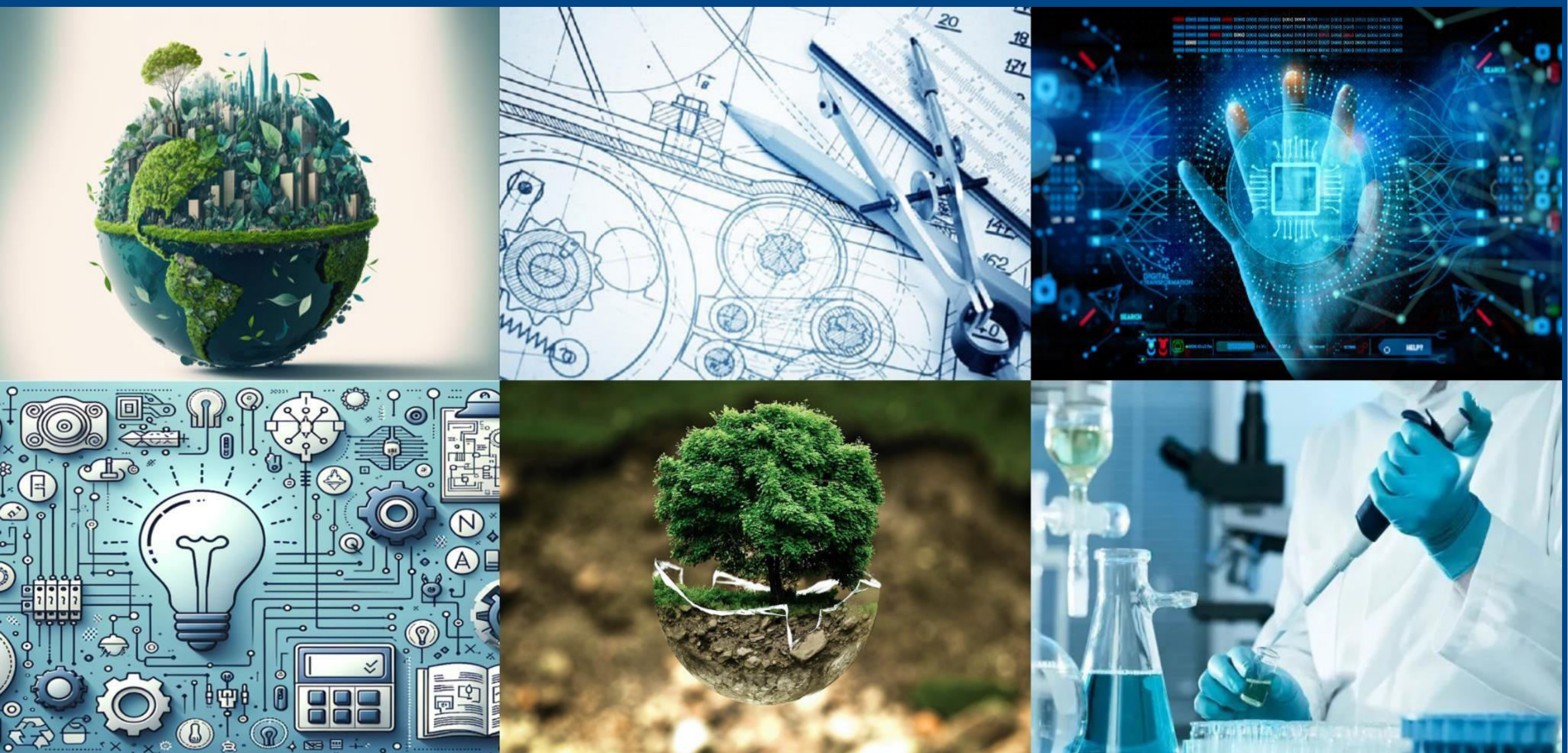
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 12, December 2025

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Skin Cancer Detection and Prediction using Machine Learning

