Skin Cancer or Skin Diseases Detection Using Machine Learning

Subhash. G. Patil[1], Prajakta Thormise[2], Apeksha Abhale[3], Rutuja Rokade[4]

[1] *Professor, Computer Engineering, SKN Sinhgad Institute Technology And Science, Lonavala, Maharashtra, India*
[2] *Student, Computer Engineering, SKN Sinhgad Institute Technology And Science, Lonavala, Maharashtra, India*
[3] *Student, Computer Engineering, SKN Sinhgad Institute Technology And Science, Lonavala, Maharashtra, India*
[4] *Student Computer Engineering, SKN Sinhgad Institute Technology And Science, Lonavala, Maharashtra, India*

ABSTRACT

Skin cancer, a concerning public health predicament, with over 5,000,000 newly identified cases every year, just in the United States. Generally, skin cancer is of two types: melanoma and non-melanoma. Melanoma also called as Malignant Melanoma is the 19th most frequently occurring cancer in women and men. It is the deadliest form of skin cancer. In the year 2015, the global occurrence of melanoma was approximated to be over 350,000 cases, with around 60,000 deaths. The most prevalent non-melanoma tumors are squamous cell carcinoma and basal cell carcinoma. Non-melanoma skin cancer is the 5th most frequently occurring cancer, with over 1 million diagnoses worldwide in 2018. As of 2019, greater than 1.7Million new cases are expected to be diagnosed. Even though the mortality is significantly high, but when detected early, survival rate exceeds 95%.

This Motivates us to come up with a solution to save millions of lives by early detection of skin cancer. Convolutional Neural Network (CNN) or ConvNet, are a class of deep neural networks, basically generalized version of multi-layer perceptron's. CNNs have given highest accuracy in visual imaging tasks. This project aims to develop a skin cancer detection CNN model which can classify the skin cancer types and help in early detection. The CNN classification model will be developed in Python using Keras and Tensor Flow in the backend. The model is developed and tested with different network architectures by varying the type of layers used to train the network including but not limited to Convolutional layers, Dropout layers, pooling layers and dense layers. The model will also make use of Transfer Learning techniques for early convergence.

Keywords: *Medical imaging, skin cancer, melanoma classification, dermoscopy, deep learning, network fusion.*

1. INTRODUCTION:

The human skin is the body's largest organ, covering roughly 20 square feet. Its primary functions include temperature regulation, protection against UV rays and microbes, and the sensation of touch, heat, and cold. Skin cancer rates are on the rise, with melanoma causing the majority of skin cancer-related deaths. The World Health Organization reported approximately 232,000 cases worldwide, and early detection is crucial for improved survival rates. The skin has three main layers: the epidermis, dermis, and hypodermis. Skin cancer is a significant public health concern, with over 5,000,000 new cases annually in the United States, divided into melanoma and non-melanoma types. Efforts are being made to develop a skin cancer detection model using Convolutional Neural Networks (CNNs) in Python, leveraging Keras and Tensor Flow. Various network architectures, including convolutional, dropout, pooling, and dense layers, will be explored and tested using a dataset from the International Skin Imaging Collaboration (ISIC) archives. The goal is to save lives through early skin cancer detection. The human skin is the body's largest organ, covering roughly 20 square feet. Its primary functions include temperature regulation, protection against UV rays and microbes, and the sensation of touch, heat, and cold. Skin cancer rates are on the rise, with melanoma causing the majority of Skin cancer-related deaths.

The World Health Organization reported approximately 232,000 cases worldwide, and early detection is crucial for improved survival rates. The skin has three main layers: the epidermis, dermis, and hypodermis. Skin cancer is a significant public health concern, with over 5,000,000 new cases annually in the United States, divided into melanoma and non-melanoma types. Melanoma, also known as Malignant Melanoma, ranks 19th in cancer incidence for both men and women. It is the deadliest form of skin cancer, with over 350,000 cases and 60,000 deaths globally in 2015. Non-melanoma skin cancers, mainly squamous cell carcinoma and basal cell carcinoma, rank 5th in cancer incidence, with over 1 million diagnoses worldwide in 2018. Early detection greatly improves survival, exceeding 95%.Efforts are being made to develop a skin cancer detection model using Convolutional Neural Networks (CNNs) in Python, leveraging Keras and Tensor Flow. Various network architectures, including convolutional, dropout, pooling, and dense layers, will be explored and tested using a dataset from the International Skin Imaging Collaboration (ISIC) archives. The goal is to serve lives through early

