

Skin Cancer Detection Application

Tushar Gharge [1], Narayan Gupta [2], Krutik Raut [3], Prof. Varsha Salunkhe [4]

[1], [2], [3] Student, Department of Computer Engineering, Atharva College of Engineering, Mumbai

[4] Professor, Department of Computer Engineering, Atharva College of Engineering, Mumbai

Abstract - The Project Entitled As "Skin Cancer Detection Application". Skin cancer is regarded as one of the most dangerous types of cancer seen in humans. Every year, tens of thousands of people in the United States are diagnosed with cancer. As the disease spreads, the rate of survival drops dramatically. Melanoma, Basal and Squamous cell Carcinoma are the most common types of skin cancer, with Melanoma being the most unpredictable. Early detection and identification of the type of cancer can aid in its cure. Many technologies have demonstrated that AI and computer science can play a key role in picture scanning. We describe a mobile technique for detecting Melanoma Skin Cancer utilising Image Processing and basic portable equipment in this technical study. The cancer image is fed into the system, which analyses it using unique image processing algorithms to determine the existence of skin cancer. The cancer picture analysis software looks for asymmetry, border, colour, diameter, and other Melanoma properties, as well as size, frequency, and form, to identify the image and feature phases. The scanned feature parameters are used to distinguish between normal skin and melanoma cancer lesions in the image.

Key Words: Melanoma, TensorFlow, Dermatology, Image Processing, Transfer Learning,

1. INTRODUCTION

In the age of apparatuses, machine learning is finding its application in all fields and clinical diagnosis is the newest calculation to it. Emerging machine learning (ML)-based analytical tools are exciting to build, but they must be chosen carefully. To improve diagnostic accuracy, choose the correct decision-making system. Face identification, fingerprint recognition, tumour detection, and segmentation are just a few of the applications of image processing and machine learning. Fungal infection, bacteria, allergies, and viruses, among other things, can cause skin problems.

The project's purpose is to create and test robust, automated textural feature extraction and selection algorithms, as well as learning algorithms for identifying pigmented skin lesions, in order to improve skin cancer diagnosis in primary care settings. The goal of this initiative is to enhance the detection and treatment of skin cancer. One of the project's main objectives is to develop an interactive support tool with a simple user interface to assist primary care practitioners in accurately screening pigmented skin lesions in order to improve early detection, reduce unnecessary hospital referrals, and reduce the number of malignant melanomas missed.

The impact of the proposed project is twofold significantly improved understanding of correspondence between automated extracted features The ABCD rule of thumb and computerised automatic identification (such as size, shape, colour, and texture information) physicians used to pre-screening the early skin cancer; and advances integration of feature extraction and classification in bio-medical imaging processing. The overarching goal is to provide transparent support for skin cancer screening by primary care physicians based on advanced bio-medical image processing and decision-making techniques.

1.1 MOTIVATION OF PROJECT

The primary motivation that drives the project is to use the advanced image classification technology for the well-being of the people. Computer vision has made good progress in machine learning and deep learning that are scalable across domains. The model's improved accuracy and efficiency may help detect melanoma in its early stages and prevent unneeded biopsies.

1.2 BASIC CONCEPT

We would gather images from user's devices via camera or system storage and identify skin cancer diseases like melanoma to their closest prediction in this application. It would give a description and symptoms of the condition after successful identification. The user could also look up the contact information for nearby doctors.

It would also provide chart data as well as approximations of other nearby conditions via a user-friendly interface, allowing consumers to be directed to specific skin disease results.

2. BACKGROUND STUDY

Sheha, Mariam A. On a collection of dermoscopy photos, Mai S.Mabrouk and Amr Sharawy [1] proposed an automated method for melanoma diagnosis. So, while the first exercise, Automatic iteration counter, is faster, the second practise, Default iteration counter, is more precise, with a precision of 100% for the training set and 92 percent for the test set.

Jonathan Blackledge et al. [2] proposed the Moletest online skin growth screening framework.

