

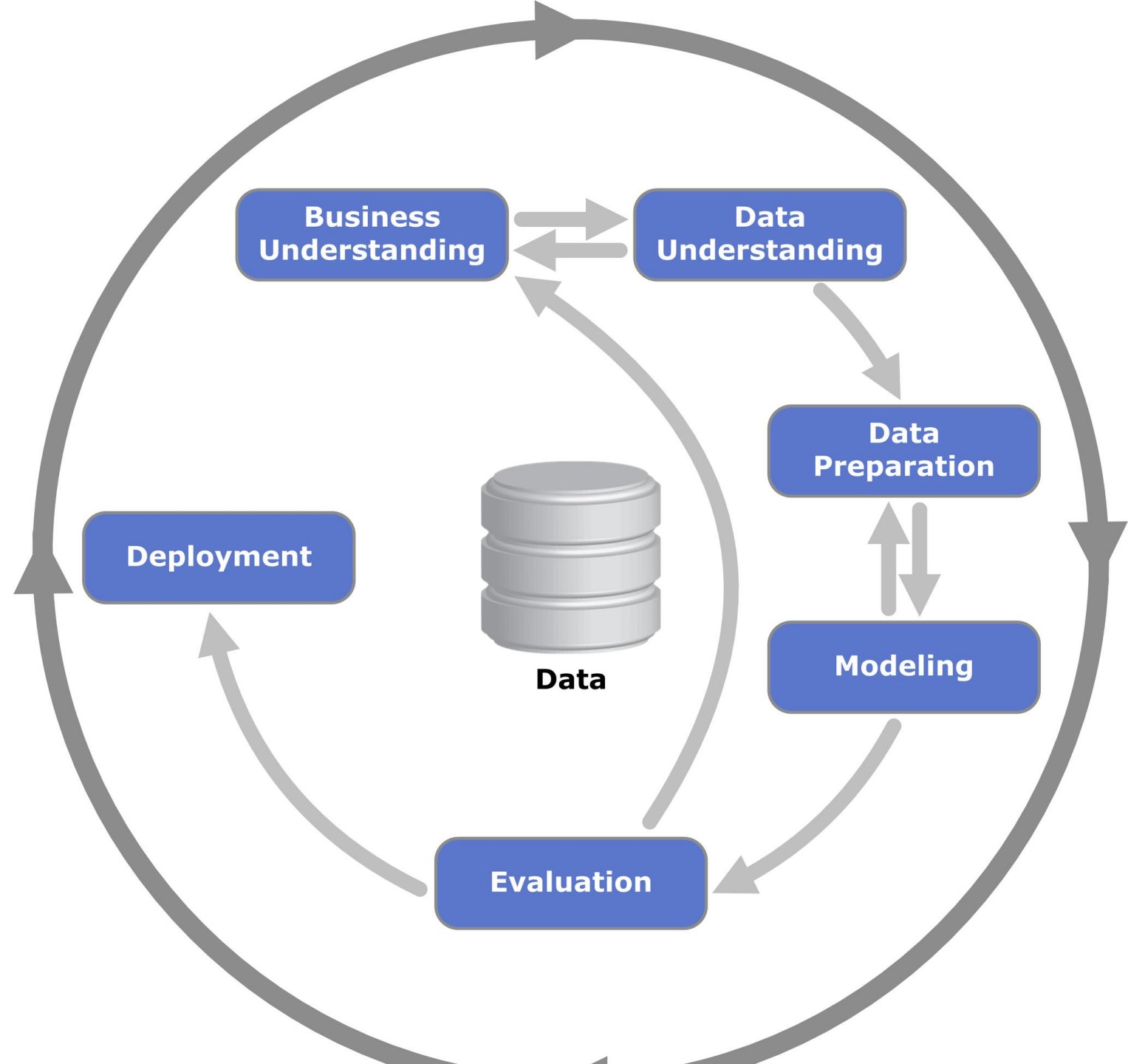
# **CLASSIFICATION OF SKIN DISEASES**



## Group 3 members

- Antony Kanai
- Brenda Kinoti
- Jared Bii
- Leah George
- Phelix Okumu
- Sharon Kimutai
- Janet Khainza

# Project Scope





# Business Understanding

- Skin disorders can vary widely, **impacting overall health**, and some, **like skin cancer can be life-threatening**.
- **Early and accurate identification** of the types of skin diseases is of great importance.
- Cutting edge technologies can revolutionize dermatological diagnostics, **enhancing efficiency, reducing errors, and ultimately improving patient outcomes**.

# Problem Statement

---

- Dermatologists at Flatter Dermatological Clinic struggle with **accurately categorizing skin diseases from medical images.**
- Current **manual inspection and personal judgment** - time-consuming, prone to errors leading to delayed or inaccurate diagnoses.
- This inefficiency can result in **missed patterns** and **life-threatening consequences.**





# Objectives

- **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;**
  - i. To **explore the distribution** of the different types/class of skin images in the dataset.
  - ii. To **assess the quality and consistency** of images in the dataset.





# Data Understanding

---

- Dataset : **2357** images
  - **2,239** images for Training
  - **118** images for Testing
- The data set contains **9 skin diseases**:  
Actinic keratosis, **Basal cell carcinoma**, Dermatofibroma, **Melanoma**, Nevus, Pigmented benign keratosis, Seborrheic keratosis, **Squamous cell carcinoma** , Vascular lesion.

