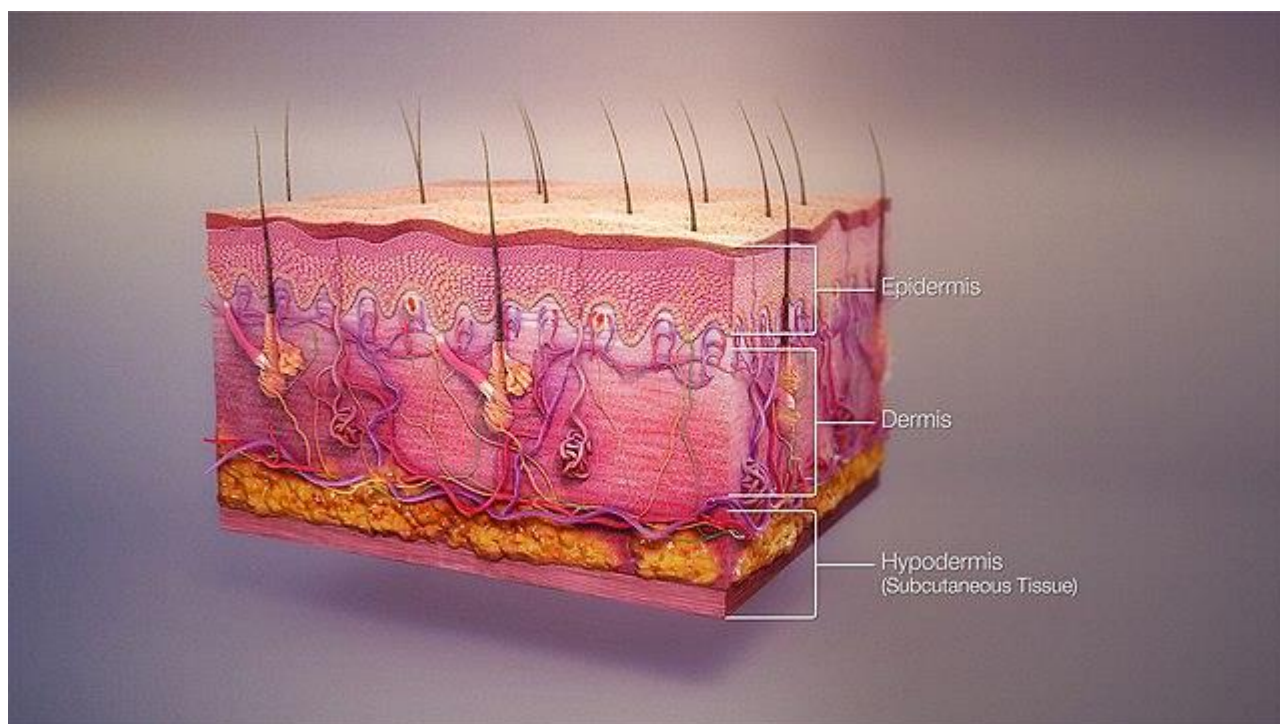


# **PROJECT SUMMARY**

## **MULTI-CLASS IMAGE CLASSIFICATION OF SKIN DISEASES USING CONVOLUTIONAL NEURAL NETWORKS**



## Contents

1.0 Business Understanding .....	3
1.1 Problem Statement .....	3
1.2 Objectives .....	3
2.0 Data Understanding.....	4
3.0 EXPLORATORY DATA ANALYSIS (EDA) .....	4
3.1 Image Count .....	4
3.2 Cancerous and Non-Cancerous images .....	5
3.3 Image Sample per Class .....	5
3.4. Class Distribution.....	6
3.5 Height and Width Distribution of Images and Density Distribution (Height Vs Width) .....	7
3.6 Class Separation.....	9
3.7 Texture Analysis.....	10
3.8 RGB Visualization .....	11
3.9 Pixel Intensity .....	11
4.0 Data Preparation and Preprocessing .....	12
4.1 Splitting Data .....	12
4.2 Rescaling and Resizing .....	13
5.0 Modeling.....	13
5.1. Baseline Model .....	14
5.2 Models with Data Augmentation .....	15
<b>5.2.1 Model 2 .....</b>	<b>15</b>
5.2.2 Model 3 - Pre-trained Model (VGG16) .....	16
5.3 Best Model and Model Evaluation .....	18
6.0 Model Deployment:.....	20
7.0 Conclusions and Recommendations.....	20
7.1 Data Limitations:.....	20
7.2 Conclusions.....	21
7.3 Recommendations.....	21

## **1.0 Business Understanding**

Skin diseases have become a major global health problem in recent years, affecting millions of people worldwide. The prevalence of various skin conditions, such as dermatitis, acne, and skin cancers, has increased. Environmental factors such as pollution and exposure to harmful UV rays have contributed to this rise in skin-related issues.

However, advances in medical research have led to a better understanding, diagnosis, and treatment options for many skin diseases. Dermatologists now use cutting-edge technologies such as AI-powered diagnostics and targeted therapies to improve the accuracy of diagnoses and the effectiveness of treatments.

Despite these advancements, there are still challenges such as limited access to specialized care in certain regions and the emergence of new, complex skin disorders. These challenges necessitate ongoing research and healthcare efforts.

Accurate skin disease diagnosis is crucial for patient well being. Deep learning, particularly convolution neural networks (CNN) offers a promising solution unlike manual diagnosis methods which are time-consuming and prone to errors. Such learning algorithms can revolutionize dermatological diagnostics, improving efficiency and patient outcomes.

### **1.1 Problem Statement**

Flatter Dermatological Clinic grapples with the accurate categorization of skin conditions from medical images. Manual inspection, prone to errors and time constraints, results in delayed diagnoses. The clinic seeks to deploy a CNN model for swift, precise, and automated classification of skin diseases, mitigating delays and enhancing diagnostic accuracy.

### **1.2 Objectives**

**Main objective:** To develop a CNN model capable of classifying 9 skin disease types 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

## **2.0 Data Understanding**

In our pursuit to revolutionize skin conditions diagnostics, we utilized a dataset comprising 2357 images of malignant and benign oncological diseases in 9 different classes sourced from the International Skin Imaging Collaboration (ISIC).

No.	Classes	Train Images	Test Images
1.	Actinic keratosis	114	16
2.	Basal cell carcinoma	376	16
3.	Dermatofibroma	95	16
4.	Melanoma	438	16
5.	Nevus	357	16
6.	Pigmented benign keratosis	462	16
7.	Seborrheic keratosis	77	3
8.	Squamous cell carcinoma	181	16
9.	Vascular lesion	139	3

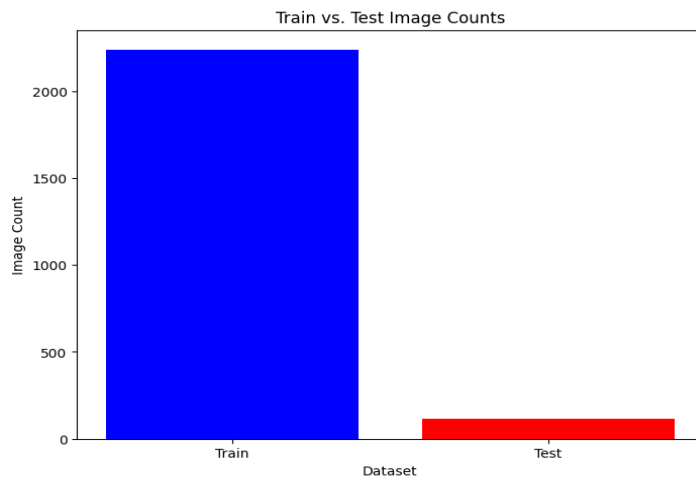
These images are split into Train and Test images, 2239 and 118 respectively. Below is a table showing the distribution of the train and test images in the 9 classes;

## **3.0 EXPLORATORY DATA ANALYSIS (EDA)**

We looked closely at the data to help us develop a better way to classify skin diseases. We used different types of visualizations to understand the data, such as density distributions and t-SNE analysis.

