Skin Cancer Detection

Cher Wee Zheng, Ng Wei, Ryan Yong

We Are

You Are

SingHealth

Defining Tomorrow's Medicine

Table of contents

01

Introduction & Problem Statement

Joseph Liu

Joseph Liu

Who is Joseph?

Joseph is a PR construction foremen and father of 2

Joseph Liu

Who is Joseph?

Joseph is a PR construction foremen and father of 2

What are his goals?

Joseph wants to pay off his 30 year hdb loan.

Joseph Liu

Who is Joseph?

Joseph is a PR construction foremen and father of 2

What are his goals?

Joseph wants to pay off his 30 year hdb loan.

What are his concerns?

Joseph noticed a large black spot on his forearm

Joseph Liu

Who is Joseph?

Joseph is a PR construction foremen and father of 2

What are his goals?

Joseph wants to pay off his 30 year hdb loan.

What are his concerns?

Joseph noticed a large black spot on his forearm

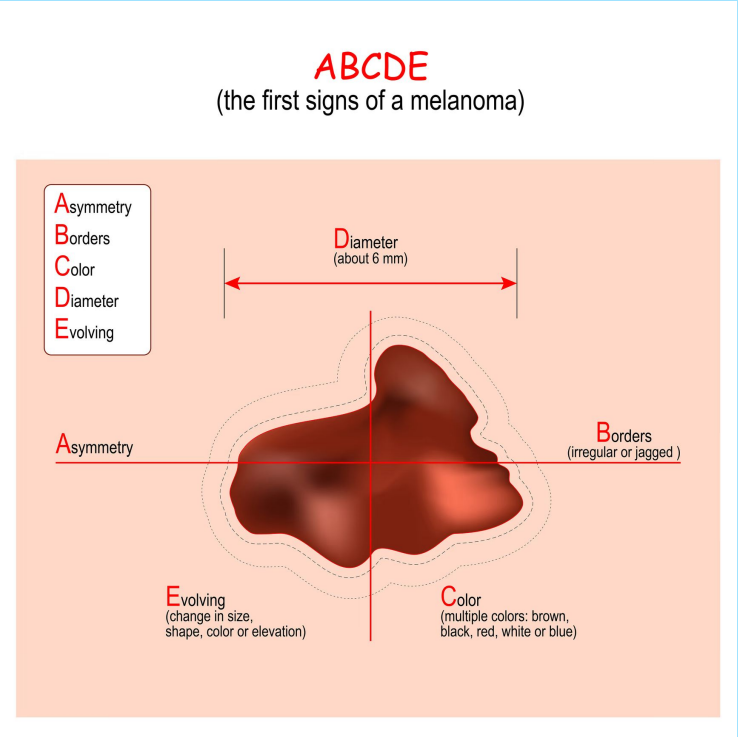
What's the issue?

Joseph does not want to spend unnecessary money

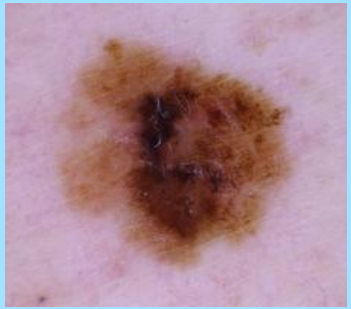
What should Joseph do?

Skin Cancer

Source: CNA



Cancerous

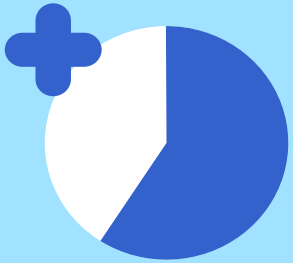


Cancerous

Non-cancerous

Types of Examination

Visual

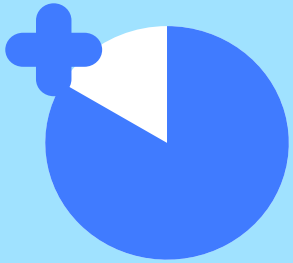


60%

Instant

$35 - $200

Dermoscopy

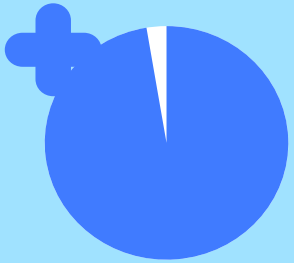


85%

Instant

From $250

Biopsy



95%

3 weeks

$500 - $3000

Source: nsc, CNA & MSIG

Problem Statement

Can we develop a model to be able to **effectively** and **accurately** distinguish between a benign or malignant skin cancer spot visually, improving visual tests and **reducing the necessity** for more **invasive** biopsies?

Key Objectives

01

Maximise

Accurate diagnosis/treatment at polyclinic level

02

Reduce

Unnecessary specialist referrals

03

Reduce

Costs

Modelling Process

Exploration	Modelling	Tuning	Implementation
An overview of the dataset gathered and how the data is preprocessed	Using our dataset, develop an effective and accurate model	Improve our chosen model by altering its hyperparameters	We work together with you to upload our app in your systems

