

Skin Cancer Detection

Cher Wee Zheng, Ng Wei, Ryan Yong



We Are



You Are





Table of contents

01

**Introduction &
Problem Statement**

02

EDA & Preprocessing

03

Model

04

App Demo

05

Conclusion

01

Introduction & Problem Statement





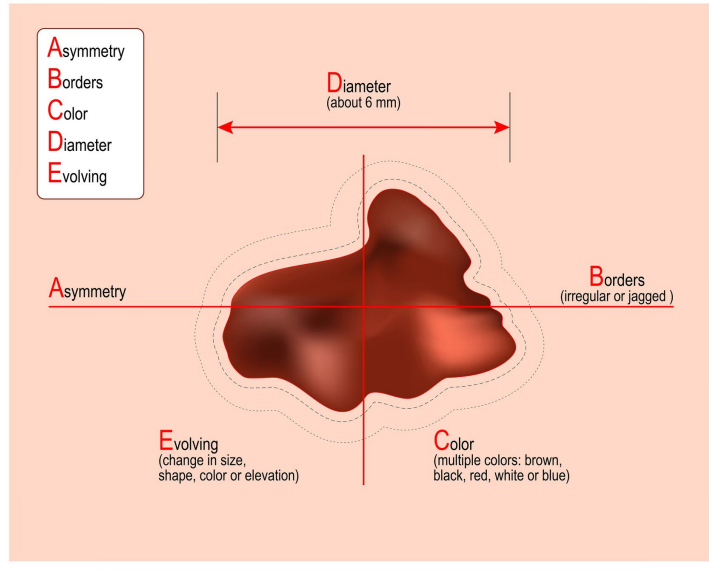
Trigger Warning

The following presentation contains material that may be **disturbing, harming** or **traumatizing** to some audiences. Viewer discretion is advised.

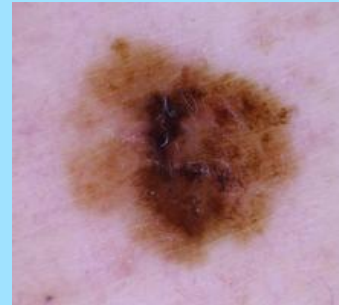
Skin Cancer

Source: CNA

ABCDE (the first signs of a melanoma)



Cancerous



Cancerous



Non-cancerous

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 foreman 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 foreman 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 foreman 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?

Video Consultation





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

Can you move the camera for me to take a closer look?

Ok Ok





Hi Joseph, due to the video quality, I am unable to properly assess the spot. Can you come down to the nearest polyclinic for me to perform a dermoscopy?

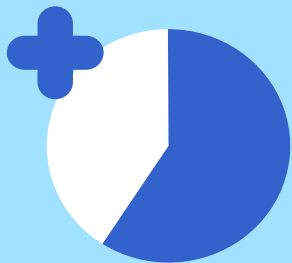
Huh? More tests? Then I pay money and call you for me to come down. Waste my money! Waste my time! Forget it la!

–End of Call–




Types of Examination

Visual

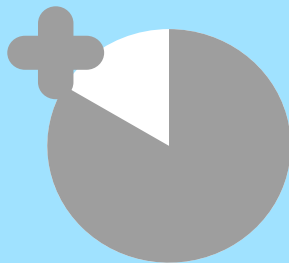


60%


 Instant


 \$35 - \$200

Dermoscopy

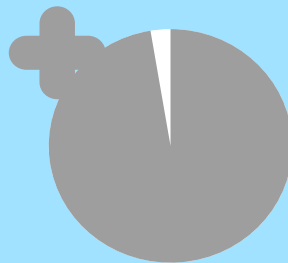


85%

 Instant


 From \$250

Biopsy



95%

 3 weeks

 \$500 - \$3000



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

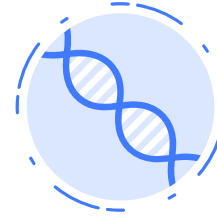
Accurate
diagnosis/treatment
at polyclinic level



02

Reduce

Unnecessary
specialist
referrals



03

Reduce

Costs

Development Process

Exploration

An overview of the dataset gathered and how the data is preprocessed

Modelling

Using our dataset, develop an effective and accurate model

Tuning

Improve our chosen model by altering its hyperparameters

Implementation

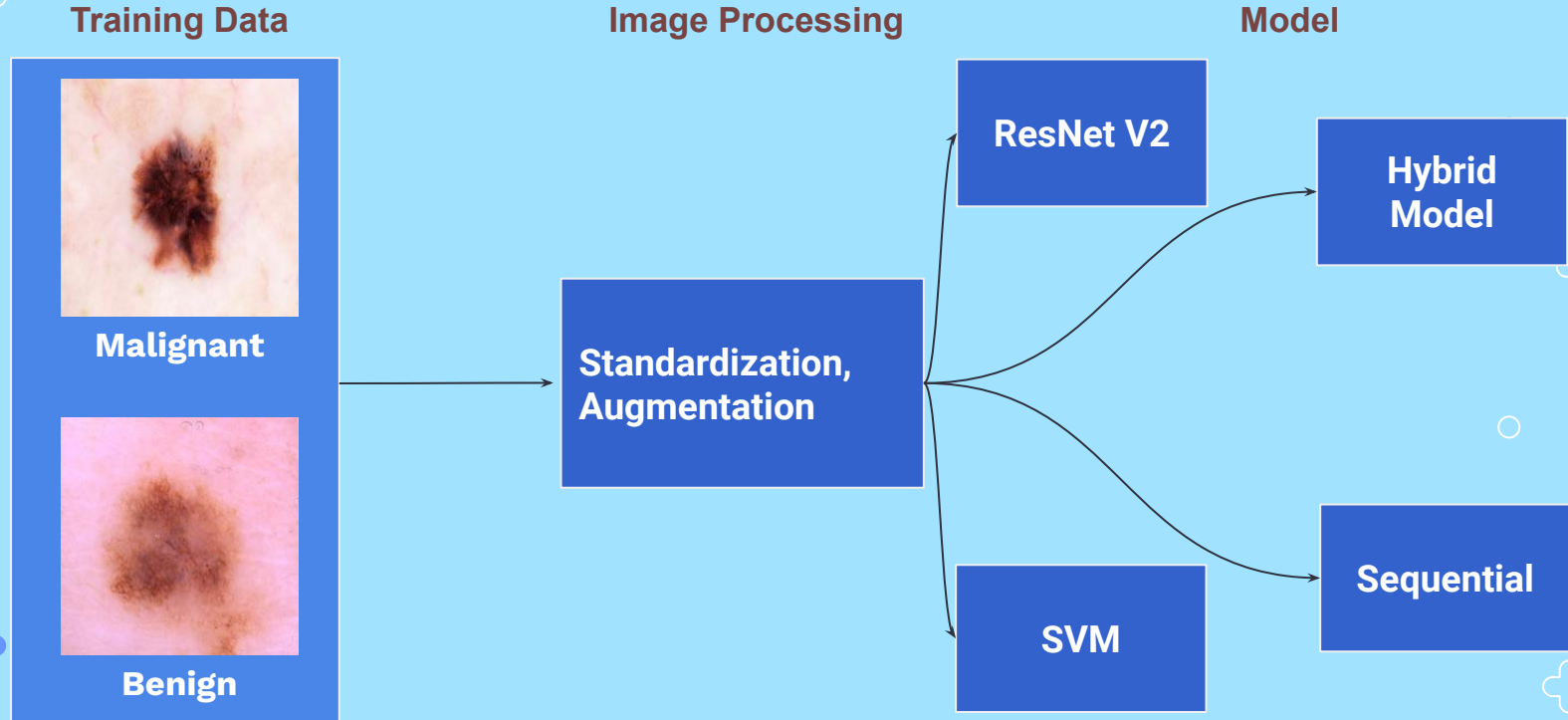
We work together with you to upload our app in your systems



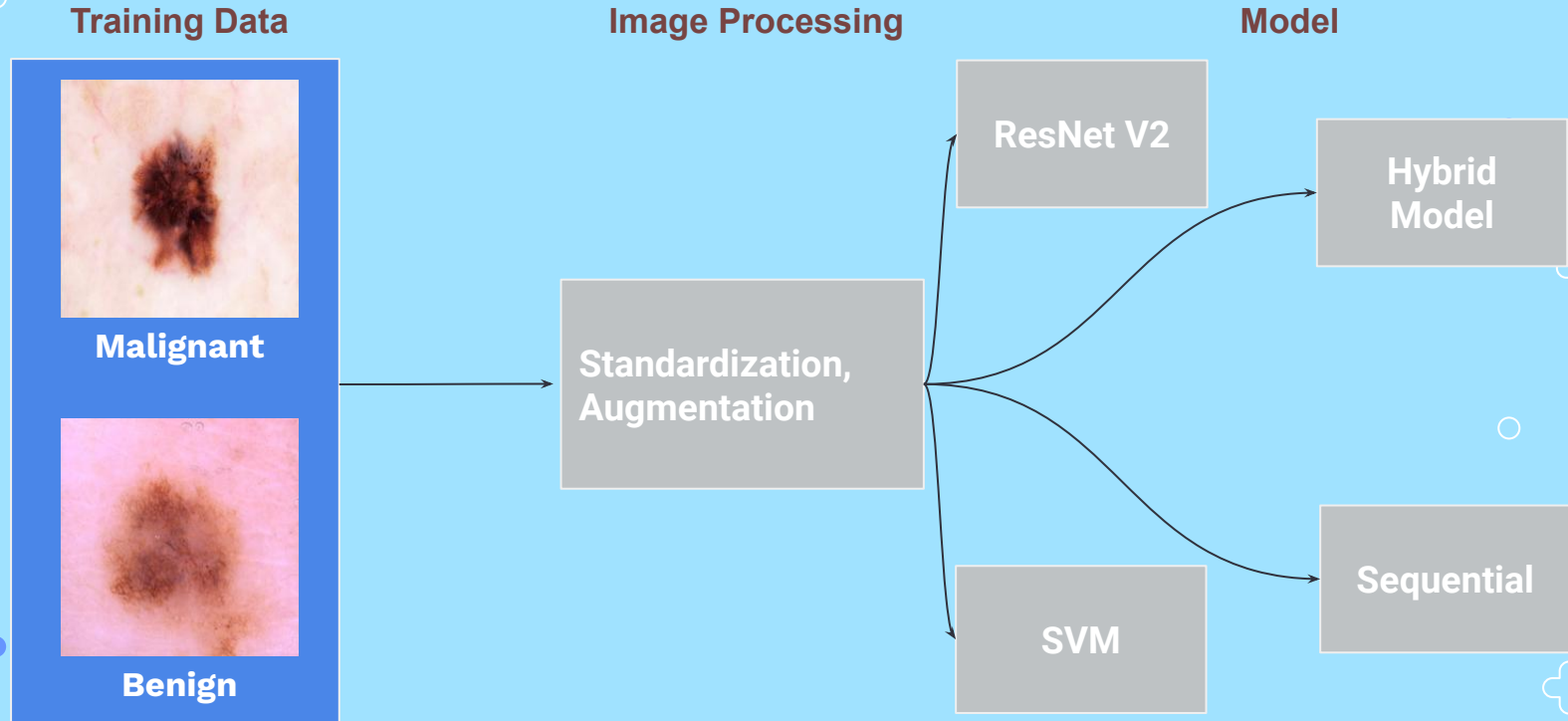
02

EDA & Preprocessing

Modelling Process



Modelling Process



Our Data are all images!

