AI-BASED LOCALIZATION AND CLASSIFICATION OF SKIN DISEASE WITH ERYTHEMA

NALAIYA THIRAN PROJECT BASED LEARNING
on
PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP

SATHISH . K
HARISH RAJ . A 412519106042
ARUN KUMAR . C 412519106302
KARTHIKEYAN P 412519106060

BACHELOR OF ENGINEERING
IN
ECE

SRI SAIRAM

ENGINEERING COLLEGE

Approved by AICTE, New Delhi, Accredited with 'A+' Grade by NAAC
(An Autonomous Institution, Affiliated to Anna University, Chennai)
CHENNAI - 600044

NOVEMBER 2022

1.INTRODUCTION

1.1 Project Overview

Utilization of artificial intelligence coupled with the underlying technology of deep learning for diagnosis is made reasonable by the fact that each of these conditions has a unique visual feature.Cases of skin diseases that are prevalent in the Philippines and potentially can be identified by image processing technologies are chicken pox, acne, eczema, pityriasis rosea, psoriasis, tinea corporis, and vitiligo. Given an image of the skin, we decompose the image to normalize and extract high-level features. Using a neural network-based segmentation model to create a segmented map of the image, we then cluster sections of abnormal skin and pass this information to a classification model. We classify each cluster into different common skin diseases using another neural network model. Our segmentation model achieves better performance compared to previous studies, and also achieves a near-perfect sensitivity score in unfavorable conditions. Our classification model is more accurate than a baseline model trained without segmentation, while also being able to classify multiple diseases within a single image.

Technologies such as Python,Cloud,Deep Learning,CAD,YOLO and Image Processing are used in this model. Although computer-aided diagnosis (CAD) is used to improve the quality of diagnosis in various medical fields such as mammography and colonography, it is not used in dermatology, where noninvasive screening tests are performed only with the naked eye, and avoidable inaccuracies may exist. This study shows that CAD may also be a viable option in dermatology by presenting a novel method to sequentially combine accurate segmentation and classification models.

The image processing techniques are involved in the following ways, namely, the given input data sets go through the preprocessing techniques, these techniques are handled by using the median filter in our proposed approach, the preprocessing techniques are helping to remove the noise in the images, the median filter removes the salt and the pepper noise in the given input images.

Deep learning models are efficient in learning the features that assist in understanding complex patterns precisely. This study proposed a computerized process of classifying skin disease through deep learning .Machine Learning is that branch of computer studies that gives the potentiality to the computer to grasp without being characteristically programmed.

Artificial intelligence (AI) algorithms for automated classification of skin diseases are available to the consumer market. Studies of their diagnostic accuracy are rare. We assessed the diagnostic accuracy of an open-access AI application for recognition of skin diseases. The AI algorithm classified the images giving 5 differential diagnoses, which were then compared to the diagnoses made clinically by the dermatologists and/or histological.

1.2 Purpose

Skin diseases are characterized as disorders that often begin inside the body or start from the skin, and outwardly show on the skin. Some of them are extremely uncommon, however, others are commonly occurring. They bring the person itch, and pain, as well as emotional and social impacts because of its visibility. We implemented an Artificial Neural Network ANN-based single level system as well as a multi-model, multi-level system for eczema detection. Moreover, ANN was applied in to detect certain circulatory diseases through the color of the fingernails. A similar implementation was applied in disease detection using tongue images . A method of diagnosing Alzheimer's disease is by considering the retina images of the patients. Meanwhile, a system was proposed using artificial neural network as well as digital image processing in detecting BCC disease . The detection is based on special characteristics of basal cell carcinoma. The system will be capable to correctly identify the occurrence of carcinoma using the proper threshold values with percent reliability of 93.33%.

Various systems have developed on methods of image processing and feature extraction that help predict and

detect disease type. There are other systems designed to identify specific types of skin disease through clinical features and features obtained from tissue analyzes after a skin biopsy of the affected area. The aim of this project is to develop machine learning based classifiers for predicting skin infections for three classes from a clinical dataset. Convolutional neural network (CNN) has been proved to perform well in image classification. The performance of the neural network is compared with a benchmark multiclass SVM classifier.

2.Literature survey

2.1 Existing Problem:

Now a day's people are suffering from skin diseases, More than 125 million people suffering from Psoriasis also skin cancer rate is rapidly increasing over the last few decades especially Melanoma is most diversifying skin cancer. If skin diseases are not treated at an earlier stage, then it may lead to complications in the body including spreading of the infection from one individual to the other.

Diagnosis of skin disease from an image is a challenging problem as there exist many skin diseases. Researchers reported following problems during skin disease classification: 1) A disease may have many lesion types. 2) Many diseases may have a similar visual characteristic, which is often confusing for the dermatologist as well to identify the disease by visual inspection. 3) The varying skin colors and skin type (age) introduce more difficulty in computer-based diagnosis. Therefore, relevant feature selection for such diseases is very important in computer-based diagnosis in order to identify it correctly. The success of an automatic system rely on how accurately the system performs and does needed image processing as well as machine learning tasks.