2.LITERATURE SURVEY:

Nazia Hameed [1] that statistics and results from the most important implementations reported to date. They compared the performance of several classifiers specifically developed for skin lesion diagnosis and discussed the corresponding findings. Whenever available, indication of various conditions that affect the technique's performance is reported. They suggest a framework for comparative assessment of skin cancer diagnostic models and review the results based on these models. The deficiencies in some of the existing studies are highlighted and suggestions for future research are provided. Nazia Hameed et.al.

Maciej ogorzalek [2] state of the art in computer aided diagnosis systems and examined recent practices in different steps of these systems. Statistics and results from the most important and recent implementations are analyzed and reported. They compared the performance of recent work based on different parameters like accuracy, dataset, computational time, color space, machine learning technique etc. and summarized them in table format for better understanding of emergent researchers in the field of computer aided skin diagnosis systems. Research challenges regarding the different parts of computer aided skin cancer diagnosis systems are also highlighted. Maciej Ogorzałek et.al.

Fabio Santos[3] methodology for Computer Aided Detection and classification of skin lesions for diagnostic support merges medical experience with several cutting-edge technologies: image processing, pattern classification, statistical learning, assembling techniques of model based classifiers, also proposed approach proved to give excellent results giving correct classification of up to 98% of cases. Fabio Santos et.al.

Shetu rani guha [4] on the current state of automated skin lesion diagnosis, while also providing a comprehensive view into the challenges and opportunities in dermatology care. Shetu Rani Guha et.al.

A.D Mengistu [5] a machine learning based technique using convolutional neural network (CNN) for classifying seven types of skin diseases. Transfer learning, along with CNN, has been used to improve the classification accuracy on the International Skin Imaging Collaboration 2018 (ISIC) dataset. Evidence of an 11% increase in the accuracy by using transfer learning than using only CNN has been found. Compared to some existing works, performance of this proposed method is promising. A D Mengistu et.al.

Uzma Bano Ansari et.al [6] proposed a skin cancer detection system using svm for early detection of skin cancer disease. The diagnosing methodology uses Image processing methods and Support Vector Machine (SVM) algorithm. The dermoscopy image of skin cancer is taken and it goes under various pre-processing techniques for noise removal and image enhancement. Then the image is subjected to segmentation using the Thresholding method. Some features of the image have to be extracted using GLCM methodology. These features are given as the input to the classifier. Support vector Machine (SVM) is used for classification purposes. It classifies

3. Problem description:

Skin cancer is a prevalent and potentially deadly disease, with increasing incidence rates worldwide. Early detection is critical for successful treatment and improving patient outcomes. However, the current methods of skin cancer detection, such as visual examination by dermatologists or self-assessment, are limited

by subjectivity, access to healthcare, and the potential for human error. This poses a significant public health challenge, resulting in late-stage diagnoses, increased treatment costs, and preventable deaths.

The need for a more reliable, accessible, and objective method of skin cancer detection is evident. To address this problem, a skin cancer detection project aims to develop and implement advanced technology, such as artificial intelligence-driven image analysis and telemedicine, to enhance early detection and improve the overall management of this disease.

4. Objectives:

1. Develop a system to detect skin cancer or skin diseases at an early stage when treatment is most effective.
2. Improve the accuracy of diagnoses for skin conditions through the use of technology and data analysis.
3. Minimize the occurrence of misdiagnoses and false negatives in the detection of skin disorders.
4. Create an automated system that assists dermatologists in diagnosing skin conditions, reducing the burden of manual examination.
5. Develop a user-friendly and accessible tool for individuals to perform preliminary skin health screenings at home.
6. Gather a comprehensive dataset of skin images to train and validate machine learning models for skin disease detection.
7. Raise awareness and educate the public about the importance of skin health and early detection of skin diseases.
8. Integrate the technology into the existing healthcare system to improve patient care and outcomes.

