Skin Lesion Classifier

Tapping on Machine Learning to identify malignant melanoma

Melanoma is a public health concern globally, with over 325,000 cases and over 57,000 deaths due to melanoma in 2020.

Early detection plays a crucial role in improving treatment outcomes and reducing mortality rates associated with melanoma.

However, differentiating between melanoma and benign skin lesions is difficult, even for clinicians.

Can you tell?

Benign

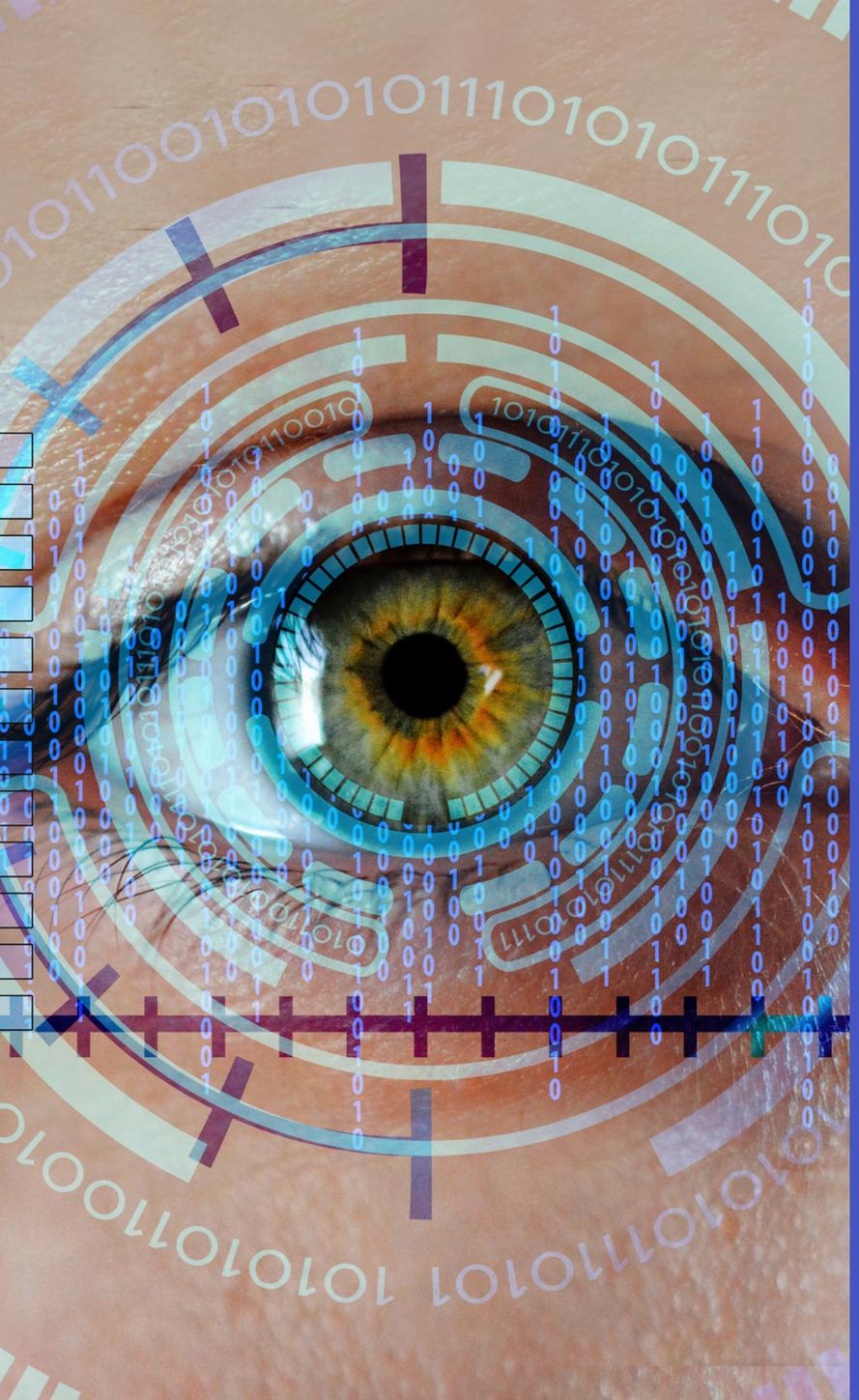
Malignant

What do we do?

Visit a doctor for every skin lesion?

Only visit when it's severe?

How can the public be certain that a visit to the doctor is required?



Skin Lesion Classifier App

1

Use transfer learning to develop a neural network model capable of accurately identifying skin lesions.