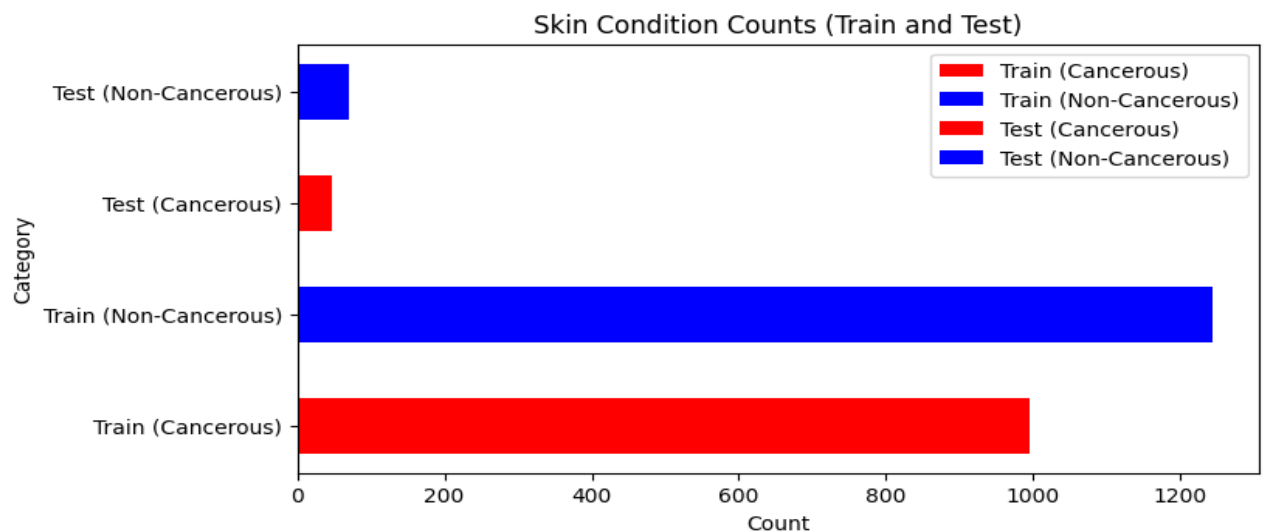
### **3.1 Image Count**

The bar-graph below shows the distribution of the skin diseases image dataset with 5% of dataset reserved for testing purposes and 95% of dataset for training the CNN Model



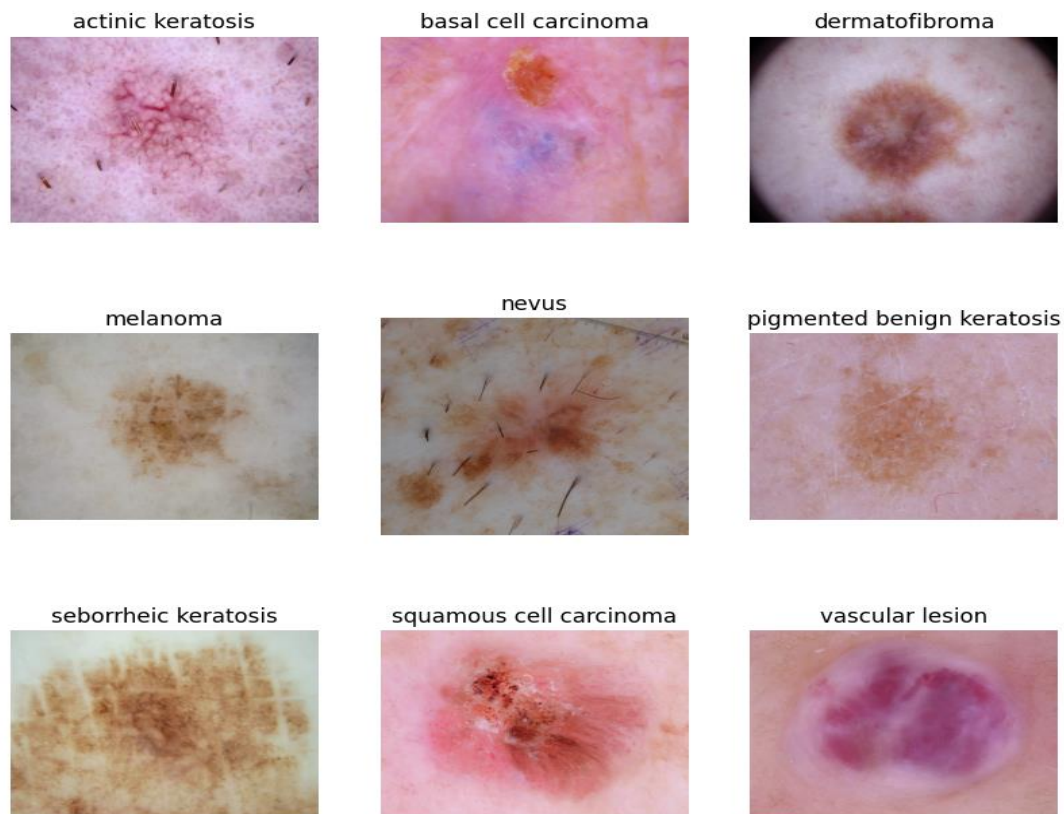
### **3.2 Cancerous and Non-Cancerous images**

We specifically examined cancerous and non-cancerous skin images for visualization purposes as shown in the graph below. This provided essential insights into the dataset, enhancing our understanding of diverse skin conditions.



