International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Skin Cancer disease analysis and detection using machine learning

Akanksha Kulkarni[1, 2, a)] , Nalini Jagtap[1, 2, b)] , Nikhil Kadus[3, c)] ,Devashish Dumbhare [3, d)] , Tejas Katkade[3, e)] , Yash Mohan[3,f)]

[1]Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune

[2]D Y Patil International University

[3] Department of Computer Engineering, Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune

[a)] akankshajkulkarni@gmail.com

[b)]nalinisjagtap@gmail.com

[c)] nikhilkadus1908@gmail.com

[d)]devashishdumbhare26@gmail.com

tejasbkatkade01@gmail.com

yashmohan793@gmail.com

Abstract

Skin cancer, particularly melanoma, poses a serious global health concern due to its increasing incidence and potential lethality when not detected early. Standard diagnostic techniques frequently necessitate the expertise of a dermatologist, which may not be easily accessible to everyone. This research paper presents an automated system for the early detection and classification of skin cancer using convolutional neural networks (cnn). The model is trained using a dataset of labeled dermatoscopic images to distinguish between benign and malignant skin lesions. Image preprocessing techniques such as resizing and normalization are applied to enhance model performance. In addition to the classification module, a chatbot is integrated to provide users with informative responses to queries related to skin cancer symptoms, causes, and prevention. The system aims to assist healthcare professionals in diagnosis and increase public awareness through early, accurate, and accessible detection. Experimental results demonstrate promising accuracy, highlighting the potential of CNNbased approaches in supporting dermatological diagnostics.

Keywords : Skin Cancer Detection, Melanoma, Image Processing, Machine Learning,CNN.

INTRODUCTION

Over the past three decades, the global prevalence of melanoma, the most severe form of skin cancer, has steadily increased. While early detection significantly enhances treatment success rates, the survival rate for melanoma remains lower than that of non-melanoma skin cancers due to its high potential for metastasis. Melanoma can appear on any area of the skin and commonly affects different regions in men and women. In men, it is often observed on the head, neck, or upper torso, whereas in women, it frequently occurs on the lower legs or torso. While less common in individuals with darker skin tones, when it does manifest, it is usually observed in less pigmented regions like the palms, soles, or under the nails.

Skin cancer is one of the most common and potentially dangerous forms of cancer if not identified in its early stages. Traditional diagnostic procedures rely on expert visual examination, which can be timeconsuming and subject to human error. In recent times, machine learning has gained significant attention for potential solution for improving the accuracy and efficiency of skin diagnostic procedures. By leveraging large datasets of dermoscopic images, machine learning algorithms can automatically learn to distinguish between benign and malignant lesions. This capability not only reduces diagnostic time but also supports healthcare professionals in making timely and informed decisions.

LITERATURE SURVEY

TABLE 1. Literature Survey

Sr.No	Paper Title	Advantages
1	Skin Cancer Detection and Classification using Deep Learning	Multi-class classification using CNNs Uses dermoscopic image inputs Evaluated with precision and recall metrics

2	Detection of Skin Cancer: A Deep Learning Approach	Uses large curated dataset Incorporates contrast enhancement and segmentation Applies data augmentation to prevent overfitting
3	Skin Cancer Classification Using Deep Learning Algorithms	Hybrid model combining CNN and SVM Hierarchical feature extraction via CNN Effective classification with SVM layer
4	Skin Cancer Prediction using Machine Learning and Neural Networks	Implements neural networks for prediction Uses publicly available datasets (e.g., ISIC) Enhances accuracy through tuning techniques

Study of existing techniques

1. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are among the most widely used deep learning architectures for image-based classification tasks, including skin cancer detection. The system automatically extracts and learns hierarchical features from dermoscopic images using multiple convolutional and pooling layers, allowing for precise classification of benign and malignant lesions. Their ability to learn directly from raw image data without manual feature engineering has made them a preferred choice in recent studies. Approaches based on CNNs have shown exceptional performance on benchmark datasets like isic, making them highly effective in clinical decision support systems.