The half-examination of a photograph in terms of its fractal structure and the fractal qualities that characterise it.

So, which of these two perspectives, or a combination of both, has been used to describe a handling and image test engine that is unique in its customary technique and entirely nonexclusive in terms of the applications to which it may be connected.

In light of form descriptors, Messadi M. et al. [3] developed an interpretable order technique for skin cancers in dermoscopic images. Their research demonstrates the use of a fuzzy rule-based classifier to distinguish between melanoma and other cancers. A flexible Neuro Fuzzy inference System (ANFIS) has a specific purpose in mind: to uncover the fuzzy rules that prompt the correct classification. The technique used in previous stages of development was to reduce the impact of small structures, hairs, bubbles, and light refraction. In the second step, an unsupervised lesion segmentation algorithm is proposed. In a transparent way, repeating thresholding is linked to the instate level specified. They've also addressed the requirement to focus all the ascribes employed to add to a depiction strategy that enables professionals to make the best decision.

The "ELM 7-point agenda," developed by G. Di Leo et al. [4], is a new analytic approach that describes an arrangement of seven components, including texture parameters, that convey the malignancy of a lesion. It has been shown to be faster and more precise than the traditional ABCD criteria in the detection of melanoma.

Capdehourat et al. [5] from Germany suggested a machine learning method for dealing with order melanocytic lesions in hazardous and generous dermoscopic images. There are 433 benign sores and 80 benign lesions in the image database. Following a step of image pre-processing that includes hair evacuation filtering, each image is then fractured using well-known image segmentation techniques. Yogendra Kumar Jain et al.

[6] focuses on the development of a skin cancer screening framework that may be used as part of routine practise by non-specialists to distinguish between normal and abnormal instances. Highlight Discovery and Order Procedure are two steps in the advancement process. The components are separated using wavelet change to decompose images into multiple recurring sub-groups. As a result, the characterization framework is based on the use of a Probabilistic Neural System and a Grouping Classifier.

Omar Abuzaghle et al. [7] suggested an ingenious and entirely practical advanced mobile phone-based programme for determining the type of early disease and its prevention. The first segment is a constant alarm to provide customers with help in preventing skin smoulder caused by sunshine, and the second segment is a new counted statement to know the most of the possibility for

skin to points is presented along these lines. The proposed framework is effective, according to the output data, grouping typical, atypical, and melanoma images with 96.3 percent, 95.7 percent, and 97.5 percent accuracy, respectively.

V. Jeya Ramya et al. [8] proposed an automated approach with typical and anomalous classes for skin cancer identification. The great division execution is accomplished by active contour segmentation. In an order approach with two classifications (threatening and favorable sores), an affectability of 90%, precision of 95% and a specificity of 85% is watched.

3. IMPLEMENTATION

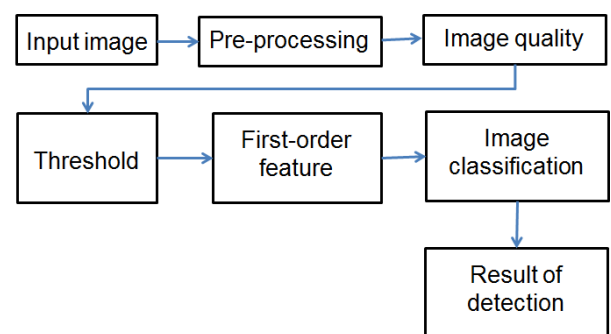


Fig-1: System Architecture

3.1 DATA COLLECTION

We start by gathering the best dataset containing photographs of skin conditions. We started by looking for melanoma-related photos on Kaggle and dataverse. While searching for skin disorders, there were numerous datasets with a large number of good photographs and models, so we choose the one with the most likes or recommends. We discovered Kaggle's SIIM-ISIC Melanoma Classification, which included over ten GB of data, and Harvard's data verse HAM10000 images part 1 & HAM10000 images part 2 dataset image when searching for the top datasets. We chose these two datasets because of the enormous number of datasets and recommendations. Normal and day-to-day living duties, educators, and psychiatrists might benefit from knowing the distinction between benign and malignant types. Choosing the best photographs from the two datasets, merging them, and categorizing them. Data was framed and stored in csv files from the obtained photos and datasets.