# Data Preparation

1

**Exploratory Data Analysis:** image counts, sampling per class, Class distributions, Class separation, Texture Analysis, RGB color channels and Pixel Intensity

2

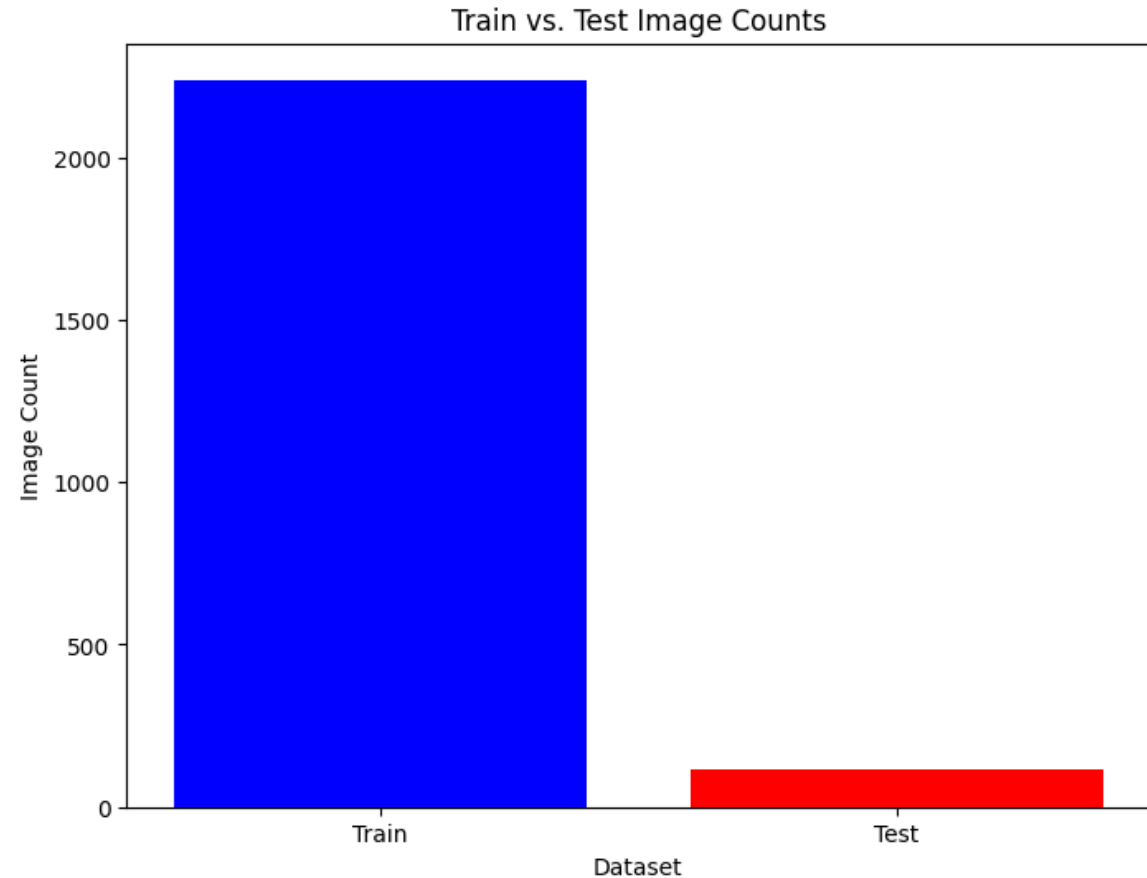
**Data Preprocessing** – Splitting the data (Training Validation sets), Rescaling and Resizing, Dealing with imbalance and transforming images through augmenter





# Image counts

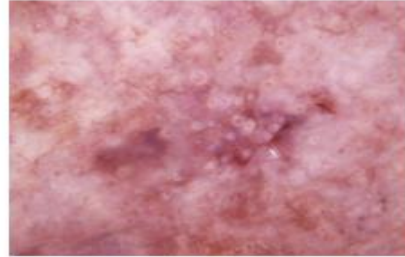
- Distribution of the skin diseases image dataset with **5%** of dataset reserved for **testing purposes** and **95%** of dataset for **training the CNN Model**



# Image sample per class

- This images shows a sample of **image per class**
- The images have distinct appearance from **red, pink, brown, black and purple colors**
- **Actinic Keratosis** appears as a **visible rash**, **Melanoma** appears as a **dark patch** and **Dermatofibroma** presents like **tiny brownish patches**.