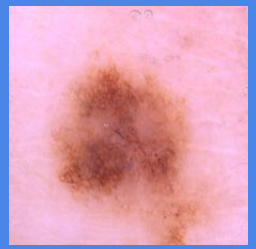
02

EDA & Preprocessing

Modelling Process

Training Data

Malignant



Benign

Image Processing

Standardization, Augmentation

Model

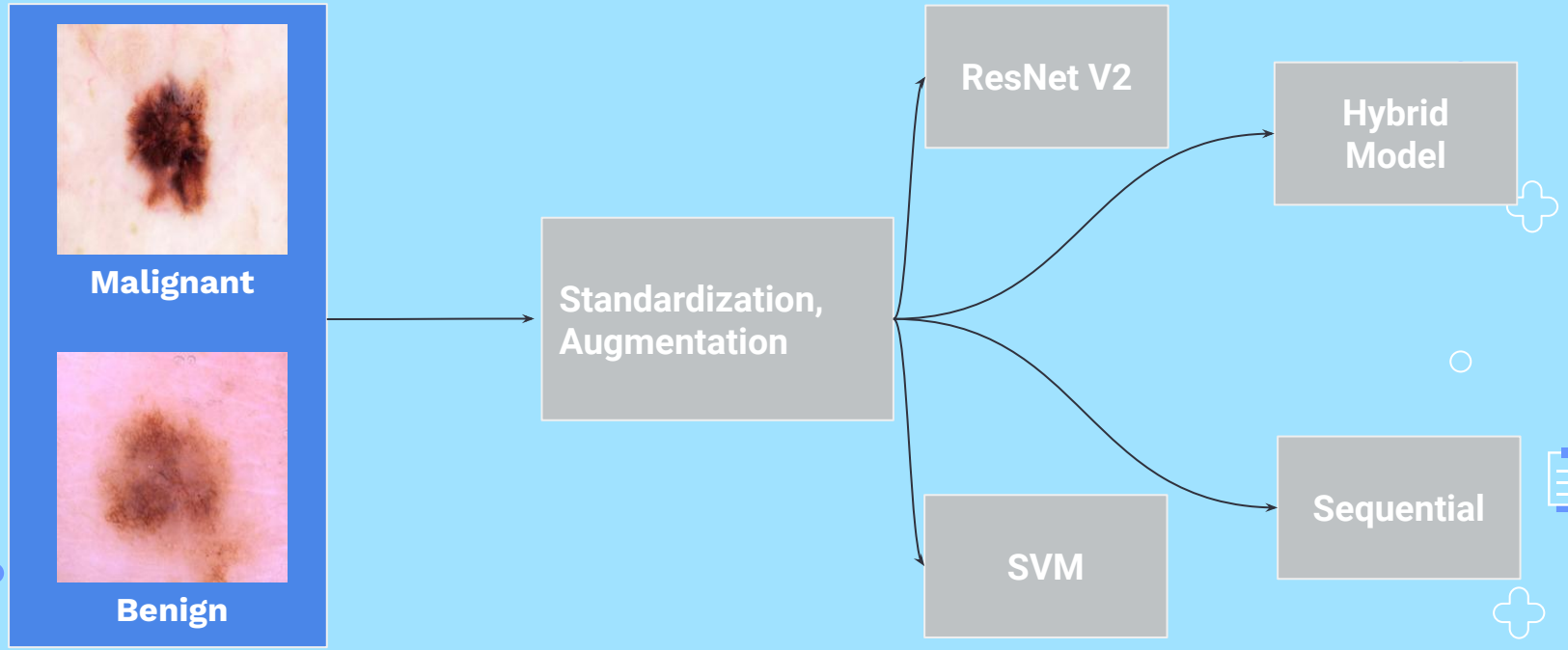
ResNet V2

Hybrid Model

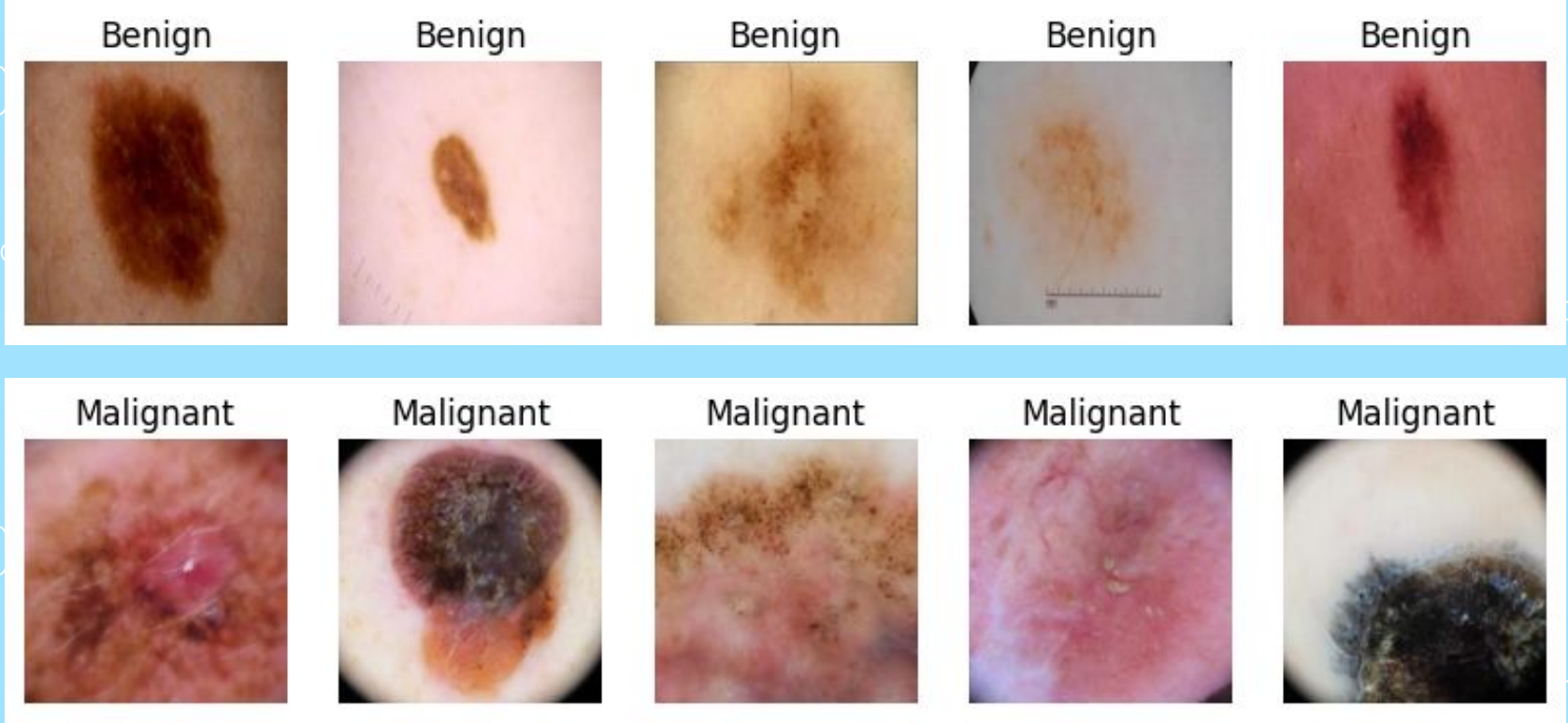
SVM

Sequential

Modelling Process



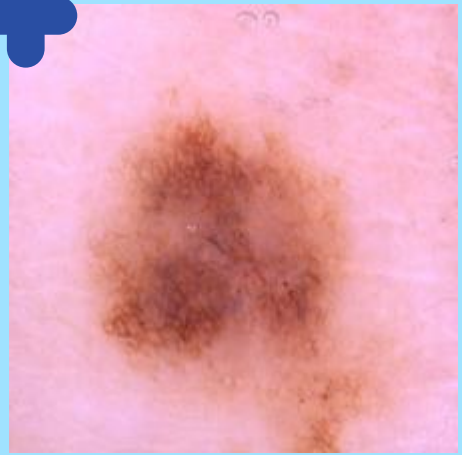
Our Data are all images!



Our Data

Malignant

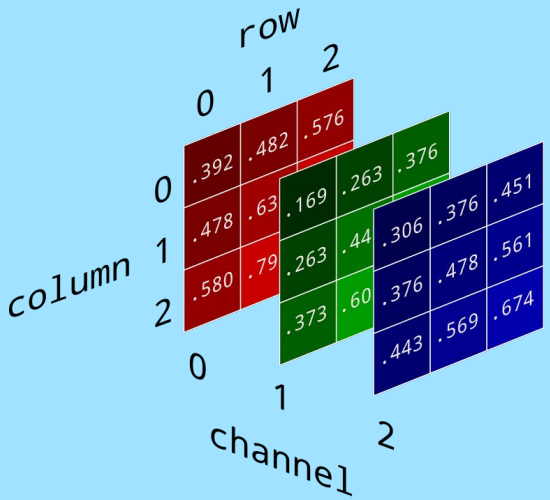
5590
Training Data
1000
Testing Data



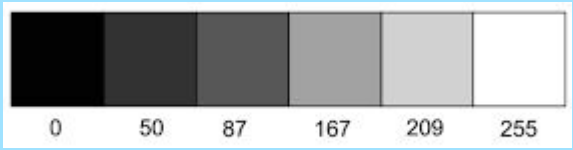
Benign

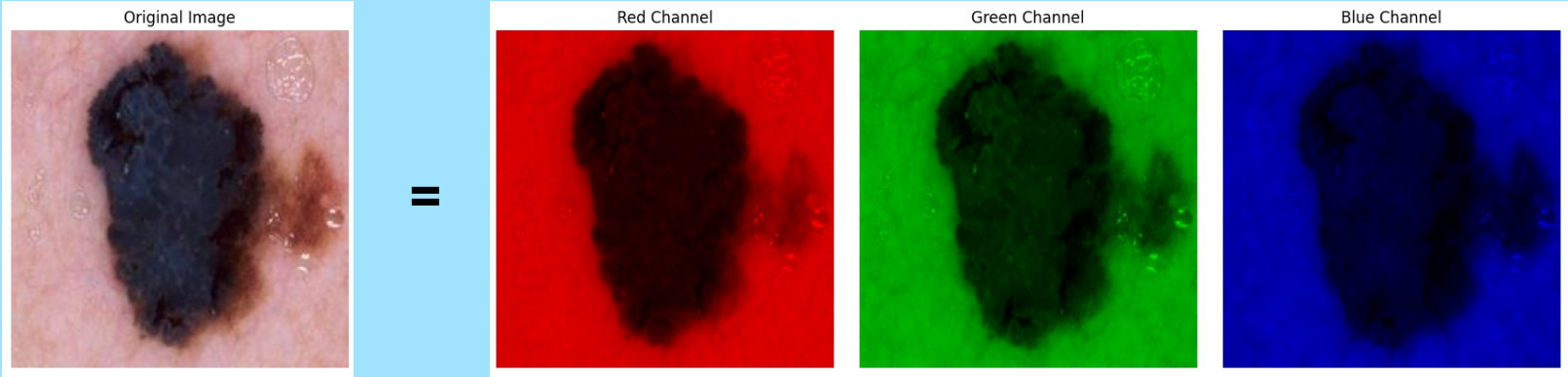
6289
Training Data
1000
Testing Data

How does models see images?

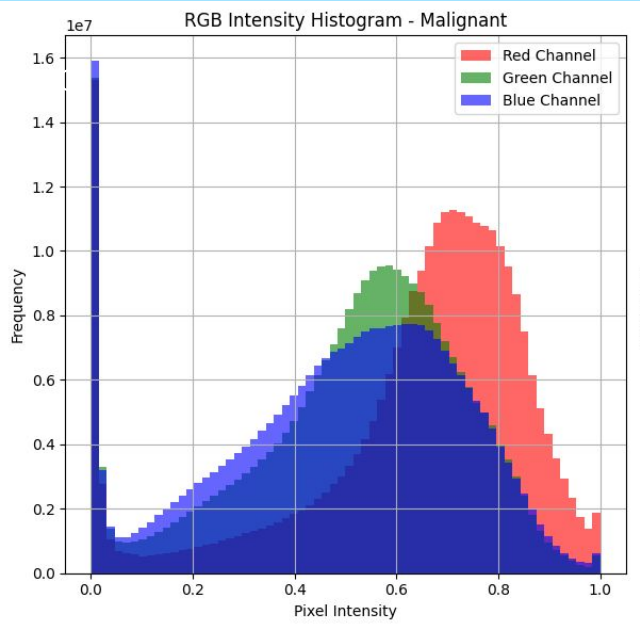


Images are **composed of pixels**
Each pixel contains **Red, Green and Blue** information.
Normally **0-1 or 0-255** to representing each **color's intensity**.





