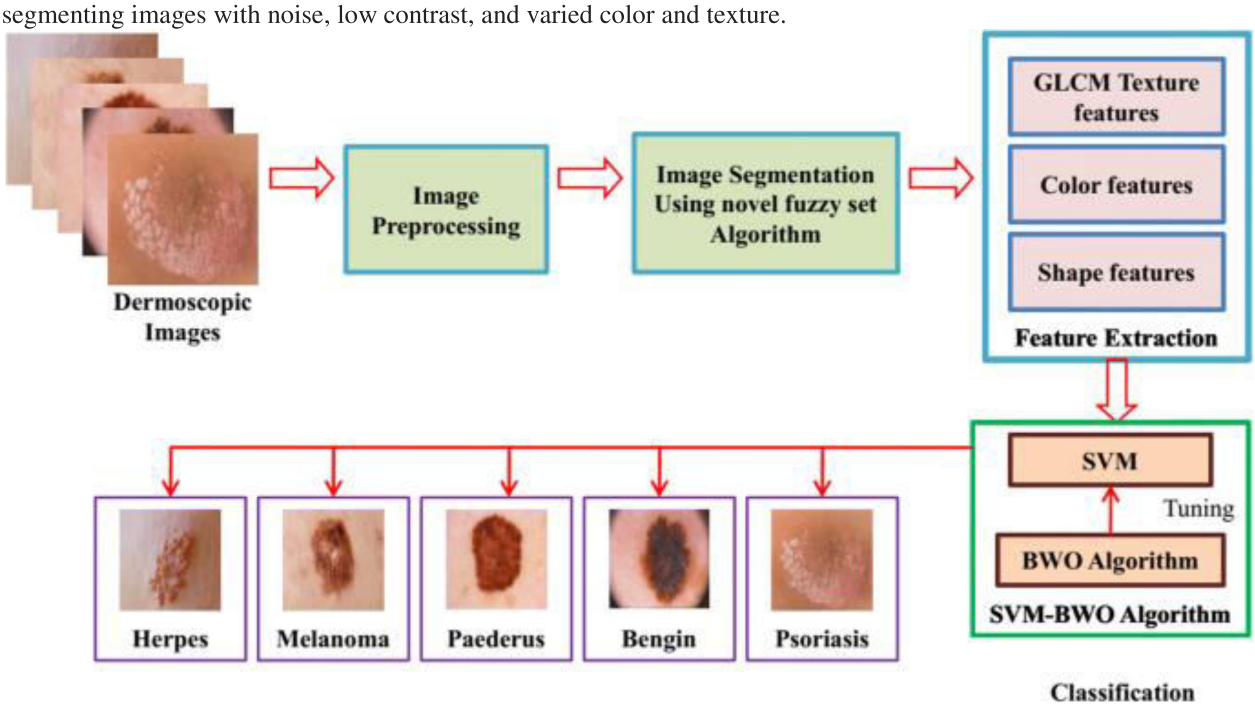

Fig 2: segmentation and black widow

Neural networks (NNs), evolutionary computation, and fuzzy logic have been proposed as methods to segment the ROI based on learning, natural evolution, and human reasoning. These methods can be used individually or in combination to achieve better performance. For example, the fuzzy method has been applied with both splitting and

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704124 |

merging techniques to segment dermoscopic images. This combination generated unsupervised perceptual segmentation using fused color and texture features. Later, Devi developed an automatic skin lesion segmentation system using fuzzy c-means (FCM) clustering together with histogram properties. The histogram property was used to select the number of clusters, and the color channel for FCM clustering (hue, saturation, value) was determined based on individual entropy. This system can effectively segment the lesion regions from normal skin automatically with an accuracy of 95.69% when compared with traditional methods. Moreover, there is still much space for advancement in terms of accuracy.

IV.RESULT ANALYSIS

Feature extraction and selection aim to identify an optimal feature set with excellent discrimination capability in this case to classify skin lesion images. Expert knowledge and clinical experience are preferred, which may effectively and efficiently guide feature extraction and selection. Image processing algorithms can generate features useful for classification, but the effectiveness of extracted features relies on the disease symptoms and diagnosis tasks at hand. To yield accurate melanoma diagnosis, various methods are explored to extract color, texture, edge, and shape features. In the popular ABCD rule, features are selected based on the knowledge that melanoma has an asymmetrical shape, an irregular border, multiple colors and shapes, and a diameter larger than 6 mm. Specifically, A stands for asymmetry, B stands for border, C stands for color, and D stands for diameter. Similar features are extracted using the CASH mnemonic.