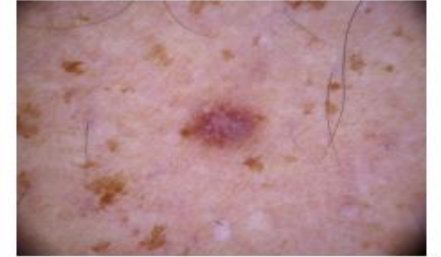
actinic keratosis



basal cell carcinoma



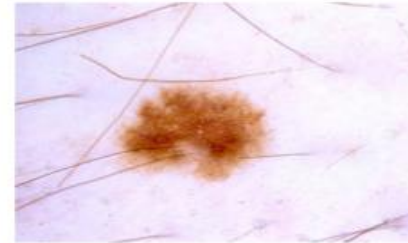
dermatofibroma



melanoma



nevus



pigmented benign keratosis



seborrheic keratosis

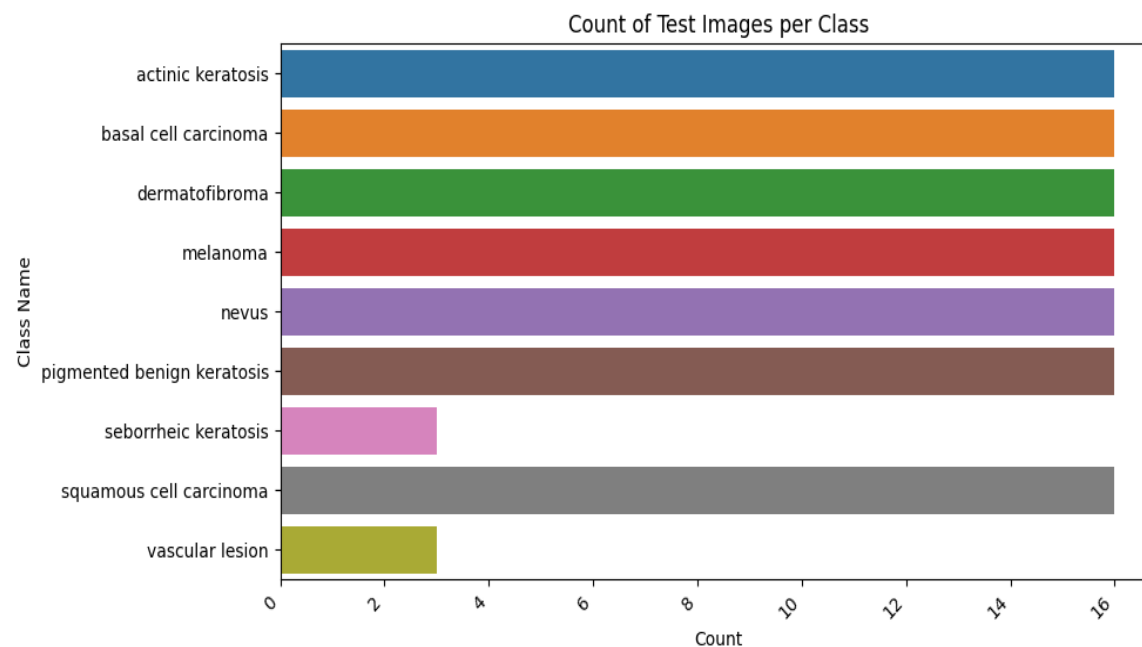
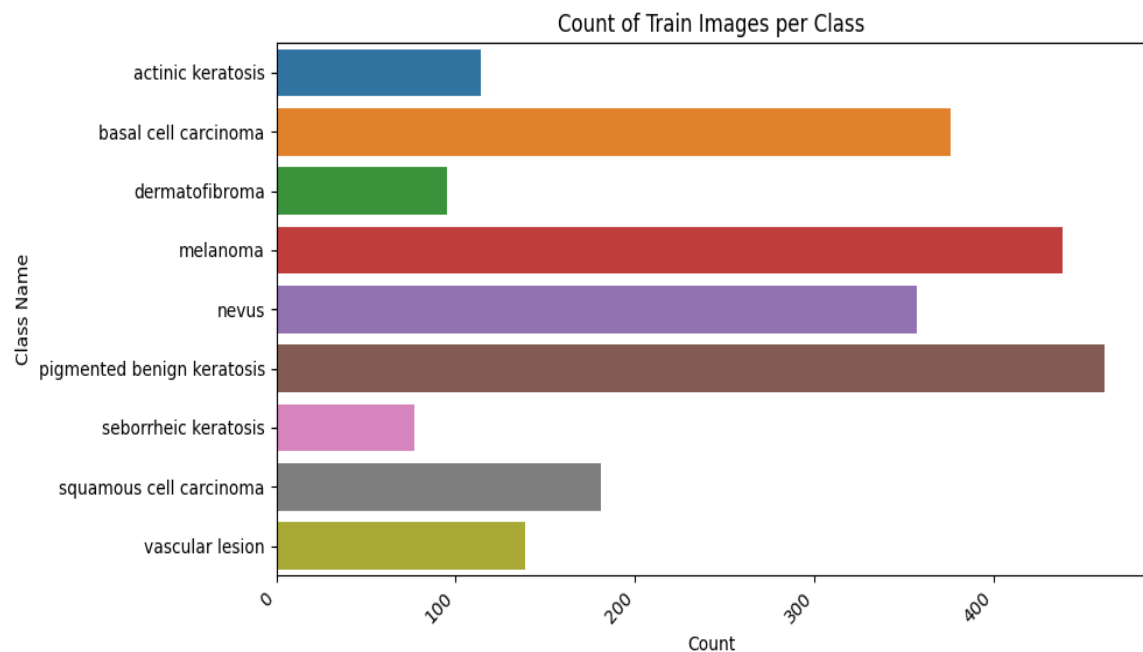