2. Support Vector Machines (SVMs)

Support vector machines (svms) are a classic yet effective machine learning technique employed in the classification of skin lesions. Usually, svms are used together with handcrafted feature extraction methods like gray-level co-occurrence matrix (glcm) or histogram of oriented gradients (hog). These features are obtained from images captured using a dermatoscope and are used as input to the svm for classifying them into two categories. SVMs are beneficial for smaller datasets and are widely recognized for their capacity to handle intricate feature spaces and exhibit strong generalization capabilities.

3. Transfer Learning (Pre-trained Deep Networks)

Transfer learning has gained significant attention in the field of medical image analysis, particularly in the area of skin cancer detection. It involves employing pre-trained deep learning models such as resnet, mobilenet, or inceptionv3, which were initially trained on vast datasets like imagenet, and then finetuning them on datasets specifically curated for skin lesion images. This approach significantly reduces training time and improves classification performance, especially when the available labeled data is limited. Transfer learning combines the strengths of deep learning and data efficiency, making it a widely adopted strategy in contemporary skin cancer diagnostic systems.

Advantages of existing techniques

- **High Accuracy in Diagnosis:**
  Deep learning-based systems, especially those using Convolutional Neural Networks (CNNs), have shown excellent performance in detecting and classifying skin cancer with high accuracy, often rivaling or surpassing dermatologists in controlled tests.

- **Automated Feature Extraction:**
  Existing systems can automatically extract important features (such as texture, color, and shape) from dermoscopic images, reducing the need for manual intervention and domain expertise in feature engineering.

- **Early Detection Capability:**
  These systems aid in the early detection of skin cancer by scrutinizing even the slightest irregularities in skin lesions, thereby enhancing the likelihood of successful treatment and lowering mortality rates.

- **Support for Remote and Scalable Diagnosis:**
  Numerous existing models can be utilized in mobile or cloud-based applications, simplifying the provision of diagnostic assistance in remote or underserved regions where access to expert dermatologists may be limited.

Disadvantages of existing techniques

- **Dependence on High-Quality Image Data:**
  Most existing systems require high-resolution and well-lit dermoscopic images. Poorquality or noisy images can significantly reduce the accuracy of diagnosis and lead to false predictions.

- **Limited Dataset Diversity:**
  Numerous models are trained using datasets that are limited in terms of diversity, including variations in skin tones, lesion types, and

demographic characteristics. This limits the generalizability of the system across different populations and geographic regions.

- **Overfitting in Deep Learning Models:**
  Deep learning systems, particularly convolutional neural networks (cnns), are susceptible to overfitting when trained on limited or imbalanced datasets. This leads to high performance during training, but is less accurate in actual testing.