3.2 DATA CLEANING AND PREPROCESSING

Maintaining the data is very important so that bad data does not coincide with the training and modelling of our main and working model. We focus to pull out the important data that is required for the process. Before

beginning with the cleaning, we concatenated the two sets and added a special column to indicate the image is from which type of cancer it is. The first step in cleaning was to reduce the number of unnecessary columns from our data framed sets. Most of the sets had about numerous amounts of columns, of which only few we considered were important. This included the title of the disease, date of the content, resolution of the content, and the probability or confidence column that we added.

Each image must be unique and clear in way of visibility, while maintaining its original dimensions and data. For even better results, .dcm files: DICOM files were also added. Most of the images were saved in the "Digital Image and Communication in Medicine" format. It contains an image from a medical scan, such as an ultrasound or MRI + information about the patient. In order to prepare the sets for our classifier pre-processing of these images was required. This includes proper lighting, removing any other visual discrepancies. After the data was pre-processed and cleaned, it was added into our column in our csv files.

size of 224-by-224.

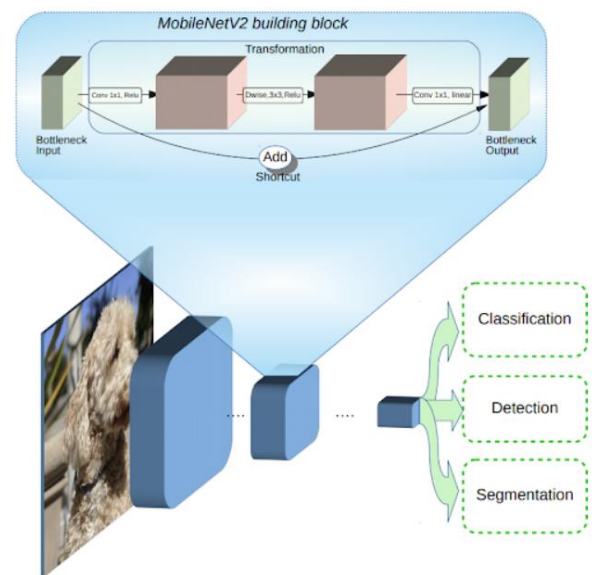


Fig-3: MobileNet-v2

3.3 DATA MODELLING

To decide on training model, we used TensorFlow lite transfer learning with MobileNet-v2 architecture model. training a TensorFlow Lite model with your own custom dataset decreases the number of trained data required and will shorten the overall time. So, the pre-trained neural network would help you to create your own classes, and so you can sort of picture that your class are becoming the last layer or step of our neural net.

MobileNet-v2 is a CNN that is 53 layers deep. People can load a pretrained version of the network trained on more than a million images from the ImageNet database. The pre-trained CNN can classify images into most of the real time objects categories, such as keyboard, mouse, pencil, and many test data. Such kind of result is the which has the network and learned rich feature representations for a wide range of images. The output has an image input

In the context of our project for training model, we would use the teachable machine's image classification to train the model for ease of having better accuracy and result. We would do the following steps:

Added the cleaned and preprocessed dataset for classification as group of classes based on the disease name of the respective data. After all the data was uploaded, it was ready for the next process of training.

For training the model with nearest accuracy possible, we tried to adjust the models' parameters of epochs (for number of times dataset fed for training), batch size (set of samples fed in each iteration), learning rate (rate to adjust the models understanding) accordingly.

So, after going multiple trial and run cases, we achieved our accuracy highest accuracy at epochs of 89, batch size of 128 and learning rate of 0.0001. After working our model on multiple permutations of the above parameters, we saved the result match and tried to deploy it for the testing stage.