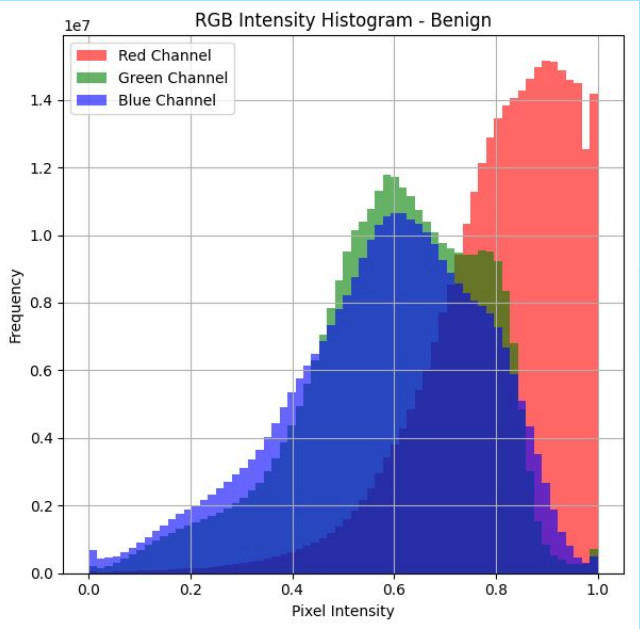
Histogram – Malignant vs Benign



0.63

0.51

0.49

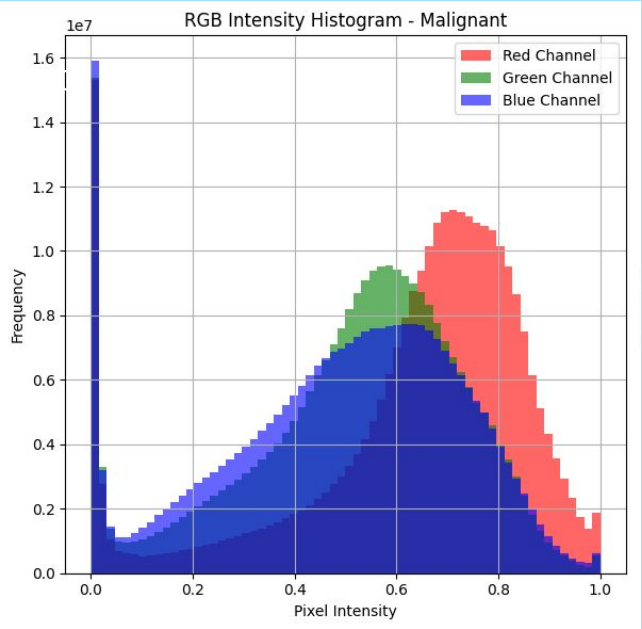


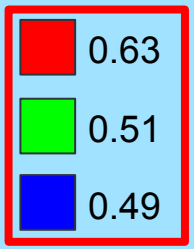
0.806

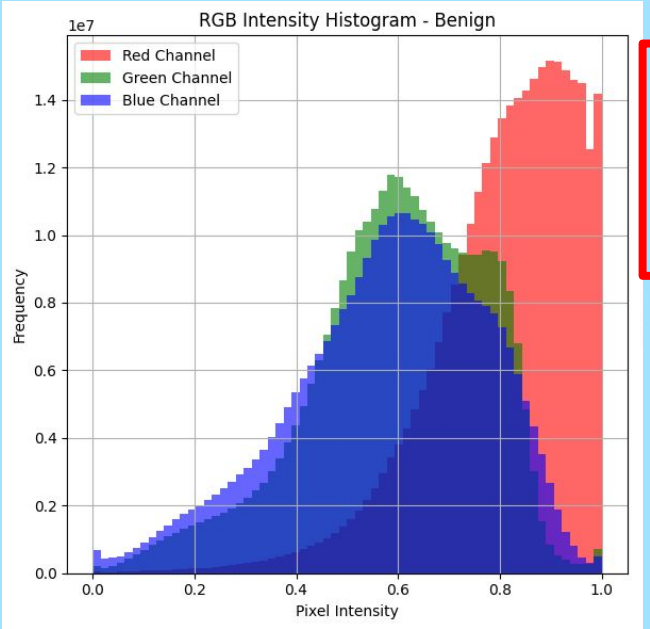
0.596

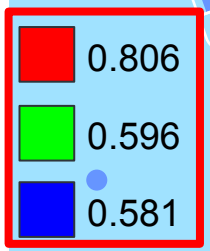
0.581

Histogram – Malignant vs Benign



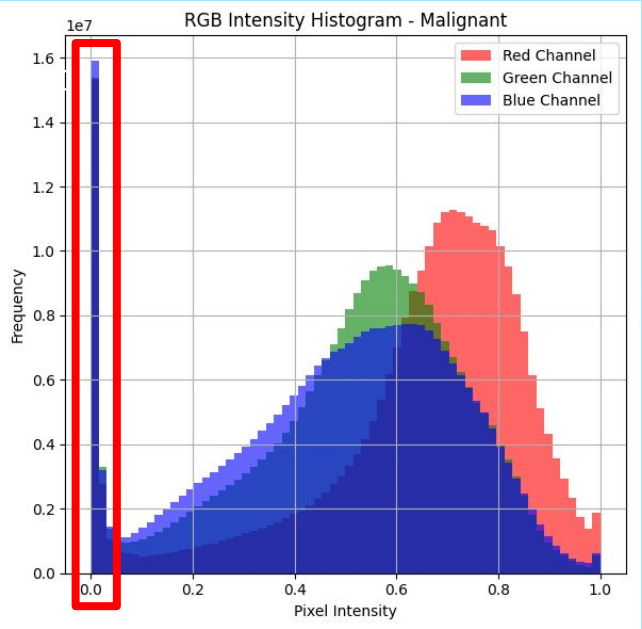






Finding 1:
Based on the average of RGB value, benign seem like brighter than malignant