Likith G R[1], Mallikarjun D A[1], Mallikarjuna K O[1], Vinayak C P[1], Prof. Damini T K[2]

UG Student, Dept. of ISE, PES Institute of Technology and Management, Shivamogga, Karnataka, India[1]

Assistant Professor, Dept. of ISE, PES Institute of Technology and Management, Shivamogga, Karnataka, India[2]

ABSTRACT: Skin cancer is rising worldwide, and many people still overlook early warning signs. Early detection makes treatment easier, but not everyone can see a dermatologist right away. To help with this issue, our project proposes a Machine Learning-based Skin Cancer Detection System that can determine if a skin lesion might be cancerous by analyzing an image.What sets this project apart is that the system doesn't just stop at detection—it also recommends medications for basic skin treatment and shows nearby pharmacies where these medicines are available. By combining image-based cancer detection with a feature that suggests local medicine options, this system aims to support users with both diagnosis help and practical next steps. The goal is to make early skin cancer awareness and access to medications easier for everyone, especially for those in rural or remote areas.

I. INTRODUCTION

Skin cancer is a disease that involves the uncontrolled growth of skin cells due to sun exposure, genetic, and environmental factors. It is one of the most frequent types of cancers, but fortunately, it is highly curable when treated on time. The big problem here is that many people either ignore the initial symptoms or cannot visit a specialist as quickly as possible.In today's world, machine learning has grown powerful enough to analyze images and identify patterns in them that may be uncatchable by human eyes. This technology finds widespread usage in healthcare for disease prediction, image classification, and medical assistance.The project emphasizes the use of ML to analyze an image of a skin lesion and predict whether it looks cancerous. The system then helps the user after detection by Suggesting commonly used medicinesShowing nearby pharmacies where those medicines are available It is not meant to replace doctors but rather to provide early awareness, support informed decisions, and ensure access to medical help. This solution can definitely help people in small towns or villages who may not have immediate access to dermatologists.

II. LITERATURE REVIEW

[1]. Esteva et al. in 2017 trained deep convolutional neural networks on around 129K clinical and dermoscopic images. They pretrained the networks on general images and then fine-tuned them for the task. The main goal was to classify benign versus malignant lesions. They reported performance that matched the level of dermatologists. That paper really helped popularize end-to-end CNNs for skin cancer screening. It focused on architectures like the Inception and ResNet family with transfer learning. They also made a strong case that mobile screening could work well in practice.

[2]. The HAM10000 dataset came from Tschandl et al. in 2018. They put together 10,015 multi-source dermatoscopic images from various places. It has turned into a key benchmark for training and testing ML models in this area. Plenty of studies now use HAM10000 for multiclass classification tasks. That includes things like nevus, melanoma, and BCC. Researchers also look at it to explore class imbalance issues and different augmentation strategies.

[3]. The ISIC Challenge series started with ISBI and ISIC from 2016 to 2018 and kept going after that. They ran large public competitions focused on lesion segmentation, attribute detection, and disease classification. Top teams relied on ensembles of CNNs for their approaches. They included sophisticated preprocessing steps like removing hair and markers. For segmentation, they often used U-Net variants. Test-time augmentation helped boost results too. Those challenges really showed up some generalization gaps in the models. They also helped set up standardized ways to evaluate everything.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[4]. Many pipelines combine segmentation and classification steps for better results. The classical U-Net from Ronneberger et al. in 2015 still holds up as state-of-the-art for lesion segmentation. Lots of U-Net variants have come along since then. Accurate masks from segmentation make downstream classifiers work better overall. They cut down on background artifacts and improve cropping. Recent papers have added things like attention mechanisms, multiscale fusion, and residual blocks. All that aims to push up the segmentation metrics even more.