Proposed Methodology

SYSTEM ARCHITECTURE

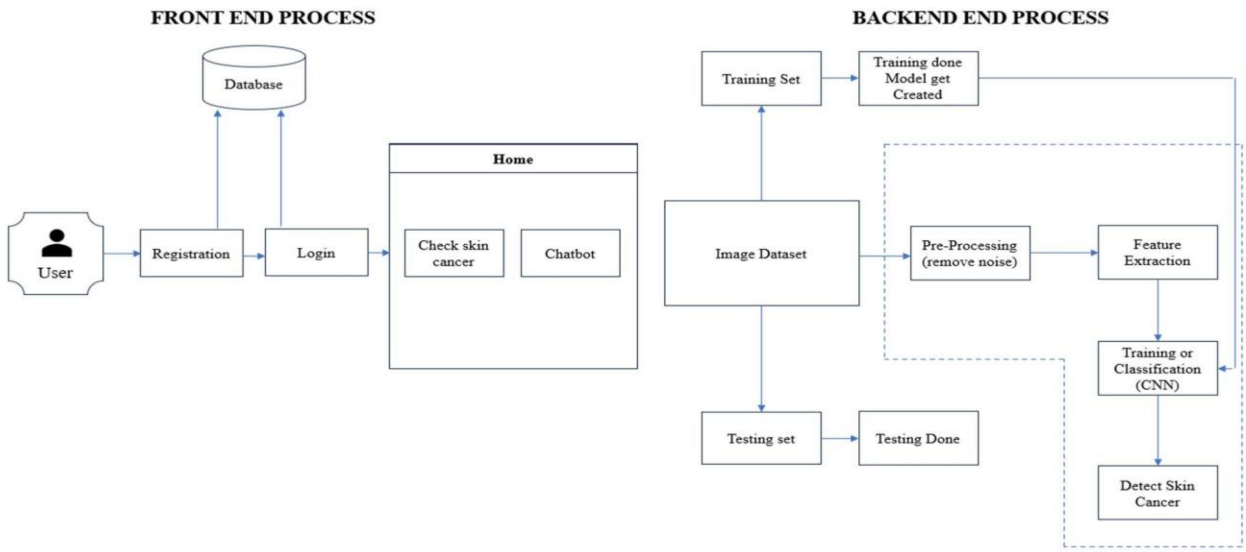

FIGURE 1. Architecture of the Proposed Model

Fig 1.0 The image depicts the entire process of a skin cancer detection system. The front-end process enables users to register, log in, and utilize features such as checking for skin cancer and interacting with a chatbot. The back-end process handles image dataset processing, where images undergo preprocessing to remove noise, followed by feature extraction. CNN-based classification is used to recognize skin cancer.. The system is trained and tested on datasets to ensure accurate diagnosis and model effectiveness.

Block Diagram

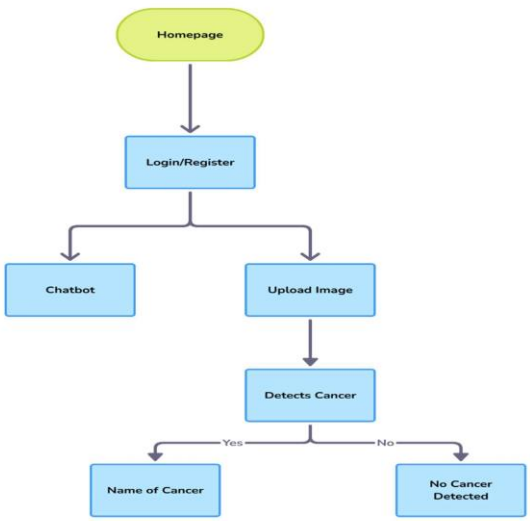

FIGURE 2. Application Structure

Fig 2 The image depicts the structure of a skin cancer detection system. The process commences at the homepage, where users can either log in or create an account. After being verified, users are presented with two choices: engage in a conversation with a chatbot for assistance or submit an image of a skin lesion for examination. After the image is uploaded, the system examines it and detects any indications of cancer. If cancer is identified, the system notifies and highlights the precise name of the cancer type (e.g., melanoma, basal cell carcinoma). If no signs of cancer are found, the system provides the user with the appropriate information. This flow offers a user-friendly interface, automates diagnosis, and provides additional support through the chatbot, enhancing early detection and user engagement in a streamlined manner.

Algorithm

Step 1: Initialization

- Import necessary libraries and modules.
- Set environment variables (e.g., API keys).
- Determines the useful features for processing data and text.

Step 2: User Registration and Login

- User accesses the web interface.
- Registers or logs in using credentials.
- User details is stored/retrieved from database when needed.

Step 3: Image Input and User Interaction User uploads images of skin lesions.

- Chatbot is available for general guidance.

Step 4: Image Dataset Preparation

- Image is added to the dataset (training or testing).
- Training set and test set are separated in data records.

Step 5: Image Preprocessing Apply filters to remove noise.

- Resize and normalize images.
- Optional: Segmentation technique is used to isolate lesion area.

Step 6: Feature Extraction

- Extract relevant features from preprocessed images.
- outlines the traits that were taken from the classification model.

Step 7: Model Training (CNN)

- Labeled data is used to train the CNN model.
- CNN learns to distinguish between benign and malignant lesions.

Step 8: Testing and Evaluation

- Evaluate model performance on test data.
- metrics: Accuracy, Precision, Recall, F1 Score.

Step 9: Skin Cancer Detection

- Preprocess and extract features from the user’s image.
- The image is classified using a trained CNN model.
- Display diagnosis result: Benign or Malignant.