2.2 References

1.Skin Disease Classification from Image - A Survey
March 2020
DOI:10.1109/ICACCS48705.2020.9074232
Conference: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)

2. Multi-type skin diseases classification using OP-DNN based feature extraction approach

Arushi Jain, Annavarapu Chandra Sekhara Rao, Praphula Kumar Jain & Ajith Abraham,*Multimedia Tools and Applications*

3. The Classification of Six Common Skin Diseases Based on Xiangya-Derm: Development of a Chinese Database for Artificial Intelligence, J Med Internet Res. 2021 Sep; 23(9): e26025 ,Published online 2021 Sep 21. doi: 10.2196/26025

4. Skin Disease Recognition Method Based on Image Color and Texture Features Li-sheng Wei,Quan Gan, and Tao Ji.

5. Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM
by Parvathaneni Naga Srinivasu,Jalluri Gnana SivaSai,Muhammad Fazal Ijaz,Akash Kumar Bhoi,Wonjoon Kim ,James Jin Kang

2.3 Problem Statement Definition

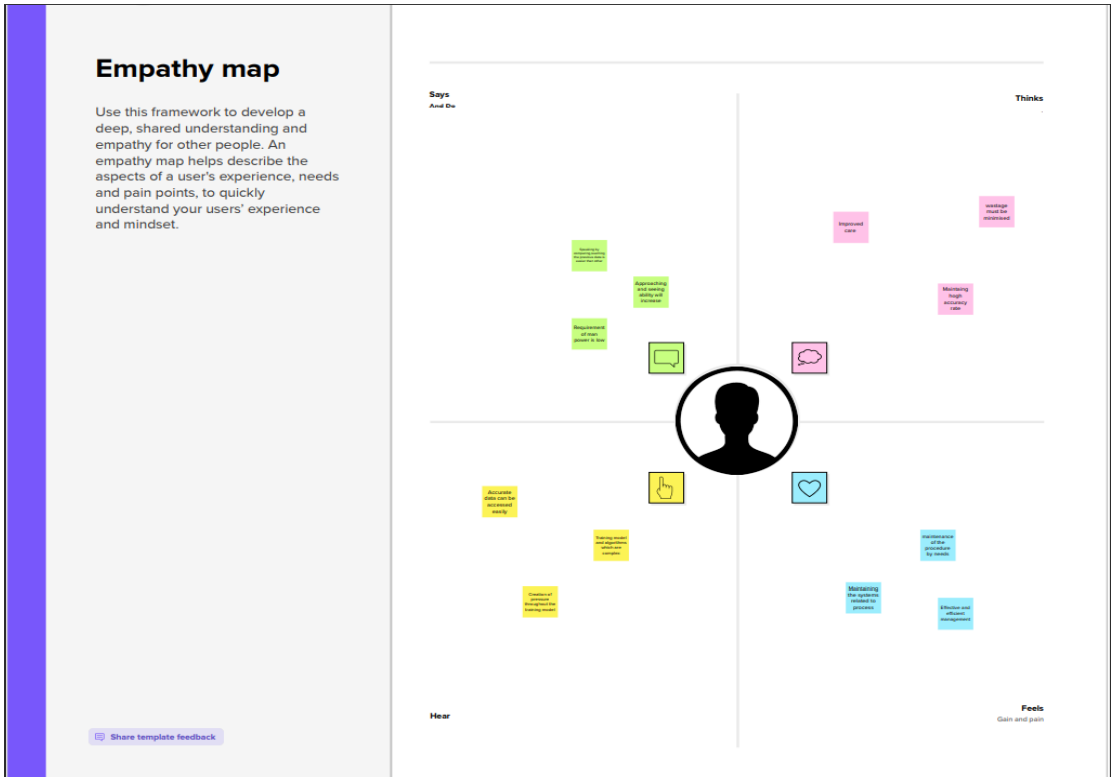
Now a day's people are suffering from skin diseases, More than 125 million people suffering from Psoriasis also skin cancer rate is rapidly increasing over the last few decades especially Melanoma is most diversifying skin cancer. If skin diseases are not treated at an earlier stage, then it may lead to

complications in the body including spreading of the infection from one individual to the other. The skin diseases can be prevented by investigating the infected region at an early stage. The
characteristic of the skin images is diversified so that it is a challenging job to devise an efficient and robust algorithm for automatic detection of skin disease and its severity. Skin tone and skin colour play an important role in skin disease detection. Colour and coarseness of skin are visually different. Automatic processing of such images for skin analysis requires quantitative discriminator to differentiate the diseases.

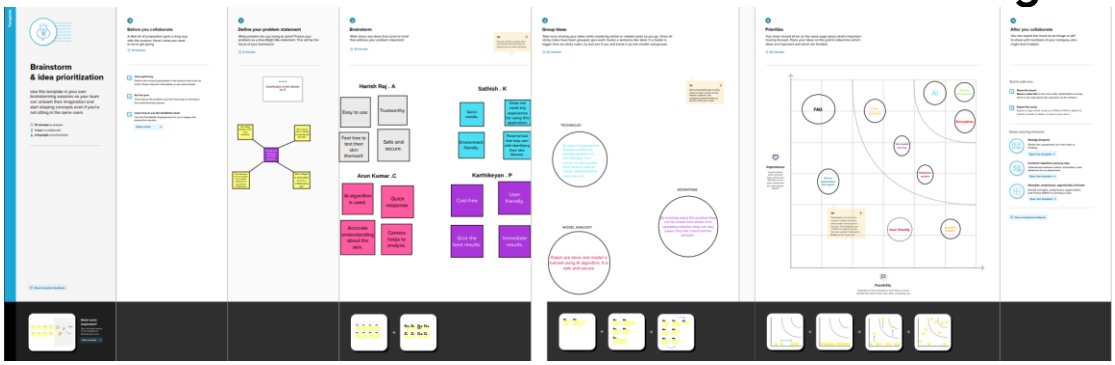
To overcome the above problem we are building a model which is used for the prevention and early detection of skin cancer, psoriasis. Basically, skin disease diagnosis depends on the different characteristics like colour, shape, texture etc. Here the person can capture the images of skin and then the image will be sent the trained model. The model analyses the image and detect whether the person is
having skin disease or not.