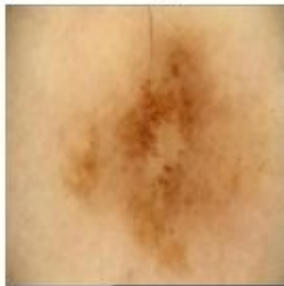
Benign



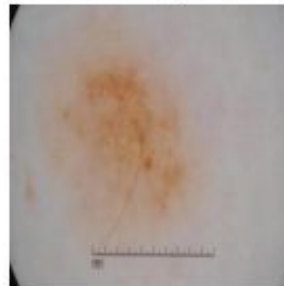
Benign



Benign



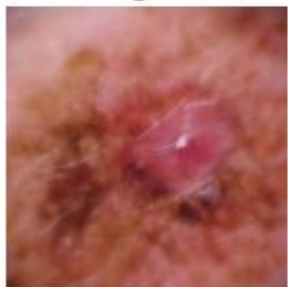
Benign



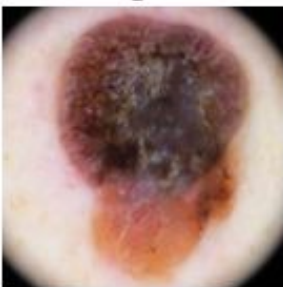
Benign



Malignant



Malignant



Malignant



Malignant



Malignant



Our Data



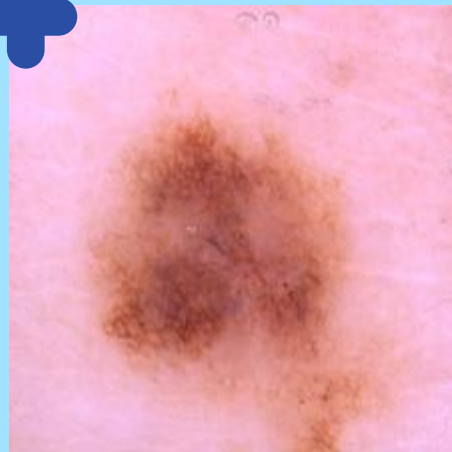
Malignant

5590

Training Data

1000

Testing Data



Benign

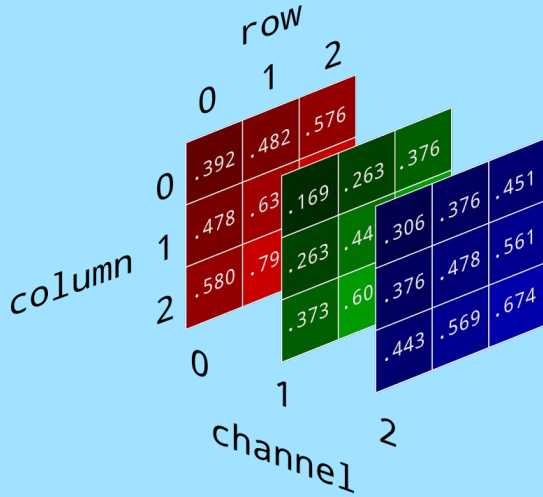
6289

Training Data

1000

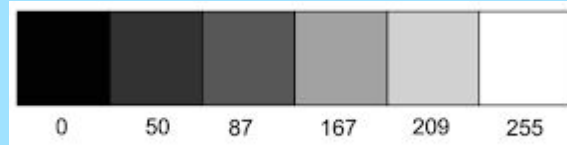
Testing Data

How does models see images?

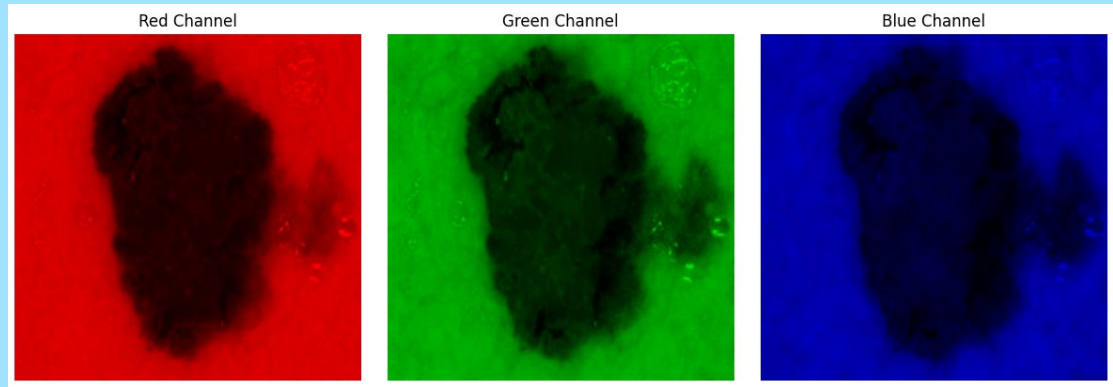


Images are **composed of pixels**
Each pixel contains **Red**, **Green** and **Blue** information.

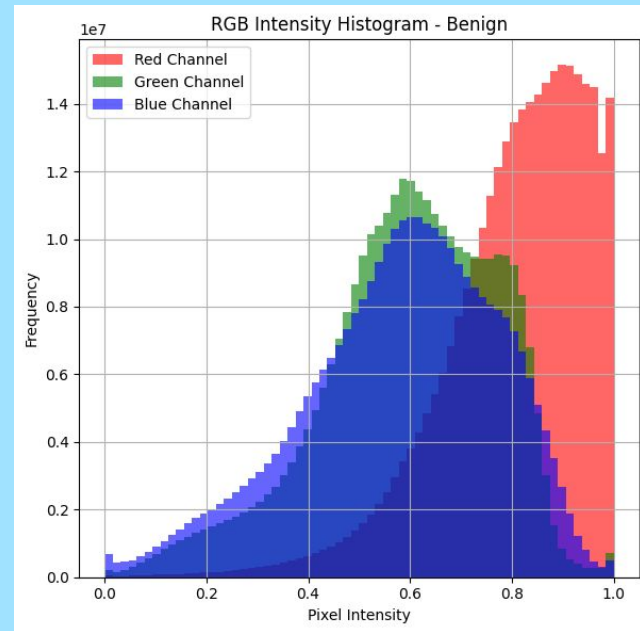
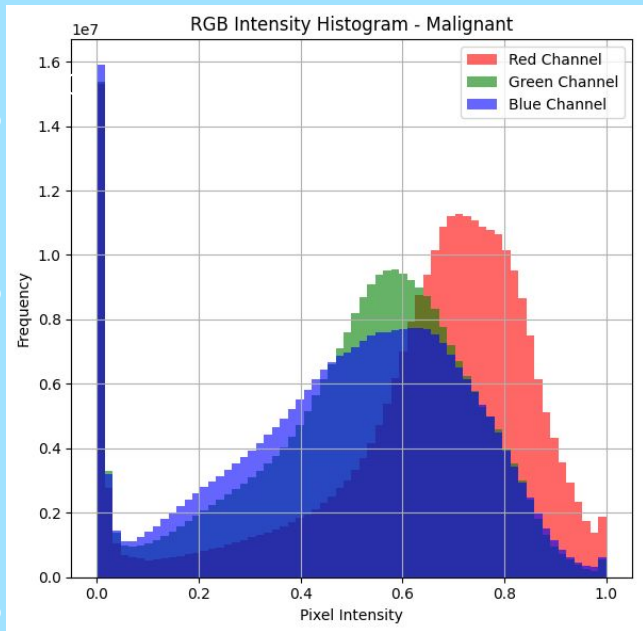
Normally **0-1** or **0-255** to representing each **color's intensity**.



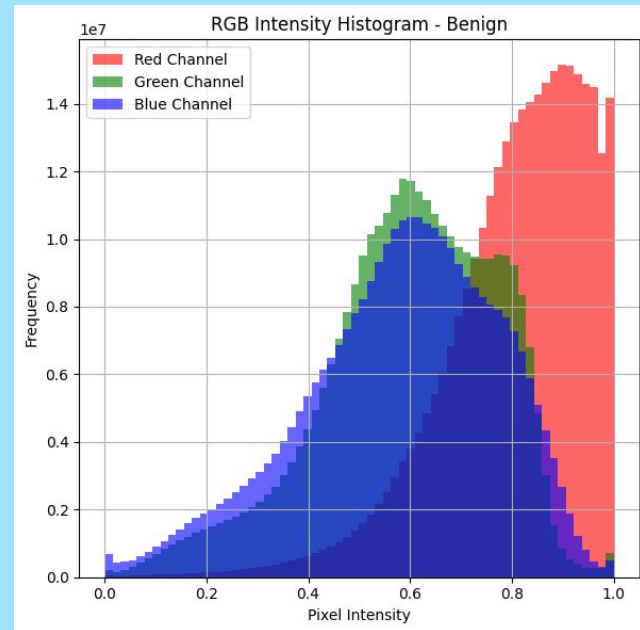
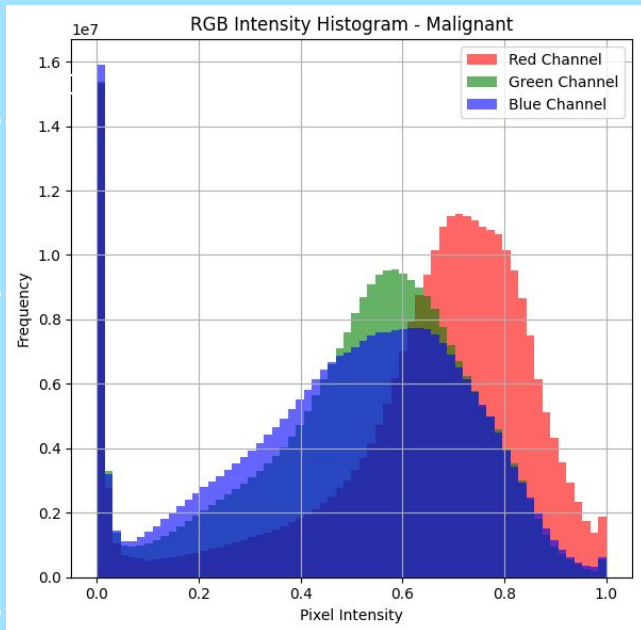
=