Step 10: Result Feedback and Support

- Provide additional information via integrated chatbot.
- If necessary, redirect to medical assistance.

Novelty of Proposed Skin Cancer Detection

- **Integration of Chatbot for User Interaction**
  The proposed system includes a chatbot interface on the front end, enabling users to interact easily with the platform, ask questions, and receive guidance, which is a unique feature not commonly found in traditional diagnostic systems..

- **Automated Pre-processing for Noise Removal**
  The pre-processing stage is designed to automatically eliminate noise from the input images. This guarantees that the dataset quality is improved before feature extraction and classification, which ultimately leads to better model performance.

- **Separation of Training and Testing Pipelines**
  The system includes a clearly defined pipeline that differentiates between the training set and testing set for the image dataset, ensuring unbiased model evaluation and reducing the risk of overfitting, which is often overlooked in simpler implementations.

- **User-Friendly Registration and Login System**
  The front-end encompasses user registration and authentication, which are vital for recording diagnostic history and usage, safeguarding data, and enabling personalized interactions in medical applications.

Details of the Proposed Skin Cancer Detection

- **User-Friendly Front-End Interface**
  The system begins with an intuitive front-end interface that allows users to easily register and log in. Upon successful authentication, users can access essential features such as the "Check Skin Cancer" module and an integrated chatbot designed to assist users by answering queries related to skin cancer detection and general information, enhancing platform accessibility even for non-technical users.

- **Secure Database Integration**
  All user-related data, including login credentials and possibly uploaded images, are stored in a centralized database. This ensures secure data management and supports future use cases like tracking diagnostic history or user-specific analytics.

- **Structured Dataset Handling**
  The dataset is divided into training and testing sets, adhering to established guidelines in machine learning. This ensures the model is trained on one portion of the data and evaluated on a separate, unseen portion, promoting better generalization and model reliability.

- **Image Pre-processing for Quality Enhancement**
  Prior to extracting features from the input images, a pre-processing phase is carried out to eliminate noise and guarantee uniform image quality. This step enhances the quality of the input data, which is crucial for the success of feature extraction and classification stages.

- **Skin Cancer Classification Using CNN**
  The backend process employs convolutional neural networks (cnns) to automatically identify and extract important features from images of skin lesions. CNNs possess the capability to acquire complex patterns and textures directly from raw image data, enabling accurate and efficient classification of different types of skin cancer without the necessity for manual feature extraction methods.

Contribution of Proposed Skin Cancer disease analysis and detection

- **Automated and Scalable Detection Pipeline :**
  The entire process from image pre-processing to classification is automated, making the system scalable and capable of handling large volumes of image data efficiently, which is ideal for deployment in real-world clinical or remote settings.

- **Interactive Interface with Chatbot Support :**
  A unique contribution is the inclusion of a chatbot in the user interface, which allows users to interactively communicate with the system, ask health-related queries, and receive guided instructions, thereby improving usability and accessibility.

- **Secure User Management System :**
  By incorporating registration and login functionalities, the system ensures secure and personalized access for users. This also creates a medical record system, which enables the tracking of user history and repeated usage.

- **Pre-processing for Improved Image Quality :**
  The proposed system includes a dedicated image pre-processing stage to remove noise and enhance image quality, which is a crucial step often overlooked in basic implementations. This significantly enhances the efficiency and effectiveness of the feature extraction and classification process.

Results

FIGURE 4. Final Result

Figure 4 shows the proposed system's interface displays a user-friendly homepage for the Skin Cancer Detection System, featuring options to register, log in, or learn about the system. A button prompts users to sign up, and a chatbot positioned nearby offers immediate support, ensuring the system is userfriendly and interactive for early detection..