squamous cell carcinoma



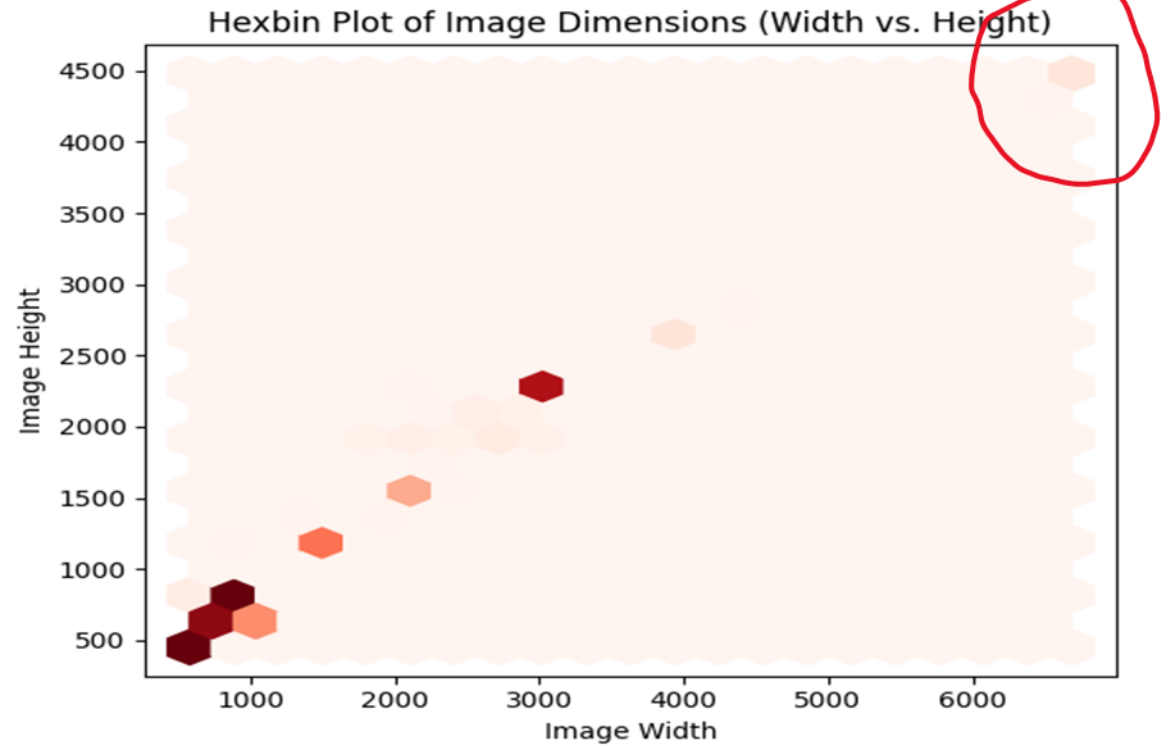
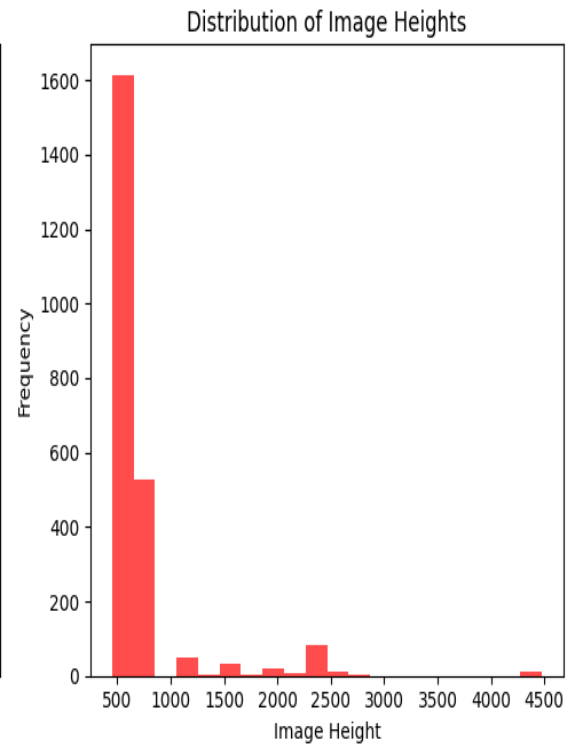
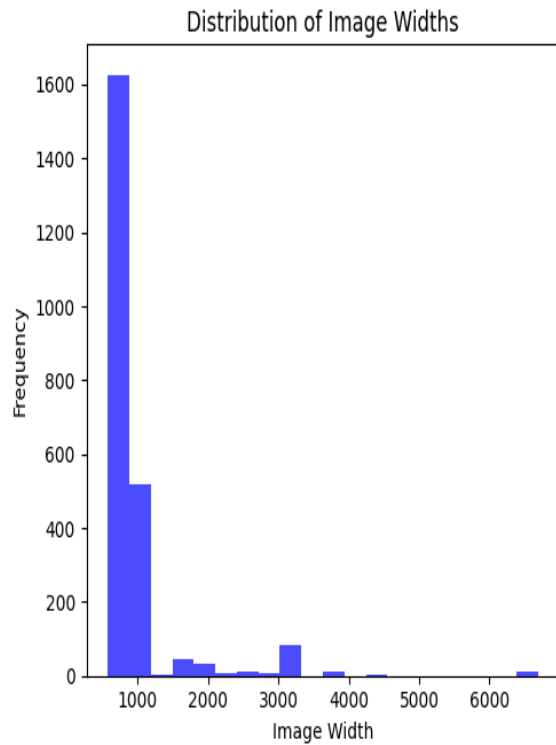
vascular lesion