3.Ideation And Proposed Solution

3.1 Empathy Map Canvas



3.2 Ideation and Brainstorming

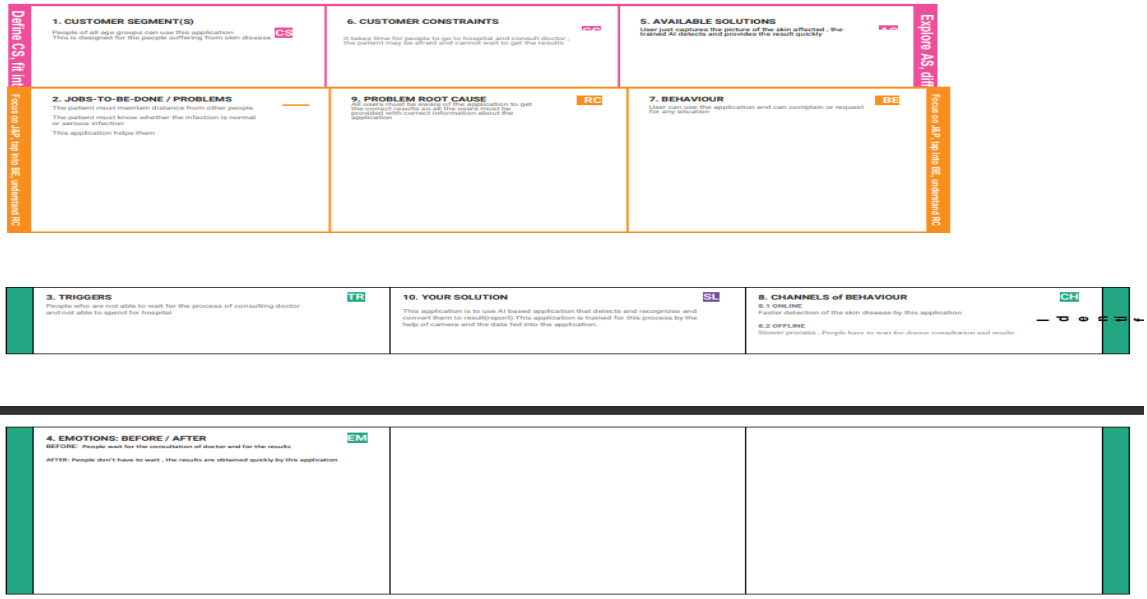


3.2 Proposed Solution

S.NO	PARAMTERS	DESCRIPTION
1	Problem Statement (Problem to be solved)	Now a day's people are suffering from skin diseases, More than 125 million people suffering from Psoriasis also skin cancer rate is rapidly increasing over the last few decades especially Melanoma is most diversifying skin cancer. If skin diseases are not treated at an earlier stage, then it may lead to complications in the body including spreading of the infection from one individual to the other. The skin diseases can be prevented by investigating the infected region at an early stage. The characteristic of the skin images is diversified so that it is a challenging job to devise an efficient and robust algorithm for automatic detection of skin disease and its severity. Skin tone and skin colour play an important role in skin disease detection. Colour and coarseness of skin are visually different. Automatic processing of such images for skin analysis requires quantitative discriminator to differentiate the diseases.
2	Idea / Solution description	To overcome the above problem we are building a model which is used for the prevention and early detection of skin cancer, psoriasis.Basically, skin disease diagnosis depends on the different characteristics like colour, shape, texture etc. Here the person can

		capture the images of skin and then the image will be sent the trained model. The model analyses the image and detect whether the person is having skin disease or not.
3	Novelty / Uniqueness	The novelty proposed in this approach is we have collected the dataset on our own. We have also annotated the images by ourselves.
4	Social Impact / Customer Satisfaction	The model which will be built by us is very useful for the users to find the type of disease quickly and get the correct the medicine as soon as possible. We ensure to the users that our model diagnoses the diseases well.
5	Business Model (Revenue Model)	Our model can be deployed in the web, and can be made like subscription based application, so that we can get the revenue from that subscription. It can be a monthly subscription.
6	Scalability of the Solution	This model is now built only for 10 specific diseases , but it can be scaled for classifying for multiple diseases. We can collect more images in the future and classify the diseases for more accurate results.

3.3 Problem Solution fit



4. Requirement Analysis

4.1 Functional requirements

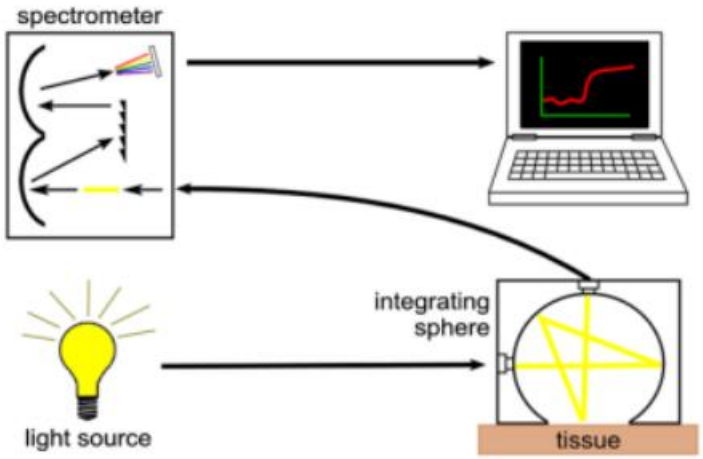
Skin Disease with Erythema:

Skin erythema, or flare, is the reddening reaction of the skin as a result of an external stimulus , immunological reaction with/out **hypersensitivity to an allergen or viral infection**

- The flare size depends on multiple parameters, for instance, distribution of the neural fibers and vascularization of the stimulated region.
- Likewise, the strength and the nature of the stimulus are factors that influence the intensity is reached shortly after the

stimulus onset. In some cases, the flare is a result of an accumulative process, such as radiotherapy treatment for cancer.