Histogram - Malignant vs Benign

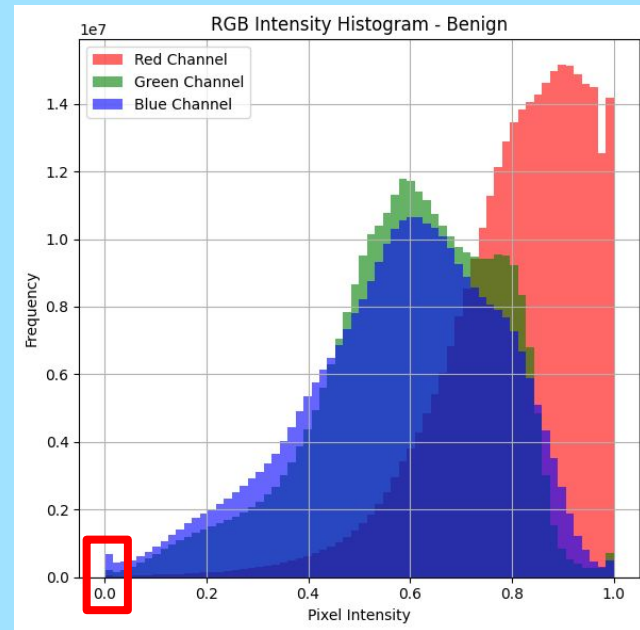
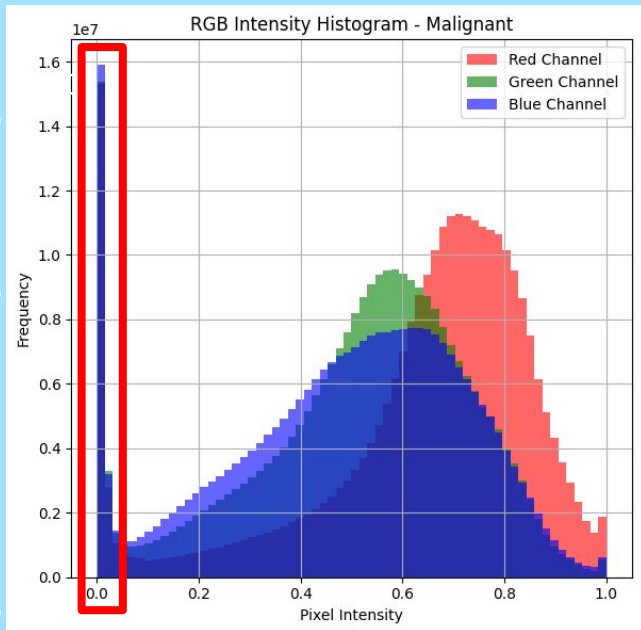


Histogram - Malignant vs Benign



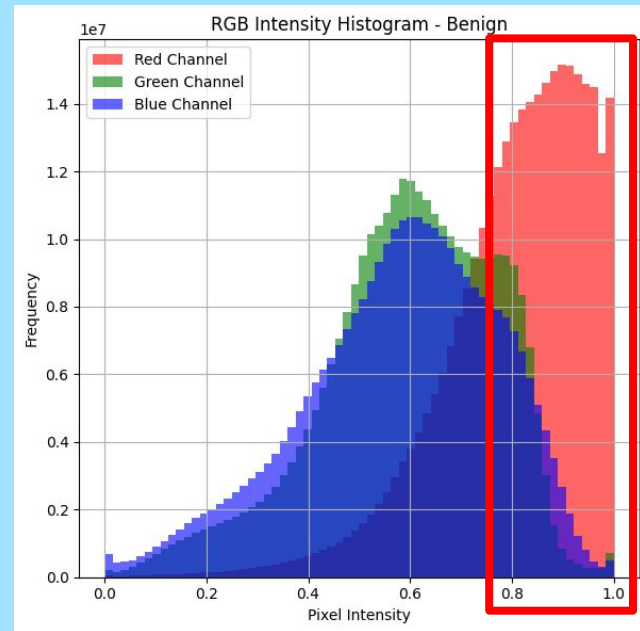
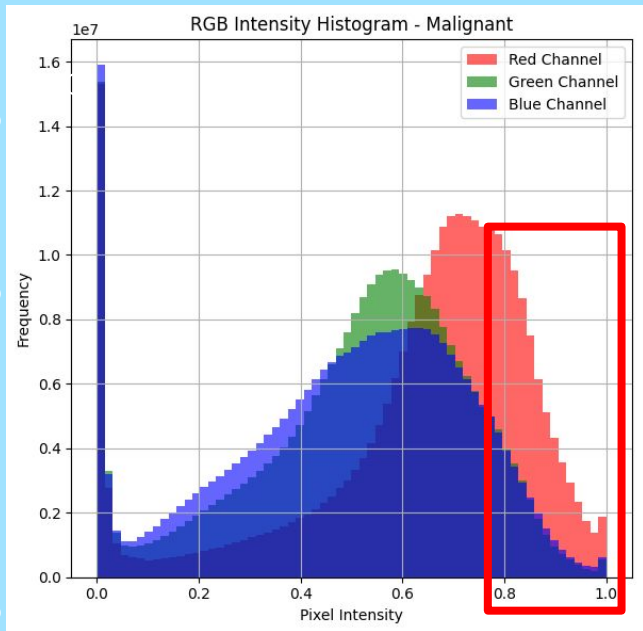
Finding 1:
Based on the average of RGB value, Benign tends to be brighter than Malignant

Histogram - Malignant vs Benign



Finding 2:
Malignant images have significantly more black pixels (R,G,B = 0)
compared to Benign images

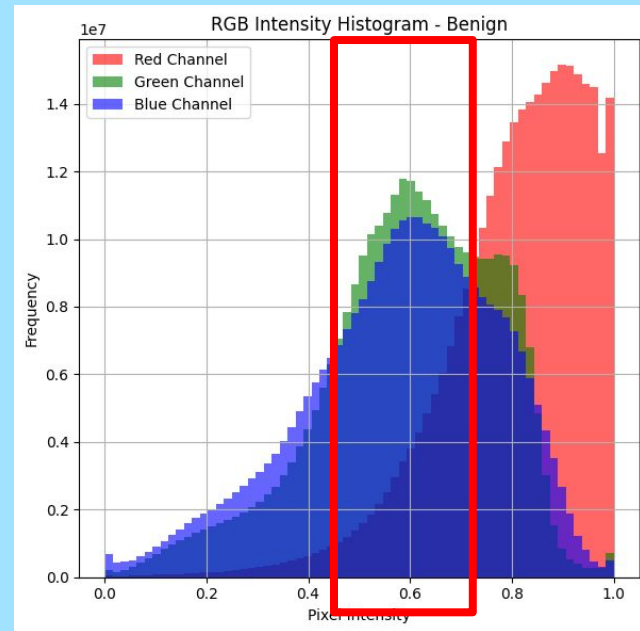
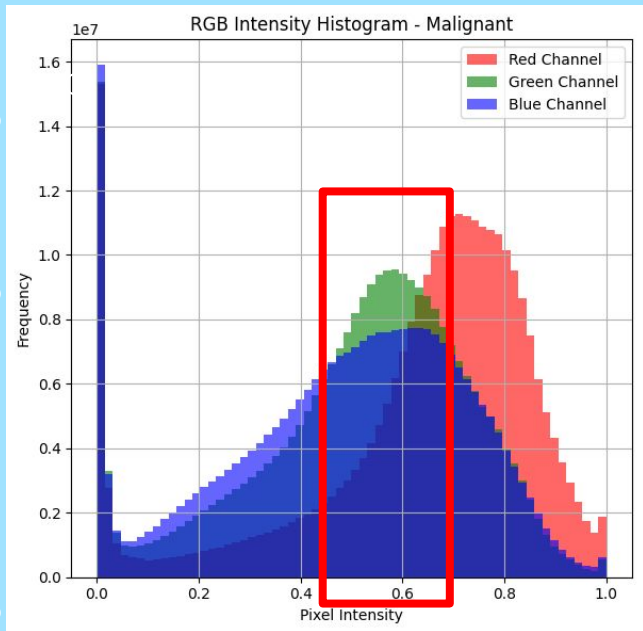
Histogram - Malignant vs Benign



Finding 3:

Benign images more “red” than Malignant, red pixels mostly > 0.75

Histogram - Malignant vs Benign



Finding 4:

Other than the black pixels from Malignant, both Green and Blue pixels mostly distributed between 0.45 - 0.7 intensities

Expected

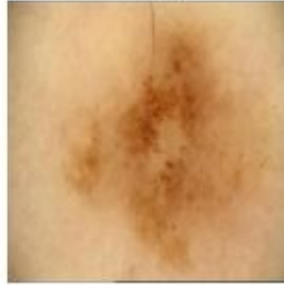
Benign



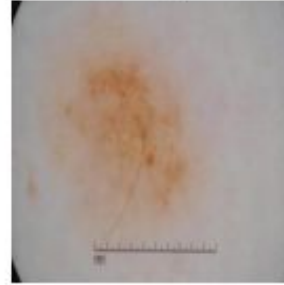
Benign



Benign



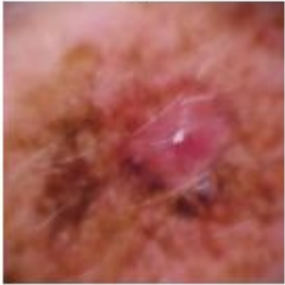
Benign



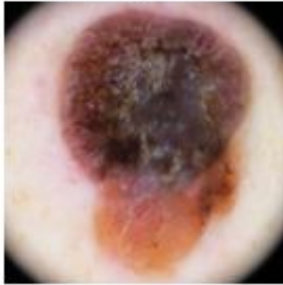
Benign



Malignant



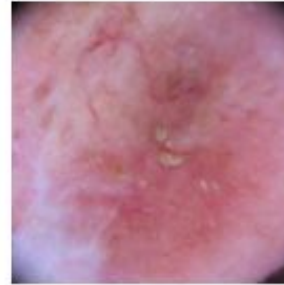
Malignant



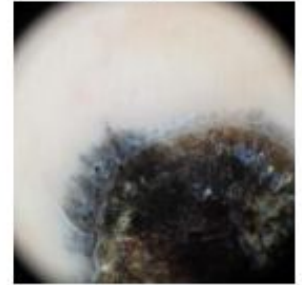
Malignant



Malignant



Malignant



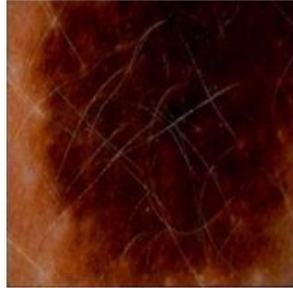
Outliers

Benign Outlier (RGB < 0.35)

