Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 14s (2025)
https://www.jneonatalsurg.com

Skin Cancer Detection Using Image Processing

Archana[1], Dr. M. Sheela Newsheeba[2], Harikrishnan V[3], Govindaraj S[4], Bharathkumar S[5], Mozhivendhan H[6], Sheik Adbul Rahman J[7]

[1]archanau2019@gmail.com

[2]Assistant Professor, Department of Computer Applications, Nehru Arts and Science, Coimbatore.

Email ID: sheela79.rincy@gmail.com

[3]harikrishhnaah@gmail.com,

[5]bharathkumarb921@gmail.com

[6]sdmozhikvltn@gmail.com

[7]sheikabdulrahmanj@gmail.com

.Cite this paper as: Archana, Dr. M. Sheela Newsheeba, Harikrishnan V , Govindaraj S, Bharathkumar S, Mozhivendhan H, Sheik Adbul Rahman J, (2025) Skin Cancer Detection Using Image Processing. *Journal of Neonatal Surgery*, 14 (14s), 487-493.

ABSTRACT

Skin Cancer is one of the most common and potentially life-threatening diseases worldwide, characterized by the abnormal growth of skin cells. This condition often arises due to prolonged exposure to harmful ultraviolet (UV) rays from sunlight. However, it can also develop in areas of the body that are not exposed to the sun. Early detection is essential for effective treatment, as skin cancer is highly manageable when identified in its initial stages. The conventional method for diagnosing skin cancer involves performing a biopsy, where a tissue sample is extracted and analyzed in a laboratory. Although this method is accurate, it is often time-consuming, invasive, and may cause discomfort to the patient. This project introduces an innovative approach that leverages Support Vector Machine (SVM) and image processing techniques to facilitate early detection of skin cancer. The process begins by capturing high-resolution dermoscopic images of the affected skin. These images undergo pre-processing steps, including noise reduction and image enhancement, to improve the quality and accuracy of subsequent analysis. Once pre-processing is complete, the enhanced images are segmented to isolate the region of interest, allowing for precise examination. The next critical step involves feature extraction using the Gray Level Co-occurrence Matrix (GLCM), a powerful statistical tool that analyzes textural patterns within the image. The extracted features, which include information about texture, contrast, and homogeneity, are then fed into the SVM classifier. The SVM, a robust machine learning algorithm known for its high classification accuracy, processes these features and classifies the image as either cancerous or non-cancerous. This streamlined approach not only reduces the time required for diagnosis but also minimizes the discomfort associated with traditional biopsy procedures, making it a faster and less invasive solution for early detection of skin cancer disases..

Keywords *Thresholding, SVM, GLCM, Skin cancer, Classifier.*

..

1. INTRODUCTION

Skin cancer is a potentially fatal condition that develops in the outermost layer of the skin, which consists of three primary cell types: squamous cells, basal cells, and melanocytes. Squamous and basal cell carcinomas, commonly referred to as non-melanoma skin cancers, usually respond well to treatment and rarely spread to other parts of the body. However, melanoma, which originates from melanocytes, is far more dangerous due to its ability to invade nearby tissues and metastasize if not detected early. The traditional method for diagnosing skin cancer is a biopsy, where a sample of skin tissue is removed and analyzed in a laboratory. While effective, this procedure is both time-consuming and uncomfortable for patients. Furthermore, delays in diagnosis can increase the risk of cancer spreading. To address these concerns, this study explores an alternative method that leverages digital image processing techniques and Support Vector Machine (SVM) algorithms to detect skin cancer early. By analyzing high-resolution images without the need for invasive procedures, this method enables faster and more accurate identification of cancerous lesions, thereby reducing the risk of unnecessary excisions and ensuring timely intervention.