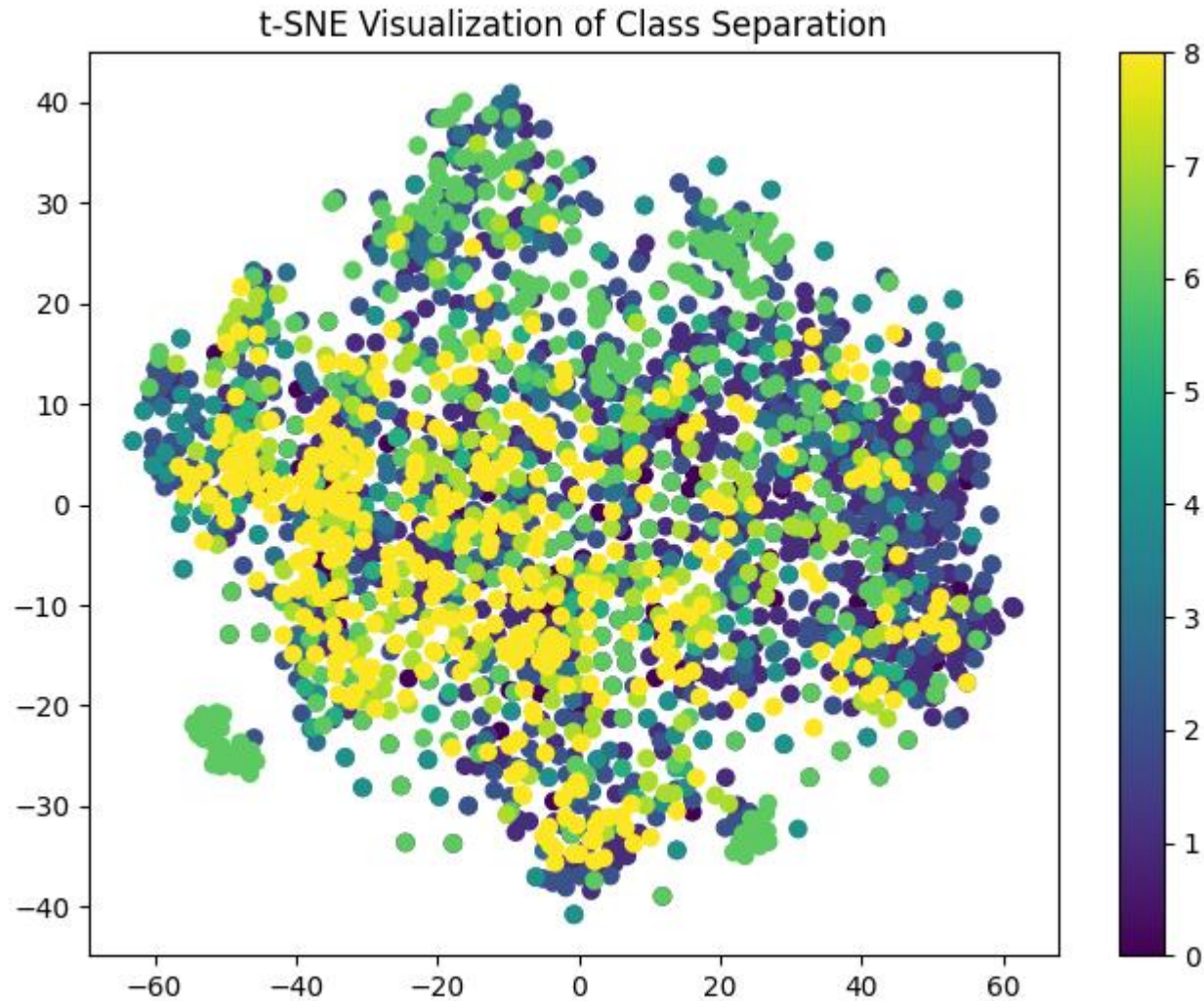
# Class distribution

- There is a significant **class imbalance on the Train set**
- **Majority of test set** class samples are **balanced**.



## Height and Width Dimensions

- Majority of images dimensions - approximately around **500 pixels in width** and **450 pixels in height**.
- Outlier **circled in red** - dimensions approximately around **6500 pixels in width** and **4500 pixels in height**.
- Standardize the image dimensions by resizing to a common size.