The optimized model scored well on our test set, scoring an AUC score of 0.65 ~ 0.75. We would try proceed to know our model a bit better before making final estimation and recommendations.

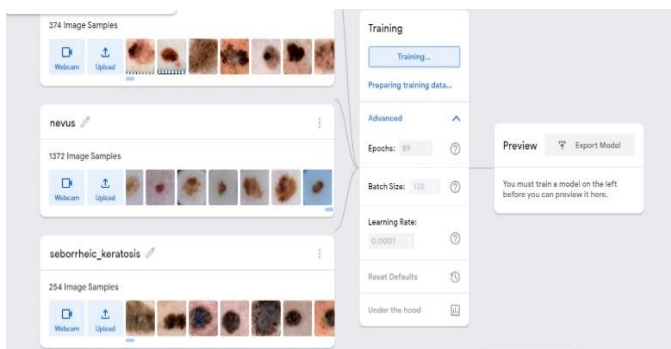


Fig-2: Training model

4. LOADING AND TESTING

The final phase was to load and test whether our trained model submitted on real time devices works accurately or not. After training, the trained model was converted into tflite quantized model as the extension of tflite. After the model was converted, it was loaded and integrated into Flutter code and was made to run on the live images from the devices. The model was made to run on threshold of 0.5, image mean and standard of 127.5.

Testing was done on many real time skin cancer diseases of melanoma from many numbers of individuals and patients. We achieved the accuracy of 0.60 ~ 0.80 approx. The images taken with better illumination and high resolutions showed the higher accuracy of around 0.70 and above, while the images having lower pixels or resolutions and less illumination had little lower result range in between of 0.40 and 0.60. Overall, the model worked accurately from the portable device like mobile trained from the MobileNet-v2 architecture. Response time was very quick and was able to identify the diseases even in the low-end devices.

```
//Load the Tflite model
LoadModel() async {
  await Tflite.LoadModel(
    model: "assets/trainedModel.tflite",
    labels: "assets/labels.txt",
  );
}

classifyImage(image) async {
  var output = await Tflite.runModelOnImage(
    path: image.path,
    numResults: 2,
    threshold: 0.5,
    imageMean: 127.5,
    imageStd: 127.5,
  );
  setState(() {
    _loading = false;
    //Declare List _outputs in the class which will be used to show the
    //classified class name and confidence
    _outputs = output;
  });
}
```

Fig-4: Loading model

Our gives average of 70.6 percent accuracy when looking at regular entries (i.e. is not including the variance of other diseases outside of melanoma). Final the level of accuracy means that the model might have the potential to be generalize for usage in day to day like schools, rural areas and even for just verifying the disease for anyone who download the application.