4627.jpg



5211.jpg



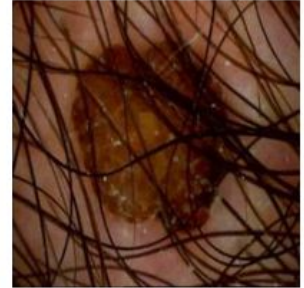
5160.jpg



3573.jpg

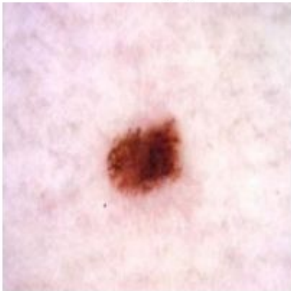


2089.jpg



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

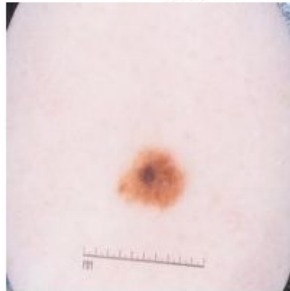
2474.jpg



5119.jpg



1028.jpg



5291.jpg



52.jpg



Modelling Process

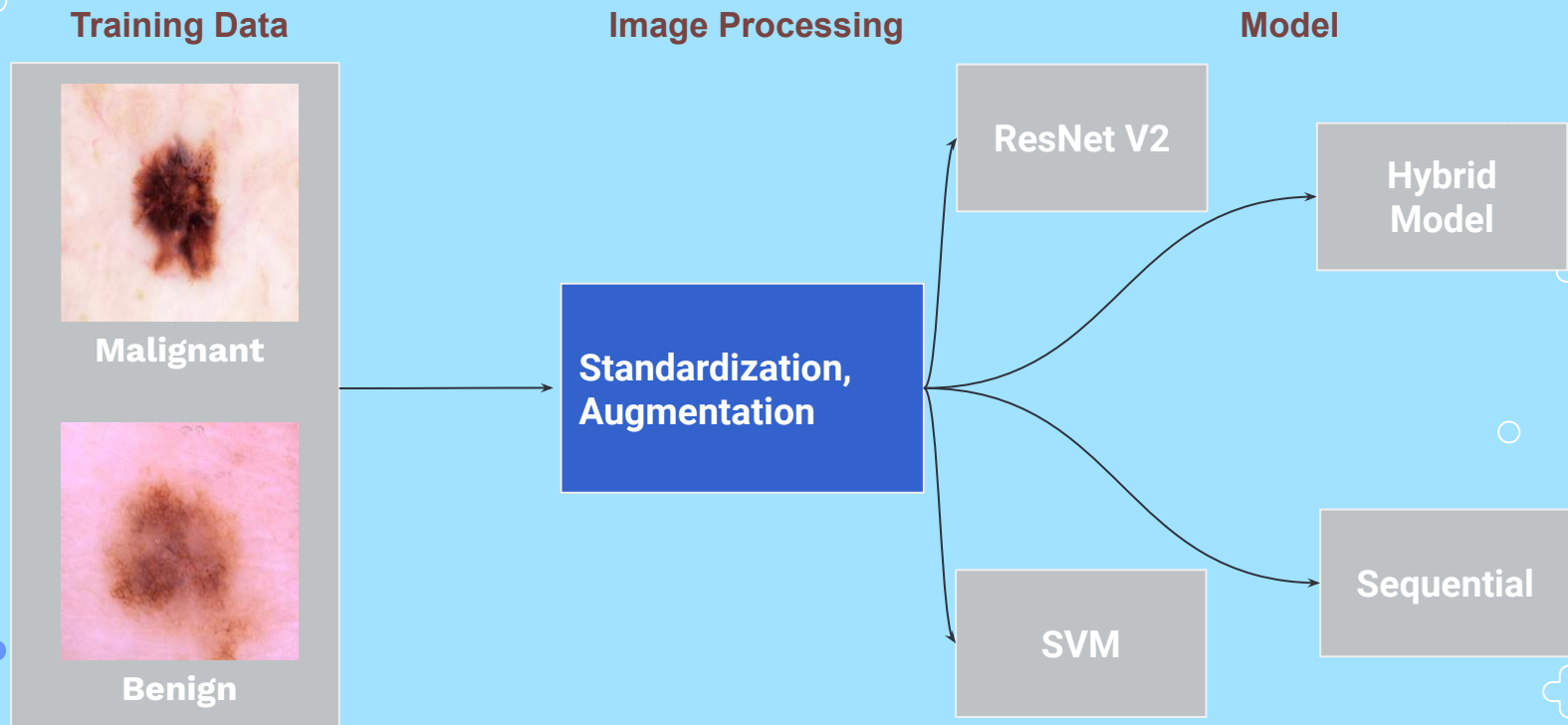
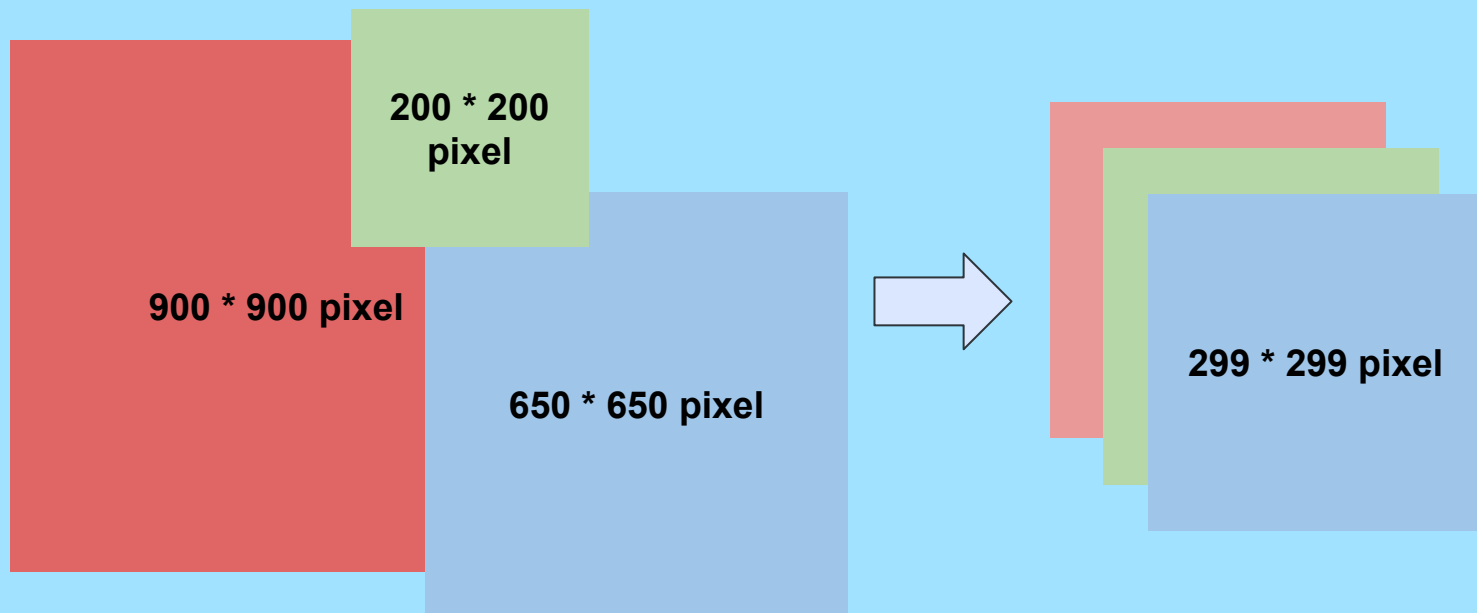


Image Preprocessing

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



In order to obtain discriminative internal features for improved performance, we also 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**

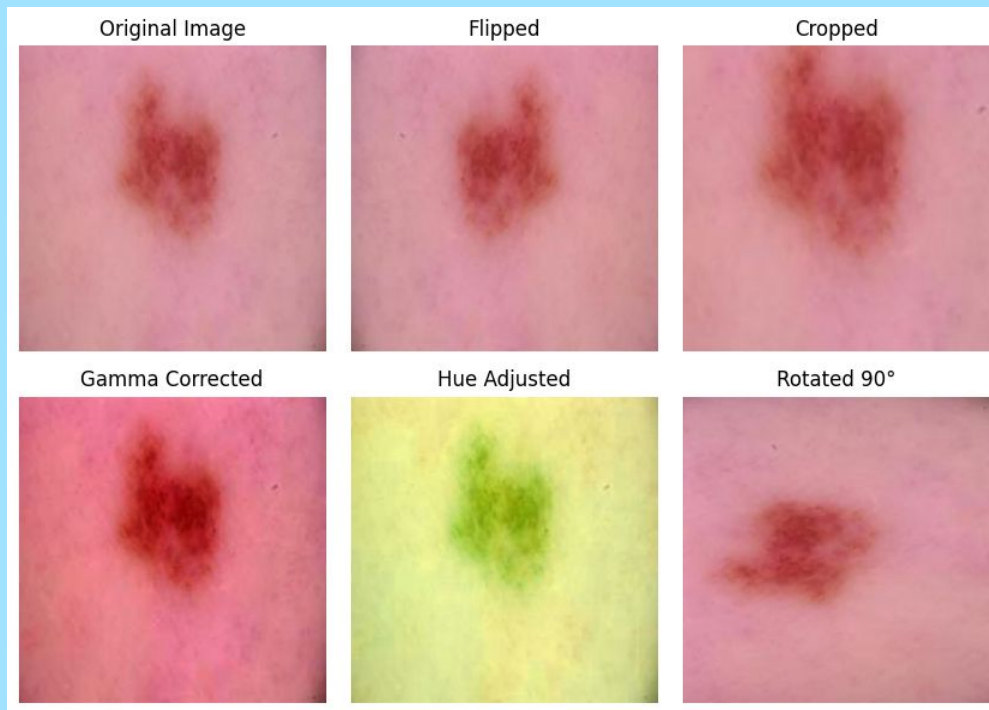


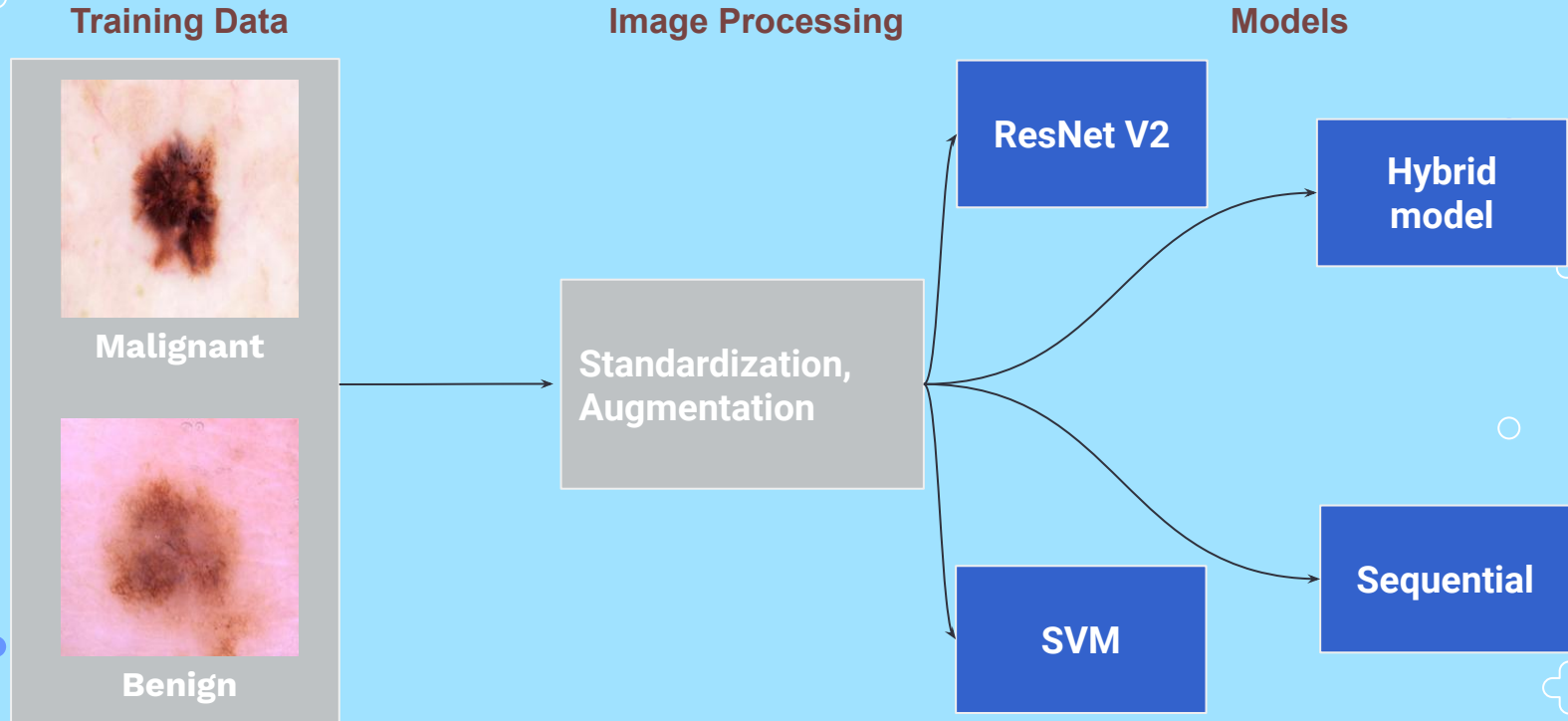
Image augmentation supplements the creation of image variations that can **help models improve their prediction accuracy** and **enhance generalization ability**.