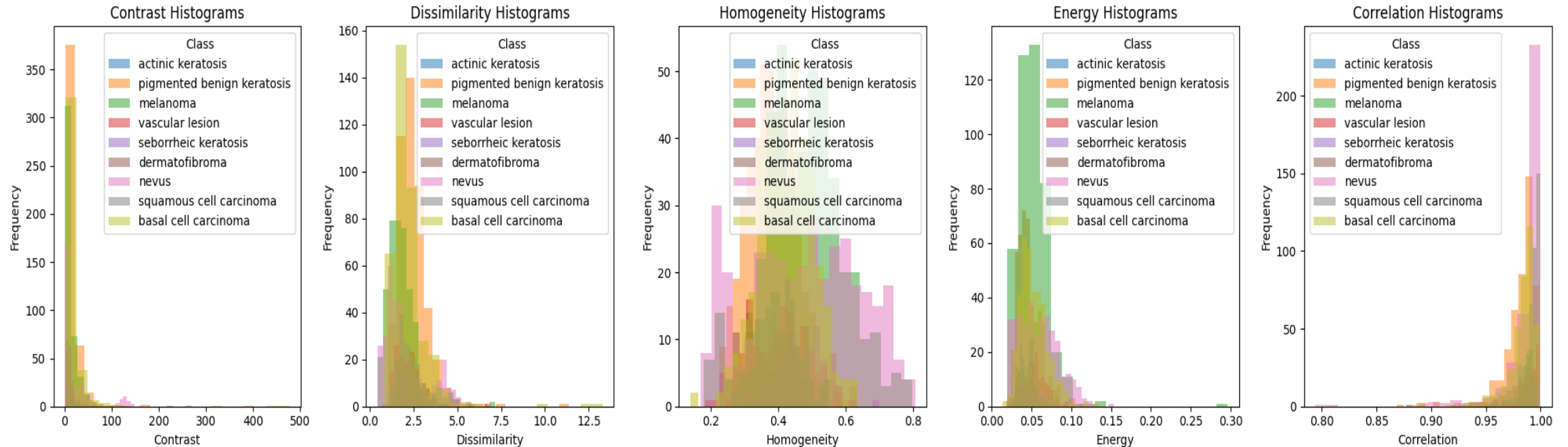
# Class Separation

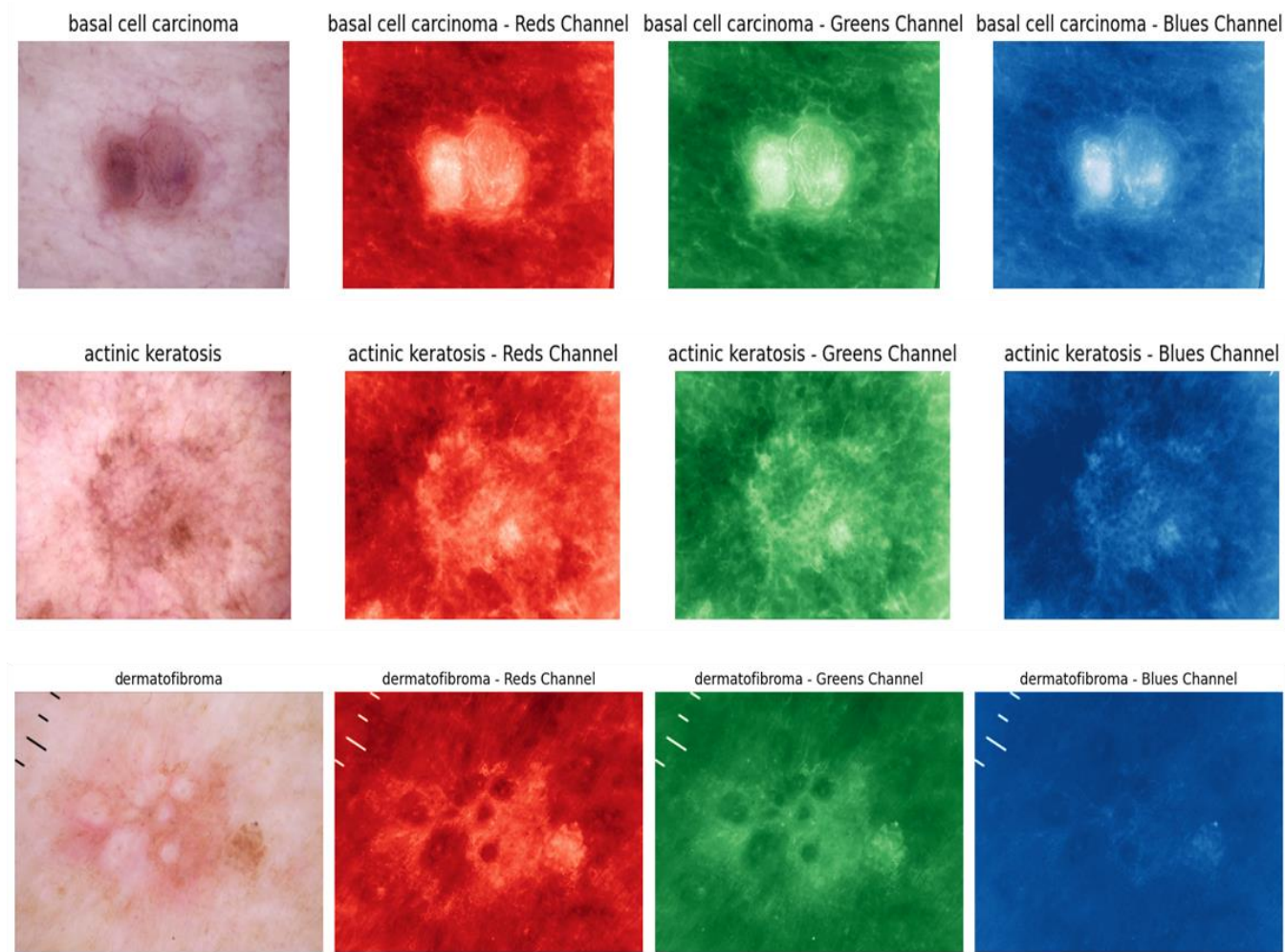
- Absence of clear class separation - images either **have complex visual characteristics** or **share visual similarities**.
- Apply augmentation to provide diverse data allowing the model to accurately distinguish between different classes.



# Texture Analysis

- Majority of images in the classes exhibit **low contrast**, **low dissimilarity**, and **low energy**, indicating **minimal intensity variations**.
- Near-normal distributed homogeneity and higher correlation suggest a **uniform and consistent texture pattern** within the images