### **3.3 Image Sample per Class**

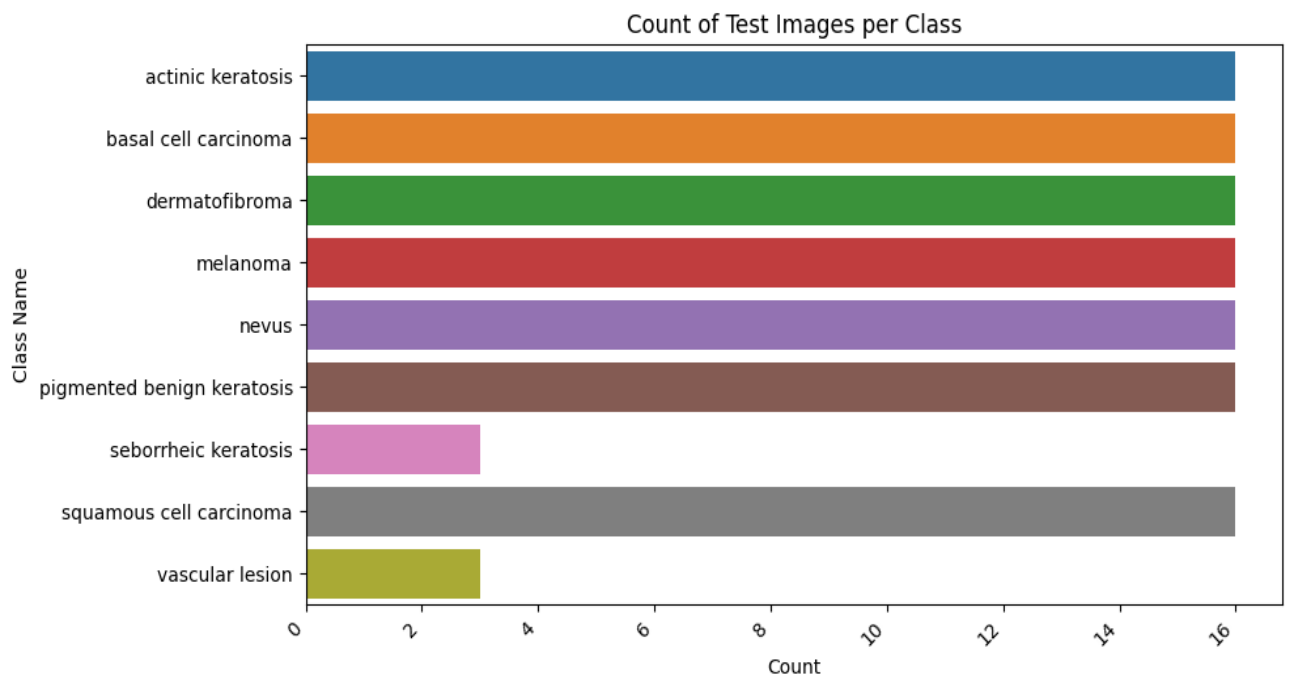
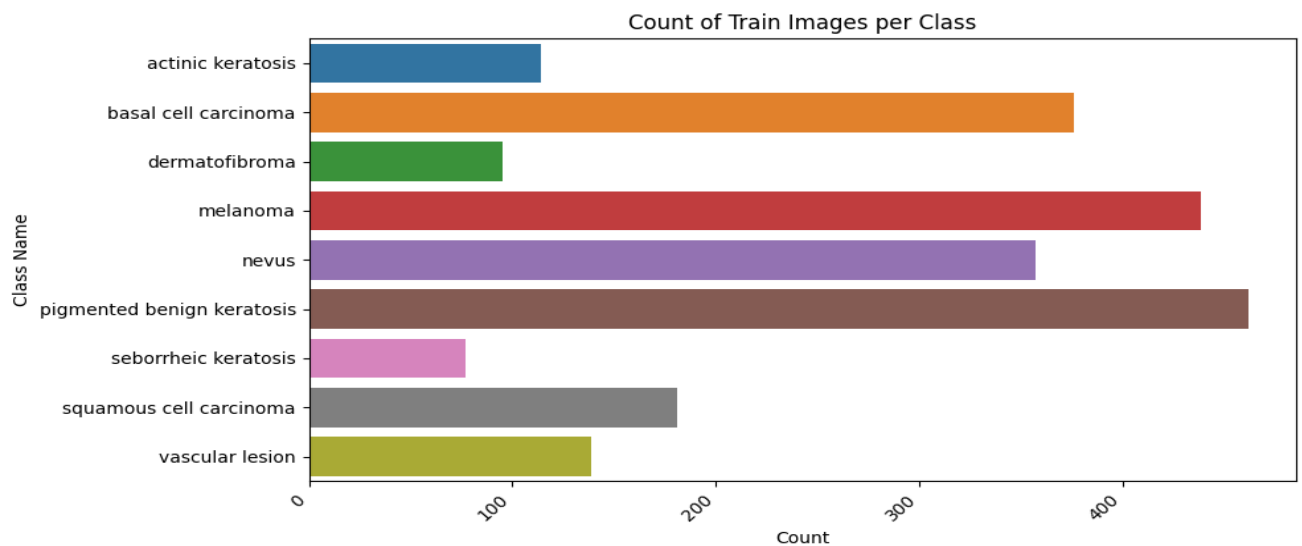
The visualization below shows a sample image per class. The images have distinct appearance from red, pink, brow, black and purple colors.



### **3.4. Class Distribution**

From the bar Graphs below, the following observations can be made:

- The train class is relatively imbalanced compared to the test set.
- There are less instances of particular skin disease types in both train and test which may impact on the model's ability to generalize.



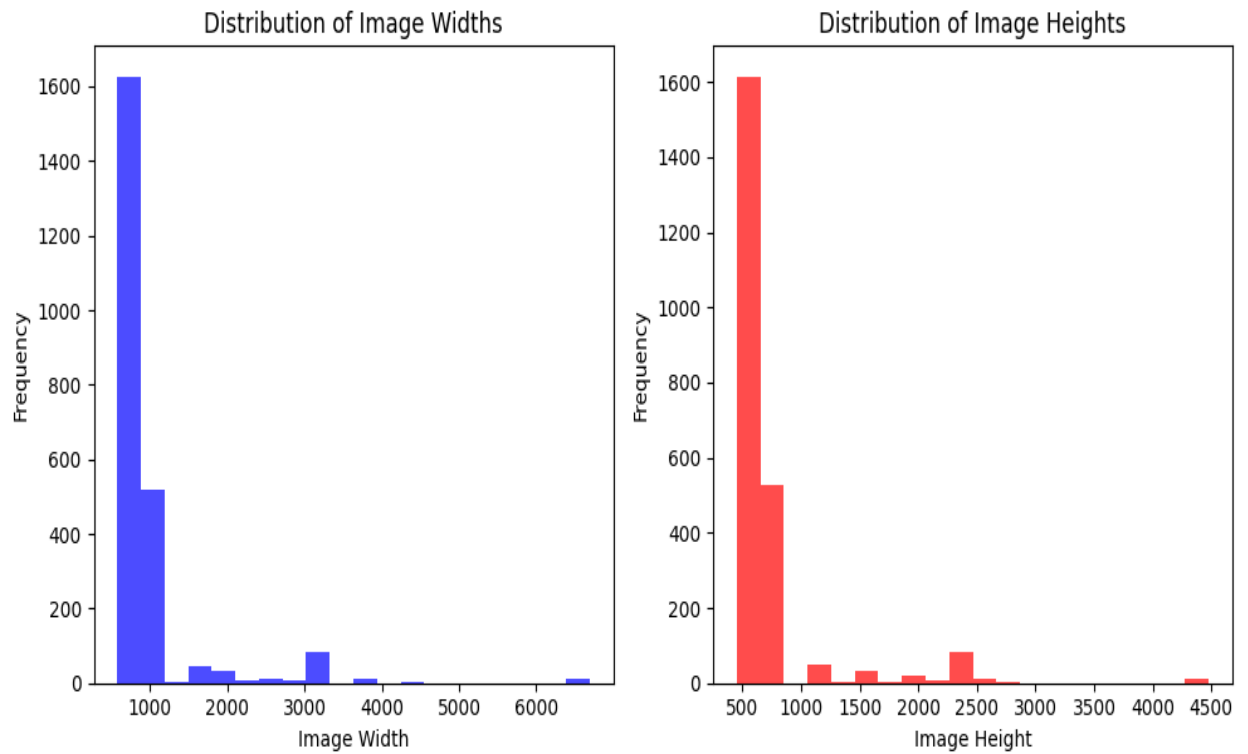
### **3.5 Height and Width Distribution of Images and Density Distribution (Height Vs Width)**