03

Model

Modelling Process





ResNetV2

Residual
Neural
Network

162 Layers



Hybrid Model

Combination
of ResNetV2
and VGG16

782 Layers



Sequential

Self-
Developed
CNN model

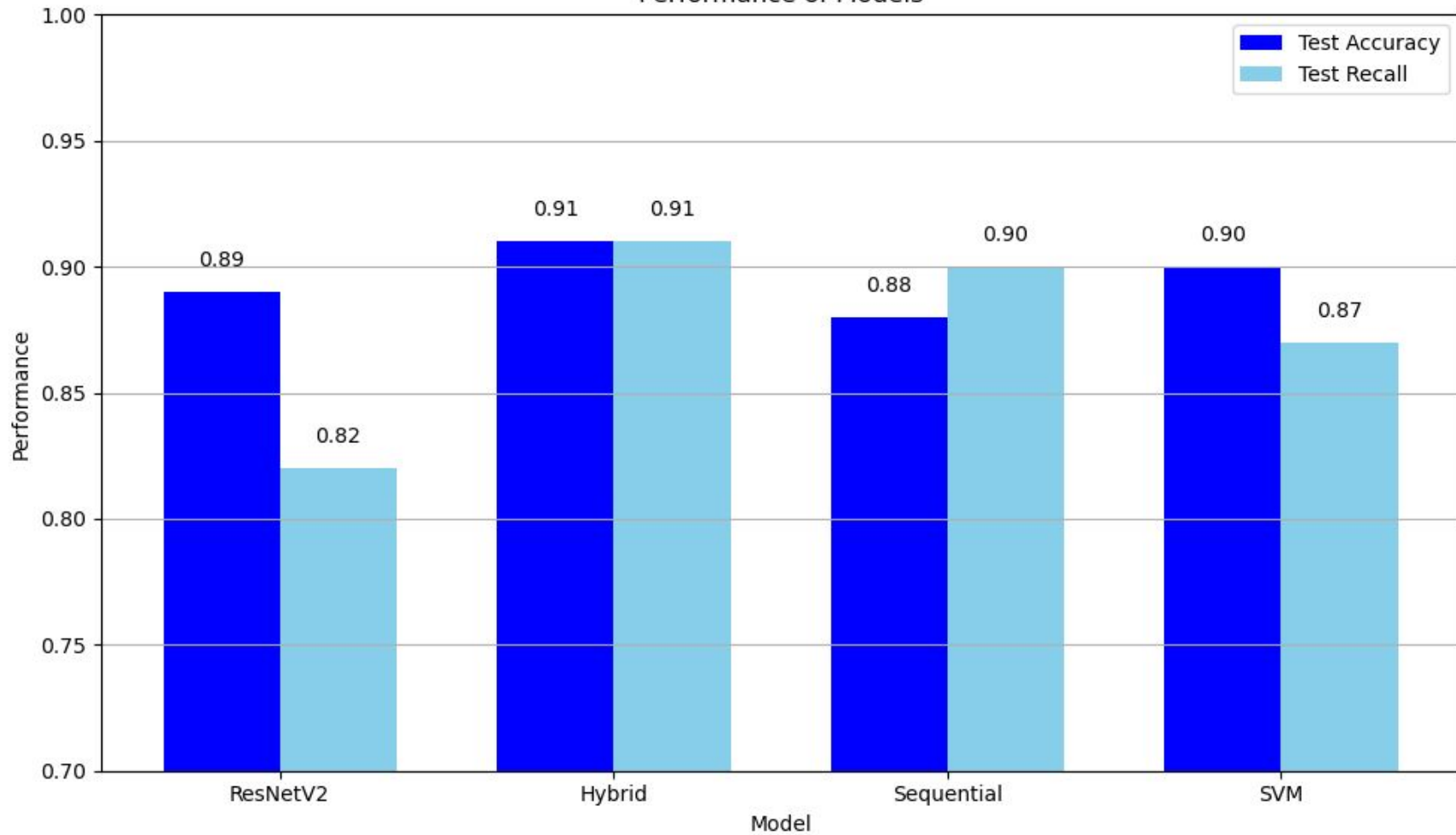
8 Layers



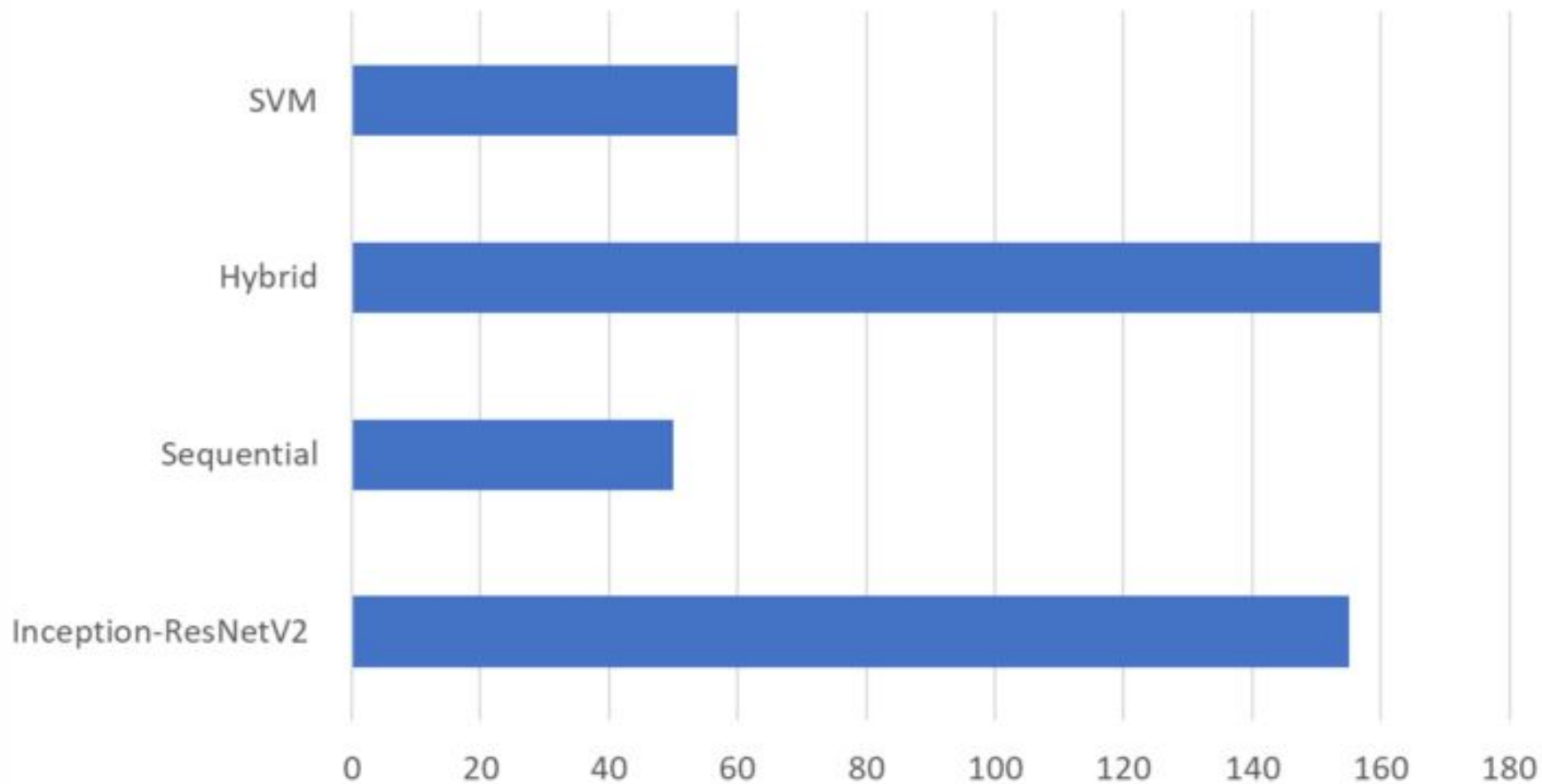
SVM

Support
Vector
Machine

Performance of Models



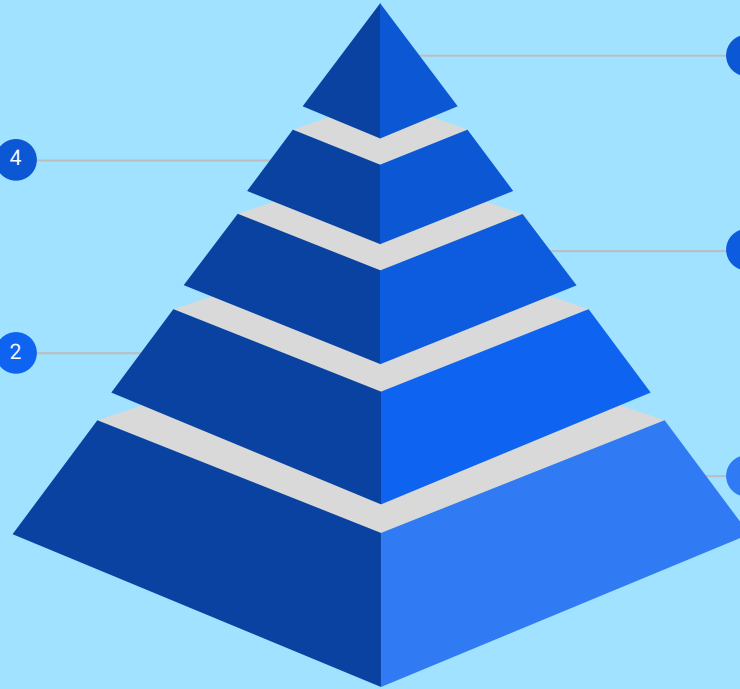
Model vs Process Time (Min)



The chosen Sequential CNN Model

Flatten()
Prepares the data for subsequent dense layers.