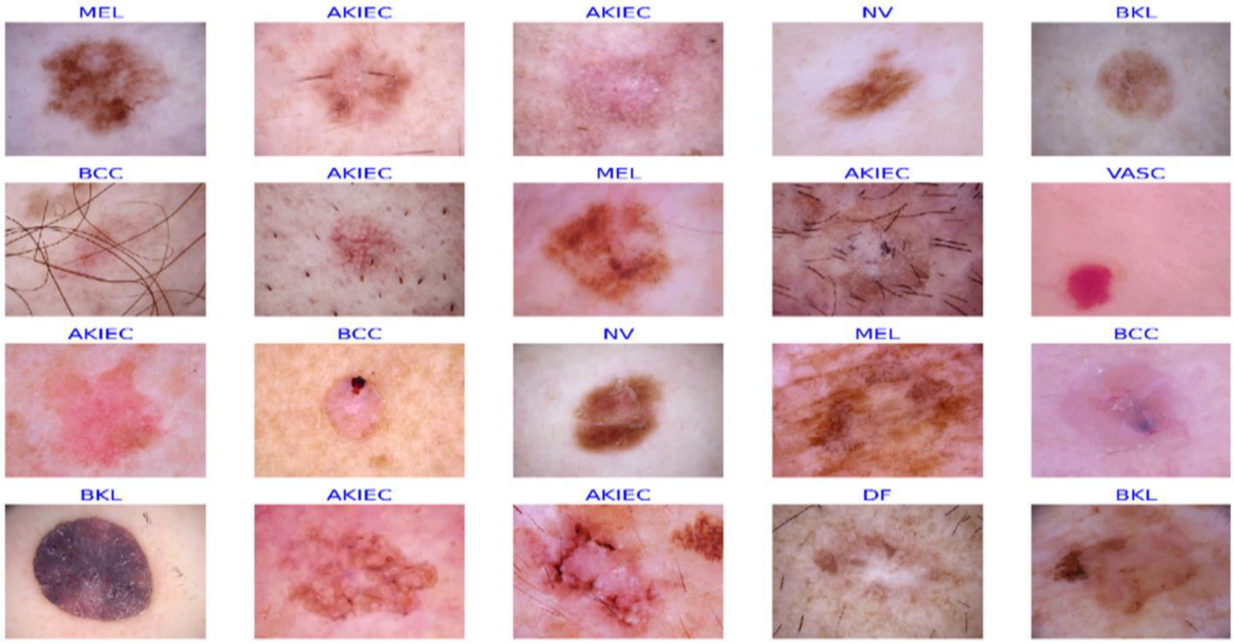

Fig 3: Result Analysis

To compare multi-class categorization performance, Hameed used two strategies to classify skin images. One is a three-category classification: healthy, inflammatory diseases (acne, psoriasis, and eczema), and non-inflammatory diseases (benign and malignant). The other is a six-category classification: healthy, acne, eczema, psoriasis, benign, and malignant. DT, SVM, K-NN, and ensemble classifiers with different kernels were applied to the two classification strategies. The result showed that the classification accuracy decreased for all classifiers when the number of categorized classes increased, and quadratic SVM achieved the highest classification accuracy in both strategies. For the same identification task, Hameed also compared the performance of multi-level multi-class classification and single-level multi-class classification.

Advances in automatic diagnostics are largely driven by the use of datasets with large repositories of digital images. Public datasets are a compilation of data accessible through an online platform, typically made available by government agencies, academic institutions, or private organizations. Due to the challenges associated with gathering skin lesion images, utilizing publicly available datasets is frequently a more practical alternative to collaborating with medical

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704124 |

institutions to obtain such datasets. Public datasets such as the ISIC contain skin disease images mostly from subjects with light-colored skin, collected in the USA, Europe, and Australia. The current recognition models are trained using these unequal distribution datasets. While such recognition models may have outstanding capabilities to recognize lesions on subjects with light-colored skin, their effectiveness is in question when dealing with skin lesion images from subjects from other geographical regions. Moreover, the prevalence and characteristics of skin disease vary with racial and ethnic groups. Traditional ML and DL models fail to offer correct recognition when a test image is from an under-represented skin color group and/or lesion type. More balanced datasets are needed, with clinical data based on gender, age, skin type, and race. This is critical when using AI diagnosis to improve healthcare in rural areas and increase global access to specialist expertise.