Visualization of the weights and heights of the images show that;

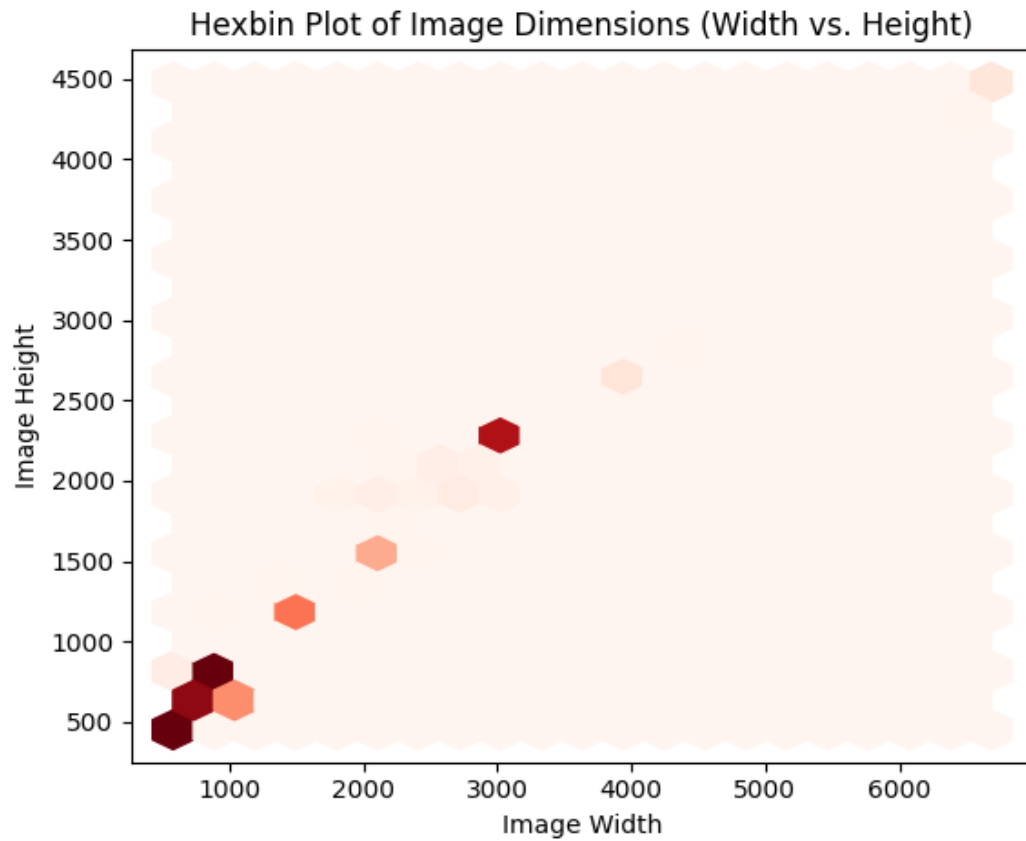
- Most of the images have widths ranging from 450 to 1100
- Most images have heights ranging from 450 -800
- The density plot reveals that most of the images are approximately sized as 500 \*450

- There is an outlier with dimensions approximately around 6500 pixels in width and 4500 pixels in height.

In modeling, we will standardize the image dimensions by re-sizing them to a common size.



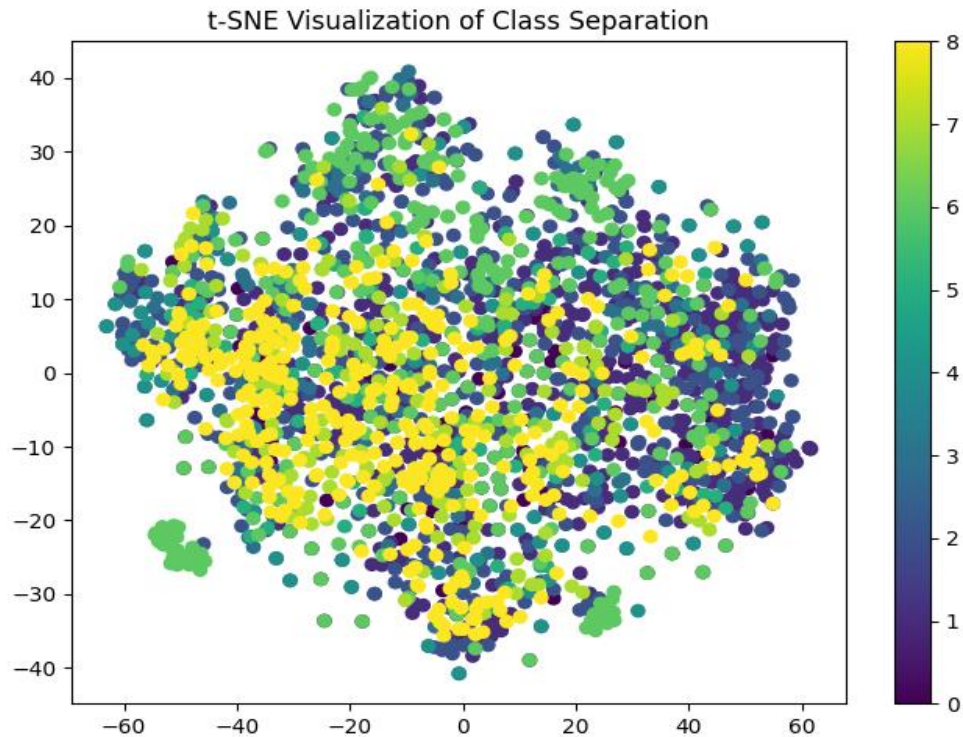




### **3.6 Class Separation**

From the t-SNE visualization on class separation below, the classes are spread all over and there are no clear-cut separation of classes.

- t-SNE showing no clear class separation implies:
  - Complex or similar visual characteristics in the data.
  - Potential challenges for image classification.