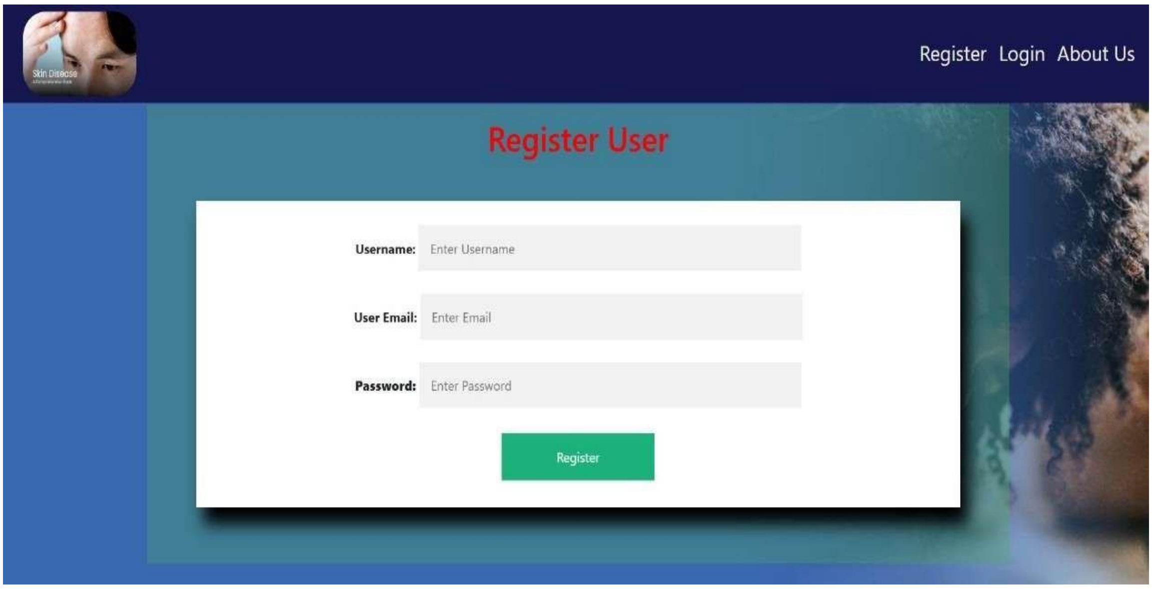

Figure 5. Signup

Figure 5 shows the diagram displays the user registration page of the Skin Cancer Detection System. It incorporates fields for username, email, and password, enabling new users to establish a secure account. The register button sends the data, granting users access to customized diagnostic tools and services within the system.

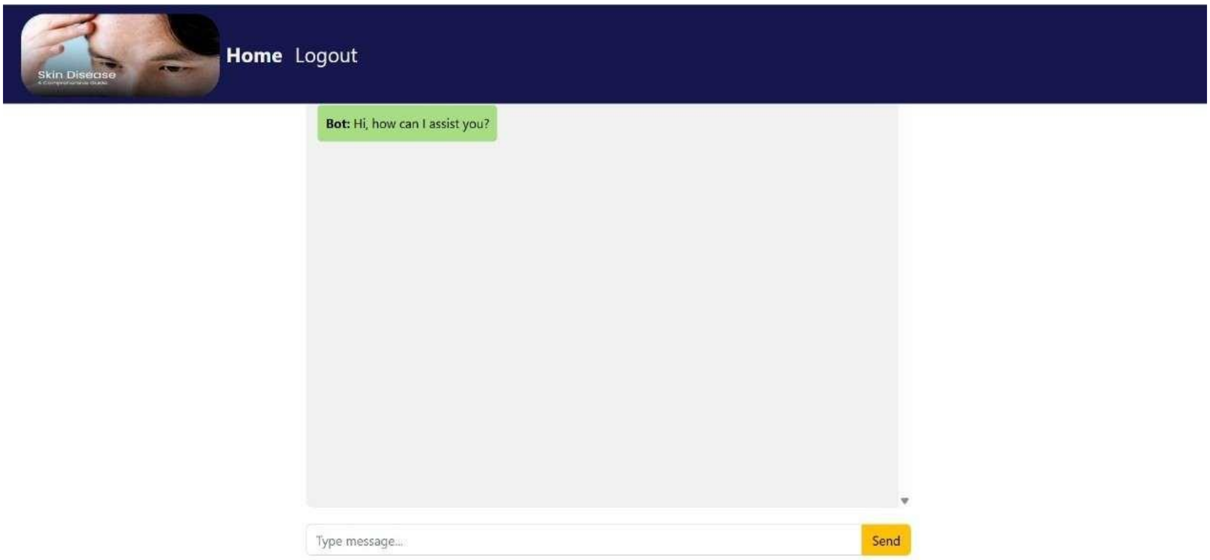

Figure 6. Chatbot

Figure 6 shows the diagram illustrates the user interface of the skin cancer detection system. It includes an interactive assistant that welcomes users and provides assistance. Users can enter their queries into the message box and click send to initiate communication. This enhances user experience by providing instant support and guidance within the system.