- Radiation dermatitis is an equivalent term to radiotherapy-induced erythema. In this case, erythema is a cancer radiotherapy treatment linked side effect.
- The dermatitis reaction is interpreted as a skin response to damage to basal cells present in the epidermal layer. To ameliorate the damaged region, deeper skin layers proliferate toreplace the impaired, superficial . The radiation dermatitis MED trigger is inconstant.



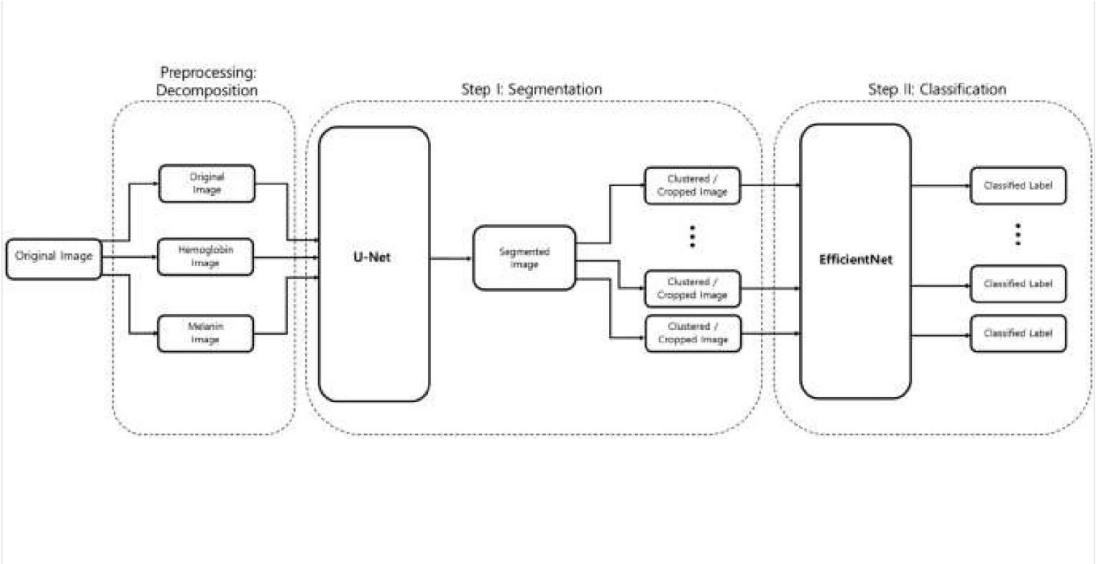
Techniques :

A major goal for any skin erythema assessment technique is to objectively quantify the redness without the need for a skin biopsy or direct contact.
A potential approach is a contactless technique that generates a real time graded redness intensity map. Moreover, it is anticipated that the erythema assessment standard device is miniaturized, easy to operate, and costeffective.

4.2 Non-Functional requirements

- Software Quality Attributes
- Prediction
- Accuracy.

5. Project Design
5.1Data Flow Diagram



5.2 Solution and Technical Architecture

Solution Architecture:

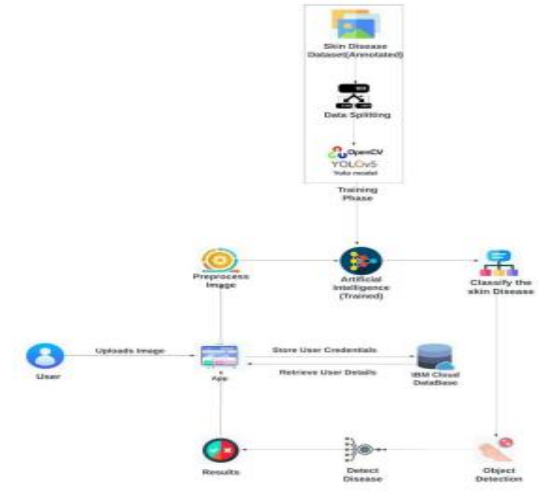
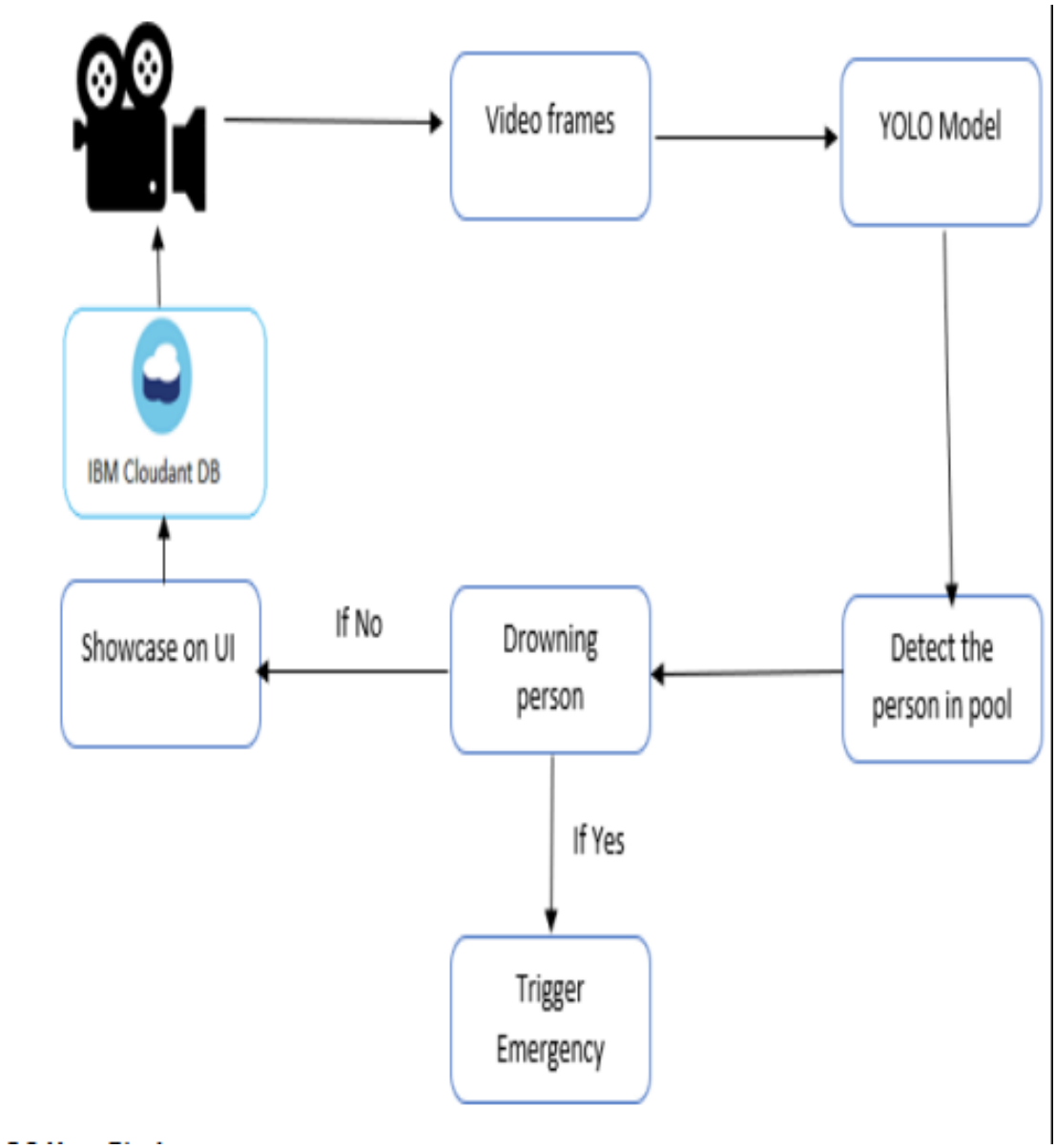


Figure 1: Architecture and data flow of the skin disease prediction application

Technology Architecture:



5.3 User Stories

Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority
Prerequisites	USN-1	Install Python IDE, Python packages, Microsoft Visual Object Tagging Tool, YoloStructure	3	High
Data Collection	USN-2	Dataset should be collected from google or using a Chrome extension such as Fatkun Batch Downloader	3	High
Annotate Images	USN-3	Create A Project in VOTT (Microsoft's Visual Object Tagging Tool)	2	Medium
Training YOLO	USN-4	train our model using YOLO weights	2	Medium
	USN-5	To Download and Convert Pre-Trained Weights	3	High
	USN-6	To Train YOLOv3 Detector	3	High
Cloudant DB	USN-7	Register & Login to IBM Cloud	3	High
	USN-8	Create Service Instant and Credentials	2	Medium
	USN-9	Launch DB and Create database	3	High
Development Phase	USN-10	To build a web application	3	High
	USN-11	Building HTML pages with python code	2	Medium
	USN-12	To run the application	3	High
Testing Phase	USN-13	As a user login to dashboard	2	Medium
	USN-14	As a user import the images with skin diseases to the software application	2	Medium
	USN-15	YOLO processes the image and give the necessary details	3	High

6.Project Planning and Scheduling

6.1Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022

7.Coding And Solutions

7.1 Feature 1

1. The system decompose the image to normalize and extract high-level features.

2.The system Uses a neural network-based segmentation model to create a segmented map of the image, we then cluster sections of abnormal skin and pass this information to a classification model.

3.The model classify each cluster into different common skin diseases using another neural network model.

4.This segmentation model achieves better performance.

CODE:

pip3 install tensorflow tensorflow_hub matplotlib seaborn numpy pandas sklearn imblearn

```
from fastapi.templating import Jinja2Templates
from fastapi.staticfiles import StaticFiles
import uvicorn
from src import const, preprocess
import os
import shutil
from pathlib import Path
import json

templates = Jinja2Templates(directory="./templates")
```

```
app = FastAPI()
```