[5]. Several external studies have done head-to-head comparisons between large CNNs and clinicians. They re-evaluated automated systems on curated test sets. Findings show that these systems match or even outperform many dermatologists in those setups. For example, Brinker et al. in 2019 pointed this out clearly. It demonstrates a lot of promise for the technology. Still, it highlights how much everything depends on data curation and how the test sets get built.

[6]. Deep CNNs tend to dominate when it comes to classification tasks in this field. That said, hybrid approaches have gotten some attention too. They mix CNN feature extractors with classical classifiers like SVM or XGBoost. Some even add temporal models for tracking follow-up studies over time. Ensembling across multiple architectures is a common way to lift AUC and accuracy scores. Transfer learning shows up a lot in practical setups as well. Recent surveys do a good job summarizing all these algorithmic trends.

[7]. A number of works have looked into practical systems for real-world use. They focus on lightweight architectures like MobileNet and EfficientNet variants. The idea is to enable on-device inference, especially for screening in low-resource settings. These efforts emphasize model compression techniques and calibration steps. They also stress good user-interface design for non-expert users who might need it.

[8]. Systematic reviews and surveys cover a lot of common challenges in this area. Things like label noise keep coming up as an issue. Class imbalance is another big one, especially with rare melanomas. Dataset bias shows through too, often from under-representation of ethnicity or skin types. Lack of interpretability remains a sticking point. Papers push for explainable AI tools like Grad-CAM and saliency maps. They call for rigorous external validation and proper auditing before any deployment happens.

[9]. From 2020 to 2025, technical improvements have kept pushing forward. Attention mechanisms have gotten more refined in models. Multiscale fusion helps handle different scales better. Loss functions tailored for class imbalance make training smoother. Synthetic data from GANs adds useful augmentation options. Advanced segmentation like MRP-UNet and attention U-Nets have stepped up. All these changes target better generalization on all sorts of heterogeneous real-world images.

[10]. The literature points out several open problems and future directions pretty consistently. One is collecting diverse datasets with solid annotations. That means covering more Fitzpatrick skin types and mixing in clinical photos with dermoscopy. Another involves federated or privacy-preserving learning for multi-center training setups. Multimodal models could combine clinical metadata like age and lesion history right with the images. Few-shot or zero-shot learning might help with rare lesion types that do not show up much. Stronger external prospective validation comes up a lot too. So do regulatory and ethical frameworks before clinical deployment. Recent reviews and challenge reports lay out these recommendations in detail.

III. METHODOLOGY

3.1 Data Description

The proposed skin cancer detection system is developed using publicly available and clinically validated dermatological datasets, primarily HAM10000 (Human Against Machine with 10,000 training images) and ISIC (International Skin Imaging Collaboration). The HAM10000 dataset consists of 10,015 dermoscopic images categorized into seven skin lesion classes, namely melanoma, basal cell carcinoma, melanocytic nevi, benign keratosis, actinic keratosis, vascular lesions, and dermatofibroma. Each image is annotated by medical experts, making the dataset suitable for supervised deep learning–based skin cancer classification.