V.CONCLUSION

AI-based skin lesion diagnosis is an increasingly attractive research area, which has been largely driven by the availability of appropriate methods and continually updated abundant datasets. Although relevant topics have been addressed over the last decade, there are still many aspects for investigation and room for improvement.

This paper reviews public skin lesion datasets, the applied image preprocessing methods, and the subsequent skin lesion segmentation and classification methods. The current status, challenges, and outlook in ML-driven skin disease diagnosis are also discussed. Such studies can empower the development of advanced concepts and methodologies. In conclusion, future trends regarding image segmentation and classification of skin lesions require the development of more comprehensive datasets, investigation of more robust models, particularly for macroscopic image recognition, and methods for increasingly reliable automated diagnosis.

REFERENCES

1. ALKolifi-ALEnezi, N.S. A Method Of Skin Disease Detection Using Image Processing And Machine Learning. *Procedia Comput. Sci.* **2019**, *163*, 85–92. [**Google Scholar**] [**CrossRef**]
2. Skin Disorders: Pictures, Causes, Symptoms, and Treatment. Available online: **https://www.healthline.com/health/skin-disorders** (accessed on 21 February 2023).
3. ISIC Archive. Available online: **https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main** (accessed on 20 February 2023).
4. Sun, J.; Yao, K.; Huang, K.; Huang, D. Machine learning applications in scaffold based bioprinting. *Mater. Today Proc.* **2022**, *70*, 17–23. [**Google Scholar**] [**CrossRef**]
5. Haenssle, H.A.; Fink, C.; Schneiderbauer, R.; Toberer, F.; Buhl, T.; Blum, A.; Kalloo, A.; Hassen, A.B.H.; Thomas, L.; Enk, A.; et al. Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann. Oncol.* **2018**, *29*, 1836–1842. [**Google Scholar**] [**CrossRef**] [**PubMed**]
6. Rotemberg, V.; Kurtansky, N.; Betz-Stablein, B.; Caffery, L.; Chousakos, E.; Codella, N.; Combalia, M.; Dusza, S.; Guitera, P.; Gutman, D.; et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Sci. Data* **2021**, *8*, 34. [**Google Scholar**] [**CrossRef**] [**PubMed**]
7. Melanoma Skin Cancer Rreport. Melanoma UK. 2020. Available online: **https://www.melanomauk.org.uk/2020-melanoma-skin-cancer-report** (accessed on 20 February 2023).
8. Mendonça, T.; Ferreira, P.M.; Marques, J.S.; Marcal, A.R.; Rozeira, J. PH 2-A dermoscopic image database for research and benchmarking. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; pp. 5437–5440. [**Google Scholar**]
9. Tschandl, P.; Rosendahl, C.; Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data* **2018**, *5*, 1–9. [**Google Scholar**] [**CrossRef**] [**PubMed**]
10. Combalia, M.; Codella, N.C.; Rotemberg, V.; Helba, B.; Vilaplana, V.; Reiter, O.; Carrera, C.; Barreiro, A.; Halpern, A.C.; Puig, S.; et al. Bcn20000: Dermoscopic lesions in the wild. *arXiv* **2019**, arXiv:1908.02288. [**Google Scholar**]
11. Dermnet. Kaggle. Available online: **https://www.kaggle.com/datasets/shubhamgoel27/dermnet** (accessed on 20 February 2023).
12. Giotis, I.; Molders, N.; Land, S.; Biehl, M.; Jonkman, M.F.; Petkov, N. MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert Syst. Appl.* **2015**, *42*, 6578–6585. [**Google Scholar**] [**CrossRef**]
13. Yap, J.; Yolland, W.; Tschandl, P. Multimodal skin lesion classification using deep learning. *Exp. Dermatol.* **2018**, *27*, 1261–1267. [**Google Scholar**] [**CrossRef**] [**Green Version**]
14. Dermofit Image Library Available from The University of Edinburgh. Available online: **https://licensing.edinburgh-innovations.ed.ac.uk/product/dermofit-image-library** (accessed on 20 February 2023).