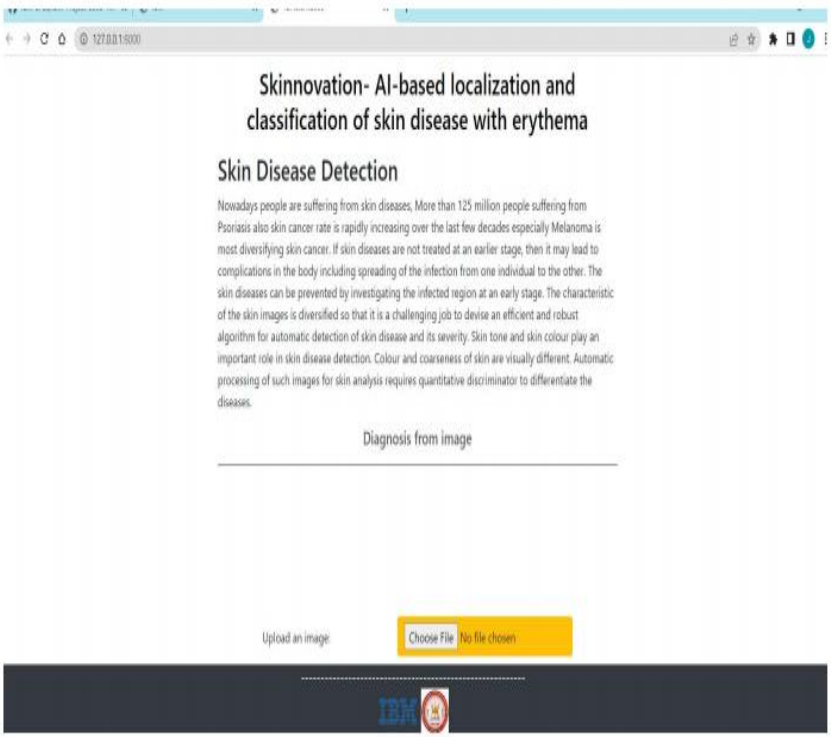
```
app.mount(
    "/static",
    StaticFiles(directory=Path(__file__+'Skin-Pigment-Analysis/').parent.absolute() / "static"),
    name="static",
)
```

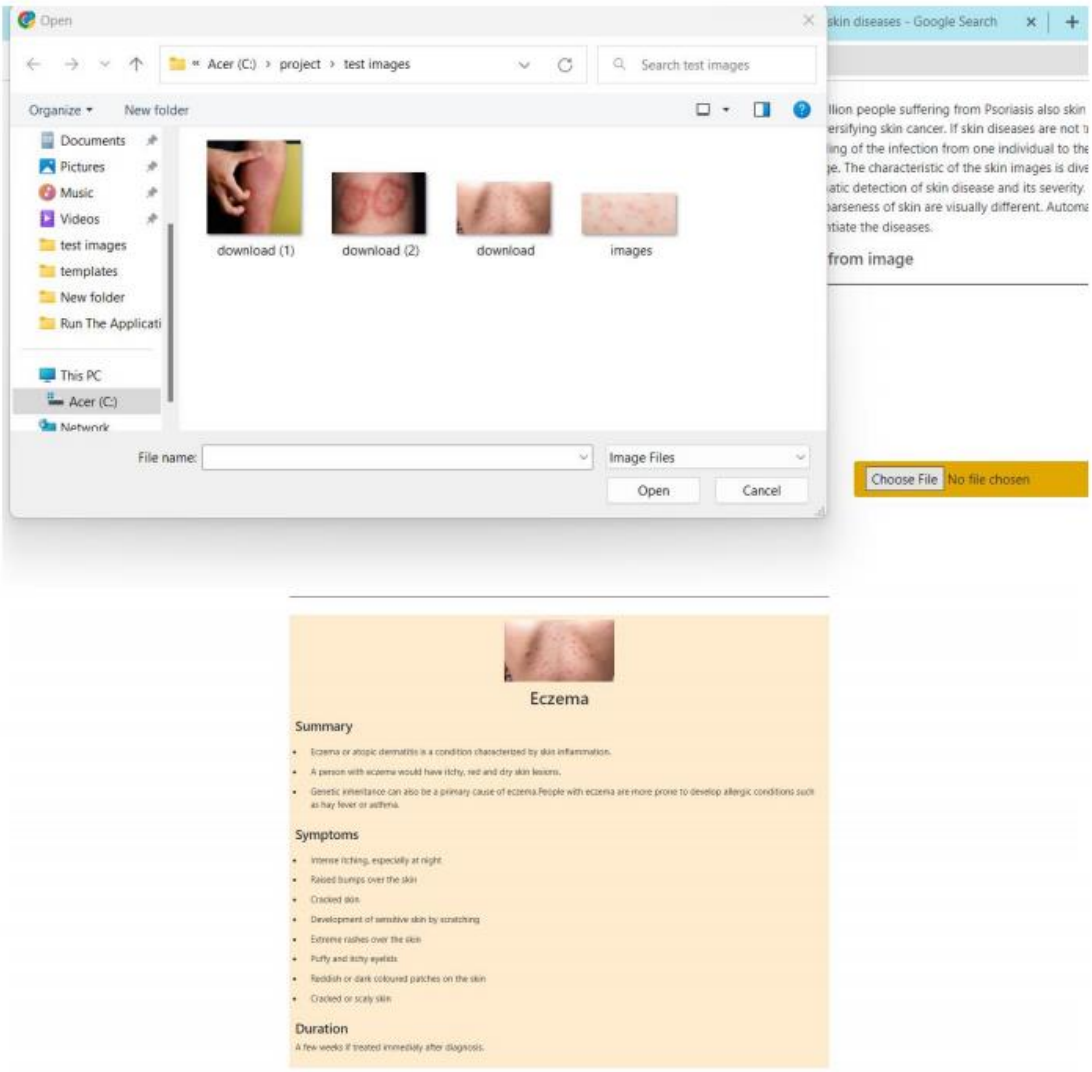
```
@app.get('/')
async def home(request: Request):
    return templates.TemplateResponse("index.html", {"request": request})
```

```
@app.post('/predict')
async def predict(image: UploadFile = File(...)):
    temp_file = save_to_disk(image, path="temp", save_as='temp')
    result = preprocess.predict(temp_file)
    with open(const.diagnosis_dir + const.diseases[result]+".json", 'r', encoding='utf-8') as f:
        diagnosis = json.load(f)
    return diagnosis
```

```
def save_to_disk(uploadedfile, path='.', save_as='default'):
    extension = os.path.splitext(uploadedfile.filename)[-1]
    temp_file = os.path.join(path, save_as+extension)
    with open(temp_file, 'wb') as buffer:
        shutil.copyfileobj(uploadedfile.file, buffer)
    return temp_file
```