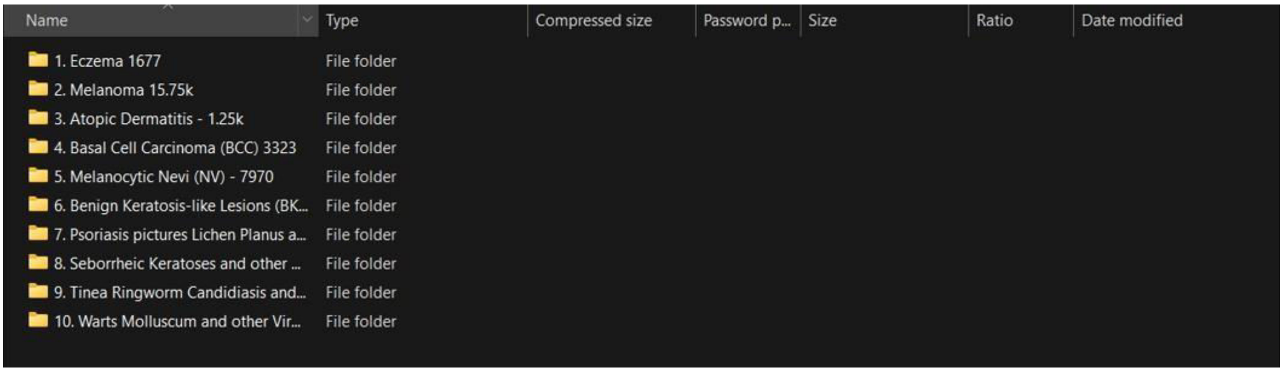


Fig 3.1 Data Description

3.2 Image Preprocessing

Dermoscopic images often suffer from noise, illumination variations, contrast differences, and class imbalance. To address these challenges and ensure consistent model input, several preprocessing steps are applied. All images are resized to a fixed resolution and normalised by scaling pixel values to the range [0, 1], which improves numerical stability and accelerates model convergence. Noise reduction techniques are applied to eliminate irrelevant artifacts present in the images.

To further improve model robustness and reduce overfitting caused by dataset imbalance, data augmentation techniques such as rotation, horizontal and vertical flipping, translation, zooming, and scaling are employed. Additionally, histogram equalisation is used to enhance lesion contrast, allowing better extraction of discriminative features. These preprocessing steps significantly improve feature learning and generalisation capability.

3.3 Model Building and Training

The proposed system employs Convolutional Neural Networks (CNNs) for automated feature extraction and classification of skin lesions. CNNs are well suited for medical image analysis due to their ability to learn hierarchical spatial features directly from image data.

The CNN architecture consists of:

- Convolutional layers for extracting low-level and high-level features such as edges, textures, and lesion boundaries
- Max-pooling layers for dimensionality reduction while preserving salient features
- Fully connected layers for mapping extracted features to output classes

The final classification layer uses a softmax activation function to produce probability scores for multi-class lesion classification.

3.4 Transfer Learning Strategy

To enhance classification accuracy and reduce training time, a transfer learning approach is adopted. Pre-trained deep learning models trained on the ImageNet dataset, which contains over 14 million labeled images across 1000 categories, are utilized. Among these, VGG16, MobileNet, and ResNet architectures are explored due to their proven performance in image classification tasks.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The initial convolutional layers of the pre-trained models are frozen to preserve generic visual feature representations, while additional task-specific layers are fine-tuned using the HAM10000 dataset. This strategy allows the model to leverage prior knowledge while adapting effectively to dermoscopic image characteristics.

3.5 Model Training and Evaluation

The dataset is divided into training (80%) and testing (20%) subsets. Model training is performed using the Adam optimizer with categorical cross-entropy loss. During training, performance metrics such as training accuracy, validation accuracy, and loss are monitored.

The trained model is evaluated using standard classification metrics including accuracy, precision, recall (sensitivity), and F1-score. These metrics provide a comprehensive evaluation of the system's ability to distinguish between benign and malignant skin lesions. Experimental results demonstrate that CNN-based models with transfer learning outperform traditional machine learning classifiers in terms of classification accuracy and robustness.

3.6 Prediction and User Interaction

Once trained, the model is used for real-time prediction. When a user uploads a skin lesion image, the system preprocesses the image and passes it through the trained model. The model predicts whether the lesion is benign or malignant and provides a confidence score indicating prediction reliability. This assists users in understanding whether further medical consultation may be required.

3.7 Auxiliary Modules and Deployment

To enhance usability, the system includes additional supportive modules. A medicine suggestion module provides general guidance on commonly used ointments or preliminary skin-care precautions for early-stage conditions. These suggestions are informational and do not replace professional medical advice.

A medical store locator module integrates location-based services using GPS or location APIs to identify nearby medical stores, display distances, and provide navigation assistance.