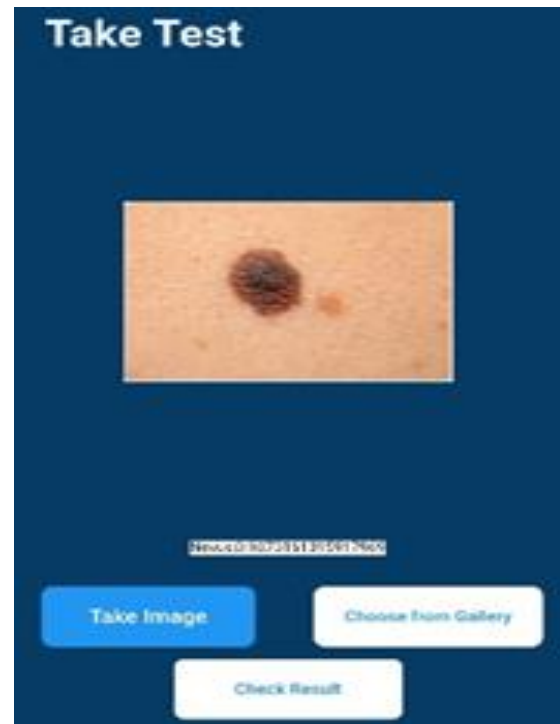


Fig-5: Testing on real time images

5. RESULTS

At last, the modelling and testing phase we can conclude that our classifier has an above average rating in terms of being correct and providing accurate results. This means in the future such a system can be implemented to check if a person or patient has been experiencing any kind of melanoma cancers that might lead to possible even bigger problems. Such a system can also be implemented on other web applications, IOS & android devices as well as embedded devices. In the end our model was able to achieve close to 70 percent rate of success, with more research and studies this rate can be increased even further and provide more accurate and dependable results. The main issue of the work was that a lot of pre-processing and collecting the data, which was required even on the sets on which the models were tested on to provide near results. While training more and more of our model from, datasets we would be able to know how our output would be able to change in different scenarios.

6. DESIGN DETAIL

6.1 E-R DIAGRAM

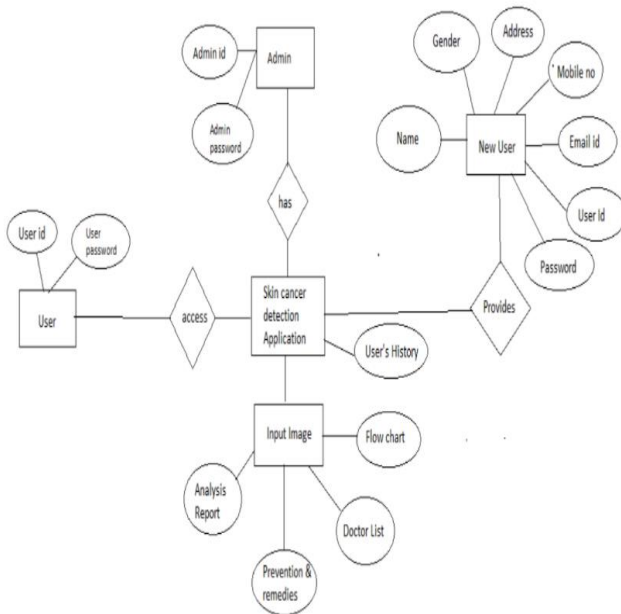


Fig-6.1: E-R diagram

6.2 USE CASE DIAGRAM

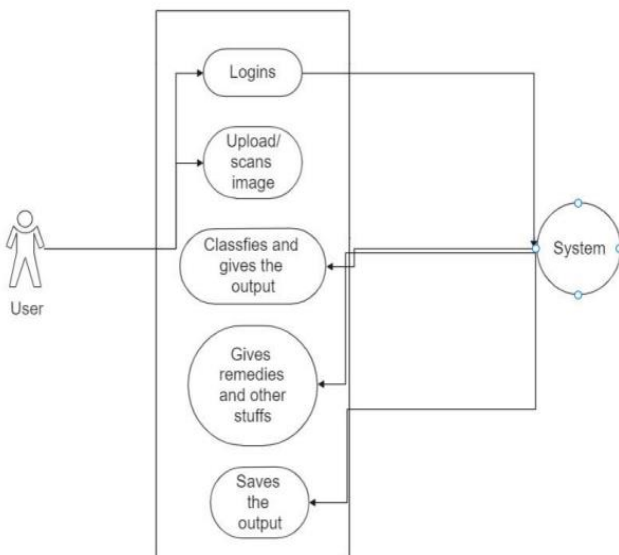


Fig-6.2: Use case diagram

7. FUTURE SCOPE

This idea and approach that we have implemented does not end here. There can be a future scope for development we can extend this project into multiple platforms.

Particularly, we can incorporate additional algorithm together to make is more appealing.

One other scope of improvement in expanding the idea hiding the data into a video format so that the application can be more secure so that different sector can use this to secure the way of communication between the user and themself. As specially, since the sensitive data can be stored without anyone knowing about it. Such that storing sensitive data won't be a problem it as we add higher level of security.