2nd Conv2D/MaxPooling 2D
Contains 64 filters of kernel size (3, 3) and a MaxPooling identical to previous



2 Dense layers
First layer with 64 nodes takes flattened feature maps and learns high-level representations. Outputting into a binary Dense layer using Softmax to represent model's confidence in each class.

3rd and Last Conv2D layer
64 filters and a kernel size of (3,3)

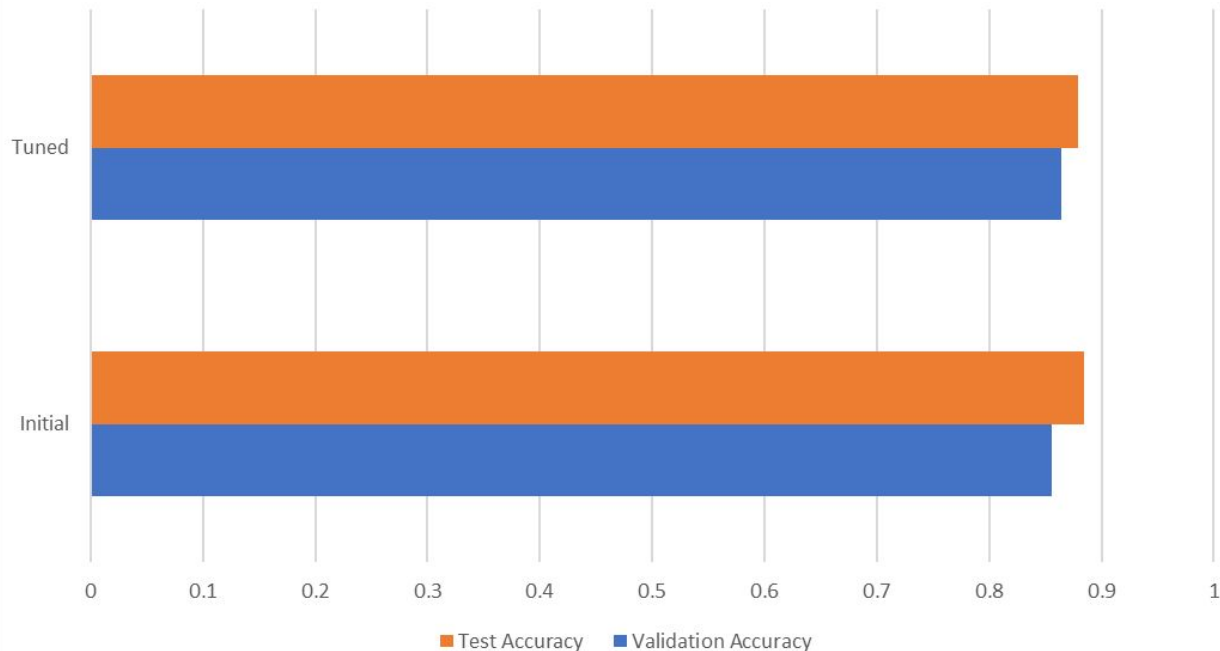
1st Conv2D/MaxPooling 2D
Contains 32 filters with a kernel size of (3,3), later downsampled by MaxPooling with a (2,2) window.



Model Tuning

3b

Initial vs Tuned (Sequential)



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>



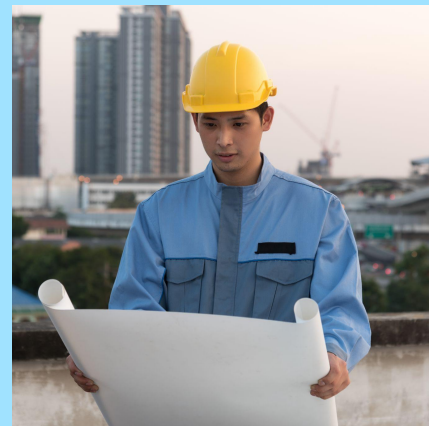
App Demo



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%

Benefits at a glance



Performance

Improvement of ~50%
compared to visual
tests



Instant

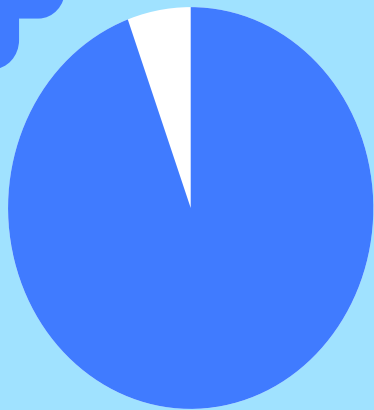
Frees up biopsy labs
for greater efficiency



Online Applicability

Able to be onboarded and
implemented with existing
video consultation services

Conclusion



Skin Cancer Detector

89%



Instant results



Minimal costs



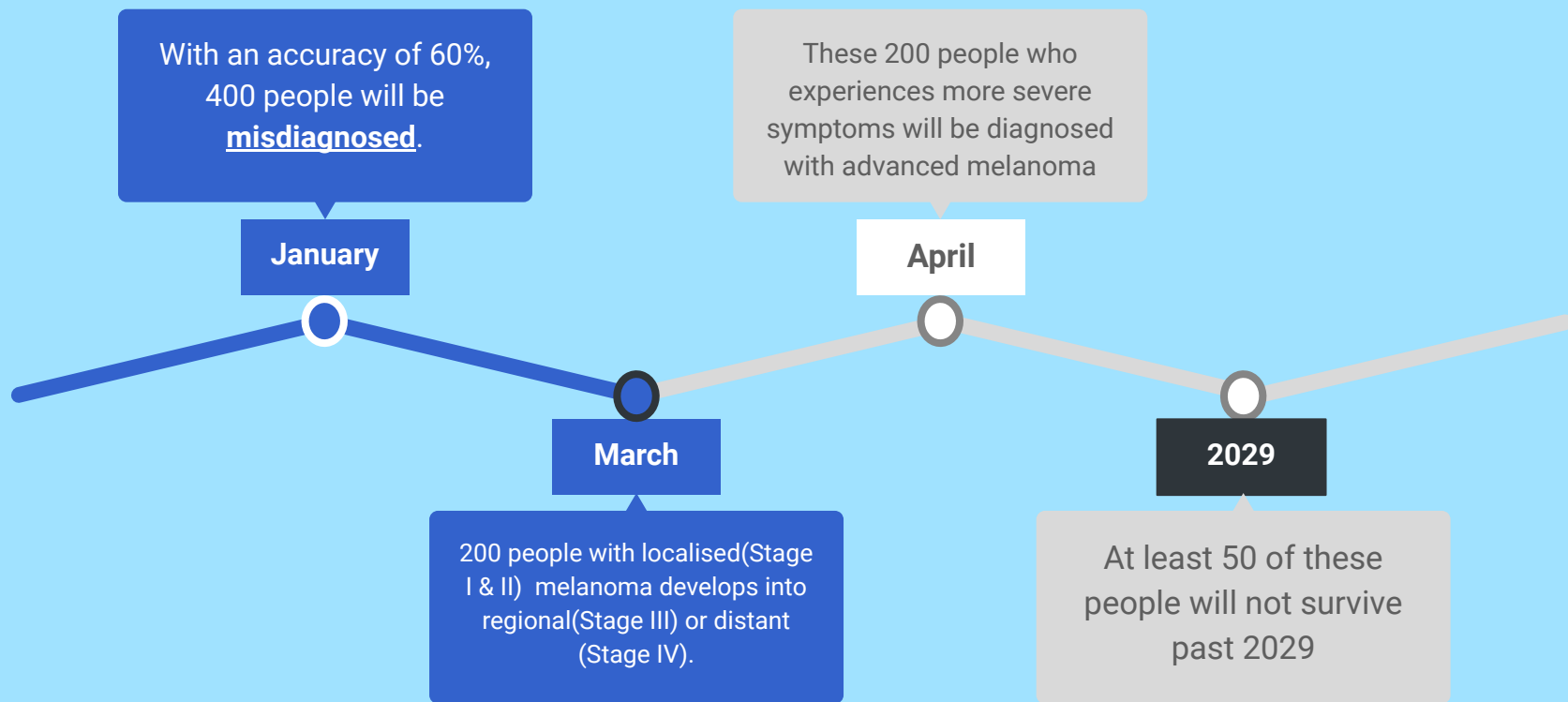
Online accessibility



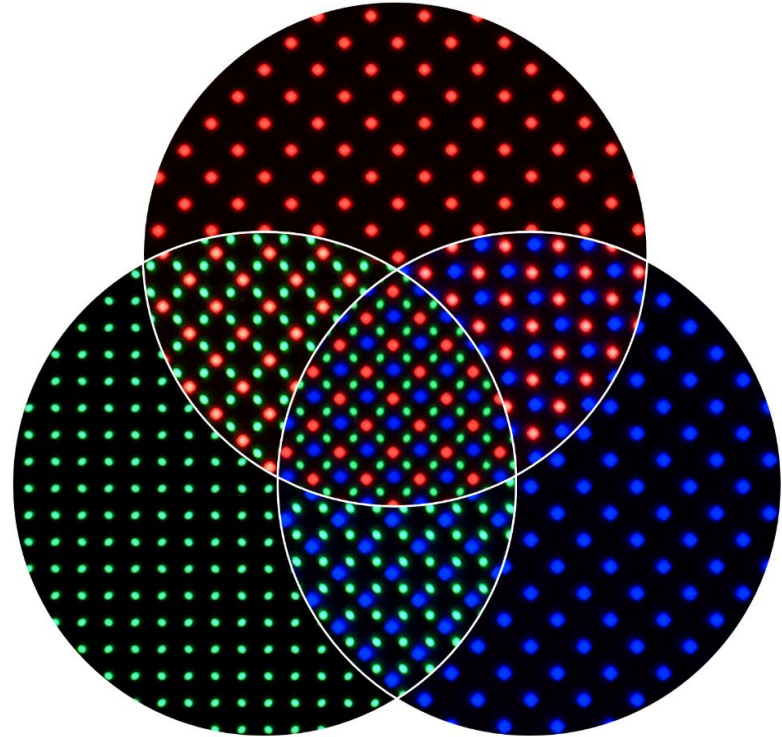
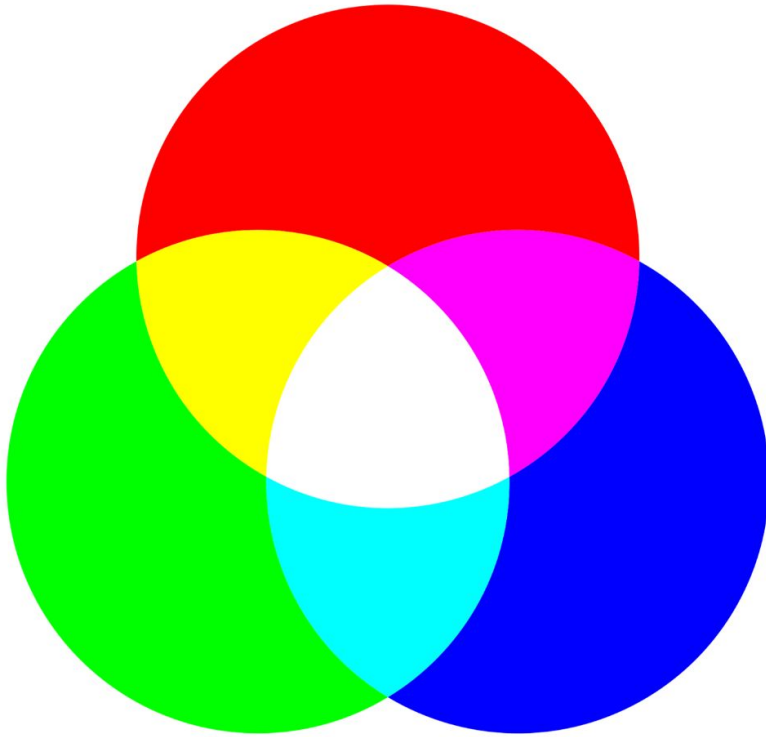
Thank You!!