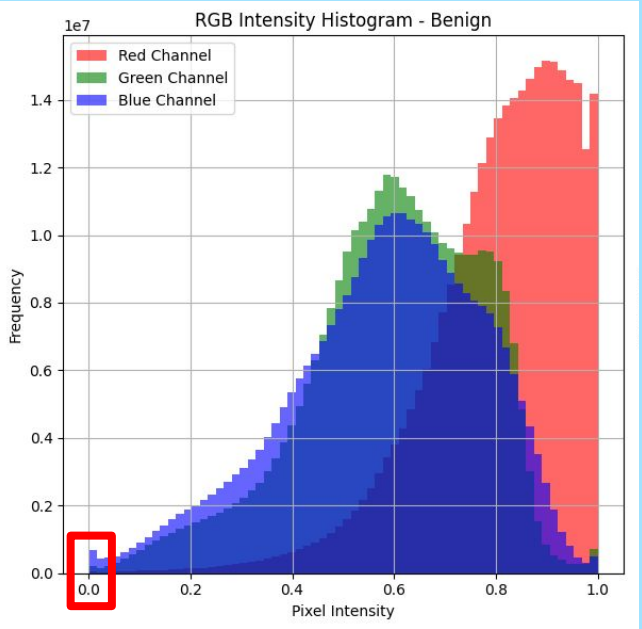
Histogram – Malignant vs Benign



0.63

0.51

0.49



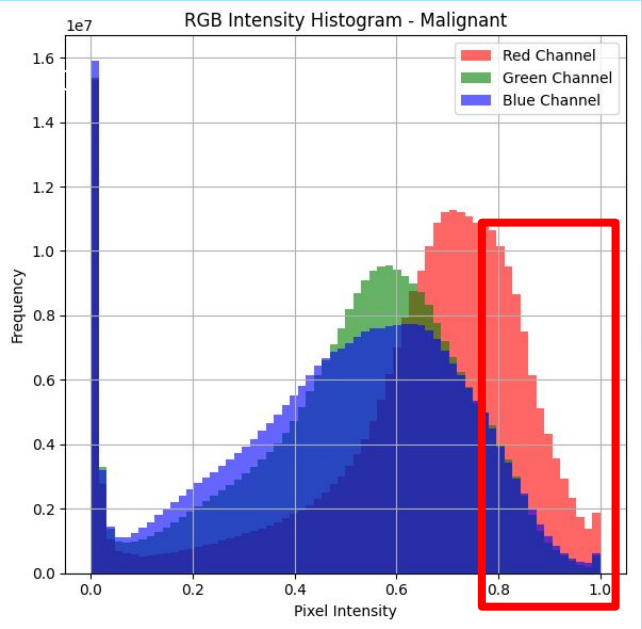
0.806

0.596

0.581

Finding 2:
Malignant image has many black pixels (R,G,B = 0) compare to Benign

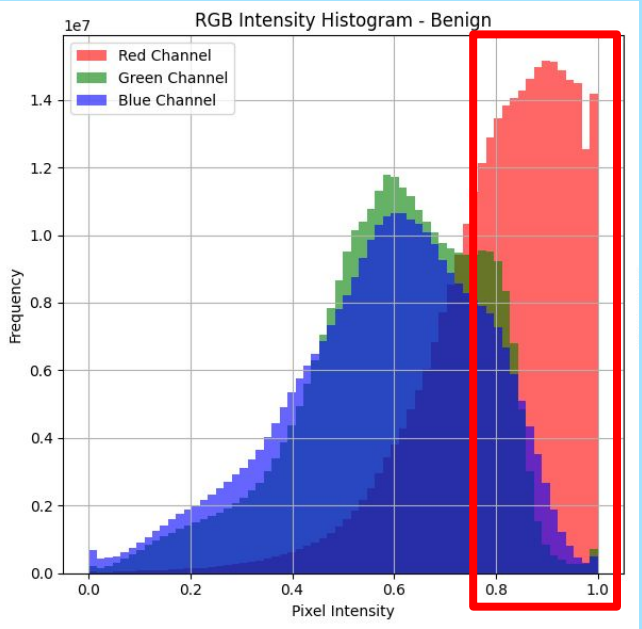
Histogram – Malignant vs Benign



0.63

0.51

0.49



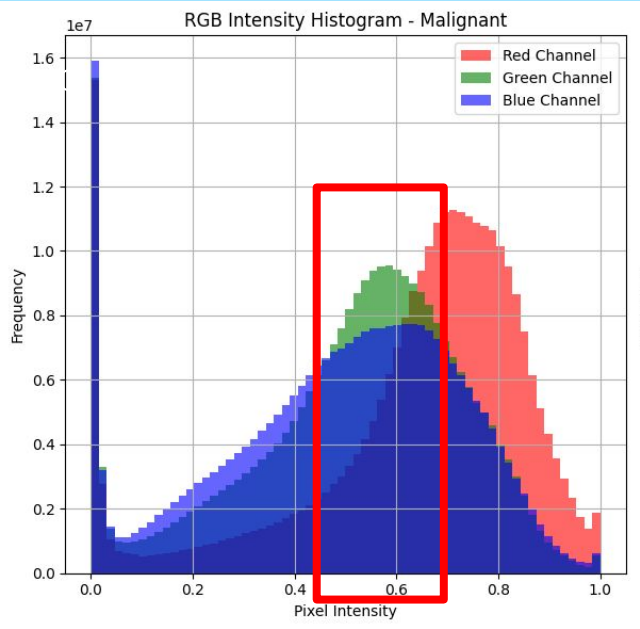
0.806

0.596

0.581

Finding 3:
Benign image is more “red” than Malignant, most red pixel > 0.75

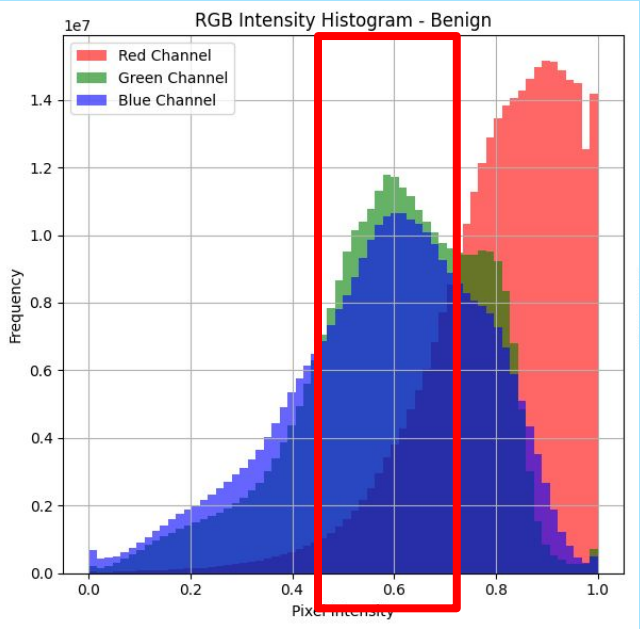
Histogram – Malignant vs Benign



0.63

0.51

0.49



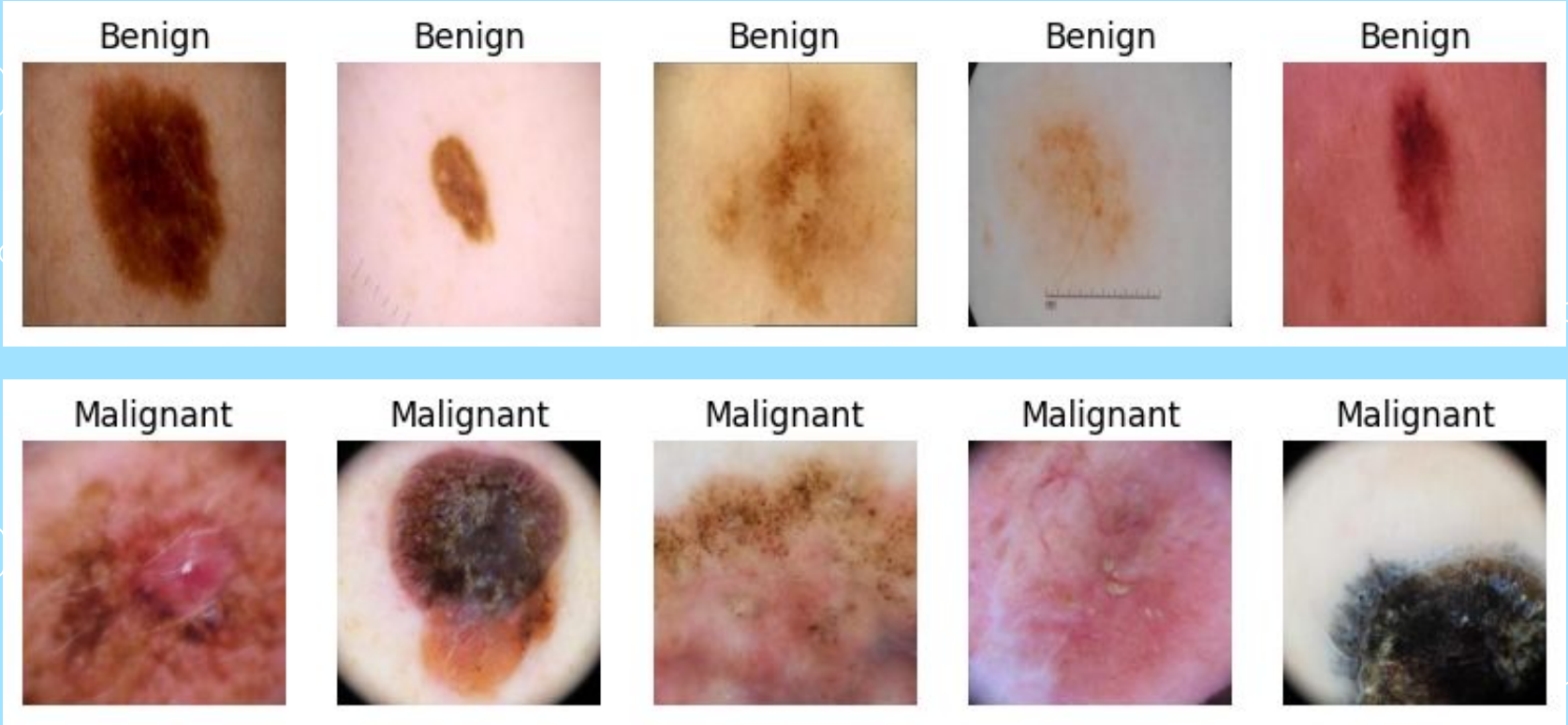
0.806

0.596

0.581

Finding 4:
Beside the black pixel from malignant, both Green and Blue mostly distributed in between 0.45 - 0.7 intensity

What we expected to see.

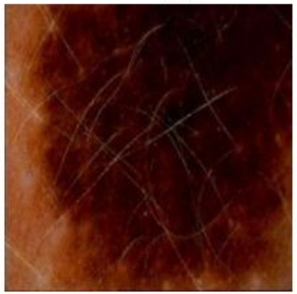


The Outliers!

Benign Outlier (RGB < 0.35)

4627.jpg

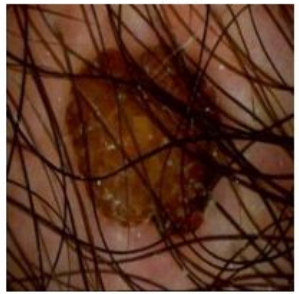
5211.jpg



5160.jpg

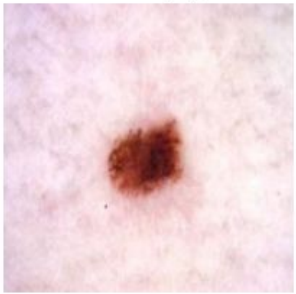
3573.jpg

2089.jpg



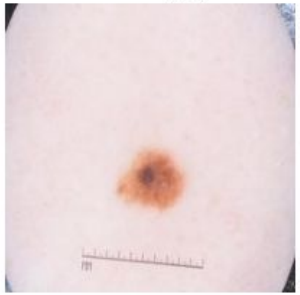
Malignant Outlier (Red > 0.75, Green and Blue > 0.45)

2474.jpg

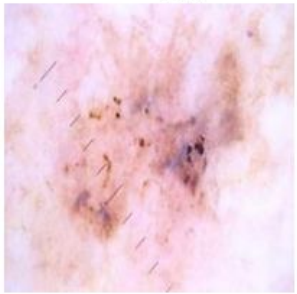


5119.jpg

1028.jpg



5291.jpg



52.jpg

Modelling Process

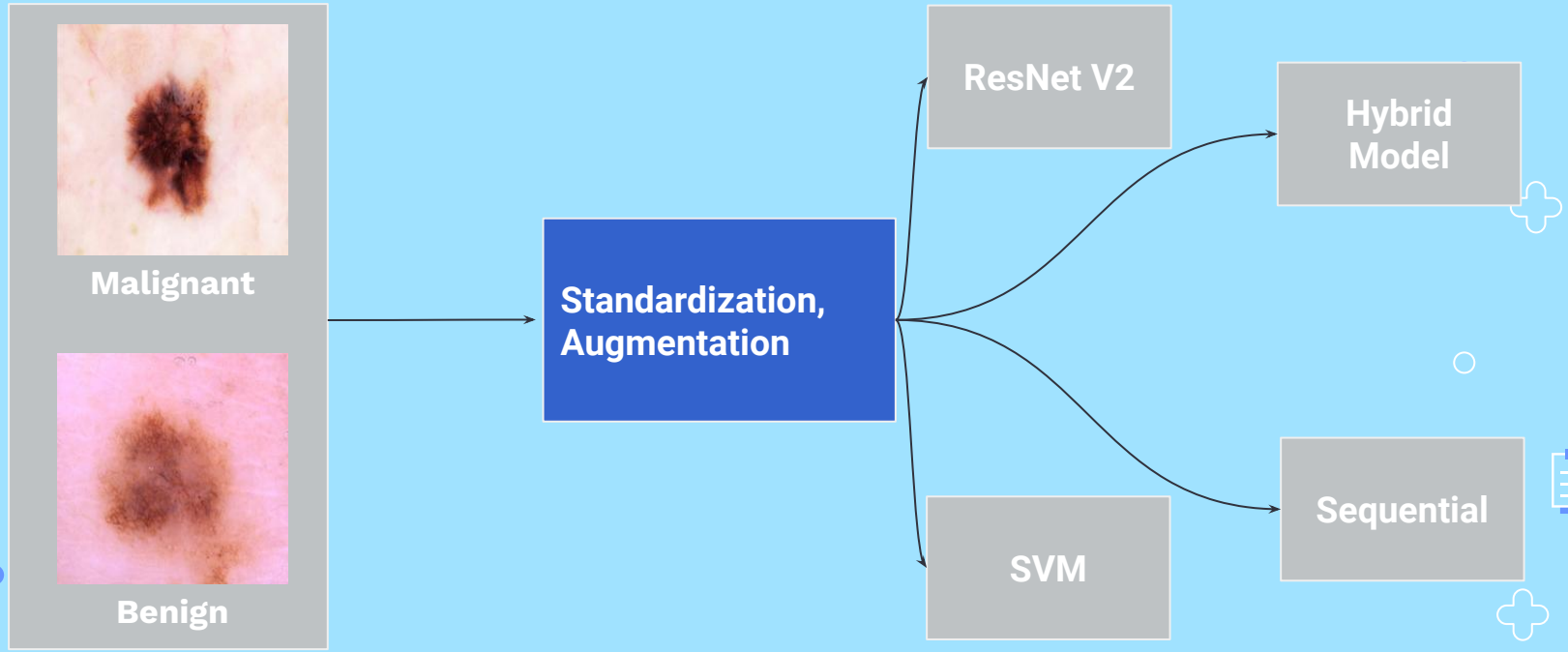
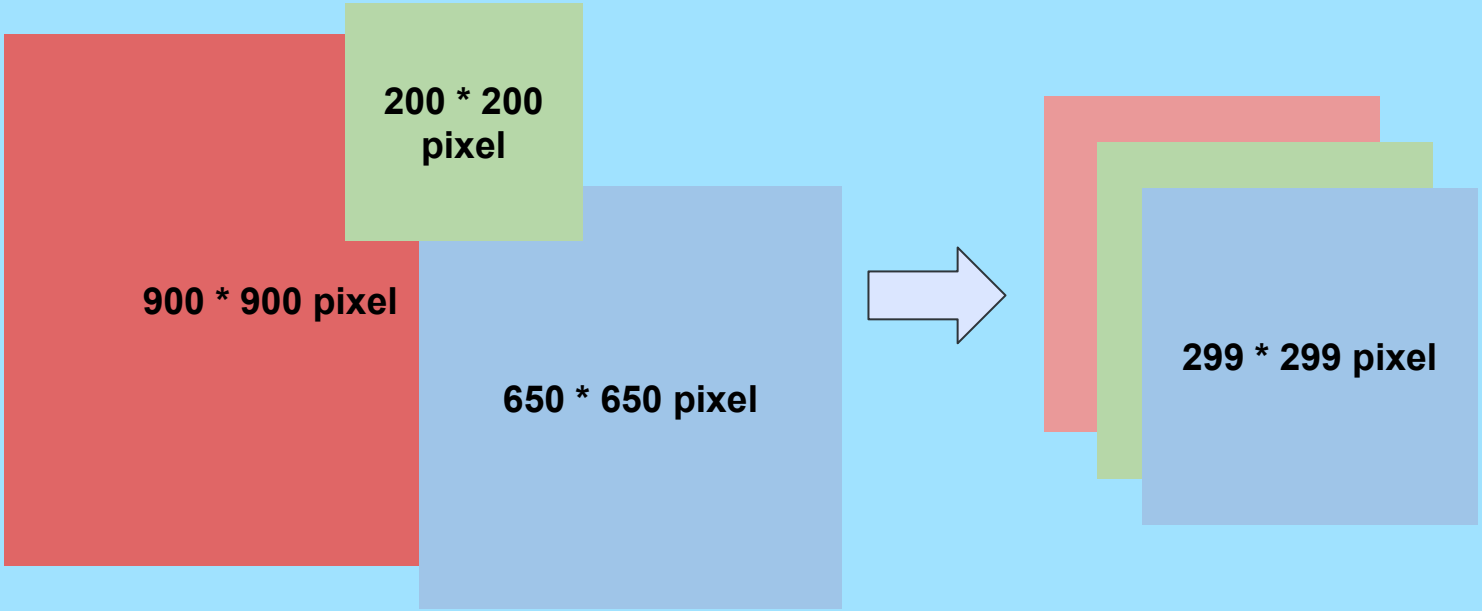


Image Preprocessing

- **Standardization - Resize image to 299*299 pixel**



In order to obtain discriminative internal features that can lead to excellent performance, we adopt an image preprocessing method to **reduce variance of images** and **enhance inherent characteristics of images**.

Image Preprocessing

- **Standardization - Scale the RGB from 0-255 to 0-1**

Original Pixel Information

0	16	32
64	4	8
128	255	2

Normalizing the Pixel Info

0/255	16/255	32/255
64/255	4/255	8/255
128/255	255/255	2/255

After Normalized Pixel Info

0.00	0.06	0.13
0.25	0.02	0.03
0.50	1.00	0.01

Instead of using StandardScaler, we can just simply divide the original intensity value with 255 to normalize the value.

