IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

ENHANCING DERMATOLOGICAL DIAGNOSIS WITH SVM-BASED SKIN CANCER DETECTION IN MATLAB

[1]Sanjana Krishna [2]Shruthi S M [3]Sparsha R [4]Suchithra R [5]Mr. Suhas Bharadwaj R
[1]B.E, Information Science & Engineering, NIE Mysuru [2]B.E, Information Science & Engineering, NIE Mysuru
[3]B.E, Information Science & Engineering, NIE Mysuru [4]B.E, Information Science & Engineering, NIE Mysuru
[5]Dept.of IS&E, Assistant Professor, NIE Mysuru

Abstract
Researchers have developed sophisticated systems capable of utilizing computer-based intelligence tools and methods to investigate human disorders, propelled by machines advancing into more skilled domains. Skin disease detection through image processing stands out as a primary focus within this research landscape. The identification of apparent skin infections from aerial photos is on the rise, highlighting the complex challenges in the field. A system of multiclass support vector machines is employed, incorporating various prominent texture and repetitive region features, such as the Grey Level Co-Occurrence Matrix (GLCM), into Support Vector Machines (SVM)-based classifiers to validate skin conditions from photos. The test set exhibited an impressive confirmation accuracy rate of 83.5%, marking a significant step towards establishing India's foremost locally automated benign system for diagnosing skin conditions.

Keywords: *image processing, machine learning, GLCM, SVM.*

I. INTRODUCTION

Skin infections rank as the sixth most dangerous form of cancer worldwide. The skin comprises cells, and within these cells lie tissues. Consequently, diseases arise from abnormal or uncontrolled cell proliferation in opposing or neighboring tissues. The skin's sensitivity to UV radiation and resilience to external impacts play crucial roles in this dynamic. This unique cell enhancement pattern can be categorized as either benign or malignant. Benign developments, like harmless moles, are typically considered non-threatening. In contrast, malignant developments are treated as life-threatening illnesses, capable of causing harm to other body tissues.

The skin's layer consists of three cell types: melanocytes, squamous cells, and basal cells, all with the potential to damage tissues. While various skin growths exist, the three most hazardous types are melanoma, basal cell carcinoma (BCC), and squamous cell carcinoma (SCC). Additionally, actinic keratosis (AK) and melanocytic nevus are noteworthy types.

Among the multitude of varieties, melanoma, the most lethal type, can recur even after treatment. Australia and the US bear the greatest impact of skin conditions. This research employs practical methods to categorize all aforementioned illness types. For image enhancement, the Gaussian filter and the dark razor technique are employed, with the central filter ensuring resolution without complications. These steps are deemed as preliminary processing. The preprocessed images undergo batching based on k-means clustering. GLCM techniques are utilized to extract constituents from segmented images, and characteristics from both approaches are reinforced for further analysis. Finally, an MSVM classifier is employed for arrangement purposes, aiming to achieve high precision.

II. RELATED WORK

In article [1], the focus was on characterizing two skin diseases: melanoma and nonmelanoma. To improve outcomes, a combination of dark images and diversity was employed, rather than relying on either alone. K-means clustering, incorporating the ABCD technique (Imbalance, Limit inconsistency, Diversity, Measurement), was used for division. The dataset comprised 150 images, with 75 each depicting melanoma and non-melanoma. Four classifiers were deployed; SVC and 1-NN achieved the highest accuracy, showcasing comparable capabilities.

In the subsequent research [2], a 3D proliferation estimation was proposed using 2D images. This method identifies the RGB and shape of the 3D picture. Pre-processing involved converting images into binary form with 0 and 1 seconds. For division, a flexible snake estimate was applied, and a 3D importance assessment restriction was employed on all components to construct the representation capacity.

To reduce the effects of melanoma, early detection is essential. Study [3] utilizes the MSVM classifier, considering five skin lesion types: actinic keratosis, squamous cell disease, basal cell malignant development, seborrheic verruca, and neurotic nevus. GLCM extracts surface elements, including homogeneity, angle, and differentiation. For every one of the five image categories, expanding zones are identified using K-means clustering. The division and grouping results are displayed using a graphical user interface (GUI).