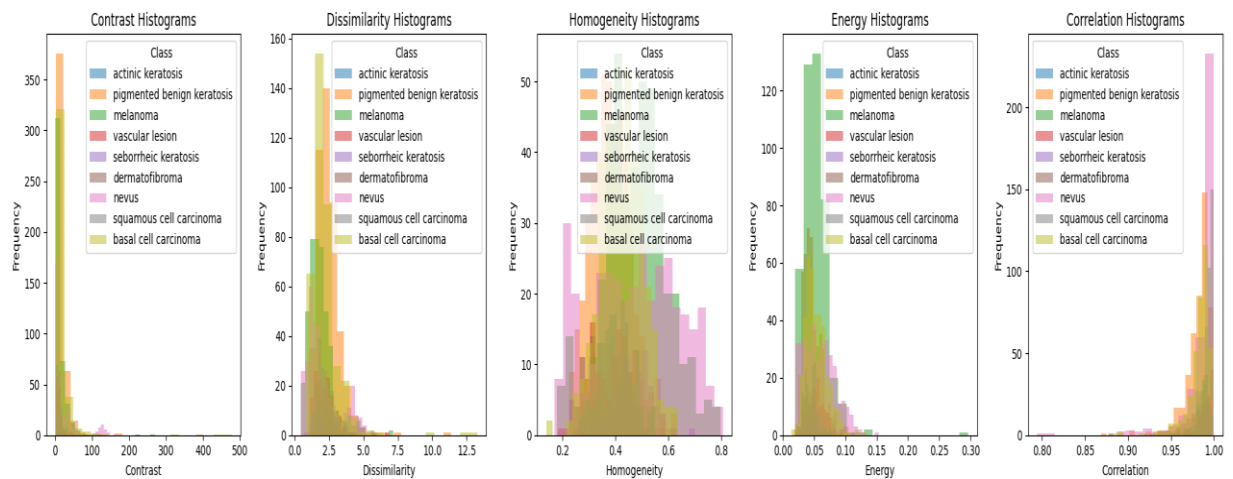


This implies we should apply data augmentation to provide diverse data allowing the model to accurately distinguish between different classes.

### **3.7 Texture Analysis**

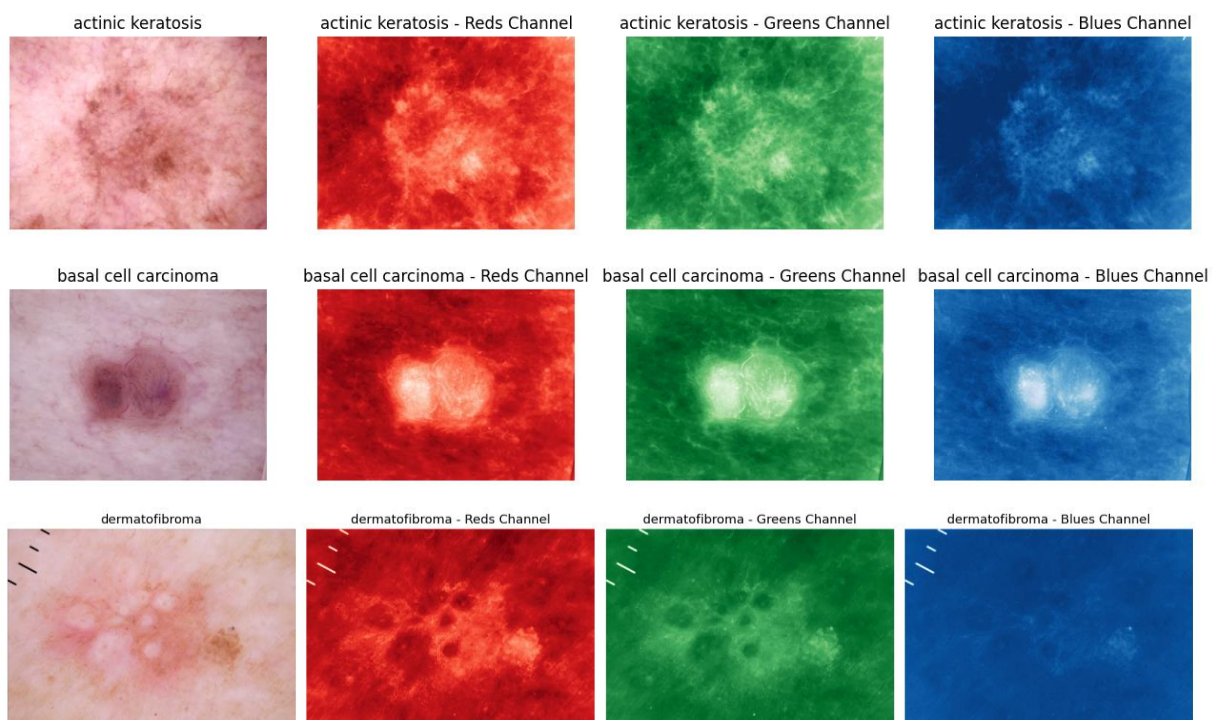
From the visualizations below, we can draw the following key insights from our texture analysis:

- Majority of the images have low contrast of below 100.
- Most classes exhibit a near normal distribution in their homogeneity or similarity. This means that there is less diversity among images in the classes.
- Most classes have their images having low energy. This means that there is less variability in pixel intensity of the images.



### 3.8 RGB Visualization

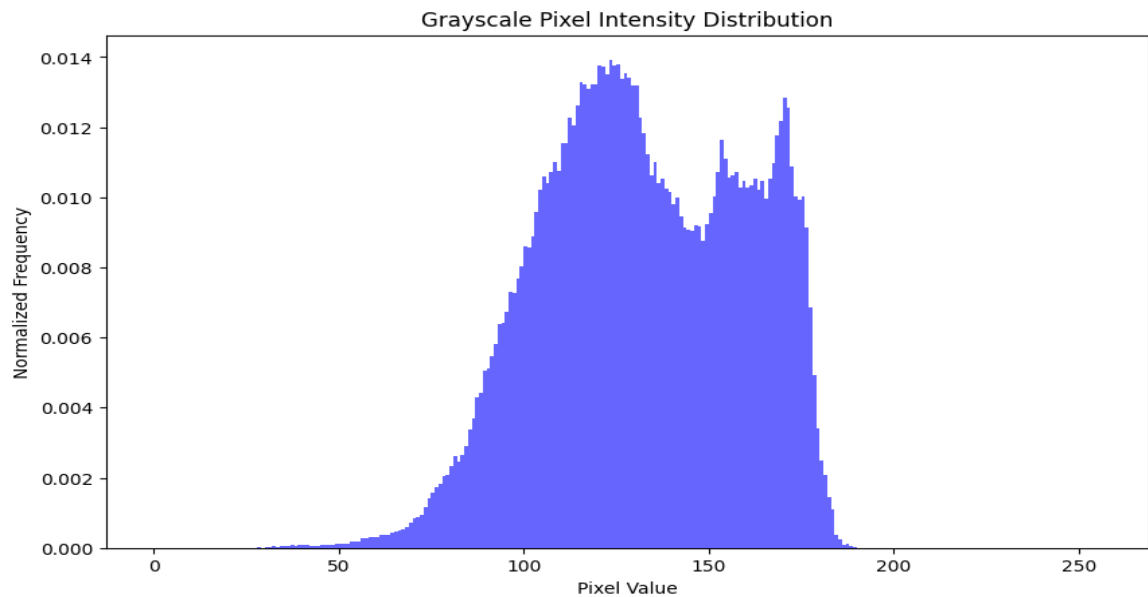
The Red Green Blue visualization below shows class images as they appears in red, blue and green color channels. The images seem to be clearer on the red color channel.



### 3.9 Pixel Intensity

The grayscale pixel intensity analysis as shown below highlights a prominent peak at pixel value 130, indicating its frequent occurrence in the images. This value represents about 1.3% of the pixels, showcasing its prevalence. The pixel values

ranging from 30 to 190 cover a wide grayscale range, depicting diverse image shades. This information is pivotal for our analysis and model development.