Output:


Skinnovation- AI-based localization and classification of skin disease with erythema
Skin Disease Detection
Nowadays people are suffering from skin diseases, More than 125 million people suffering from Psoriasis also skin cancer rate is rapidly increasing over the last few decades especially Melanoma is most diversifying skin cancer. If skin diseases are not treated at an earlier stage, then it may lead to complications in the body including spreading of the infection from one individual to the other. The skin diseases can be prevented by investigating the infected region at an early stage. The characteristic of the skin images is diversified so that it is a challenging job to devise an efficient and robust algorithm for automatic detection of skin disease and its severity. Skin tone and skin colour play an important role in skin disease detection. Colour and coarseness of skin are visually different. Automatic processing of such images for skin analysis requires quantitative discriminator to differentiate the diseases.
Diagnosis from image
Upload an image
Choose File
No file chosen



7.2 Feature 2

1.The model provides a improved quality of diagnosis in various medical fields such as mammography and colonography, it is not used in dermatology, where noninvasive screening tests are performed only with the naked eye, and avoidable inaccuracies may exist.

2. This system proposed a computerized process of classifying skin disease through deep learning .

3. This model shows that CAD may also be a viable option in dermatology by presenting a novel method to sequentially combine accurate segmentation and classification models.

4.The classification model is more accurate than a baseline model trained without segmentation, while also being able to classify multiple diseases within a single image.

8.Testing

The proposed model's results on implementation and the statistical analysis through various performance evolution metrics that include, but are not limited to, accuracy measures determine how many times the proposed MobileNet V2 model with the LSTM model is successfully classifying the skin disease.

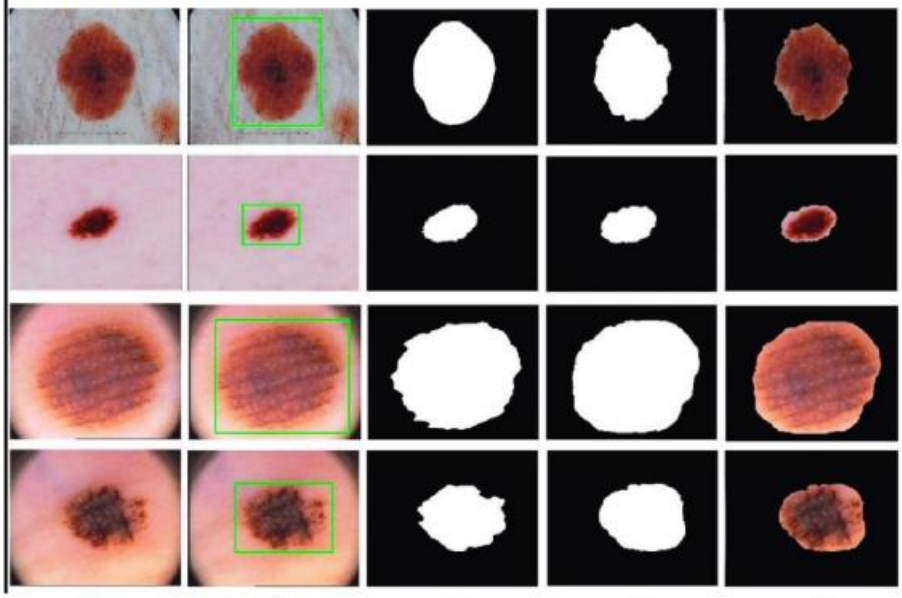
At first, the experiment was performed over several images, and the type of disease
is assessed through the proposed MobileNet V2 with the LSTM approach. The actual type of disease based on the actual ground facts is also presented.. The predicted confidence is on par with the ground reality.

The performance of the proposed Mobilenet V2 with LSTM is evaluated through various assessment metrics, and the implemented results are presented along with the graphs of the hyperparameters. It is evident from the obtained results that the proposed model's performance for lesion classification is reasonably fair with minimal computational time than the other approaches. The proposed model needs a considerable lesser computational effort in performing the classification of the images, which makes it suitable to deploy in mobility devices. The prospectus application that works with the proposed model can precisely identify the skin disease for the image that is captured.

The MobileNet model was trained on using the imbalanced dataset and the default preprocessing of input data. The system produces an accuracy of 93.6%.

9.Results

The final results are based on the accuracy results in the form of the melanoma and the non-melanomaskin diseases



(a) (b) (c) (d) (e)

classifications.

10.Advantages And Disadvantages

10.1 Advantages

1)Establishment of standardized skin disease image dataset

A large amount of data is the basis of skin disease recognition and the premise of acceptable generalization ability of the network model. However, the number of images, disease types, image size, and shooting and processing methods of the published datasets are considerably different, which leads to the confusion of different studies and the loss of the ability to quantitatively describe different models, Moreover, it is difficult to collect images of certain rare diseases. As mentioned above, there are numerous kinds of skin diseases; however, only approximately 20 datasets are available, including less than 20

kinds of skin diseases. There is an urgent requirement to expand access to medical images. For example, Indian researchers have trained neural networks to analyze images from “handheld imaging devices” instead of stationary dermatoscope devices to provide more prospects for early and correct diagnosis. However, a public database that allows the collection of a sufficient number of labeled datasets is likely necessary to truly represent projections of the population.