The complete system is deployed as a web-based application, enabling users to upload images, receive predictions, and access support features through a user-friendly interface. This deployment ensures scalability, accessibility, and real-time usability.

IV. SYSTEM ARCHITECTURE

The proposed skin cancer detection system follows a modular and layered architecture that integrates user interaction, machine learning inference, location-based services, and secure data management. The architecture ensures scalability, usability, and real-time performance.

4.1 User Interface Layer

The User Interface (UI) serves as the interaction point between the user and the system. It allows users to upload skin lesion images, grant location access, view prediction results, receive medicine suggestions, and visualize nearby medical stores on a map. The interface is designed to be simple, intuitive, and accessible to non-technical users.

4.2 Live Location Layer

The Live Location Layer captures the user's real-time geographical location using GPS services after user consent. This information is used to identify nearby medical stores and provide location-aware recommendations, making the system practical and context-sensitive.

4.3 Backend Server

The backend server acts as the central controller of the system. It receives user inputs, forwards images to the machine learning model for analysis, retrieves prediction results, processes location data, and communicates with the database. It ensures seamless coordination among all system components and delivers responses to the user interface.

© 2025 IJMRSET | Volume 8, Issue 12, December 2025| DOI:10.15680/IJMRSET.2025.0812020

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4.4 Machine Learning Model Layer

This layer represents the core intelligence of the system. It preprocesses the uploaded skin lesion images and applies a trained Convolutional Neural Network (CNN) or transfer learning model to classify lesions as benign or malignant. The model outputs prediction results along with confidence scores, and may also identify lesion types.

4.5 Medicine Recommendation Module

Based on the prediction outcome, this module provides basic guidance such as commonly used ointments, creams, and early-care precautions. The recommendations are derived from a predefined medical knowledge base and are intended to support awareness rather than replace professional medical consultation.

4.6 Medical Store Finder

The Medical Store Finder integrates with the live location layer and a store database to identify nearby pharmacies. It checks medicine availability, calculates distance, and displays store details such as name, location, and availability status, enabling users to take quick and informed action.

4.7 Database Layer

The database layer manages essential system data, including user interactions (optional), prediction logs, medicine information, medical store details, and inventory updates. It ensures reliable storage, efficient retrieval, and consistency of data across the system.

4.8 Map and Visualization Layer

This layer provides geographical visualization of nearby medical stores using mapping tools such as Google Maps API, OpenStreetMap, or Leaflet. It displays real-time distance, directions, and location markers to enhance user navigation and decision-making.

4.9 Security Layer

The Security Layer ensures data privacy and system integrity. Uploaded images are encrypted, location data is not stored permanently, and access to administrative features is restricted to authorized users. These measures protect sensitive medical and personal information.

