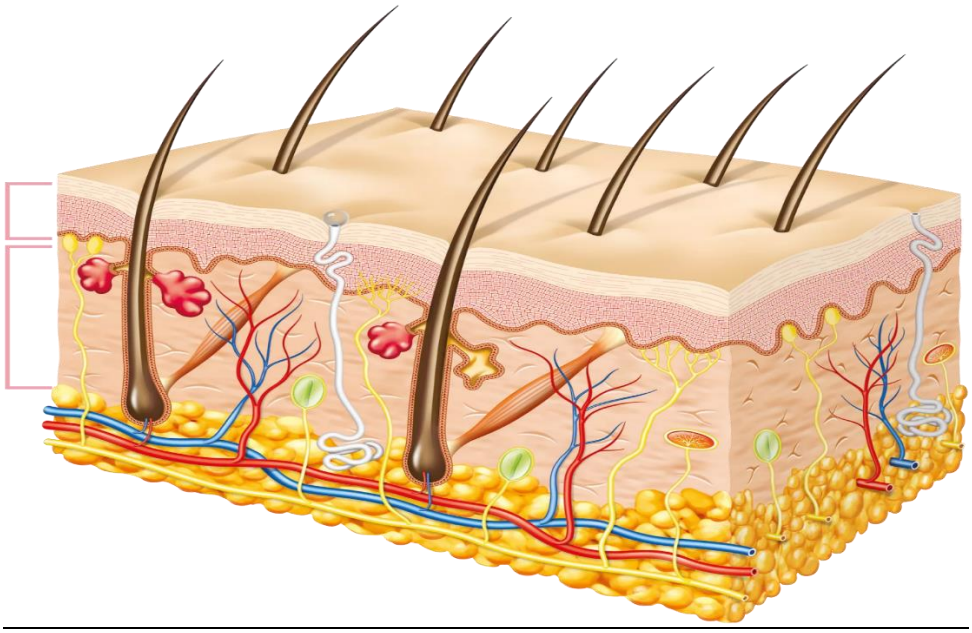
MULTICLASS SKIN IMAGE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

CANALYTICS



1. Business Understanding

Skin diseases present in many different forms impacting individuals 'overall health and well-being. Some of these skin diseases can be challenging to categorize and detect, which introduces complexity to the field of dermatology. The importance of accurate diagnosis cannot be overstated, as certain skin disorders, including various types of skin cancer, have the potential to be life-threatening. Early and accurate identification of the types of skin diseases is of great importance. The diagnostic process typically involves a range of methods, including visual image inspections, biopsies, and histopathological analyses. Distinguishing between benign and malignant lesions is particularly important as even minor or inconspicuous abnormalities can be difficult to detect. In response to these diagnostic challenges, cutting-edge technologies like deep learning algorithms offer the potential to revolutionize dermatological diagnostics, enhancing

efficiency, reducing errors, and ultimately improving patient outcomes across various skin disorders types. Read more on below links:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5817488/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6403009/

2. Problem Statement

Dermatologists at Flatter Dermatological Clinic are facing difficulties in accurately determining or categorizing various skin conditions types when examining medical skin images. Currently, this task heavily relies on manual visual inspection and personal judgment which is time-consuming, prone to human error and can result in delayed or inaccurate diagnoses. This inefficiency increases the chances of missing important skin conditions patterns and making mistakes which could have life-threatening consequences.

3. Objective

Main Objective: To build a convolutional neural network model capable of classifying the 9 different types of skin diseases with over 70% precision.

Other objectives are;

- To explore the distribution of the different types or class of skin images in the dataset.
- To assess the quality and consistency of images in the dataset per class.

4. Data Understanding

The dataset consists of 2357 images of malignant and benign oncological diseases, which were formed from the International Skin Imaging Collaboration (ISIC). All images were sorted according to the classification taken with ISIC, and all subsets were divided into the same number of images, with the exception of melanomas and moles, whose images are slightly dominant.
The data set contains the following diseases:

- Actinic keratosis
- Basal cell carcinoma
- Dermatofibroma
- Melanoma
- Nevus

- Pigmented benign keratosis
- Seborrheic keratosis
- Squamous cell carcinoma
- Vascular lesion

5. Data Preparation

Exploratory Data Analysis: Explore the dataset through statistical analysis and visualizations to understand its characteristics, class distribution, and potential challenges.

Data Preprocessing: Clean and preprocess the data, including resizing, normalizing, and augmenting.

6. Modeling

Below outlines the steps of how we will develop the model.

- We shall use Convolutional Neural Networks (CNNs) a deep learning model for image classification.
- Secondly, we will split the dataset into training, validation, and testing sets.
- Then train the neural network on the training data, using appropriate optimization techniques and loss functions.
- We shall then optimize hyperparameters, such as learning rate and batch size, to improve model performance.
- Conduct regularization techniques like dropout and batch normalization to prevent overfitting.
- Choose an appropriate data augmentation strategy to resolve underfitting/overfitting
- Then evaluate the model's performance on the validation set using relevant metrics, such as accuracy, precision, recall, and F1-score
- We then make necessary adjustments to the model and hyperparameters based on validation results.
- We will thereafter assess the final model's performance on the test set to ensure it generalizes well.

7. Model Evaluation

The success of our skin disease image classification project shall be measured through the following metrics:

- **Accuracy:** The percentage of correctly classified images.
- **Precision:** The ratio of true positive predictions to the total positive predictions, measuring the model's ability to avoid false positives.
- **Recall**: The ratio of true positive predictions to the total actual positive cases, measuring the model's ability to find all positive cases.
- **F1-score:** Balance the trade-off between false positives and false negatives.
- **Confusion Matrix:** To determine insights into the types of errors made by the model (e.g., false positives and false negatives).

8. Deployment

Model deployment will be executed using Streamlit which offers a user-friendly interface with seamless access to accurate skin disease predictions.