Utilizing these critical insights gathered from our detailed EDA, we are now poised to construct a model tailored for the precise identification of various skin diseases. The ultimate goal is to revolutionize dermatological diagnostics, ensuring swift and accurate diagnoses.

## **4.0 Data Preparation and Preprocessing**

We employed Python libraries including PIL and scikit-learn for data preparation. Images were preprocessed by re-sizing, normalizing and augmenting ensuring consistency and enhancing the quality of the dataset.

### **4.1 Splitting Data**

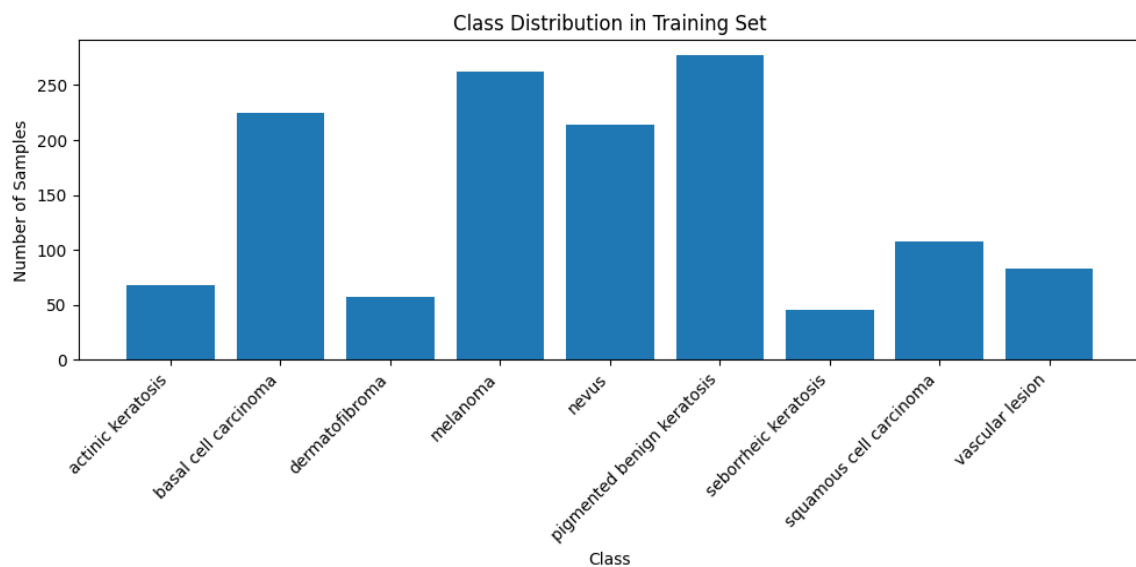
We chose a 60:40 split to balance training data sufficiency and a sizable validation set for effective model evaluation, considering our separate test set. This division ensures thorough performance assessment without significantly reducing the training dataset size.

## **4.2 Rescaling and Resizing**

We efficiently preprocessed the image datasets by re-sizing and re-scaling the images. This crucial step ensured memory optimization. The resulting generators provided data batches and labels directly usable for model training, validation, and testing.

## **4.3 Class Imbalance**

The visualization revealed notable class imbalances. Classes like Seborrheic keratosis, Dermatofibroma, and Actinic keratosis had fewer samples, while melanoma, pigmented benign keratosis, and basal cell carcinoma had more. To address this, we applied augmentations in our second model.



These steps were imperative to provide the neural network with standardized inputs, optimizing its learning process. By ensuring uniformity in image sizes and pixel values, we enhance the model's ability to generalize patterns effectively. Standardized inputs are fundamental for the neural network to learn efficiently and make accurate predictions, essential elements in our pursuit of precise skin disease classification.

## **5.0 Modeling**

For our deep learning approach, we chose TensorFlow and Keras libraries. We constructed a neural network with convolutional layers tailored to the complexities of the medical images. Our model architecture included convolutional layers for feature extraction, max-pooling layers for dimensionality reduction, and dense layers for

classification. We employed the Adam optimizer and utilized the categorical cross-entropy loss function in our models.

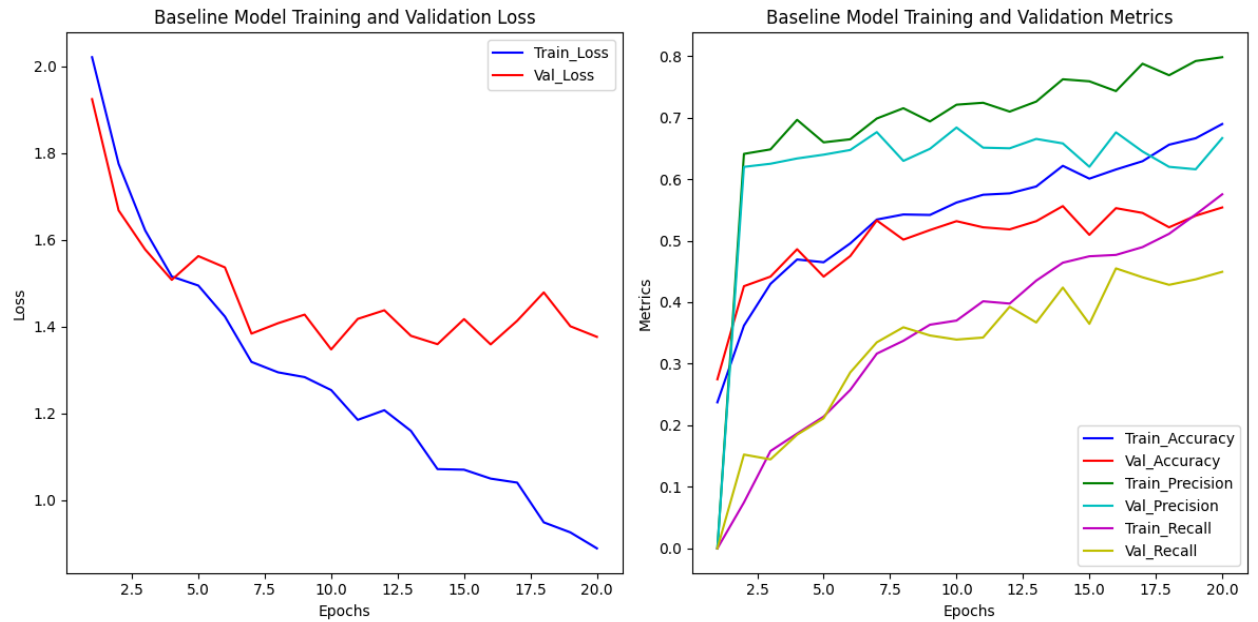
In our model evaluation, we strategically chose precision, recall, and accuracy as our primary metrics to assess the performance of our classification model. Precision, in particular, took precedence as our main evaluation metric. Precision gauges the model's ability to correctly identify positive cases, emphasizing the importance of accurate positive predictions. By incorporating precision, recall, and accuracy into our evaluation strategy, we aimed for a comprehensive understanding of our model's effectiveness in classifying skin diseases accurately and reliably.

**5.1. Baseline Model**

The Baseline Model demonstrated signs of over-fitting, evident in the training precision of approximately 75.27% compared to the validation precision of 68.39%. Notably, the training loss (1.1437) surpassed the validation loss (1.3475), indicating over-fitting challenges. To mitigate these issues and tackle class imbalances, we introduced data augmentation techniques in the next model.