2)Interopreability of skin disease recognition

The progress of deep learning in skin disease recognition depends on a highly nonlinear model and parameter adjustment technology. However, the majority of the neural networks are “black box” models, and their internal decision-making process is difficult to understand. This “end-to-end” decision-making mode leads to the weak explanatory power of deep learning. The internal logic of deep learning is not clear, which makes the diagnosis results of the model less convincing. The interpretability research of skin disease classification could allow the owner of the system to clearly know the behavior and boundary of the system, and ensure the reliability and safety of the system. Moreover, it could monitor the moral problems and violations caused by training data deviation and provide a superior mechanism to follow the requirements within an organization to solve the bias and audit problems caused by AI.

3)Intelligent diagnosis and treatment of skin diseases

Deep learning can be used to address the increasing number of patients with skin disease and relieve the pressure of limited dermatologists. With the popularity of mobile phones, mobile computers, and wearable devices, a skin disease recognition system based on deep learning can be expected to be available to intelligent devices to serve more people. Using a mobile device camera, users can upload their own photos of the affected area to the cloud recognition system and download the diagnosis results at any time. Through simple communication with the “skin manager”, diagnosis suggestions and possible treatment methods could be available. Furthermore, the “skin manager” could monitor the user’s skin condition and provide real-time protection methods and treatment suggestions.

10.2 Disadvantages

- Medicine is an area that is not yet fully understood. Information is not completely transparent. The characteristics of dermatology determine that the majority of the data cannot be obtained. At the same time, the AI technology route is immature, the identification accuracy of which must be improved owing to the uncertainty of manual diagnosis. There is no strict correspondence between the symptoms and results of a disease and no clear boundary between the different diseases. Thus, the use of deep learning for disease diagnosis continues to require considerable effort.

- Before systematic debugging, extensive simulation, and robust validation, flawed algorithms could harm patients, which could lead to medical ethical issues, and therefore require forward-looking scrutiny and stricter regulation [98]. As a “black box”, the principle of deep learning is unexplained at this stage, which could result in unpredictable system output. Moreover, it is possible that humans could not truly understand how a machine functions, even though it is actually inspired by humans . Hence, whether or not patient care can be accepted using an opaque algorithm remains a point of discussion.

- There is a problem with the change in the error rate value in a dataset, which is caused by the change in the size of the dataset used in different skin cancer experiments. Therefore, the lack of a standard dataset can lead to serious problems; the error rate values are considered in many experiments. In addition, the collection of datasets for numerous studies depends on individual research, leading to unnecessary effort and time. When the actual class is manually marked and compared to the predicted class to calculate one of the parameter matrices, pixels are lost when the background is cut from the skin cancer image using Adobe Photoshop .At this

point, the process influences the results of all the parameter reliability groups (matrices, relationships, and behaviors), which are considered controversial. High reliability and low rate of time complexity cannot be achieved simultaneously, which is reflected in the training process and is influenced by conflicts between different standards, leading to considerable challenges .A method that works for the detection of one skin lesion could possibly not work for the detection of others . Numerous different training and test sets have been used to evaluate the proposed methods. Moreover, for the parameters in the training and evaluation, different researchers are interested in different parts. This lack of uniformity and standardization across all papers makes a fair comparison virtually impossible . Although these indicators in the literature have been widely criticized, studies continue to use them to evaluate the application to skin cancer and other image processing fields.

- The data used for evaluation are frequently overly small to allow a convincing statement regarding a system's performance to be made. Although it is not impossible to collect an abundance of relevant data through the Internet in this information age, this information, with significant uncertainty, apparently cannot meet the requirements of independent and identical distribution, which is one of the important prerequisites for deep learning to be successfully applied. For certain rare diseases and minorities, only a limited number of images are available for training. To date, a large number of algorithms have demonstrated prejudice against minority groups, which could cause a greater gap in health service between the "haves" and the "have-nots" . Numerous cases are required for the training process using deep learning techniques. In addition, although the deep learning technique has been successfully applied to other tasks, the developed models in skin are valid in only specific dedicated diseases and are not applicable to common situations. Diagnosing dermatology is a complex process that, in addition to image recognition, must be supplemented by other means such as palpation, smell, temperature change, and microscopy.

11.Conclusion

The proposed model based on the MobileNet V2 and LSTM approach proved efficient for skin disease classification and detection with minimal computational power and effort. The outcome is promising, with an accuracy of 85.34% when experimented with and compared with other methods over the real-time images . The model is computationally effective, and would produce the prediction accuracy by maintaining the previous timestamp data. The information related to the current state through weights optimizations would make the model robust. It is also compared against various other conventional model like CNN.

The project achieved 94.4% accuracy in determining the skin diseases. Using under sampling method and the default preprocessing of input data achieved an 84.28% accuracy on the test dataset. While, using the imbalanced dataset and the default preprocessing of input data achieved a 93.6% accuracy. In order to enhance the accuracy of the model different techniques and preprocessing of input data can be explored.

12. Future Scope

This implementation of the Structural Co-Occurrence matrices for feature extraction in the skin diseases classification and the pre-processing techniques are handled by using the Median filter, this filter helps toremove the salt and pepper noise in the image processing; thus, it enhances the quality of the images, and normally, the skin diseases are considered as the risk factor in all over the world. Our proposed approach provides 97% of the classification of the accuracy results while another existing model such as FFT + SCM gives 80%, SVM + SCM gives 83%, KNN + SCM gives 85%, and SCM + CNN gives 82%. Future work is dependent on the increased support vector machine's accuracy in classifying skin illnesses, and SCM is used to manage the feature extraction technique.

13.Appendix

Github Link :

https://github.com/IBM-EPBL/IBM-Project-23847-1659932123