Melanoma is the most prevalent skin cancer. In research [4], a method for organizing melanoma is proposed using the Naive Bayes classifier and shear let transform coefficients. Shear let transform reduces the dataset using predetermined coefficients (50, 75, and 100). The resulting coefficients integrate into the Naive Bayes classifier for further analysis.

III. PROBLEM STATEMENT

Skin disease stands out as one of the most widely recognized forms of cancer globally, and effective treatment hinges significantly on early detection. Traditional methods for identifying skin malignancies predominantly rely on visual examinations by dermatologists — a process that can be both time-consuming and subjective.

The objective of this project is to formulate an AI-based framework utilizing image processing techniques capable of accurately categorizing skin lesions as either benign (non-dangerous) or malignant, thereby enhancing the precision and efficiency of skin disease detection.

Dataset: The project will utilize a curated dataset comprising photographs of skin lesions. This dataset encompasses images depicting both benign and malignant cases, ensuring robustness and generalizability through variations in skin types, lesion types, and diverse image properties.

Image Preprocessing: Employ image processing techniques to refine images of skin lesions. Tasks may include scaling, noise reduction, standardization, and image enhancement, contributing to improved image quality for subsequent analysis.

Feature Extraction: Utilize feature extraction techniques to capture relevant information from the preprocessed images. Components may involve surface analysis, edge identification, and other pertinent characteristics of skin lesions.

Segmentation: This technique aims to isolate and emphasize the region of the image containing the skin anomaly, facilitating further analysis and classification. Segmentation provides a well-defined zone for feature extraction and classification, elevating the precision and efficiency of machine learning models.

Machine Learning Model: Train an artificial intelligence model using the extracted features and corresponding labels (benign or malignant) from the dataset.

IV. OBJECTIVES

Information Collection:
Establish an extensive dataset comprising images of skin lesions, encompassing both benign and malignant cases. Dermatologists or other experts should meticulously sort, label, and appropriately name the dataset.

Preprocessing:
Normalize the size, resolution, and color spaces to prepare the image data. Preprocessing may also include standardization, noise reduction, and the removal of artifacts.

Highlight Extraction:
This phase involves extracting crucial details from the preprocessed images. It transforms raw pixel data into meaningful component representations, including texture, variety, shape, and intensity.

Feature Extraction:
Utilize feature extraction techniques to extract relevant information from the preprocessed images. Components may include surface analysis, edge identification, and other pertinent characteristics of skin lesions.

Model Preparation:
Train the selected model using the designated dataset. This process involves using one portion of the data for training and another for validation to assess the model's performance.

Model Assessment:
Evaluate the prepared model's performance using various metrics, including precision, accuracy, and recall.