Components used:

1. Recommendation Engine
2. Natural Language Processing (NLP)
3. Deep Learning (Tensor Flow/Keras)
4. User Interface (AngularJS)
5. Chatbot Integration
6. Data Preprocessing
7. User Feedback Mechanism
8. System Monitoring and Maintenance
9. Product Dat(JSON/CSV)

Methodology:

1. Data Preprocessing: Resize and standardize the images to a consistent resolution.
2. Feature Extraction: For traditional machine learning algorithms, extract relevant features from the images.
3. Model Training: Train the selected model using the training dataset.
4. Model Evaluation**:** Evaluate the model's performance using the validation dataset and appropriate evaluation metrics**.**
5. Model Interpretability: Implement methods for interpreting the model's predictions.
6. Deployment**:** Deploy the trained model in a real-world clinical setting, which may include integration with systems, telemedicine platforms, or mobile applications
7. Continuous Improvement: Collect feedback from healthcare professionals and use to refine the model continuously.

5. Requirement Analysis:

The Requirement Analysis for this project is a critical phase that encompasses several key elements. Firstly, it involves understanding and delineating the specific needs and expectations of the end-users, thereby ensuring that the recommendation system aligns with their preferences and shopping behaviour. Equally important is the identification of the data requirements, including the sources and formats of data, such as JSON or CSV files, which are pivotal in facilitating accurate recommendations. Functionality requirements encompass the core system features, including the recommendation algorithms, chatbot capabilities, and user interface components like product listing, details, and the checkout process. Non-functional aspects such as performance benchmarks, response times, and scalability targets are also vital for ensuring an efficient and reliable system. Additionally, technical requirements encompass the choice of technologies and frameworks, which in this case involve Angular JS, Python, Tensor Flow.

Architecture:

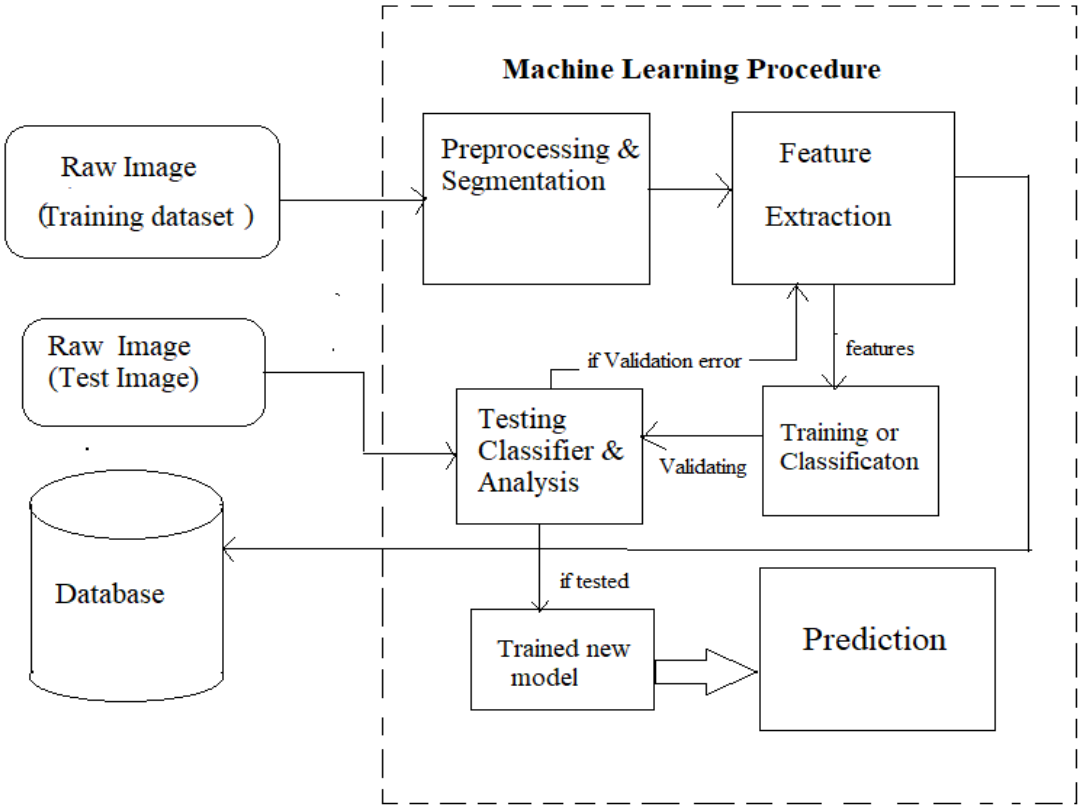
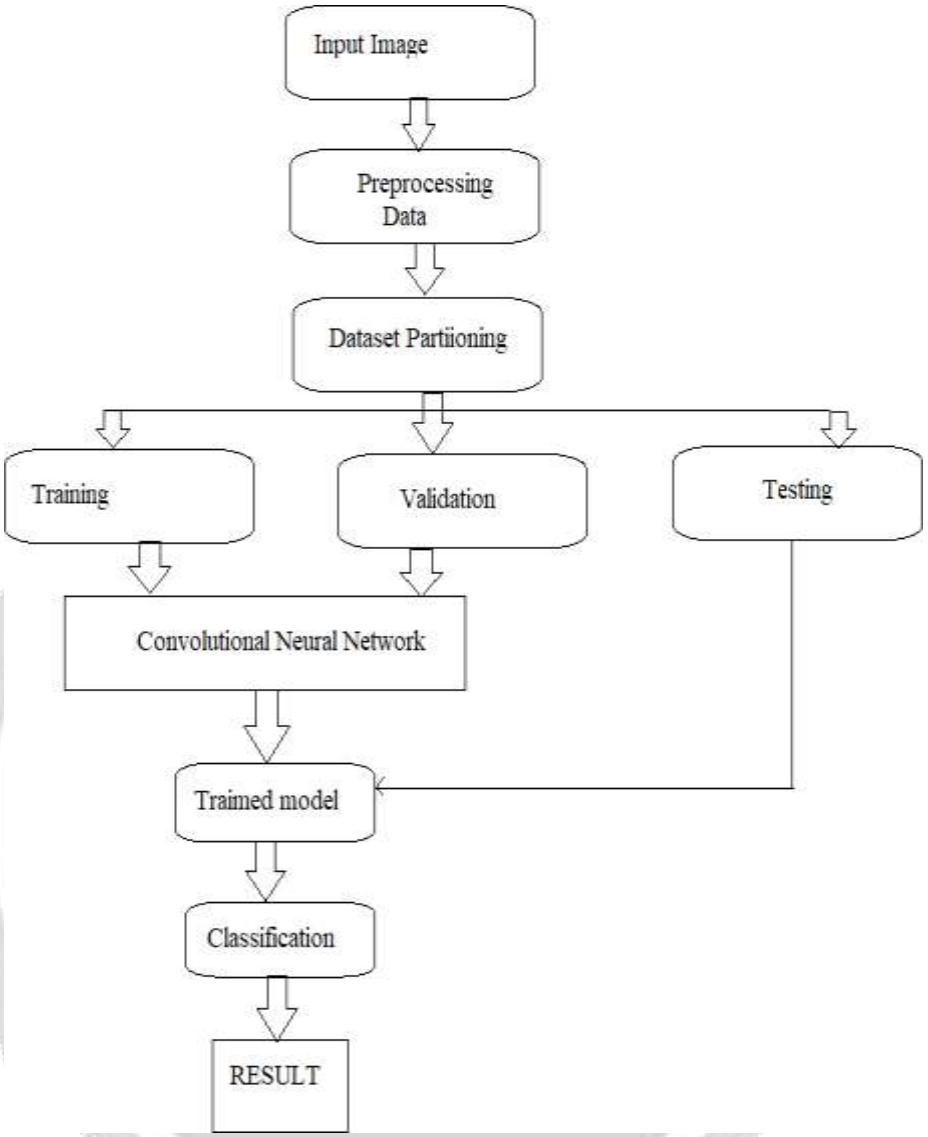


Fig. : Architecture diagram

Flowchart:



7. Requirement Specification:

User Requirements: Users expect a user-friendly interface that makes their online shopping experience enjoyable and effortless. They also demand personalized product recommendations that cater to their unique preferences and needs. The chatbot should be a proficient conversational partner, understanding and responding to their natural language queries in real-time.