2. LITERATURE REVIEW

This section highlights the contributions of various researchers in the field of skin cancer detection using image processing and machine learning techniques:

- J Abdul Jaleel (2013): This study introduced a skin cancer detection model using Maximum Entropy Thresholding for segmentation, followed by feature extraction using the Gray Level Co-occurrence Matrix (GLCM). The extracted features were then classified using an Artificial Neural Network (ANN), specifically employing the Back-Propagation Neural (BPN) Network for classification.
- M. Chaithanya Krishna (2016): The research utilized clustering techniques for segmentation and employed the ABCD (Asymmetry, Border, Color, and Diameter) method for feature extraction. This method focuses on identifying visual attributes that are commonly associated with malignant skin lesions.
- A.A.L.C. Amarathunga (2015): This system implemented a rule-based and forward-chaining approach to detect skin diseases, with a particular focus on pediatric dermatology. The system provides online diagnosis and suggests appropriate medical advice. Various classification algorithms such as AdaBoost, BayesNet, Multilayer Perceptron (MLP), and Naive Bayes were used to predict and diagnose skin diseases. However, this approach was limited to detecting only three types of skin conditions—Eczema, Impetigo, and Melanoma.
- Kawsar Ahmed (2013): The study applied different data pre-processing techniques and utilized the Maximal Frequent item algorithm for training. K-means clustering was employed for image segmentation, and significant frequent patterns were extracted to aid in classification.
- Mariam A. Sheha, Mai S. Mabrouk, and Amr Sharawy (2012): Their research proposed a method for diagnosing melanoma through digital image analysis. GLCM was used for texture-based feature extraction, and a Multilayer Perceptron (MLP) classifier was applied to distinguish between cancerous and non-cancerous images.

3. EXISTING SYSTEM

The existing system for skin cancer detection includes manual dermatological diagnosis, traditional computer-aided diagnostic (CAD) systems, and deep learning-based models like ResNet. Manual diagnosis relies on visual inspection and biopsy tests, but it is time-consuming, subjective, and not always accessible. Traditional CAD systems use image processing and machine learning but struggle with complex cases and lower accuracy. Deep learning models like ResNet have improved accuracy (82.87%) but still face misclassification risks, sensitivity to image quality, limited interpretability, data imbalance issues, and high computational requirements. These challenges make AI-based skin cancer detection less reliable for real-world applications, requiring further improvements for higher accuracy, better transparency, and cost-effective deployment.

3.1 Disadvantages

- Challenges in Accuracy for Critical Cases: While achieving an 82.87% accuracy is commendable, it implies that the system may still misclassify a significant number of cases. In critical scenarios where timely and accurate diagnoses are paramount, this level of accuracy could result in delays in administering appropriate treatment to individuals with skin cancer.
- Sensitivity to Image Quality: Deep learning systems, such as the RESNET model, can exhibit sensitivity to variations in image quality, including factors like lighting conditions, resolution, and the presence of noise. Suboptimal image quality or significant deviations from the training data may hinder the system's performance, potentially leading to reduced accuracy.
- Limited Interpretability: Understanding deep learning models like RESNET can be challenging due to their lack of transparency in explaining prediction mechanisms. This lack of clarity poses concerns, particularly in critical medical contexts where the rationale behind decision-making is crucial.
- Data Imbalance Issues: In some cases, there may be insufficient data available for all types of skin cancer, resulting in imbalanced datasets. This imbalance can introduce bias and diminish accuracy, particularly for less prevalent cancer types.
- Resource Intensiveness: The development and training of deep learning models like RESNET require substantial computational resources, such as powerful GPUs or TPUs. This demand for resources can entail significant setup and operational costs, making it challenging to deploy and maintain the system.

4. PROPOSED METHODLOGY

Deep Learning Techniques: The proposed system leverages advanced deep learning methods to enhance skin cancer prediction, with a specific focus on analyzing dermatoscopic images.