Image Preprocessing

- **Augmentation - Generate new transformed images**

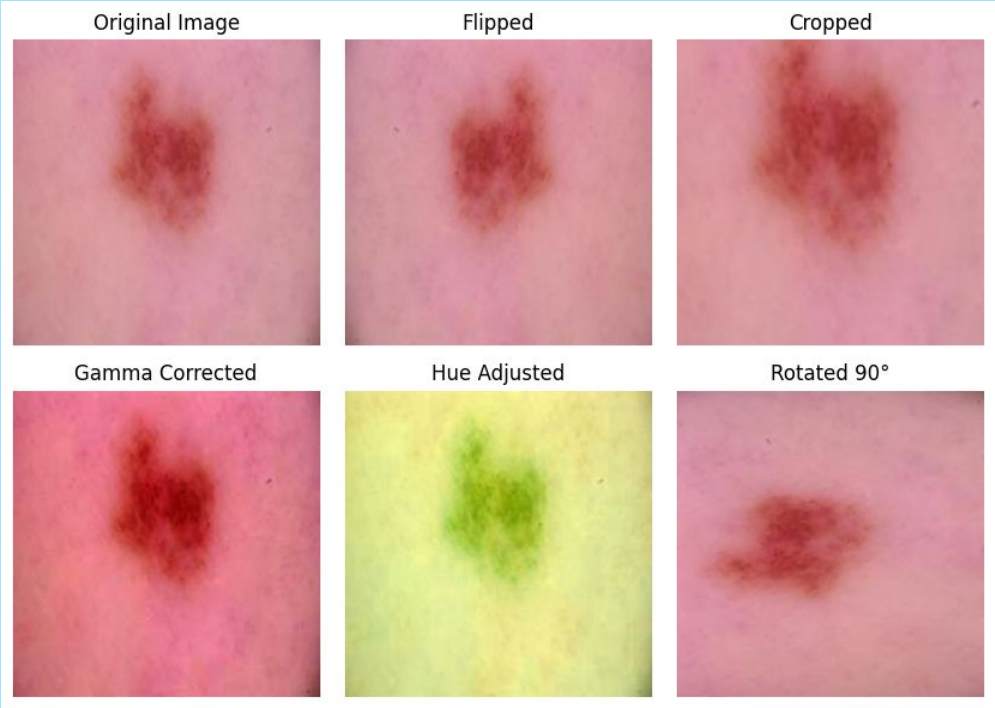
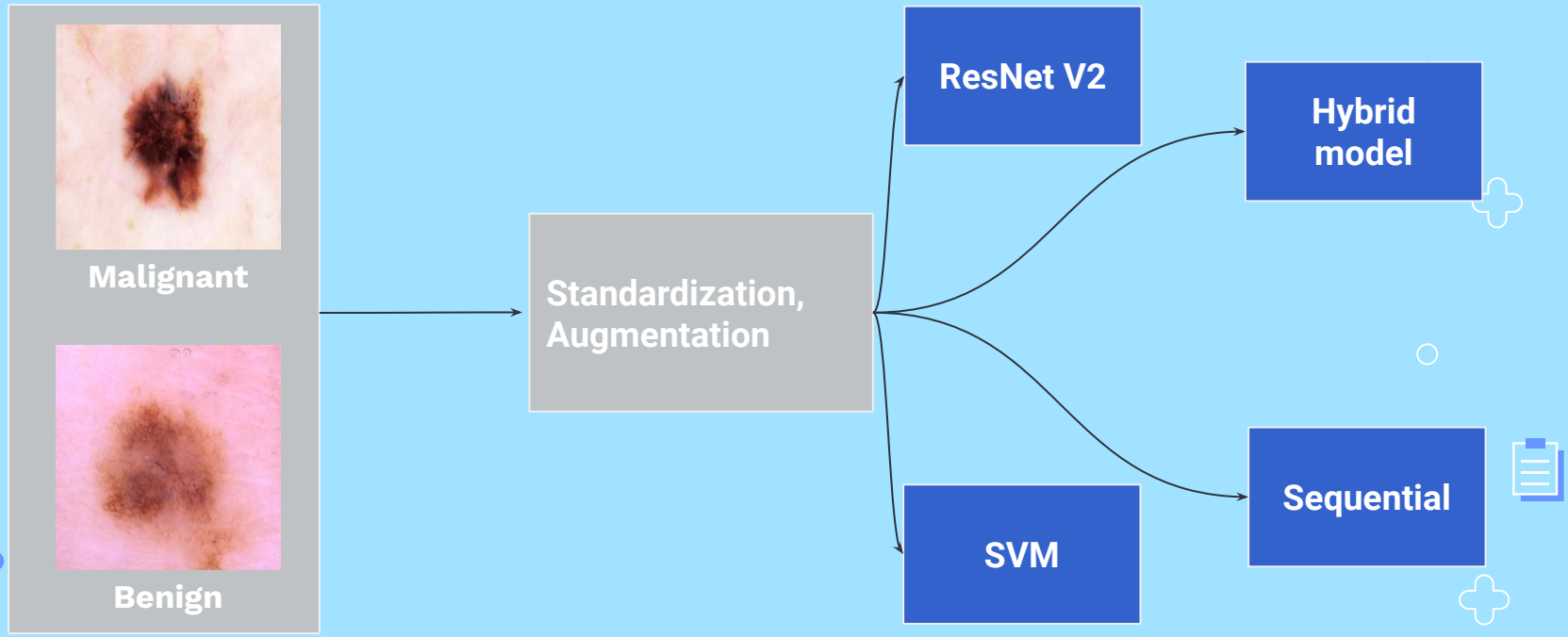


Image augmentation supplements the creation of image variations that can **help models improve its prediction accuracy** and **enhance the generalization ability of models.**