Summary of the evaluation metrics:

Training Precision	Validation Precision	Training Accuracy	Validation Accuracy	Training Recall	Validation Recall
72.27%	68.39%	68.96%	55.39%	57.54%	44.94%



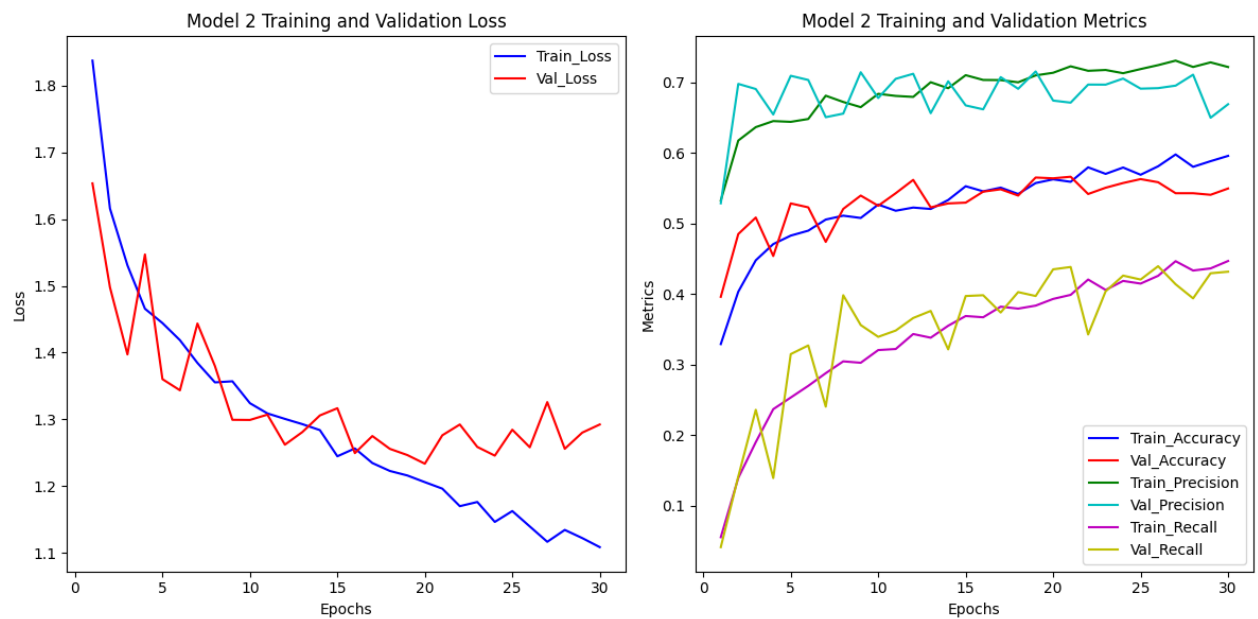
## 5.2 Models with Data Augmentation

### 5.2.1 Model 2

Despite the reduction in overfitting due to data augmentation, challenges persisted. The training precision stood at 70.67%, slightly higher than the validation precision of 67.41%. Although the model showed improvement, overfitting was not completely eradicated. This led us to consider integrating a pre-trained model.

Summary of the evaluation metrics:

Training Precision	Validation Precision	Training Accuracy	Validation Accuracy	Training Recall	Validation Recall
70.67%	67.41%	59.57%	54.95%	44.67%	43.16%



### **5.2.2 Model 3 - Pre-trained Model (VGG16)**

The introduction of the VGG16 pre-trained model resulted in enhanced metrics. Training precision reached 81.48%, with validation precision at 73.16%. While overfitting reduced, some challenges remained. The training loss (0.7620) was lower than the validation loss (1.0313), indicating our ongoing efforts to strike a balance between high performance and generalization.

Summary of the evaluation metrics:

Training Precision	Validation Precision	Training Accuracy	Validation Accuracy	Training Recall	Validation Recall
81.48%	73.16%	82%	62.29%	78.80%	59.84%



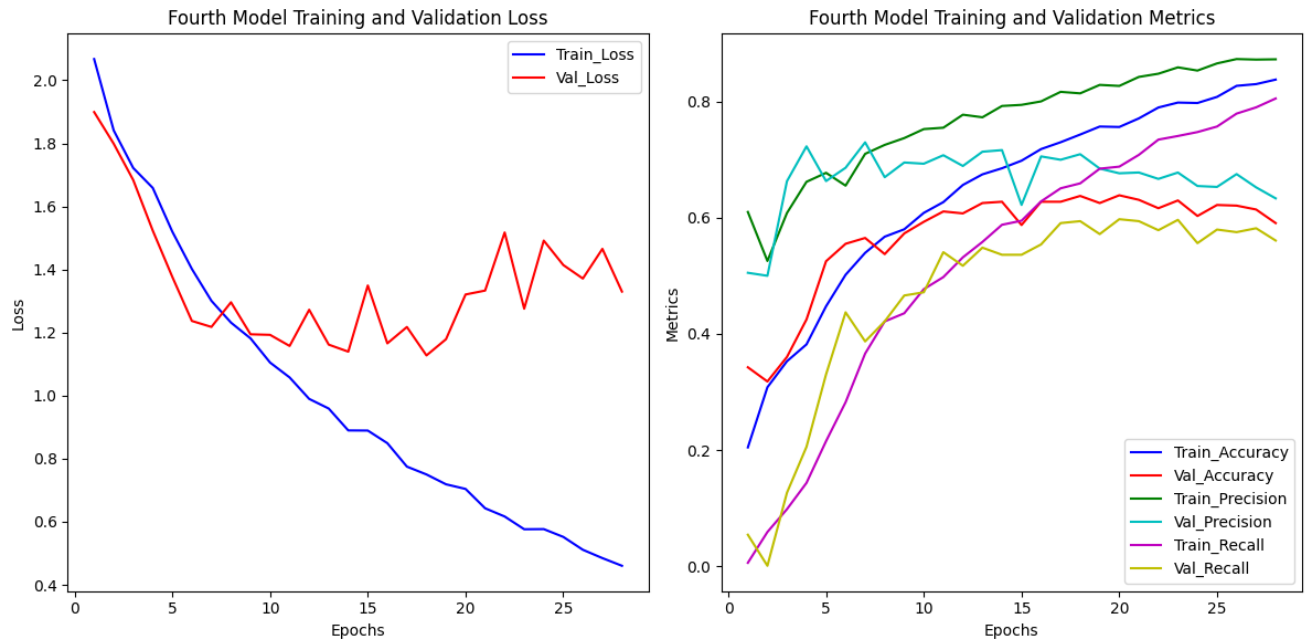


### **5.2.3 Model 4 - Tuned Pre-trained Model (VGG16 with Dropout and Diverse Learning Rates):**

Integrating dropout regularization introduced complexities, leading to a decrease in performance. Training precision reached 83.43%, while validation precision was 70.92%. Over-fitting persisted, highlighting the intricate balance between model complexity and generalization. Exploring diverse learning rates underscored our commitment to refining the model's performance.