Calculation of Relative Survival Rate



How does machine see images?



Original Image - Benign



RGB Intensity Histogram



RGB Colour Mixer - Used by Screens

rgb(200,180,180)

Red - Currently set to 200



Green - Currently set to 180



Blue - Currently set to 180



☒ Decimal ☐ Hexadecimal

rgb(255,180,180)

Red - Currently set to 255

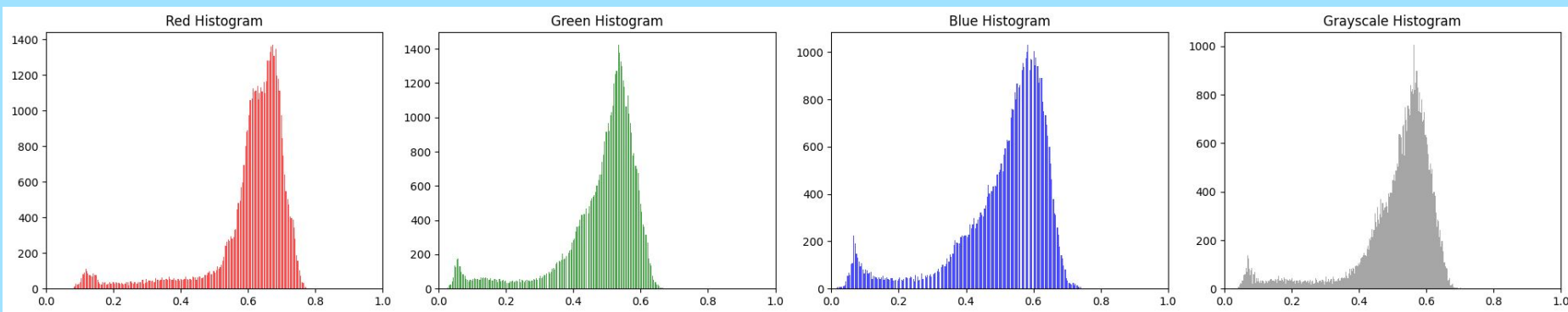
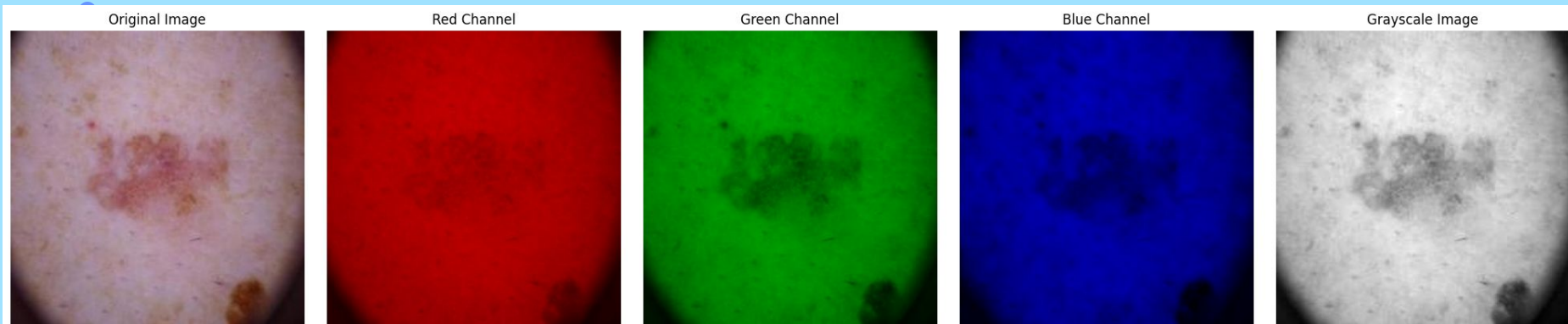