8. CONCLUSION

Skin cancer detection is very hard to for most of the people to detect without visiting doctors. Main project of our diagnosis requires experience of data, in which early stages may look identical to benign. Detecting the disease with such tool for less experience physicians. Our model's project is an automated melanoma classification system that was applied to dermoscopy pictures to aid in the early diagnosis of malignant melanoma and melanocytic nevi lesions. With main advantage that it contrasts with other methods in medical image analysis segmentation process is avoided using mobile analysis. In this project we had created a mobile application to classify skin cancer. This method consists of data classification, image processing and image classification using Mobile net model. Using this technique, we can classify images at nearest probability.

9. ACKNOWLEDGEMENT

We owe sincere thanks to our college Atharva College of Engineering for giving us a platform to prepare a project on the topic "Skin cancer Detection application" and would like to thank our principal Dr. Shrikant Kallurkar for instigating within us the need for this research and giving us the opportunities and time to conduct and present research on the topic. We are sincerely grateful for having Prof. Varsha Salunkhe as our guide and Dr. Suvarna Pansambal, Head of Computer Engineering Department, and our project coordinators during our research. So, at last, work of this research wouldn't have been impossible without the cooperation of all the members.

10. REFERENCES

- [1] Automatic Detection of Melanoma Skin Cancer Using Texture Analysis, International Journal of Computer Applications (0975 – 8887) Volume 42- No.20, March 2012. Mariam A.Sheha, Mai S.Mabrouk, Automatic Detection of Melanoma Skin Cancer Using Texture Analysis, International Journal of Computer Applications (0975 – 8887) Volume 42- No.20, March 2012.

- [2] Moletest: A Web-based Skin Cancer Screening System, Dublin Institute of Technology, 2011.] Jonathan Blackledge, Moletest: A Web-based Skin Cancer Screening System, Dublin Institute of Technology, 2011.
- [3] A hybrid machine-learning approach for segmentation of protein localization data, Oxford University Press, Vol. 21 no. 19 2005, pages 3778–3786. Peter M. Kasson¹, Johannes B. Huppa, Mark M. Davis, and Axel T. Brunger, A hybrid machine-learning approach for segmentation of protein localization data, Oxford University Press, Vol. 21 no. 19 2005, pages 3778–3786.
- [4] Automatic Diagnosis of Melanoma: a Software System Based on the 7-Point Check-List, Proceedings of the 43rd Hawaii International Conference on System Sciences, IEEE, 2010. G. Di Leo, A. Paolillo, P. Sommella, and G. Fabbrocini, Automatic Diagnosis of Melanoma: a Software System Based on the 7-Point Check-List, Proceedings of the 43rd Hawaii International Conference on System Sciences, IEEE, 2010.
- [5] German Capdehourat, Andres Corez, Anabella Bazzano, and Pablo Mus, Pigmented Skin Lesions Classification Using Dermatoscopic Images, Springer-Verlag Berlin Heidelberg, 2009, pp. 537–544
- [6] Yogendra Kumar Jain, Comparison between Different Classification Methods with Application to Skin Cancer, International Journal of Computer Applications (0975 – 8887) Volume 53– No.11, September 2012
- [7] Dr. J. Abdul Jaleel, Artificial Neural Network Based Detection of Skin Cancer, International Journal of Advanced Research in Electrical, Elcetronics and Instrumentation Engineering, Col. 1, Issue 3, 2012
- [8] DETECTION OF MELANOMA SKIN CANCER USING DIGITAL CAMERA IMAGES, ARPN Journal of Engineering and Applied Sciences, VOL. 10, NO. 7, APRIL 2015. V. Jeya Ramya, J. Navarajan, R. Prathipa, and L. Ashok Kumar, DETECTION OF MELANOMA SKIN CANCER USING DIGITAL CAMERA IMAGES, ARPN Journal of Engineering and Applied Sciences,