Summary of the evaluation metrics:

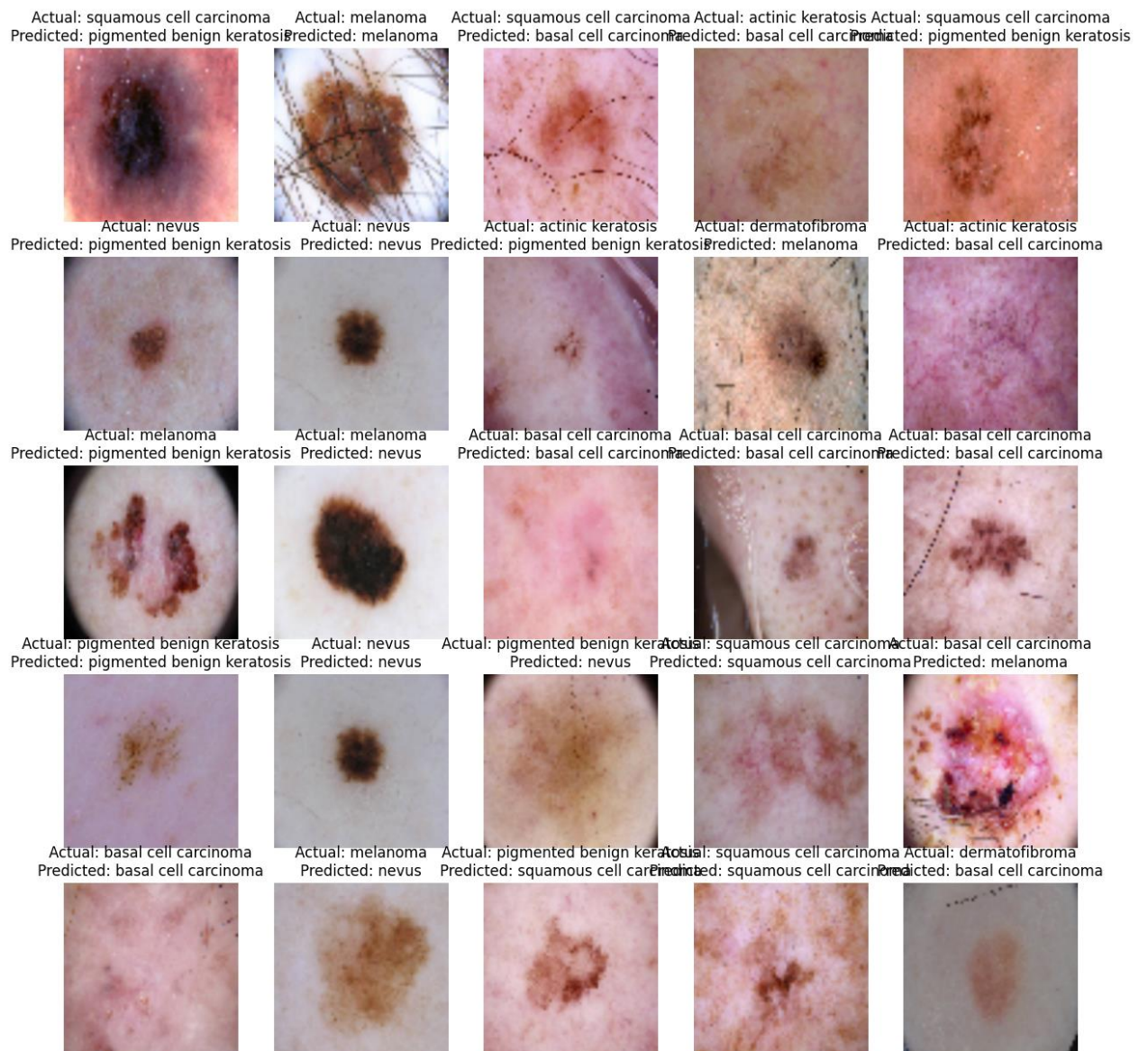
Training Precision	Validation Precision	Training Accuracy	Validation Accuracy	Training Recall	Validation Recall
83.43%	70.92%	83.75%	59.07%	80.50%	56.06%



In each model, we navigated challenges, aiming for a precise, balanced, and effective skin disease classification system. These models signify our continuous efforts to enhance accuracy while addressing the complexities inherent in our dataset.

### **5.3 Best Model and Model Evaluation**

The third model demonstrated outstanding precision scores of 81% in training and 73% in validation, establishing it as our preferred choice for predicting skin disease classes. Its remarkable performance affirmed our decision to proceed with this model for accurate predictions.



The model exhibits superior accuracy in predicting Nevus and basal cell carcinoma compared to other diseases. To further enhance its performance across all classes, additional training images for other diseases are crucial. Increasing the dataset's diversity will enable the model to generalize better, ensuring accurate predictions for a broader range of skin diseases.

## **6.0 Model Deployment:**



Model deployment was successfully executed via Streamlit which offers a user-friendly interface, empowering dermatologists with seamless access to accurate skin disease predictions.

## **7.0 Conclusions and Recommendations**

### **7.1 Data Limitations:**

#### **1. Limited Dataset Size.**

The challenge of a limited dataset size underscores the importance of striking a delicate balance between model complexity and data sufficiency. To mitigate the risk of over-fitting, efforts extend beyond augmentation. Rigorous validation techniques become essential, ensuring the model's accuracy and stability even within the constraints of the available data.

#### **2. Limited Diversity in Skin Tones.**

The dataset's lack of diverse skin tones could introduce bias, impacting diagnoses, especially for darker skin tones. Inclusivity is crucial for accurate predictions across diverse populations.

#### **3. Ethical Concerns.**

Ethical considerations about patient privacy and consent are paramount. Adherence to rigorous privacy standards and ethical guidelines is vital to safeguard patient confidentiality.

#### **4. Impact of Technological Advancements.**

Evolving imaging technology introduces variations in image quality over time. Adapting the model to diverse image qualities is necessary for consistent performance.

## **7.2 Conclusions**

A 73% precision underscores the model's potential in revolutionizing dermatological diagnostics and elevating patient care standards.

## **7.3 Recommendations**

To optimize our model's performance, we propose:

### **1. Focused Application:**

Utilize the model's high precision in predicting Basal cell carcinoma and Nevus for targeted early detection efforts.

### **2. Enhance Dataset:**

Increase dataset diversity, especially in underrepresented classes, ensuring a more robust and unbiased model.

### **3. Diversify Sources:**

Integrate data from varied regions to account for environmental factors, enriching the model's adaptability.

### **4. Capture Variability:**

Incorporate diverse age groups and data that spans different time periods to accommodate evolving skin conditions, enhancing diagnostic accuracy.

Implementing these steps will refine our model, ensuring accurate and reliable diagnoses across different skin conditions and demographics.

## **Future Endeavors:**

### **1. Advanced Architectures and Transfer Learning.**

Investigate intricate architectures and employ advanced transfer learning techniques to enhance model efficacy and accuracy.

### **2. Iterative Hyper-parameter Tuning.**

Continuously refine hyper-parameters and regularization strategies iteratively, ensuring the model maintains optimal performance over time.

### 3. Dermatologist Collaboration and Integration.

Partner closely with dermatologists to seamlessly integrate the model into clinical workflows. This collaboration ensures practical applicability, real-world effectiveness, and patient well-being.