03

Model

Modelling Process

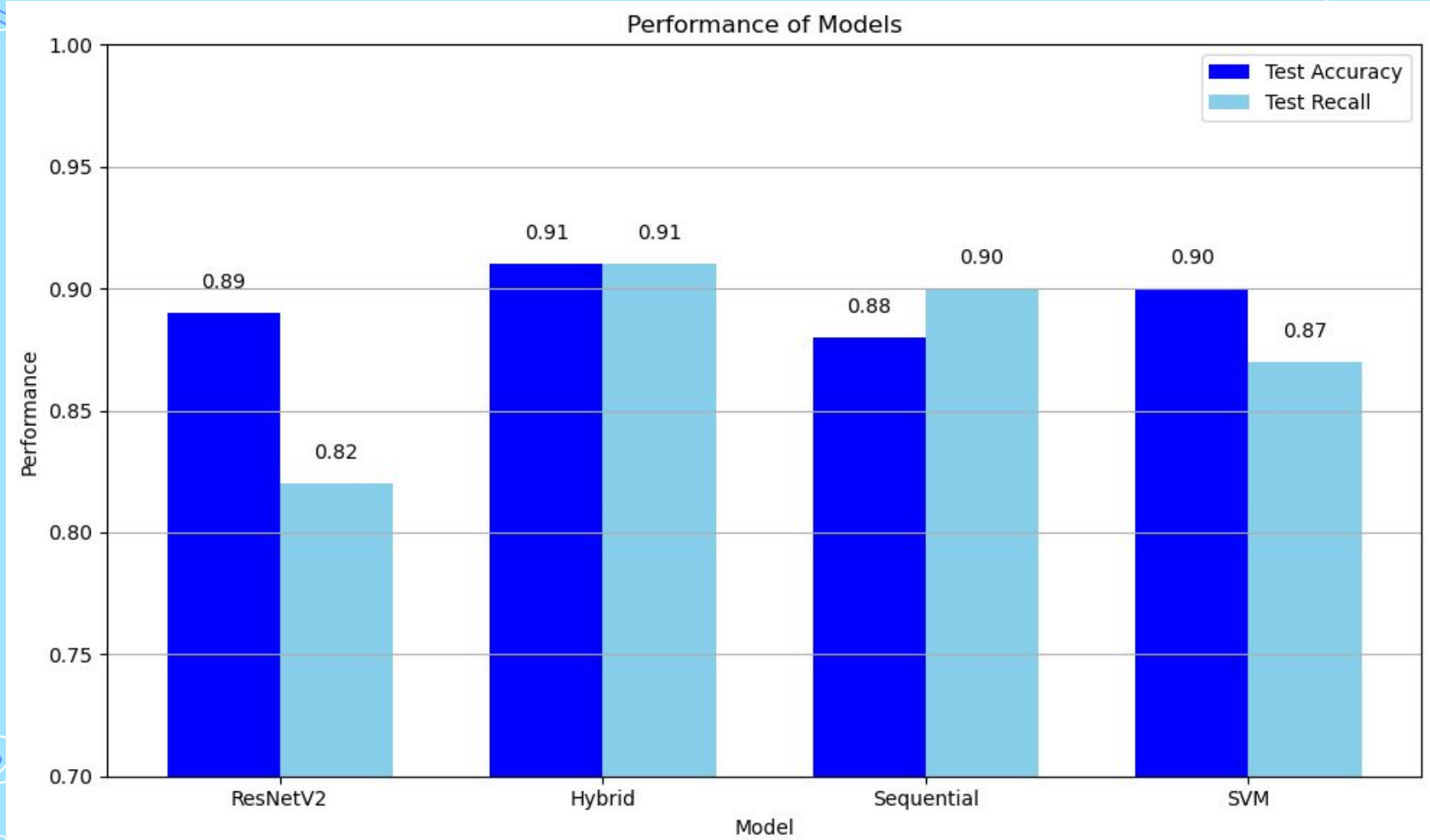


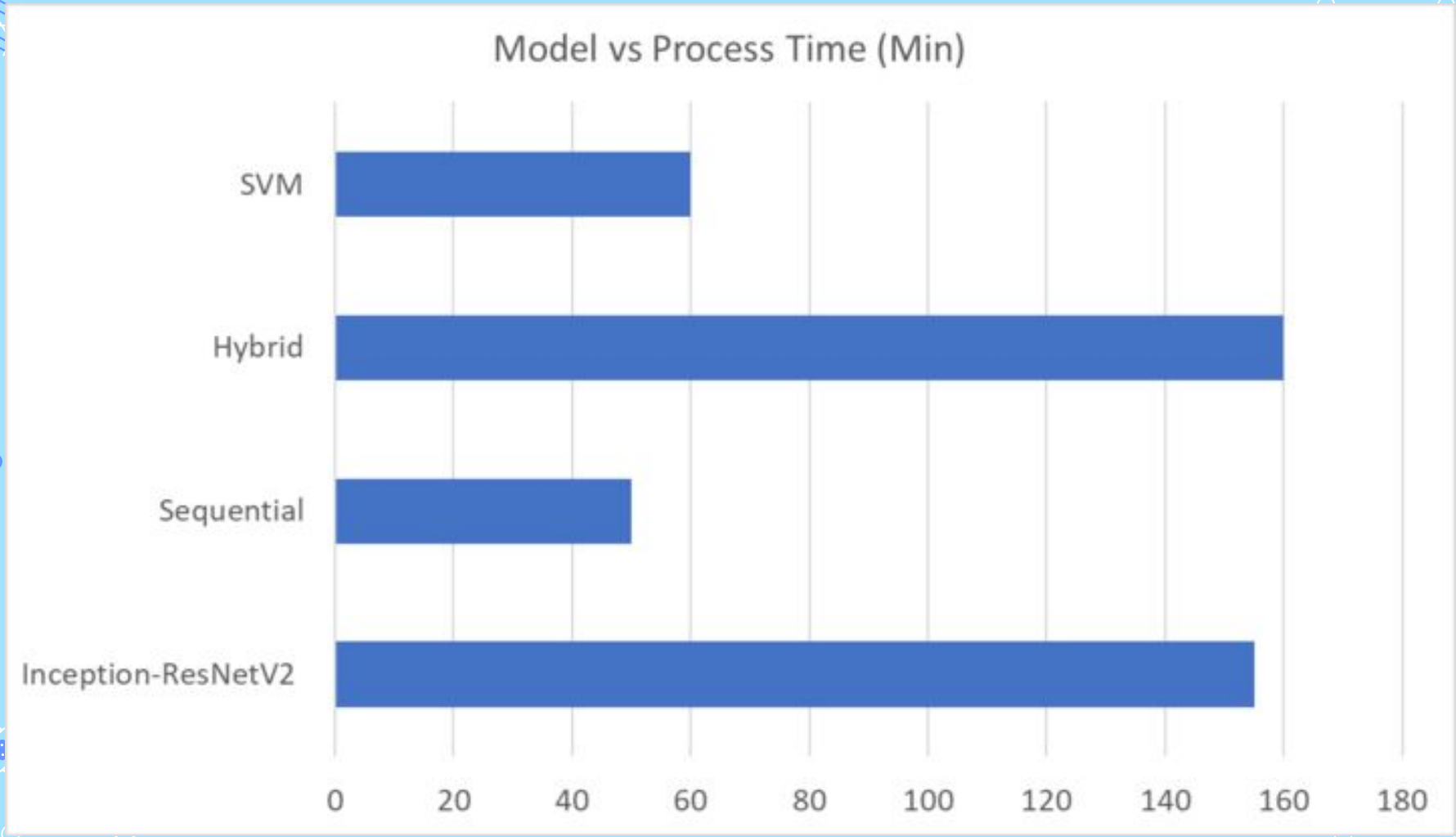
ResNetV2

Hybrid Model

Sequential

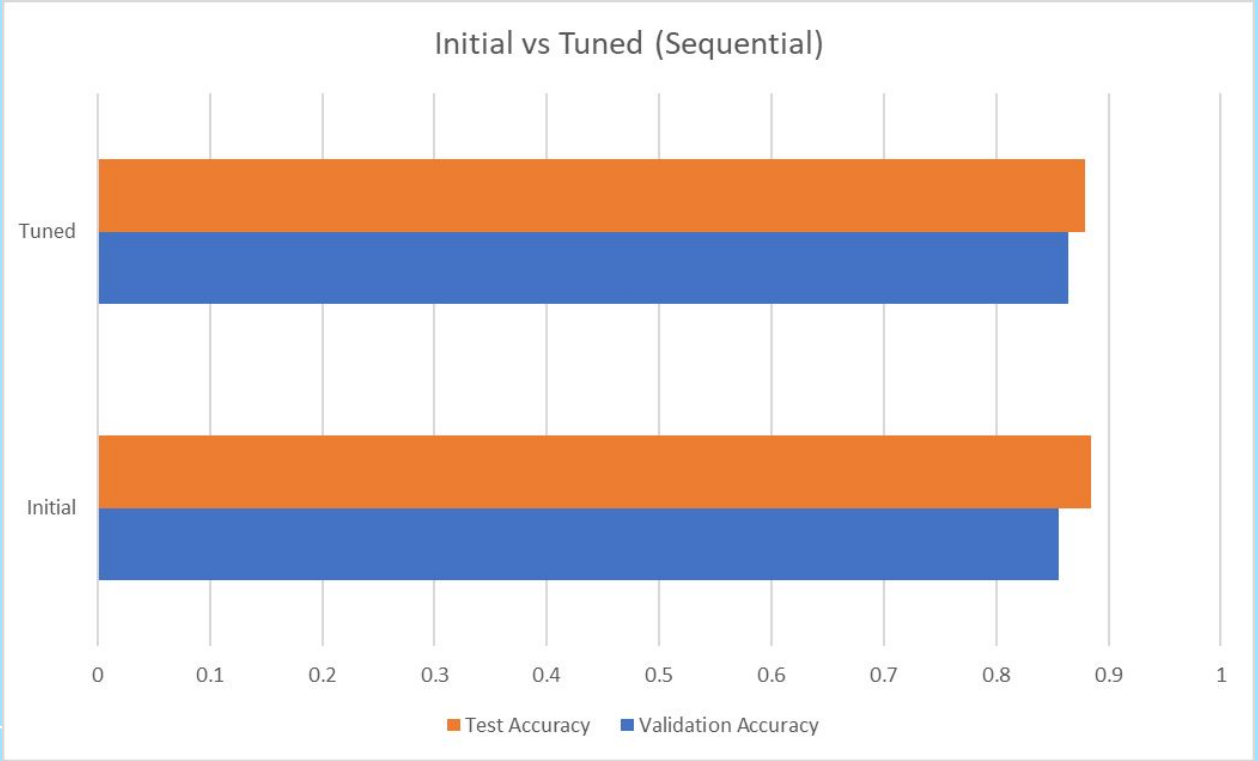
SVM


Performance of Models
Test Accuracy
Test Recall
1.00
0.95
0.90
0.85
0.80
0.75
0.70
Performance
0.89
0.82
0.91
0.91
0.88
0.90
0.90
0.87
ResNetV2
Hybrid
Sequential
SVM
Model


Model vs Process Time (Min)
SVM
Hybrid
Sequential
Inception-ResNetV2
0
20
40
60
80
100
120
140
160
180

3b

Model Tuning



Learning rate:
0.0001, 0.001, 0.01

Number of epochs:
5, 10, 20

RandomSearchCV:
6 iterations

04

App Demo

Hi Doctor King, I noticed these spots on my arm which weren't there before. Should I be worried?

Please upload to the app pictures of your spots, I'll have a look.

<image1.jpeg>
<image2.jpeg>
<image3.jpeg>

App Demo

Hi Joseph, I am going to refer you to a specialist for further tests.

Oh I see, I didn’t want to visit a specialist since I wasn’t sure if it was serious, but since you recommended, I guess I will do it now. Thank you!

You’re welcome Joseph! Glad I was able to help.

05

Conclusion

Cost Benefit Analysis

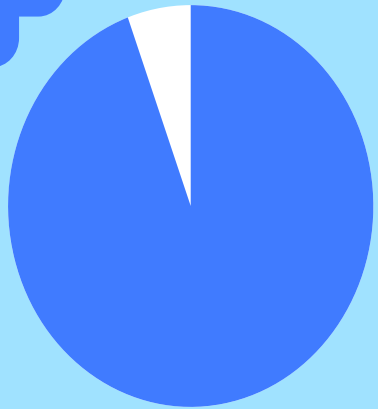
	Visual	Image Detection	Biopsy
Accuracy	60%	89%	95%
Cost	$100-$200	$20-$30	$500 - $3000
Relative Survival Rate*	~95%	~98.7%	~99.5%

SEER stage	5-year relative survival rate
Localized	>99%
Regional	74%
Distant	35%
All SEER stages combined	94%

Img: Survival Rates for Melanoma Skin Cancer

Conclusion

Skin Cancer Detector



89%

- Instant results
- Minimal costs
- Online accessibility

Thank You!!

Calculation of Relative Survival Rate

With an accuracy of 60%, 400 people will be **misdiagnosed**.

January

March

200 people with localised(Stage I & II) melanoma develops into regional(Stage III) or distant (Stage IV).

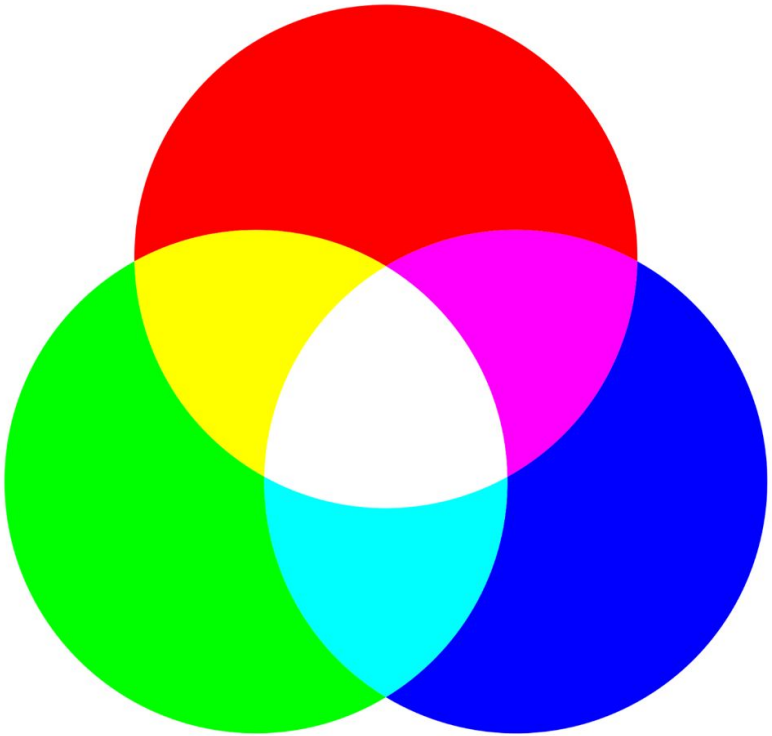
These 200 people who experiences more severe symptoms will be diagnosed with advanced melanoma