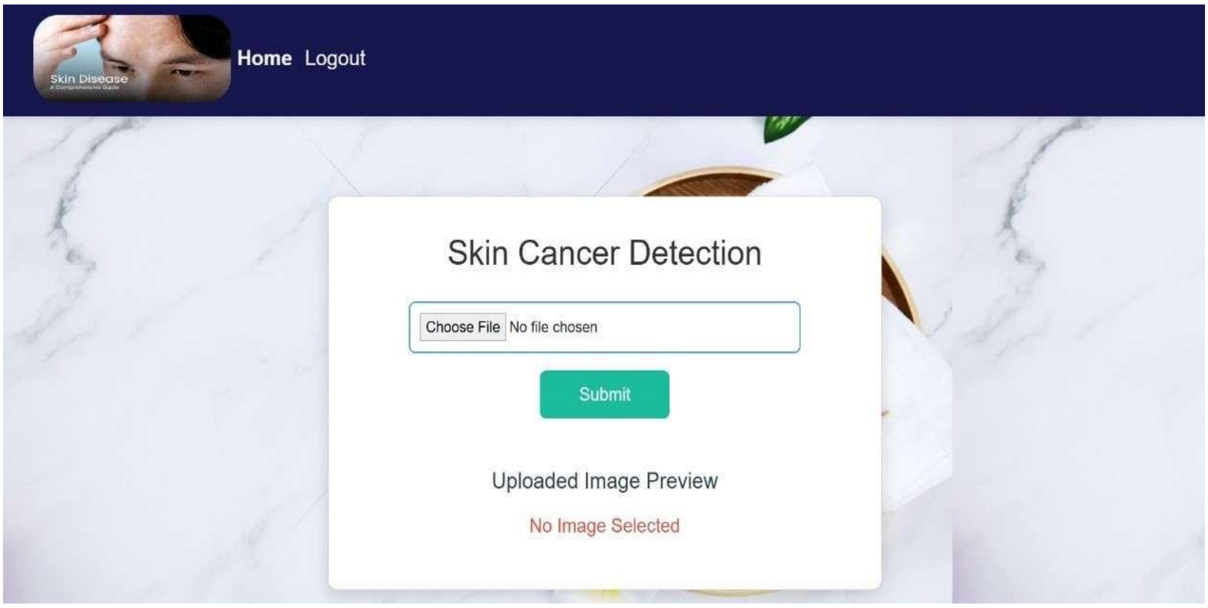

Figure 7. Upload Image

Figure 7 shows the diagram displays the Skin Cancer Detection interface. Users can upload an image of a skin lesion using the Choose File option. Once the selection is made, pressing the button triggers the image's analysis for recognition. Below, an Uploaded Image Preview section shows the chosen image, ensuring users verify their upload before submission.