- Convolutional Neural Networks (CNNs): The system utilizes CNNs, a type of deep learning architecture, along with tailored enhancements to improve accuracy, efficiency, and overall performance in classifying skin cancer.
- Cutting-edge Architecture: At the heart of the system lies a state-of-the-art deep-learning architecture, designed to effectively extract intricate patterns from dermatoscopic images and identify critical features associated with various types of skin cancer.
- Data Augmentation: The system employs data augmentation techniques, such as rotation, scaling, flipping, and cropping, to diversify the dataset and reduce the risk of overfitting, thereby enhancing model generalization.
- Addressing Class Imbalance: To tackle class imbalance within the dataset, the system implements strategies like oversampling, undersampling, or class weighting, ensuring fair representation of different skin cancer types and improving prediction accuracy.

General process of the Proposed Systems

The proposed system focuses on detecting skin cancer by analyzing dermoscopic images and classifying them using image processing techniques and a Support Vector Machine (SVM) classifier.

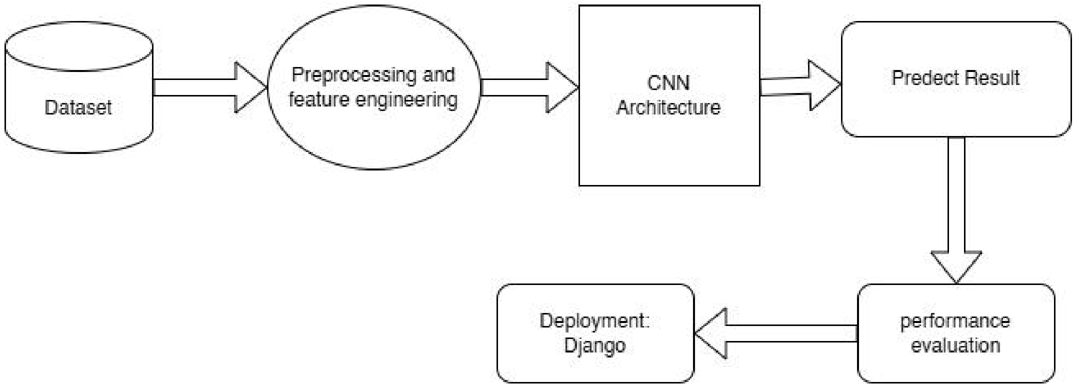


Fig : Skin Cancer Predication - System Process

The system follows a structured approach, beginning with image acquisition and progressing through several stages:

4.1.Data Collection

The input images are collected from reliable medical databases such as ISIC Archive, HAM10000, and Kaggle dermatology datasets. These datasets contain labeled images of skin lesions categorized as benign or malignant. The images may also be gathered from clinical sources, hospitals, or direct user uploads through the web application.

Fig : Input Image

4.1.2 Pre-Processing

Since raw images come in different sizes, resolutions, and formats, preprocessing is necessary to standardize them before they are fed into the model. Key preprocessing steps include:

- Resizing: All images are resized to a fixed dimension (e.g., 224×224 pixels) to ensure consistency across the dataset.
- Normalization: Pixel values are scaled between 0 and 1 to improve computational efficiency and model convergence.
- Noise Reduction: Unwanted artifacts, shadows, or distortions in images are minimized using image enhancement

techniques.

Fig : Preprocessed Image

Grayscale Conversion

Converting a color image to grayscale simplifies the processing by retaining only the brightness information. In a grayscale image, each pixel's value corresponds to the intensity of light, ranging from black (0) to white (255) in an 8-bit image. Grayscale images are computationally less intensive, making them faster and easier to process than colored images. In this system, the RGB image is converted to grayscale using the weighted sum method with the following formula:

Grayscale Intensity=0.299R+0.587G+0.114B\text{Grayscale Intensity} = 0.299R + 0.587G + 0.114BGrayscale Intensity=0.299R+0.587G+0.114B

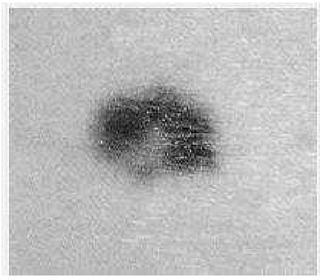


Fig 4: Noise Removal

Noise removal aims to eliminate unwanted distortions that may affect the quality of the image. Noise appears as random variations in pixel values, making it difficult to differentiate between actual features and noise. In this system, a median filter is applied to remove noise. The median filter is a non-linear technique that preserves edges while removing noise. It works by sliding a window of odd length over the image, sorting the pixel values within the window, and replacing the central pixel with the median value.