V. IMPLEMENTATION

a) Block Diagram

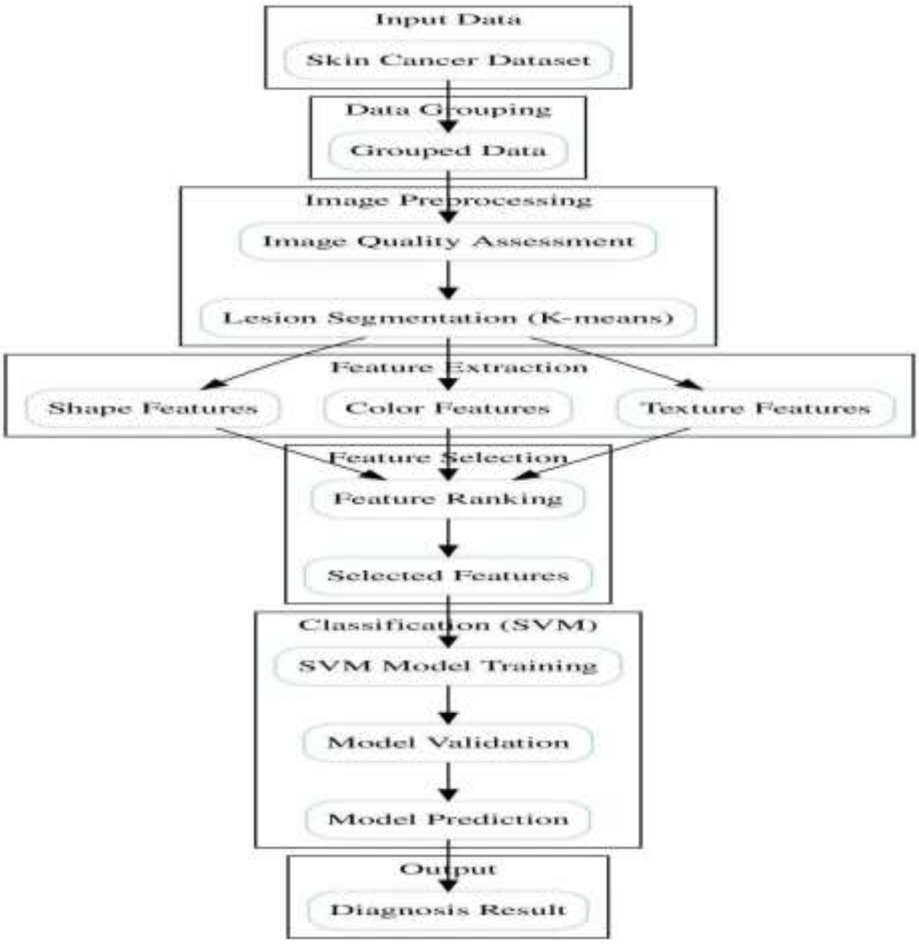

Figure1: Block Diagram

Image Input:
The proposed system utilizes a dataset comprising high-resolution dermoscopic images. Specifically, the ISIC 2019 test dataset, consisting of 800 compressed images and featuring eight distinct classes, is employed [8-10].

Pre-processing:
In brief, photo acquisition methods may vary; hence, the primary objective of the preprocessing stage is to enhance image boundaries—clarity, quality, etc.—by eliminating or reducing unwanted elements from the image or background. The dull razor approach is employed for hair removal, involving the following steps:
1. Grayscale morphological activity identifies hair locations on the lesion.
2. After locating hair pixels, it assesses if the pattern is long or delicate, using bilinear interpolation to replace hair pixels.
3. Lastly, it smoothens the replaced hair pixels with a flexible median filter.

Segmentation:
Segmentation, a key technique for isolating the region of interest in an image, treats each pixel similarly based on certain attributes. K-means clustering divides the information into approximately k groups or clusters using centroids. The process involves:
a) Selecting the number of clusters, k.
b) Choosing irregular k points to represent centroids.
c) Allocating all information to the nearest centroid to form clusters.
d) Identifying and updating each cluster's new centroid.
e) Shifting the focus to the closest new centroid.

This cycle repeats until convergence.

Feature Extraction:
Highlight extraction, considered a crucial step in characterization, involves extracting relevant features for computations, such as location and further characterization, from the dataset. The suggested framework employs GLCM to extract skin lesion highlights,

providing data for further characterization Among the characteristics are the following: deviation, diameter, mean variety channel values, standard vector, energy, entropy, autocorrelation, relationship, homogeneity, and difference.

Classification:
Multi-Class Support Vector Machine (MSVM), a component of the Support Vector Machine (SVM), is employed to address problems involving multiple classes. SVM, known for its precision, separates items into different classes by addressing the concept of decision planes. However, for multiclass problems, coordinating outputs across subclasses introduces complexity, requiring division of one class's output among multiple subclasses.

b) Hardware Specifications:

Core i5 processor from Intel
Memory: 8 GB
500GB HDD
Operating System: Windows 7 or later

c) Necessary Software:

Matlab R2023b

Matlab R2023b is a specific version of the Matlab programming language—an extremely high-level language and development environment often used for mathematical computation, data analysis, and visualization.

Some noteworthy features and updates in Matlab R2023b include:

1. Live Manager: An improved intelligent report environment with live scripts that combine output, code, and structured text into a single executable file.
2. Equitable Processing: Enhanced performance and flexibility to enable equitable registration with updates to conduct circles and similar functions.
3. Deep Learning: Comprehensive support for deep learning models, including enhancements for brain organization preparation and transmission recall.
4. Enhanced Data Input and Export: Improved options for information input and export for various formats, including JSON, Avro, Parquet, and HDF5.
5. Graphics and Visualization: Improved and new features for creating intuitive visuals and modifying storylines.
6. Application Originator: Improvements to the Application Architect tool for designing and developing unique user interfaces.