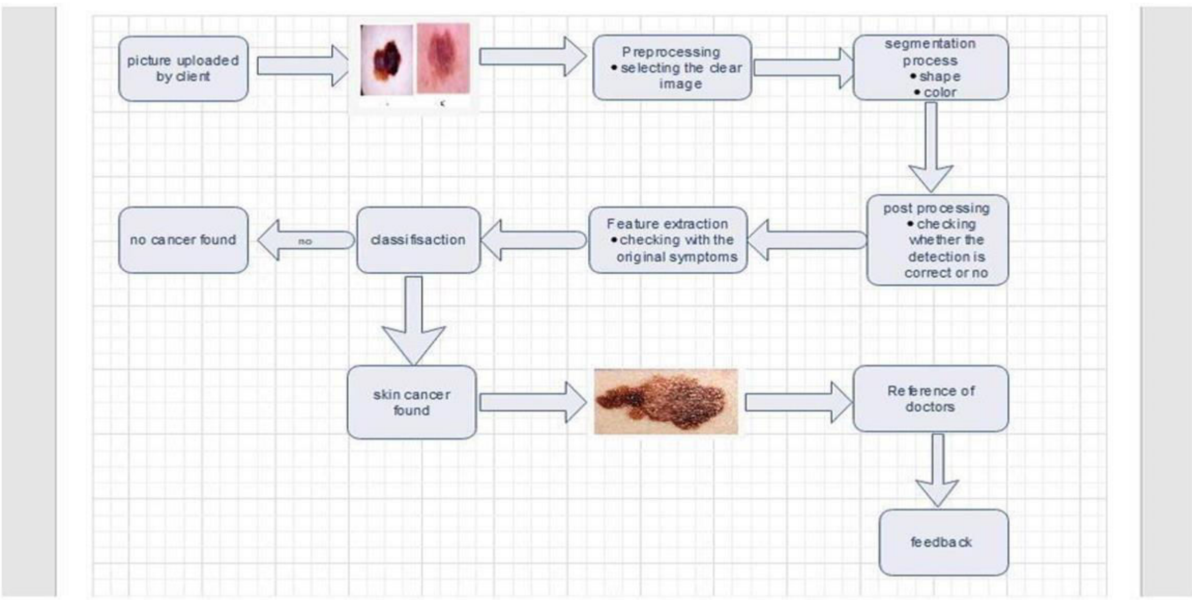


Fig 4.1 System Architecture Diagram

© 2025 IJMRSET | Volume 8, Issue 12, December 2025| DOI:10.15680/IJMRSET.2025.0812020

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

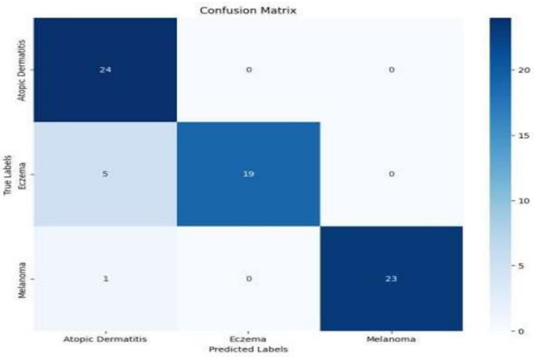
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. RESULTS

1. Skin Cancer Detection and Prediction Model

The skin cancer detection model developed using Machine Learning (ML) and the ResNet-50 deep learning algorithm achieved strong and reliable performance. By processing dermoscopic images, the model learned important visual features such as color, texture, and lesion patterns, allowing it to accurately classify skin lesions as benign or malignant. The prediction outputs were displayed through a probability-based bar graph, clearly showing the model's confidence levels for each class. This ML- based approach provides a fast, accurate, and easy-to-understand system for supporting early diagnosis and better clinical decision-making.



Metric	Value
Total Samples	24 + 19 + 23+ 5 + 1 = **72**
Correct	24 + 19 + 23 = **66**
Incorrect	5 + 1 = **6**
Accuracy	66 / 72 ≈ **91.6%**

Fig 5.1 Skin Cancer Prediction Results

2. Live Medical Assistance

The inclusion of machine learning in this skin cancer detection system helps identify the risk of cancer, but it goes a step further in guiding the patient on the next course of action. Immediately after the model analyzes the skin image and gives the prediction, the system displays nearby medical stores where needed medicines are available. It takes your live location and suggests the closest and most reliable options. By doing so, it becomes much easier for the patients to receive help at the earliest without much stress and assists them in getting the right treatment on time.