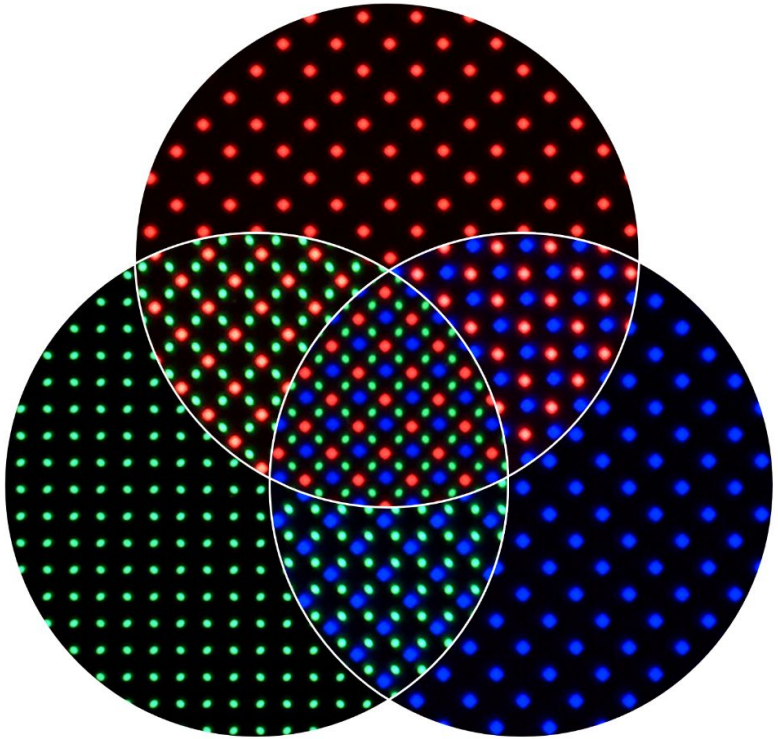
April

2029

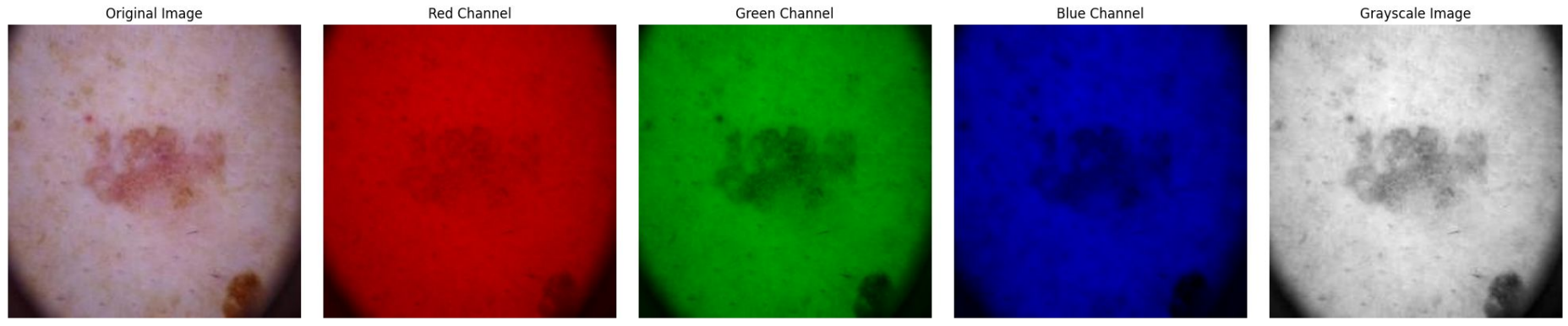
At least 50 of these people will not survive past 2029

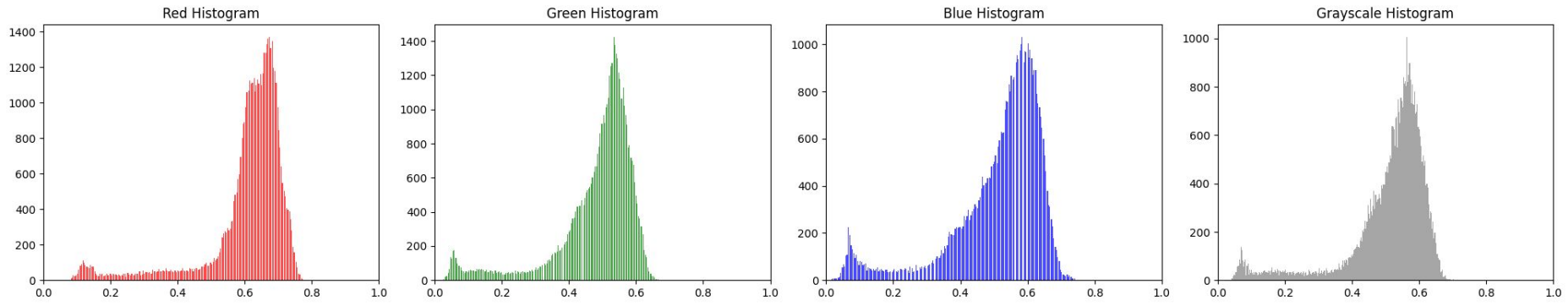
How does machine see images?