VI. RESULT AND DISCUSSION

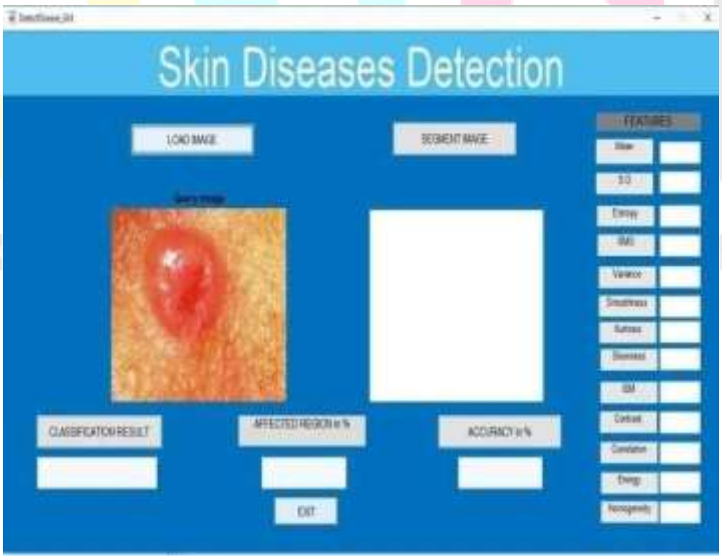

Figure1: Skin cancer detection using image processing and Machine learning

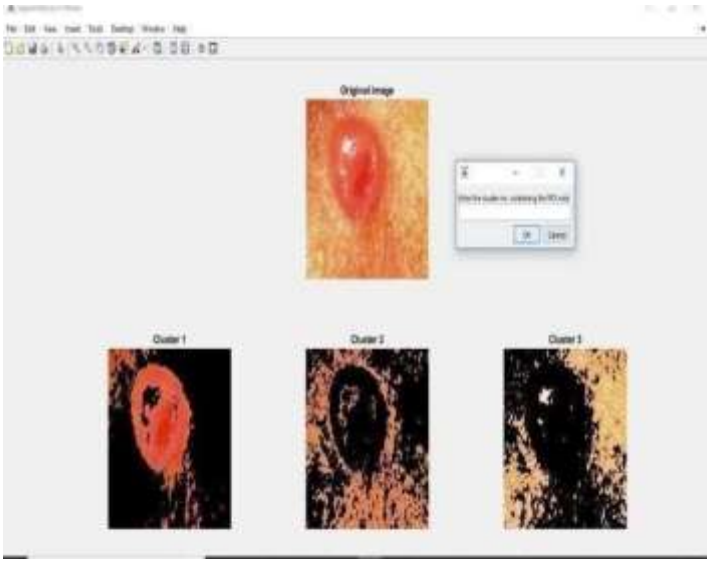

Figure2: Skin cancer clustering for selection of region of Interest (ROI) K means