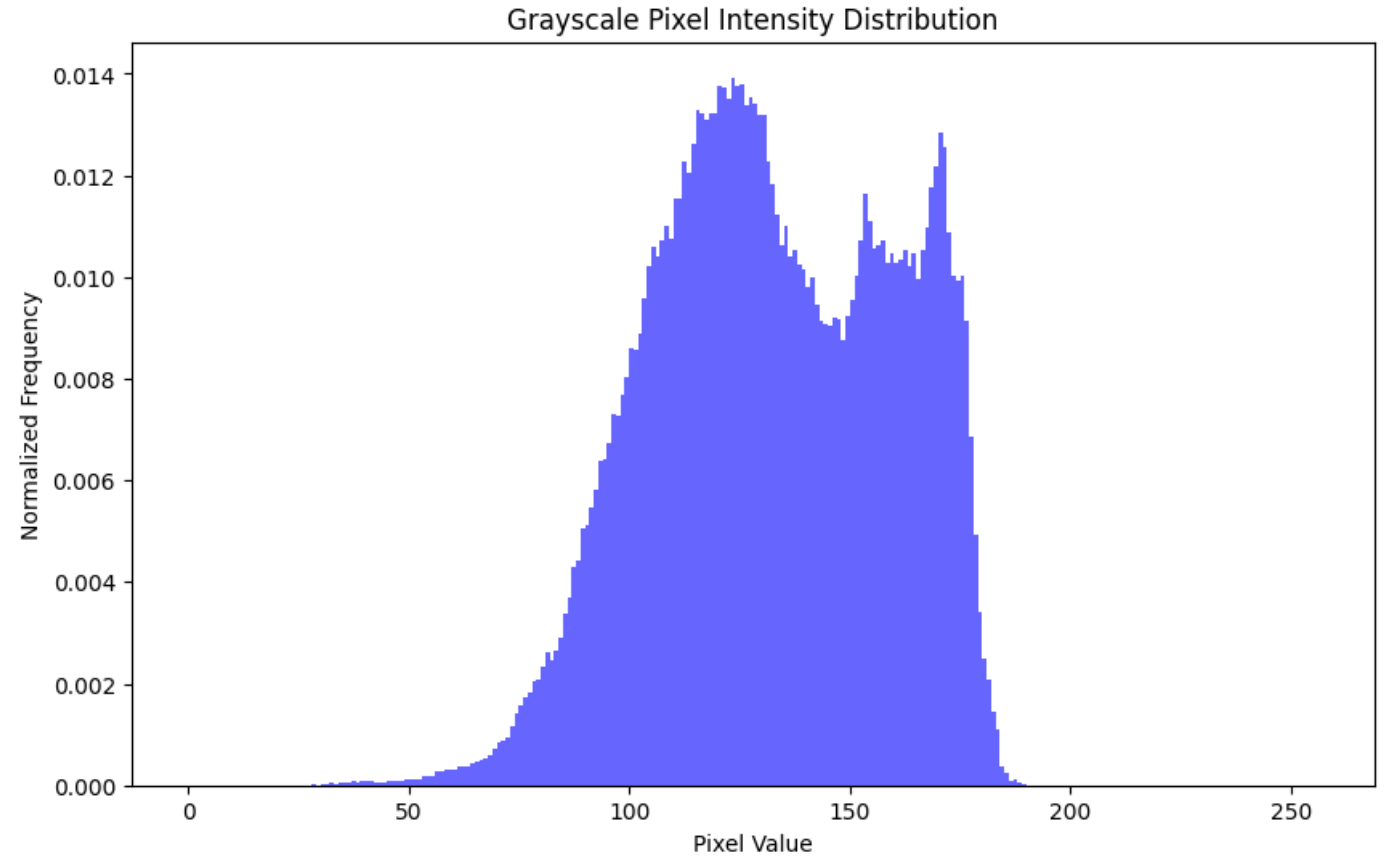


## RGB Color Channels

- Class images as they appears in **original, red, blue and green color channels.**
- The images seem to be **clearer on the red color channel.**

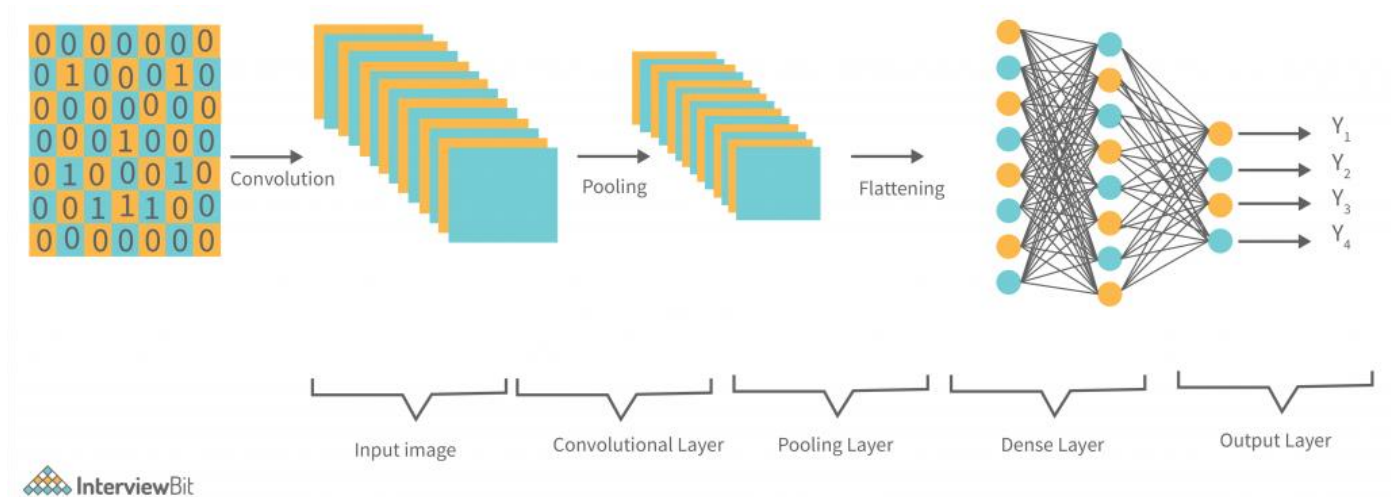
# Pixel Intensity

- 130 is the most common pixel value in the image, representing about 1.3% of the image's pixel distribution.
- Images exhibit a broad range of grayscale values, including both dark and bright regions.