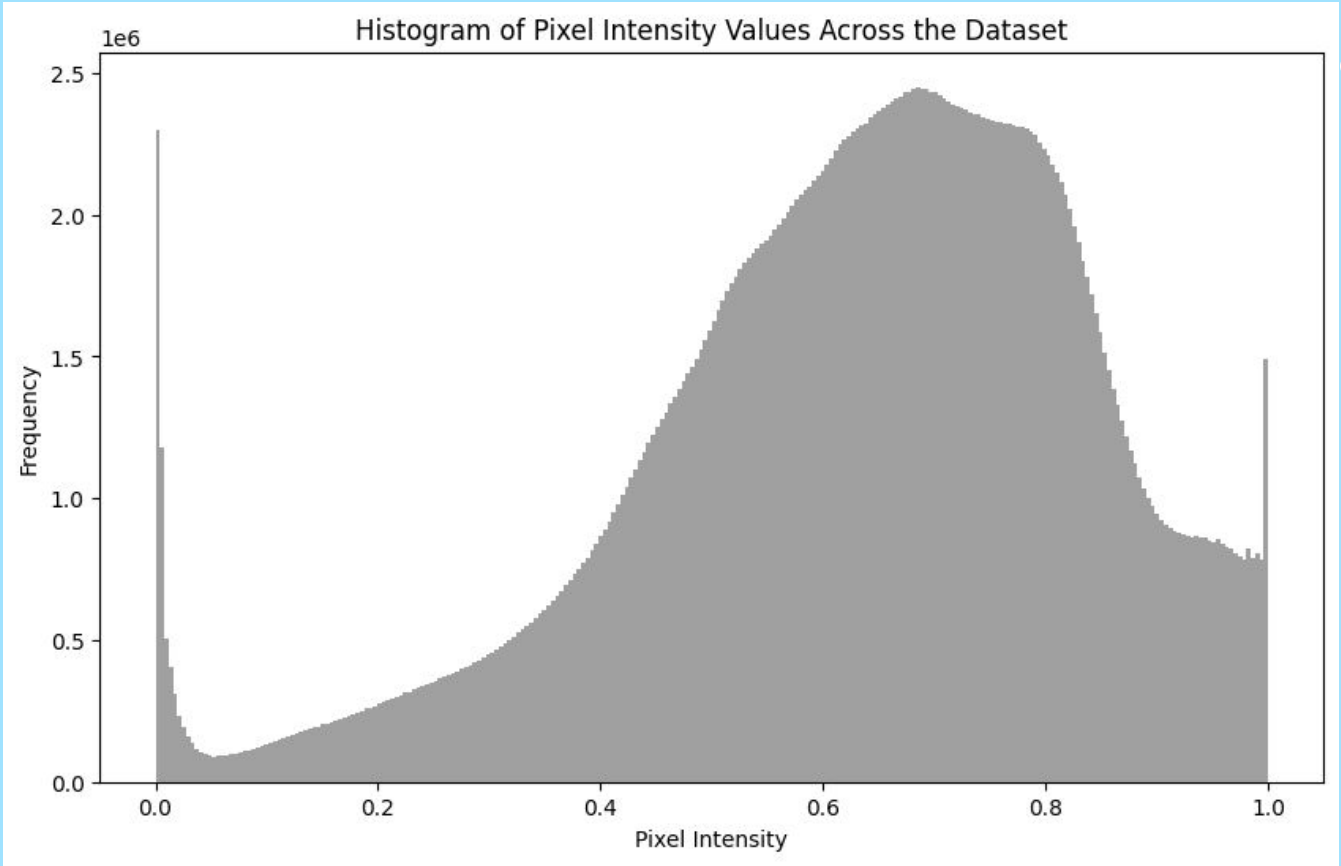


https://www.csfieldguide.org.nz/en/interactives/rgb-mixer/

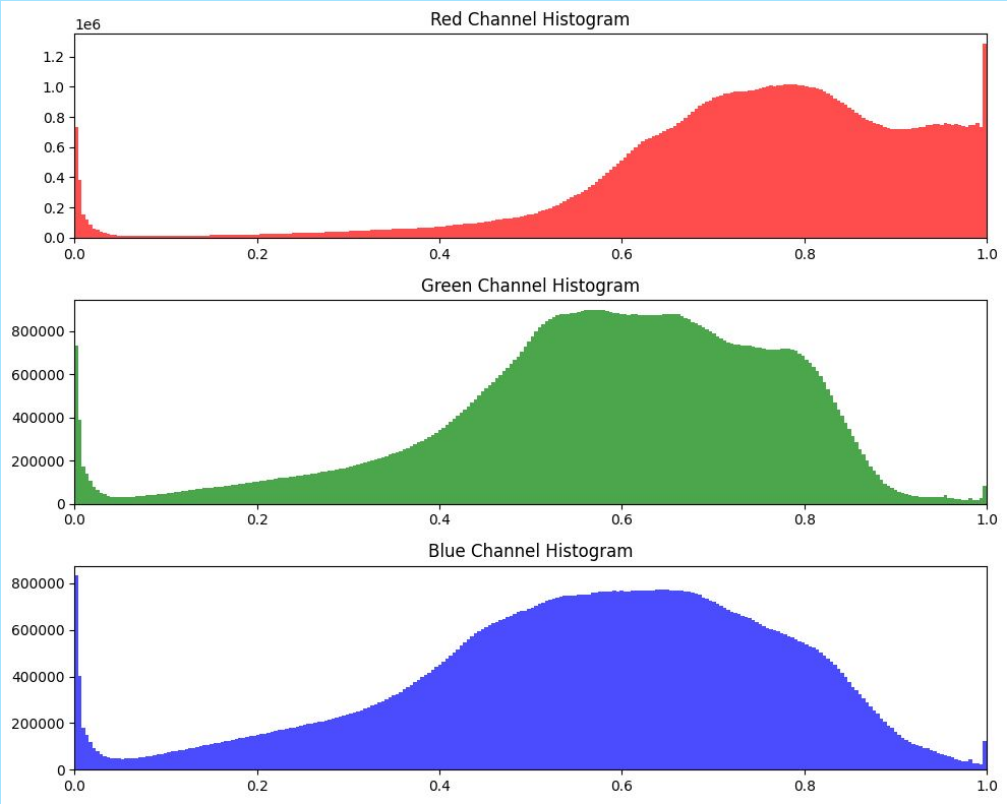

Original Image
Red Channel
Green Channel
Blue Channel
Grayscale Image


Red Histogram
Green Histogram
Blue Histogram
Grayscale Histogram

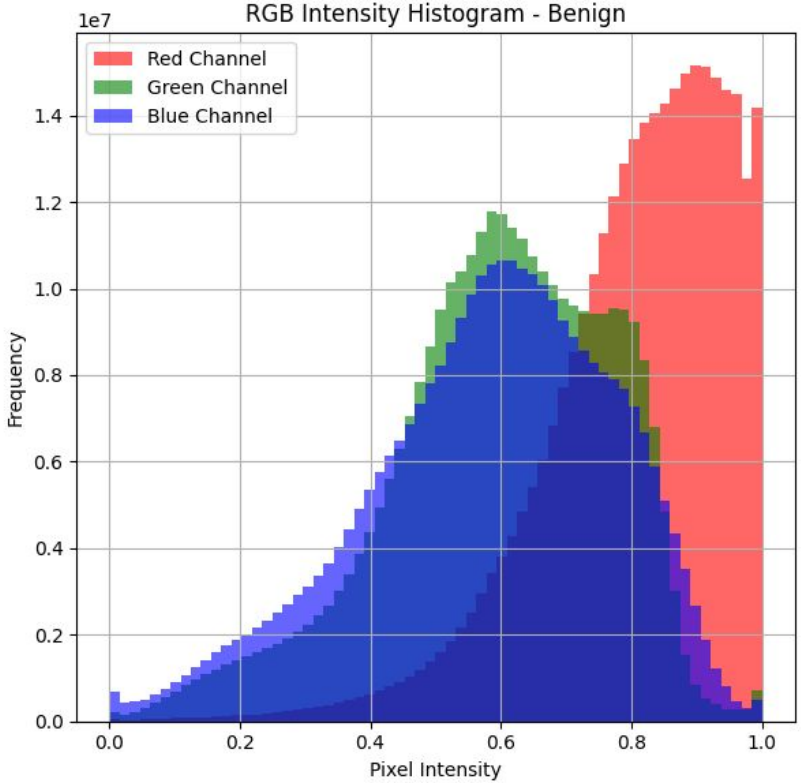
Histogram - Malignant & Benign

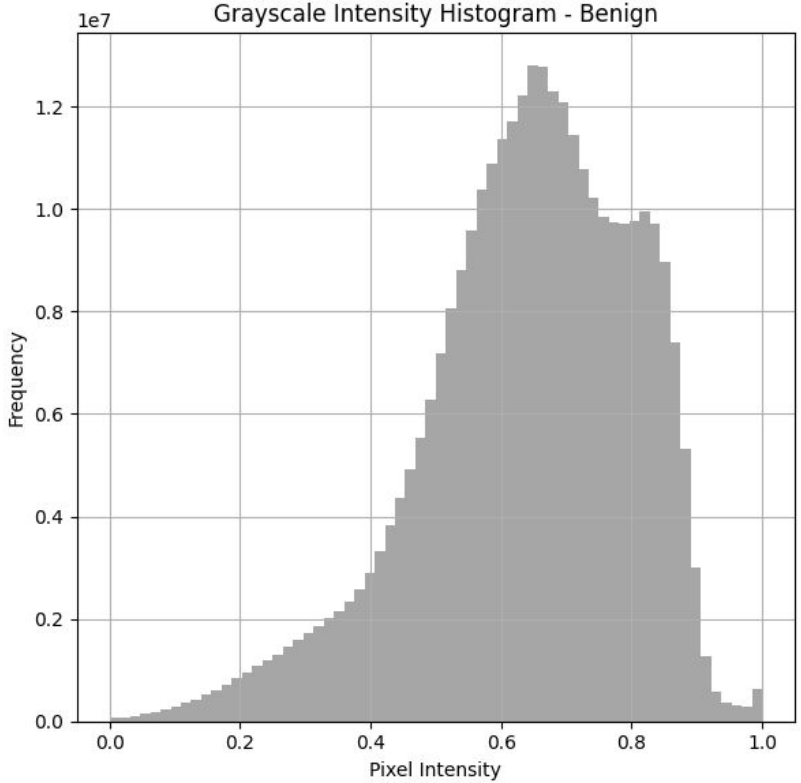


Histogram - Malignant & Benign

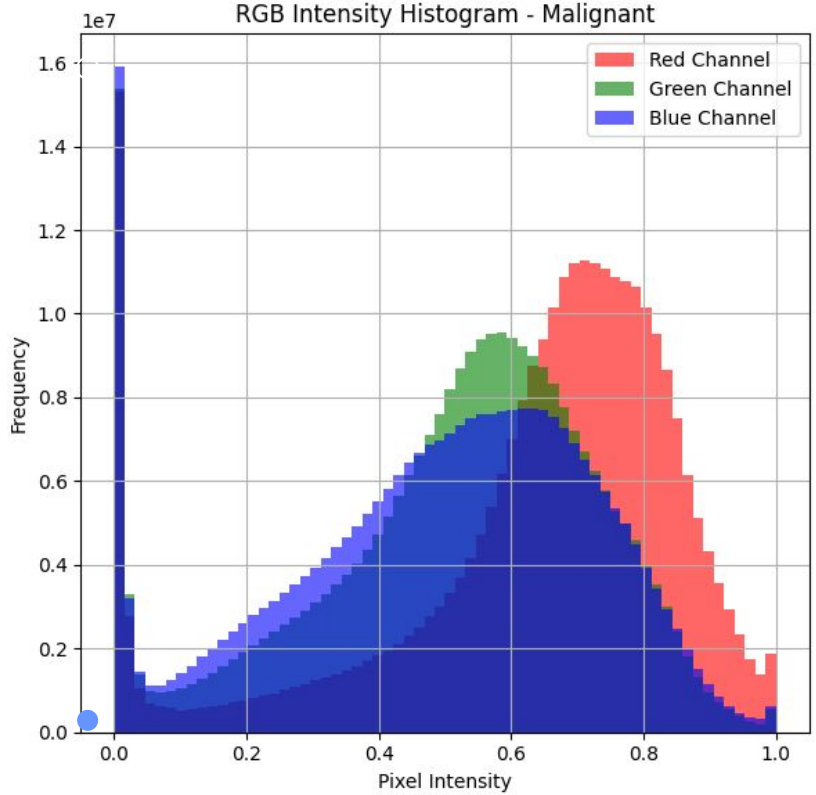


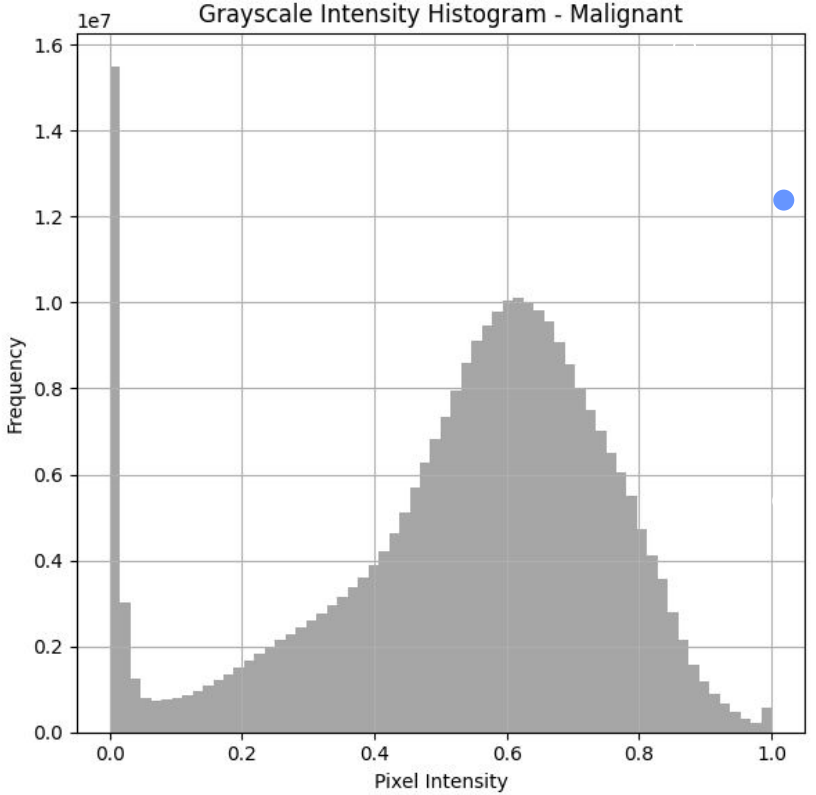
Benign Image - Statistic



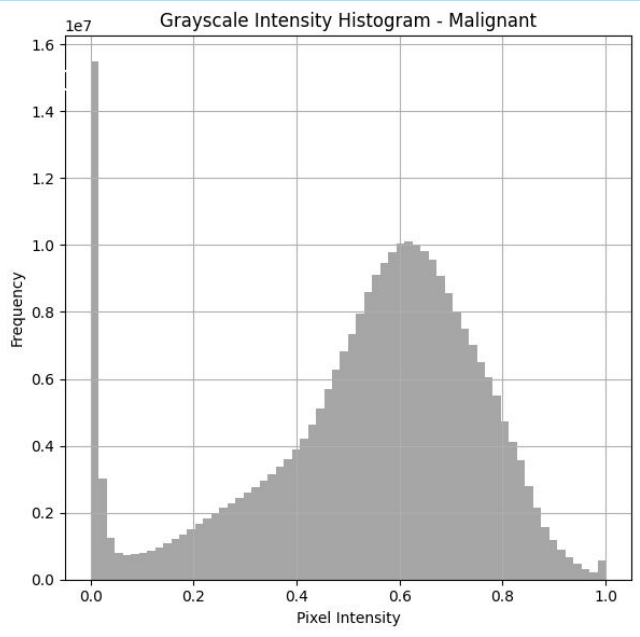


Malignant Image - Statistic





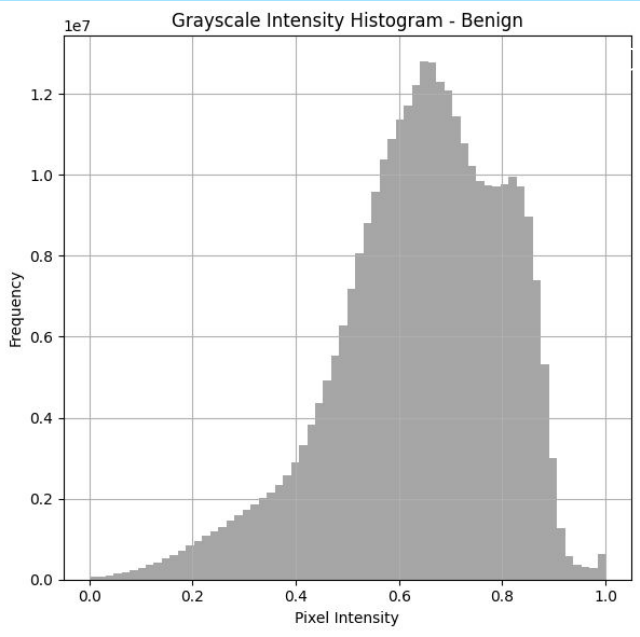
Histogram – Malignant vs Benign



0.63

0.51

0.49



0.806

0.596

0.581

“Real men don’t quit smoking,
Real men fight cancer.”

— wz