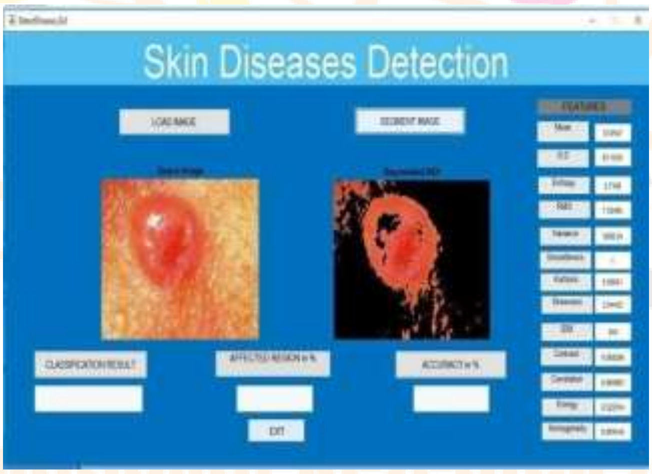

Figure3: Feature extraction on selected image of ROI

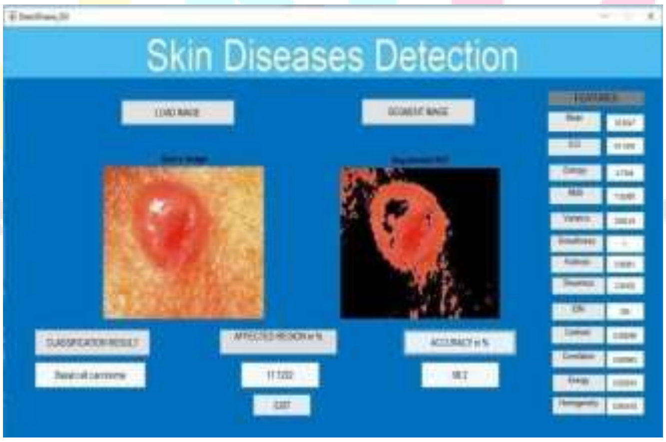

Figure4. Classification result for ROI Skin cancer, defining type of cancer

VII. CONCLUSION AND FUTURE ENHANCEMENT

a) CONCLUSION

Due to several factors, there has been a notable global increase in the incidence of skin cancer cases For prompt diagnosis and treatment, early detection is therefore essential. This paper explores a methodology based on MSVM grouping, extracting highlights through two robust algorithms, GLCM and MSVM. The achieved accuracy is approximately 83.5%. The inclusion of four types of skin cancers in the proposed framework enables efficient organization and proficient attainment of high accuracy.

b) FUTURE ENHANCEMENT

Beginning with an underlying arrangement of estimated information in AI, picture handling, and design acknowledgment, highlight extraction is carried out. It then creates determined values (highlights) that are expected to be educational and non-excess, interacting with the subsequent learning and speculation steps and occasionally leading to better human understandings. Emphasize Dimensionality reduction is related to future extraction. A reduced arrangement of highlights, also known as a component vector, is typically created when the amount of information needed for a calculation is deemed excessive and too large to handle in any way (for instance, similar estimation in feet and meters, or the dreariness of pictures introduced as pixels). The selected highlights are meant to comprise the relevant information from the data, so the optimal task can be completed by using this reduced portrayal instead of the entire initial data.

VIII. REFERENCES

[1]. Dwaipayan Choudhury 1, Avisek Naug1, Susmita Ghosh2 deals with a “Texture and Color Feature Based WLS Framework Aided Skin Cancer Classification using MSVM and ELM” [2015].

[2]. Zahra Wahee, Madeeha Zafar, et al, deals with a “An Efficient Machine Learning Approach for the Detection of Melanoma using Dermoscopic Images” 2017.

[3]. D. Mandachea, E. Dalimierb, JR. Durkinc, C. Boccarab, J.-C. Olivo-Marina, V. Meas-Yedida “Basal cell carcinoma detection fullfield oct images using convolutional neural networks” [2018].

[4]. E. Nasr-Esfahani, S. Samavi, N. Karimi, et al, Deals with a “Melanoma Detection by Analysis of Clinical Images Using Convolutional Neural Network”[2016].

[5]. Suleiman Mustafat, Akio Kimura. “A SVM-Based Diagnosis of Melanoma Using Only Useful Image Features” [2018].

[6]. Yuheng Wang, Jiayue Cai, Daniel C. Louie, et al, “Classifying Melanoma and Seborrheic Keratosis Automatically with Polarization Speckle Imaging” [2019].

[7]. Md. Zahid Hasan CSE, et al, “Integrated Use of Rough Sets and Artificial Neural Network for Skin Cancer Disease Classification”[2019].

[8]. Vidya M Department of Electronics and Communication Ramaiah Institute of Technology Bangalore, et al, “Skin Cancer Detection using Machine Learning Techniques” [2020].

[9]. MA. Ahmed Thaajwer UA. Piumi Ishanka, et al, “Melanoma Skin Cancer Detection Using Image Processing and Machine Learning Techniques” [2020]