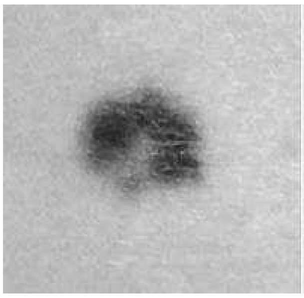


Fig 4: Image Without Noise

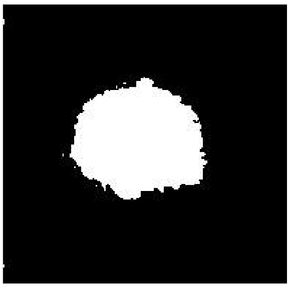
Image Enhancement

The goal of image enhancement is to improve the visibility of significant features within the image. Contrast enhancement is applied to achieve better visual quality, making the regions of interest more prominent for subsequent processing. Here contrast enhancement is used to get better-quality results shown in Fig (5).

Fig 2: Enhanced Image

4.1.3 Segmentation

Segmentation is the process of removing a region of interest from a given image. Region of interest containing each pixel similar attributes. Here we are using maximum entropy thresholding for segmentation [5]. First of all, we have to take the gray level of an original image and then calculate the histogram of the grayscale image by using maximum entropy to separate the foreground from the background. After maximum entropy, we obtained a binary image that is black and white image shown in fig (6).



Fig; Segmented Image

4.1.4 Feature Extraction

Feature extraction plays an important role in extracting information present in the given image. Here we are using GLCM for texture image analysis. GLCM is used to capture spatial dependency between image pixels. GLCM works on gray level image matrix to capture the most common features such as contrast, mean, energy, and homogeneity [2].

4.1.5 Classifier

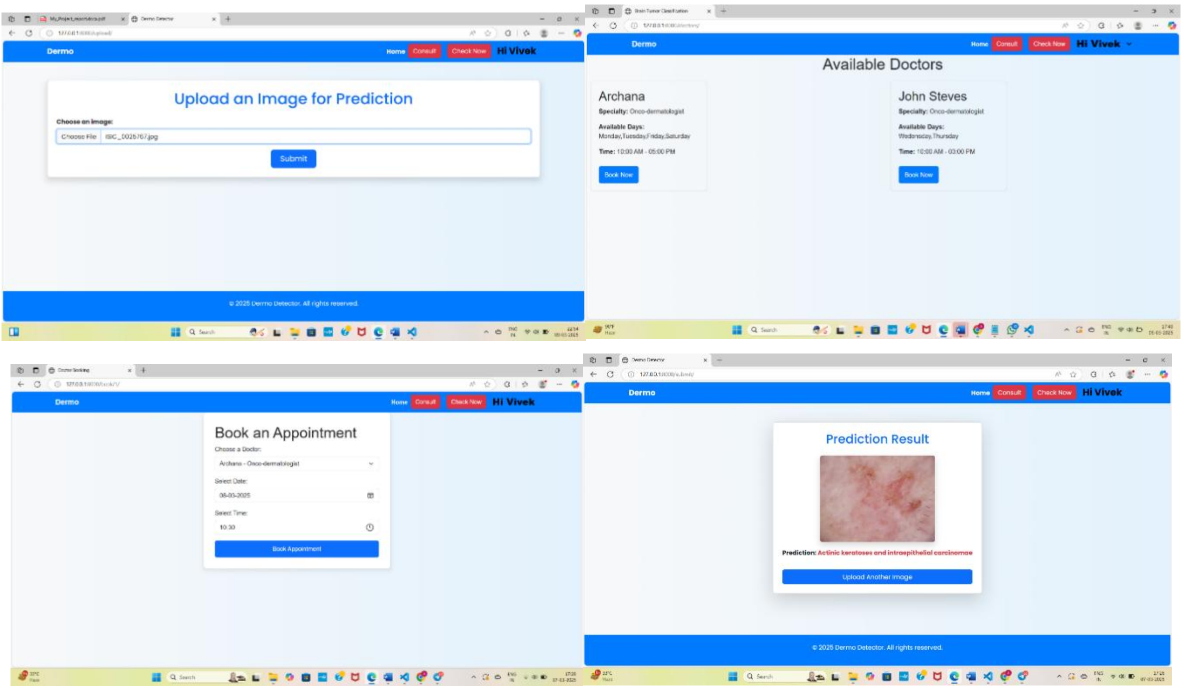
The classifier is used to classify cancerous images from other skin diseases. For simplicity, the Support Vector Machine classifier is used here. SVM takes a set of images and predicts for each input image belongs to which of the two categories of cancerous and non-cancerous classes. The purpose of SVM is to create hyper plane that separates two classes with the maximum gap between them [2]. In our proposed system output of GLCM is given as input to the SVM classifier which takes training data, testing data, and grouping information which classifies whether the given input image is cancerous or non-cancerous shown in Fig (7) [11].