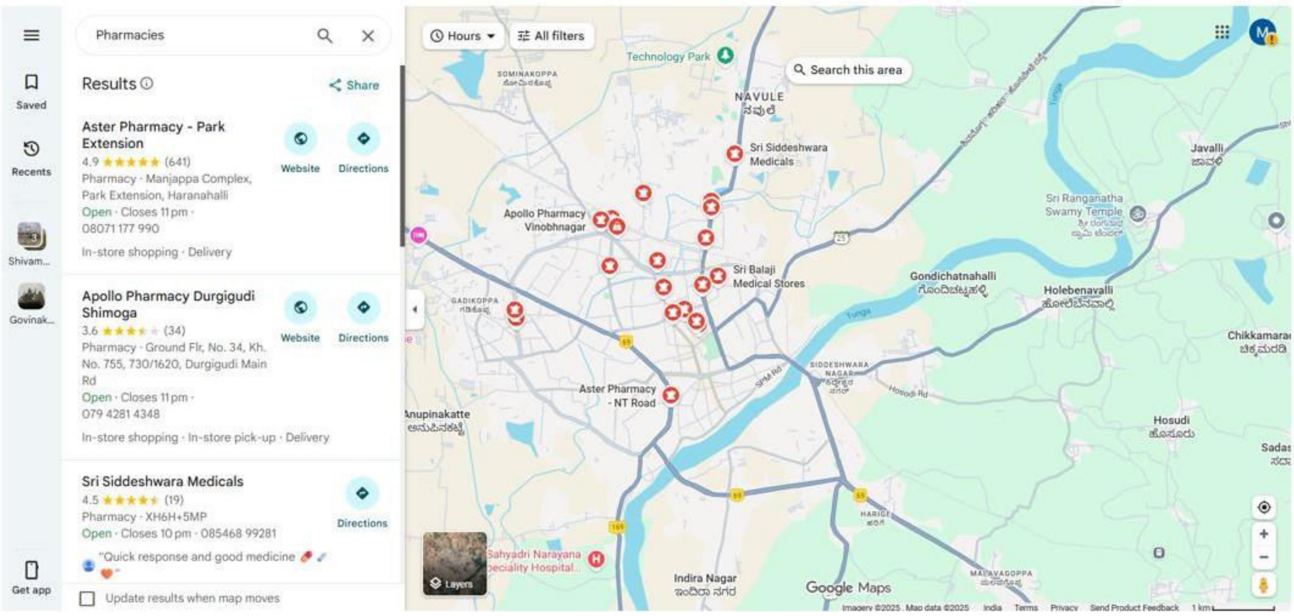


Fig 5.2 Live Medical Assistance

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. CONCLUSION

The Skin Cancer Detection System using Machine Learning aims to make early diagnosis simpler, faster, and more accessible. With image analysis based on AI, recommendations of medicines, and suggestions for nearby stores, our project shall provide comprehensive support to the user who is uncertain about the skin changes that are visible.This system is particularly helpful for people in rural or remote areas who lack immediate access to dermatologists. Although it does not substitute for professional medical advice, it serves as a kind of early warning tool that encourages users to seek proper treatment. All in all, this project demonstrates a fine collaboration of Machine Learning and healthcare in order to provide practical solutions that can save time, reduce anxiety, and potentially save lives.

REFERENCES

[1] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. *Dermatologist-level classification of skin cancer with deep neural networks*. Nature, 542, 115–118 (2017).
[2] Tschandl, P., Codella, N., Akay, B. N. et al. *The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions*. Scientific Data, 5, 180161 (2018).
[3] Brinker, T. J., Hekler, A., Enk, A. H., & von Kalle, C. *Skin cancer classification using convolutional neural networks: Systematic review*. Journal of the American Academy of Dermatology, 80(4), 1224-1232 (2019).
[4] Codella, N. C. F., Lin, C., Halpern, A., & Smith, J. R. *Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 ISIC Workshop*. IEEE International Symposium on Biomedical Imaging (ISBI), 2018.
[5] He, K., Zhang, X., Ren, S., & Sun, J. *Deep Residual Learning for Image Recognition (ResNet)*. IEEE Conference on Computer Vision and Pattern Recognition, 2016.
[6] Howard, A. G., Zhu, M., Chen, B., et al. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. Google Research, 2017.
[7] Simonyan, K., & Zisserman, A. *Very Deep Convolutional Networks for Large-Scale Image Recognition (VGG)*. arXiv preprint arXiv:1409.1556, 2014.

INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com