Green - Currently set to 180

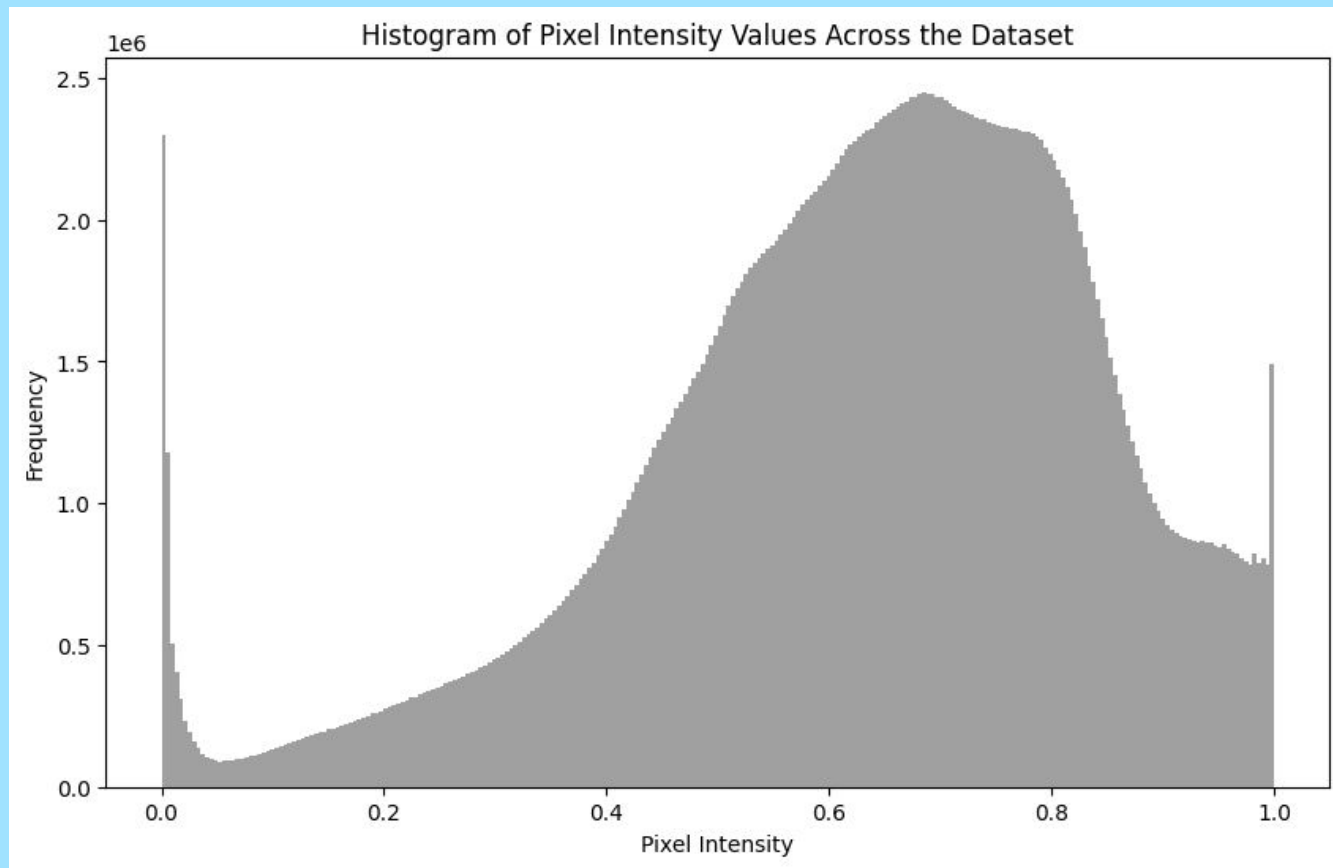


Blue - Currently set to 180

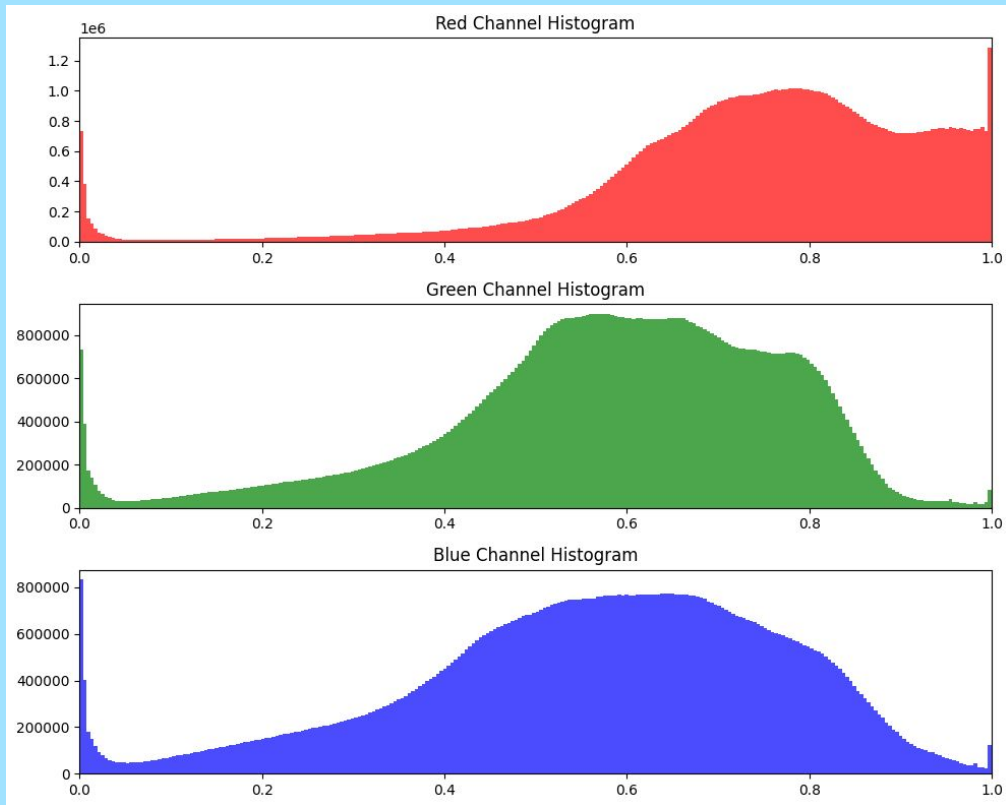




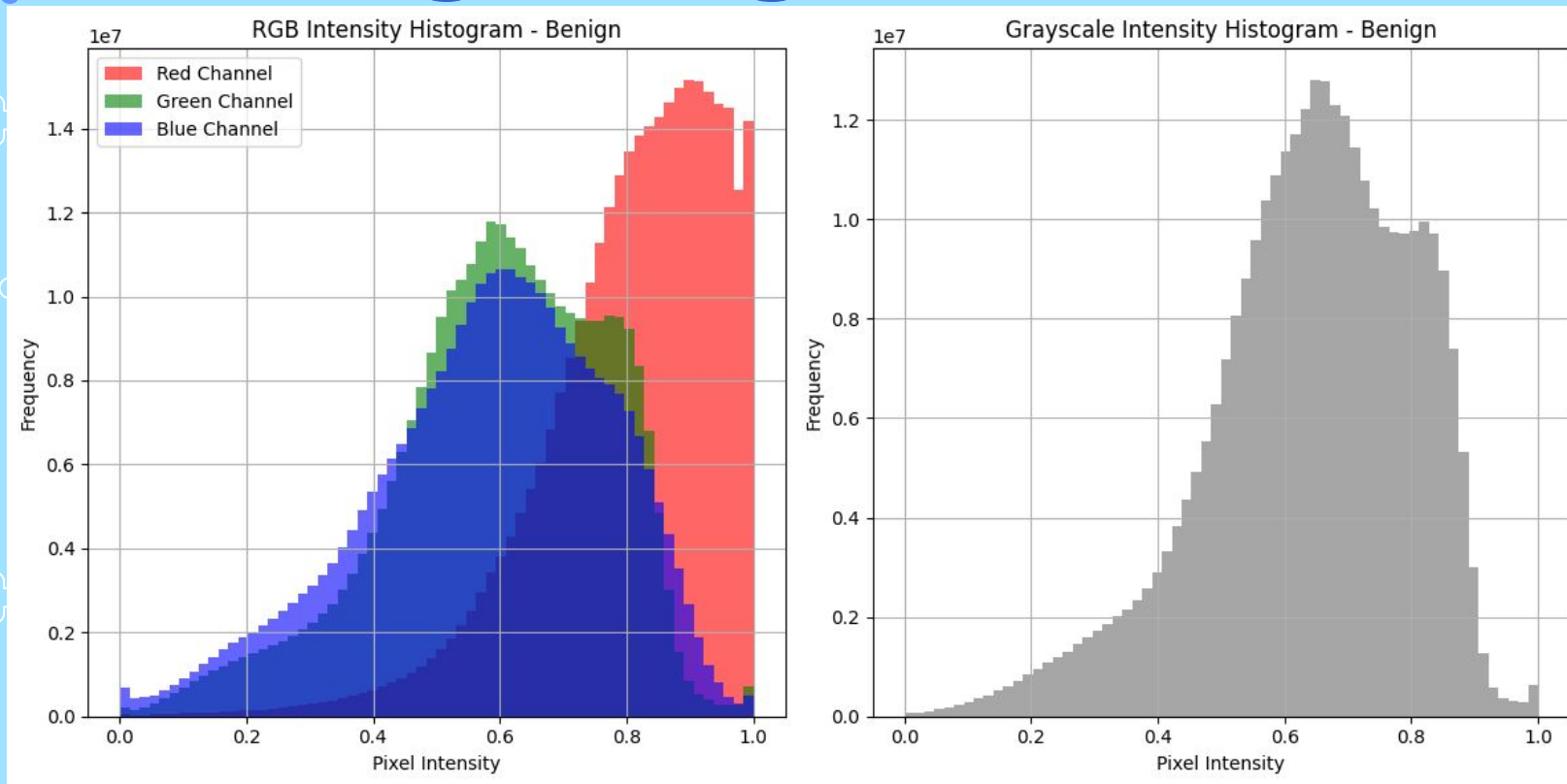
Histogram - Malignant & Benign



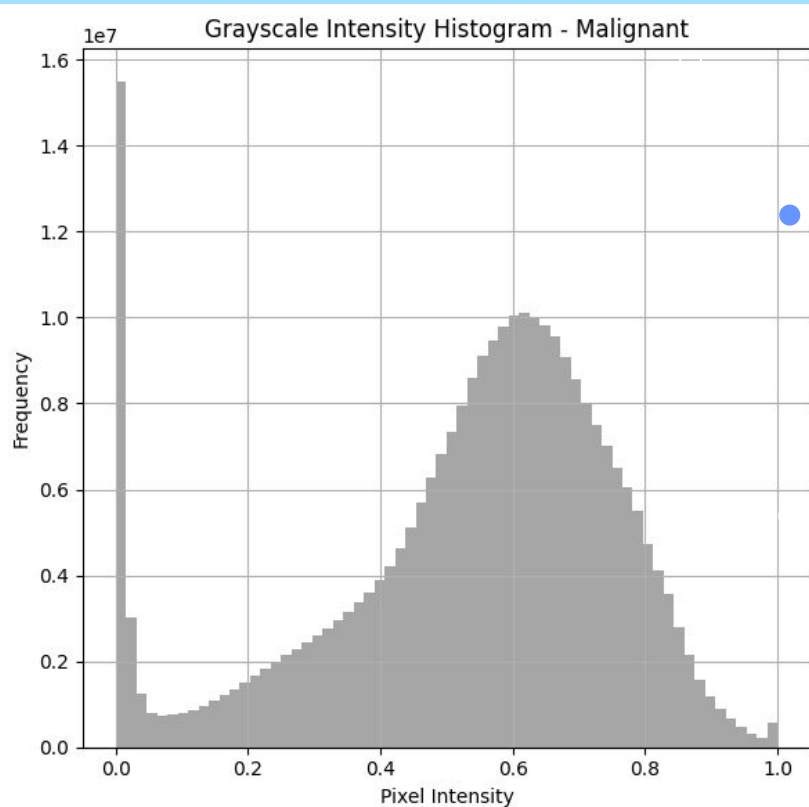
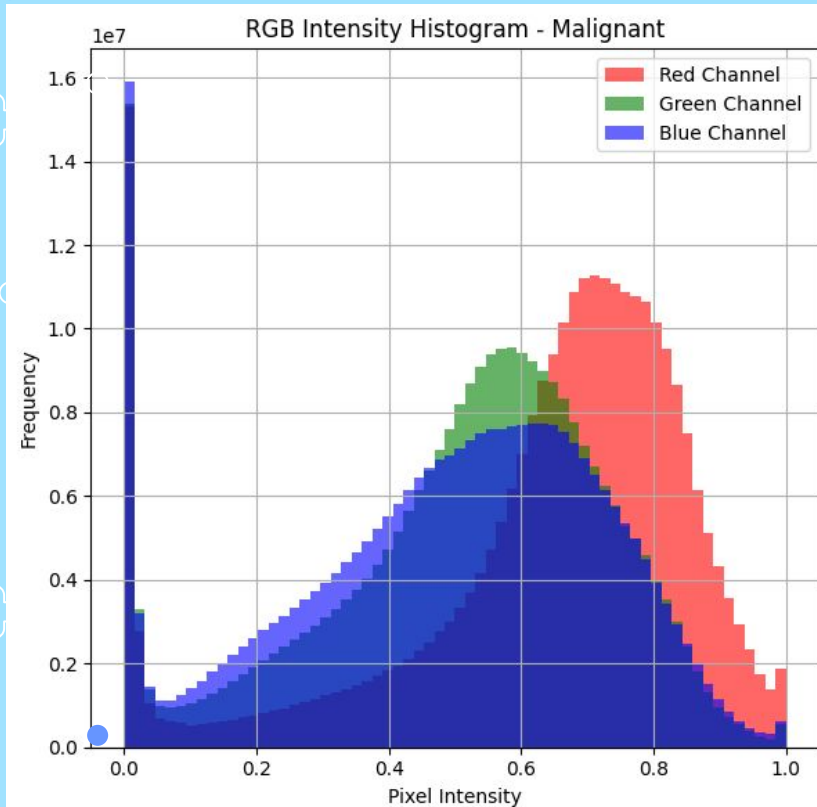
Histogram - Malignant & Benign



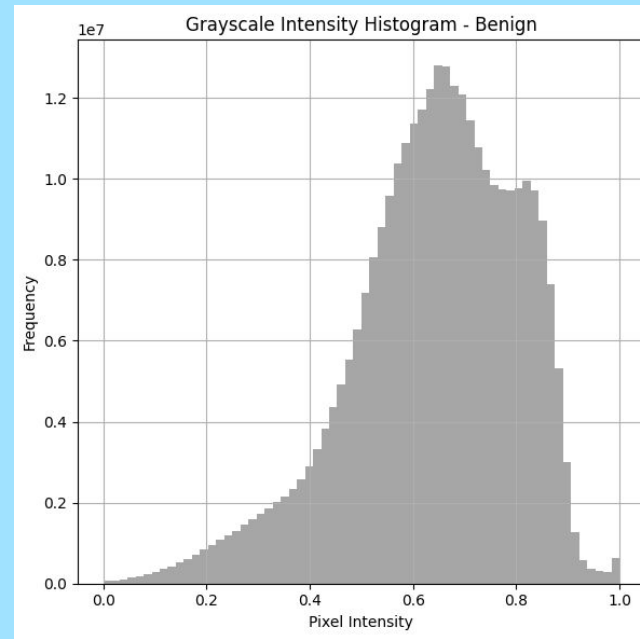
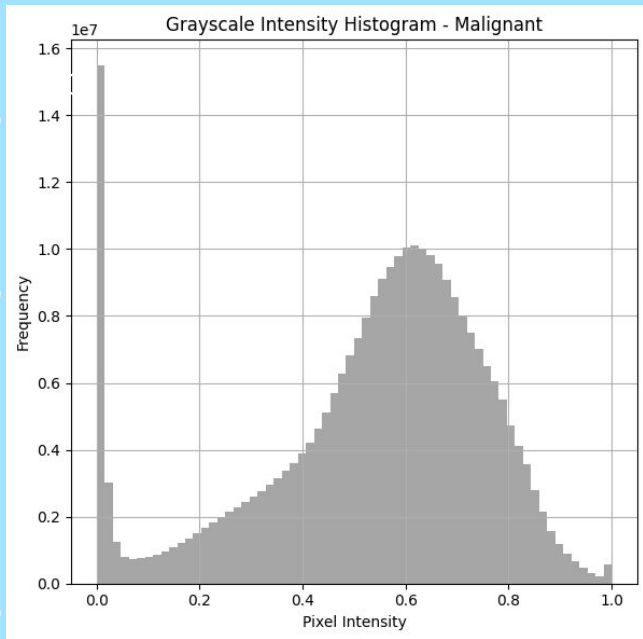
Benign Image - Statistic



Malignant Image - Statistic



Histogram - Malignant vs Benign



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

— WZ