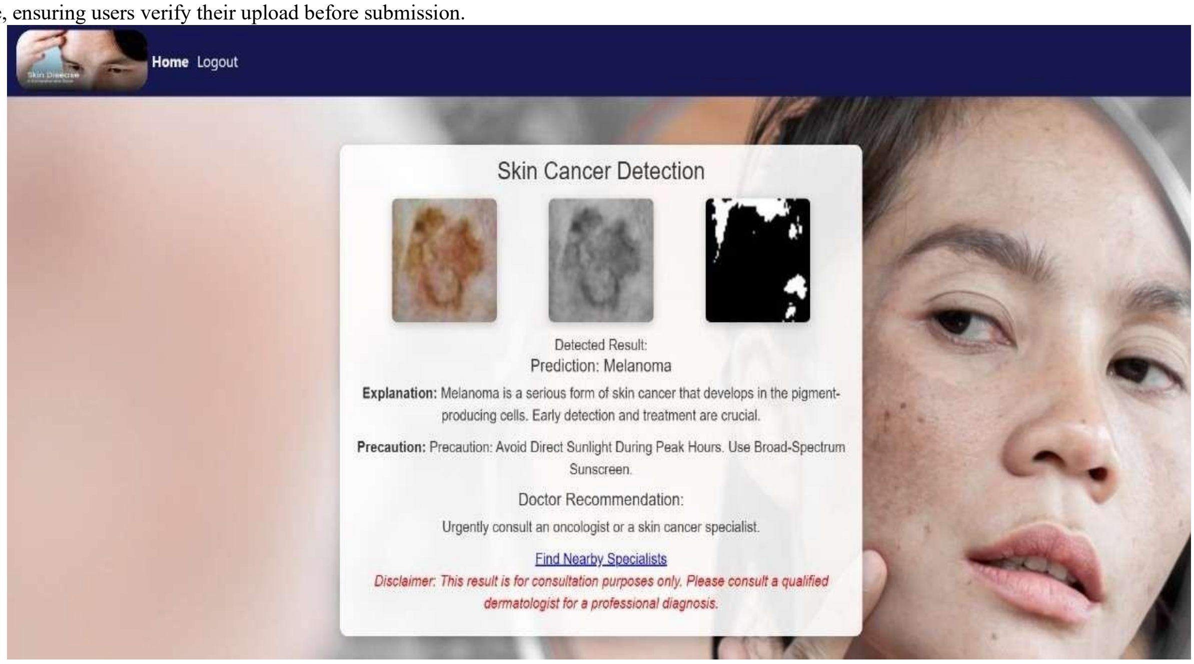

Figure 8. Result

Figure 8 shows the diagram illustrates the outcomes of the skin cancer recognition system. The image displays a processed lesion, indicating a potential melanoma. It provides a comprehensive explanation of the disease, offers advice on sun protection, and emphasizes the importance of seeking immediate consultation with a specialist. The disclaimer clarifies that the results are meant for guidance and not a specific medical diagnosis.

Conclusion

In this study, a Convolutional Neural Network (CNN)-based model was developed for the classification of skin cancer using dermoscopic images. The results demonstrate that CNNs effectively learn and extract relevant features automatically, enabling accurate classification without the need for manual feature engineering. This model has great potential to aid dermatologists in the early and accurate identification of skin cancer. Additionally, an integrated chatbot was implemented to provide users with information on the causes, prevention, and treatment measures for skin cancer, thereby enhancing user engagement and education.

Future Scope

- **Integration with Mobile Applications**
  The system can be extended into a mobile application, allowing users to capture and upload images of skin lesions directly from their smartphones. This would make early detection more accessible, especially in remote or rural areas.

- **Expansion to Other Skin Diseases**
  In addition to skin cancer, the model can be trained and adapted to identify various dermatological conditions like eczema, psoriasis, or fungal infections by augmenting the dataset and adjusting the classification layers.

- **Real-Time Teledermatology Support**
  In the future, the system can be integrated with telemedicine platforms, enabling real-time consultations with dermatologists, where AI-based results can assist doctors in making faster and more accurate decisions.

REFERENCES

1. J Abdul Jaleel, SibiSalim, Aswin.R.B," Computer Aided Detection 01 Skin Cancer", International Conference on Circuits, Power and Computing Technologies, 2013. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 04 Issue: 04 — Apr -2017 www.irjet.net p-ISSN: 2395-0072 © 2017, IRJET — Impact Factor value: 5.181 — ISO 9001:2008 Certified Journal — Page 2881

2. C.Nagesw–araRao, S.SreehariSastry 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. SantoshAchakanalli 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, 2012.

10. Anshubharadwaj, "Convucation nueral network", Indian Agriculture Statistics Research Institute.

11. Maurya R, Surya K.S,"GLCM and Multi Class Convocation neural networkbased Automated Skin Cancer Classification, "IEEE journal, vol 12, 2014.