Data Requirements: The system should seamlessly import product data from external CSV/JSON files, which should include essential attributes like product name, description, price, and category. Data integrity and consistency should be maintained in the central database.

Functional Requirements: The "Add Product" feature must facilitate the smooth addition of new products from CSV/JSON files to the system's database. An "Update Website" function should automatically refresh the product listing page, ensuring that customers have access to the latest products. The chatbot's proficiency lies in providing meaningful and prompt responses to user queries, including text and product recommendations. To enhance accuracy, the recommendation system leverages natural language processing (NLP) for a deeper understanding of both user queries and product descriptions.

Non-Functional Requirements: To meet user expectations, the recommendation system must maintain a minimum accuracy level, such as 90%, in providing relevant product suggestions. Additionally, systemresponse times should be nearly instantaneous to offer a seamless and efficient user experiece

Technical Requirements: Angular JS drives the front-end, creating an engaging user interface, while Python, Tensor Flow, and Keras provide the backend infrastructure, enhancing recommendation accuracy. The system relies on the NLTK library to effectively implement natural language processing (NLP) for understanding and responding to user queries.

Security and Privacy Requirements: The system must uphold strict data security and privacy standards. All user data, including personal information and queries, must be handled confidentially, with adherenceto data protection regulations. Secure data transmission protocols are implemented to ensure the safety of user information during interactions.

Regulatory and Compliance Requirements: Compliance with data protection laws is crucial to safeguard user privacy and ensure the responsible handling of personal data. The project also aligns with industry standards and ethical guidelines to maintain system integrity.

Budget and Resource Requirements: Adequate budget allocation is imperative for successful software development, infrastructure, and personnel resources. A skilled team of professionals in web development, NLP and database management is essential to bring this project to fruition.

Timeline and Milestone Requirements: The project must adhere to a well-defined timeline, marked by key milestones that encompass data import, recommendation model development, UI design, and chatbot integration, ensuring efficient progress and completion.

8. CONCLUSION:

In conclusion, this project addresses the need for an integrated e-commerce recommendation system that combines data management, website updates, and chatbot interactions. By seamlessly adding new products, updating the website, and enabling chatbot-driven user interactions, it enhances the online shopping experience. The project achieves its objectives of providing personalized product recommendations, leveraging natural language processing, and continuously improving accuracy. With a focus on user satisfaction and business growth, it offers a dynamic and efficient platform for tailored product suggestions, ultimately boosting customer engagement and revenue.

9. REFERENCE:

1. Esteva, A., Kuprel, B., Novoa, R. A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature,542(7639),115-118.

2. Haenssle,H.A.,Fink,C.,Schneiderbauer,R.,etal.(2018).Managainstmachine:Diagnosticperformance ofadeeplearningconvolutionalneuralnetworkfordermoscopicmelanomarecognitionincomparisonto 58 dermatologists. Annals of Oncology,29(8),1836-1842

3. Fujisawa, Y., Otomo, Y., Ogata, Y., & Nakamura, Y. (2019). Deep neural networks show an equivalentandoftensuperiorperformancetodermatologistsinonychomycosisdiagnosis: Automaticconstructionof onychomycosis datasets by region-based convolutional deep neural network. PLOS ONE, 14(12), e0225946.

4. Codella,N.C.,Gutman,D.,Celebi,M.E.,etal.(2018).Skinlesionanalysistowardmelanomadetection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). arXivpreprintarXiv:1710.05006.

5. Mandal, M., Ramakrishnan, A. G., & Gopalakrishnan, A. (2019). Skin cancer classification using deep learningandfuzzyintegral.ComputerizedMedicalImagingandGraphics,77,101636.