4.1.6 Prediction Result

Once an image is uploaded, the system runs the image through the Convolutional Neural Network (CNN) model, which processes it and classifies the lesion as benign or malignant. The prediction result is displayed in an easy-to-understand format, which may include:

- A label: "Benign" (Non-cancerous) or "Malignant" (Cancerous).
- Probability score: A confidence level (e.g., Benign: 85% or Malignant: 92%) to indicate the model’s certainty.
- Additional recommendations: The system may suggest seeking professional medical consultation if the lesion is predicted to be malignant.

The results are designed to be clear and concise, ensuring that both medical professionals and general users can interpret the output easily. Testing was performed on 20 sample images. Accuracy is calculated by using the following formula.



4.1.7 Deployment with Django Web Application

The trained model is deployed in a Django-based web application, which serves as the backend framework to handle image uploads, process predictions, and display results. The deployment architecture includes:

- Django for the backend: Manages user requests, sends images to the model, and returns predictions.
- TensorFlow/Keras for the deep learning model: Loads the trained CNN model and runs inference on uploaded images.
- HTML, CSS, and JavaScript for the frontend: Ensures a responsive and visually appealing web interface.
- Database integration (if required): Stores user uploads and prediction history for future reference.

The web application ensures that the model is accessible online and can handle multiple user requests efficiently.

5. CONCLUSION

The Machine Learning-Based Skin Cancer Prediction System represents a significant advancement in healthcare technology, particularly in the field of dermatology. With the rising number of skin cancer cases worldwide, early detection and accurate diagnosis have become essential to reduce mortality rates and improve patient outcomes. Traditional methods of skin cancer diagnosis often involve physical examination and biopsy, which can be time-consuming and sometimes lead to delayed treatment. However, the integration of Machine Learning (ML) and Deep Learning (DL) models has revolutionized the diagnostic process by offering automated and highly accurate predictions based on medical images. This technology has the potential to minimize human error, reduce diagnosis time, and increase the chances of early detection, ultimately saving thousands of lives every year.

The adoption of ML algorithms, such as Convolutional Neural Networks (CNN), has proven to be highly effective in classifying and detecting different types of skin cancer by analyzing dermoscopic images. These models are trained on vast datasets of skin lesion images, enabling them to learn and identify patterns that indicate the presence of skin cancer. By leveraging artificial intelligence, the system can differentiate between benign (non-cancerous) and malignant (cancerous) lesions with a high degree of accuracy. Additionally, the automated nature of the system makes it accessible to remote and underserved areas where expert dermatologists may not be available. With continuous advancements in machine learning and the availability of large datasets, the Skin Cancer Prediction System has the potential to revolutionize healthcare by providing early, accurate, and cost-effective skin cancer detection, ultimately reducing the global burden of the disease.