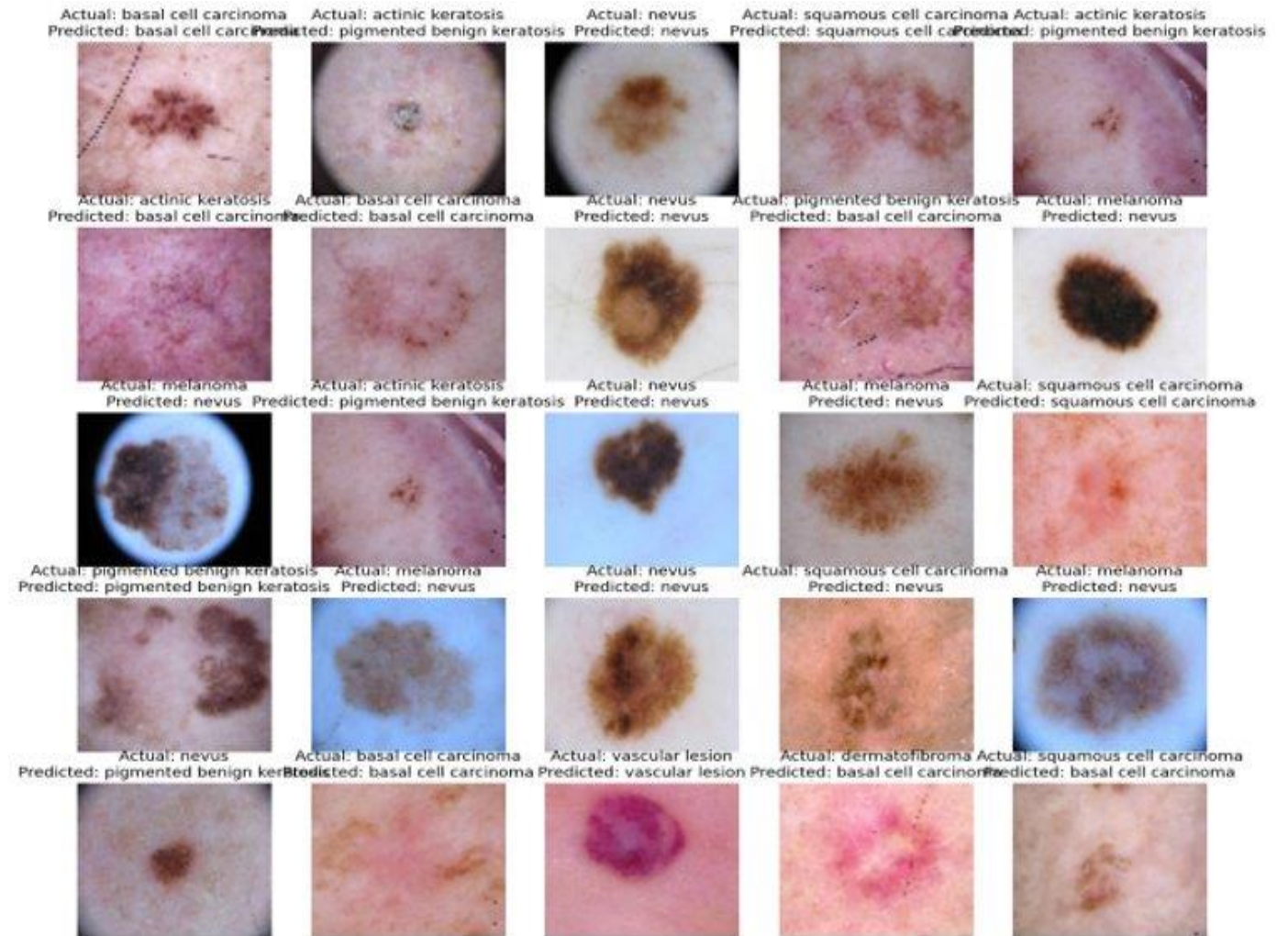
# CNN Modeling

- Among the four models , the best performer was **VGG16**, a pre-trained model originally trained on a large dataset.
- Model performance Improvement - **Enhanced the sample size** by **transforming images**, creating a more diverse set of samples and addressing class imbalance.
- Model Evaluation - Focused on three metrics: Precision, Accuracy, and Recall.



# Model Evaluation

- While all four models showed relatively high-performance metrics, **the third model demonstrated noticeable enhancements in precision scores**
- Precision – **Training 81% and Validation 73%**
- The model seems to predict **Nevus** and **Basal cell carcinoma** very well, indicating high Precision in identifying these conditions..





# Data Limitation

- The skin disease **images size was relatively small** and could have significantly impacted the model performance
- The Data may have **not been from diverse geographical locations or age groups** limiting the model's generalizability.
- The dataset **may not account for variations in skin tones and ethnicities**, which can impact the model's performance in providing accurate diagnoses across diverse patient populations.



# Conclusions and Recommendations

---

- Precision of **73%** demonstrates the potential of our **model in enhancing dermatological diagnostics** and **ultimately improving patient care**.
- To build upon our success and advance the model, we recommend the following:
  - Utilize the model's proficiency in predicting **Basal cell carcinoma** and **Nevus**.
  - Obtain a **larger image data** for training the model.
  - Seek data from **multiple sources** and **geographical regions**
  - Collect data that spans **different time periods** and **age groups**.



# Model Deployment

---

- Model deployment was done using stream lit.
- Link to upload images: [CNN · Stream lit](#)

## Skin Condition Image Classifier

This app classifies images of skin conditions into one of nine classes.

Upload an image



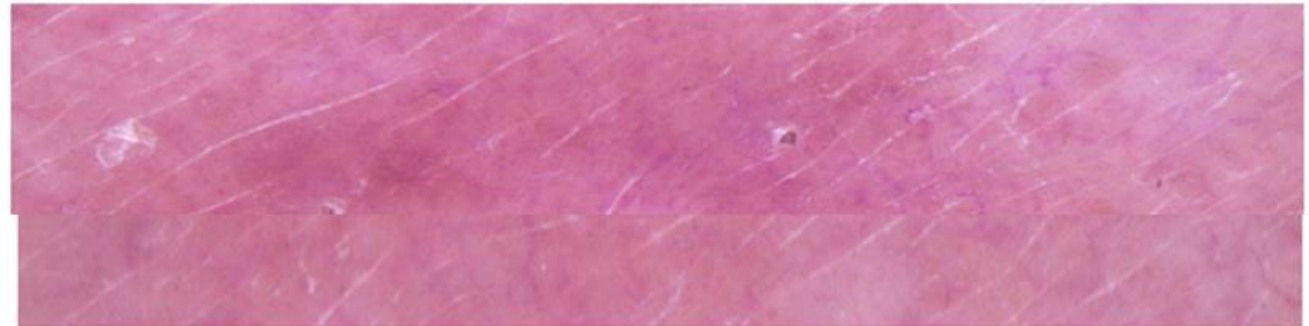
Drag and drop file here

Limit 200MB per file • JPG, JPEG, PNG

Browse files



ISIC\_0024386.jpg 346.2KB



Uploaded Image

Prediction: dermatofibroma

This is a simple skin condition image classification app using a CNN model.

*Thank  
you*