It can be easily concluded that the proposed system of skin cancer detection can be implemented using gray level cooccurrence matrix and support vector machine to classify easily whether the image is cancerous or non-cancerous. The accuracy of the proposed system is 95%. It is a painless and timeless process than the biopsy method. It is more advantageous to patients.

6. FUTURE ENHANCEMENT

To improve the accuracy and reliability of the ML-based Skin Cancer Prediction System, future enhancements can focus on advanced deep learning architectures and multi-modal data integration. Implementing Vision Transformers (ViTs) or EfficientNet can enhance feature extraction and classification accuracy. Additionally, integrating Explainable AI (XAI) techniques such as Grad-CAM, SHAP, or LIME will help dermatologists understand model predictions, improving trust and clinical adoption. Moreover, incorporating multi-modal data, including patient demographics, genetic information, and clinical history, can provide a more comprehensive assessment, leading to better diagnostic precision.

Another key improvement is enhancing accessibility and real-time deployment. Developing a mobile application that allows users to upload skin lesion images for instant analysis can enable early detection, especially in remote areas with limited access to dermatologists. Additionally, integrating the system into cloud-based APIs will enable seamless deployment in healthcare settings. Telemedicine integration can further support remote consultations by providing AI-assisted insights to dermatologists. To overcome data scarcity and improve model generalization, semi-supervised learning, self-supervised learning, and Generative Adversarial Networks (GANs) can be utilized to generate synthetic skin lesion images for model training.

Furthermore, incorporating automated disease progression prediction can enhance early intervention strategies. By analyzing sequential images of a lesion over time, the model can predict potential malignancy progression, aiding dermatologists in making informed decisions. Additionally, implementing federated learning will allow multiple hospitals and clinics to train the model collaboratively without sharing sensitive patient data, ensuring privacy and security. Lastly, integrating blockchain technology for secure and transparent medical data management can enhance trust, regulatory compliance, and ethical AI use in healthcare. These enhancements will ensure a more accurate, interpretable, and widely accessible AI-driven skin cancer detection system, ultimately improving early diagnosis and patient outcomes.

REFERENCES

[1] J Abdul Jaleel, Sibi Salim, Aswin.R.B, Computer Aided Detection Skin Cancer, International Conference on Circuits, Power and Computing Technologies, 2013.

[2] C.Nageswara Rao, S.Sreehari Sastry and K.B.Mahalakshmi Co-Occurrence Matrix and Its Statistical Features an Approach for Identification Of Phase Transitions Of Mesogens, International Journal of Innovative Research in Engineering and Technology, Vol. 2, Issue 9, September 2013.

[3] Santosh Achakanalli& G. Sadashivappa, Statistical Analysis Of Skin Cancer Image –A Case Study, International Journal of Electronics and Communication Engineering (IJECE), Vol. 3, Issue 3, May 2014.

[4] Digital image processing by Jayaraman. Page 244,254-247,270-273. (gray level, median filter).

[5] Algorithm For Image Processing And Computer Vision.Page 142-145.(Thresholding)

[6] Kawsar Ahmed, TasnubaJesmin, Early Prevention and Detection of Skin Cancer Risk using Data Mining, International Journal of Computer Applications, Volume 62– No.4, January 2013.

[7] M.Chaithanya Krishna, S.Ranganayakulu, Skin Cancer Detection and Feature Extraction through Clustering Technique, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 3, March 2016.

[8] A.A.L.C. Amarathunga, Expert System For Diagnosis Of Skin Diseases, International Journal Of Scientific & Technology Research, Volume 4, Issue 01, 2015.

[9] Mariam A.Sheha, Automatic Detection of Melanoma Skin Cancer, International Journal of Computer Applications,

[10] Anshu Bharadwaj, Support Vector Machine, Indian Agriculture Statistics Research Institute.

[11] Maurya R, Surya K.S,"GLCM and Multi-Class Support Vector Machine based Automated Skin Cancer Classification, "IEEE journal, vol 12, 2014.

.....