2

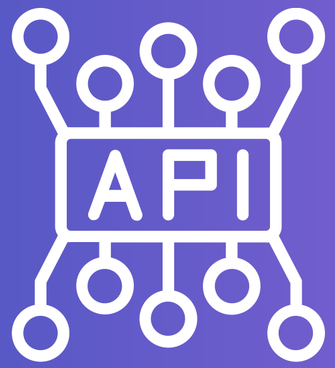
Deploy the model to an app to be used by the public.

3

Users can be sure a doctor's visit is needed based on the results of the app.

Data Collection

Data was obtained from the International Skin Imaging Collaboration (ISIC) which contains images of skin lesions with ground-truth diagnoses.



Downloaded images using their API:

- 6680 classified as malignant melanoma.
- 7500 classified as benign.

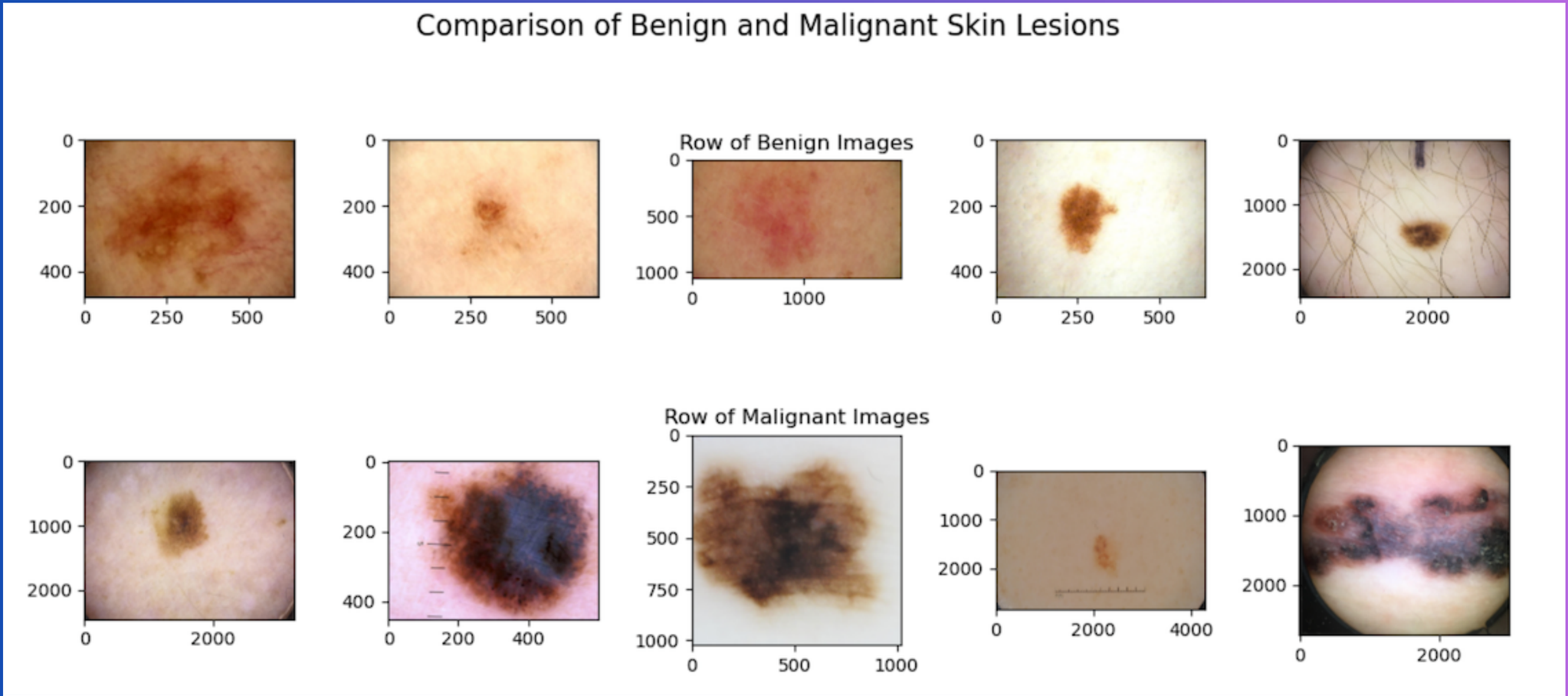
Caveat of the Dataset:

- Majority of images are of lighter skin tones.
- Results of the model may not generalise well to darker skin tones.

EDA

How to differentiate malignant melanoma vs benign skin lesions

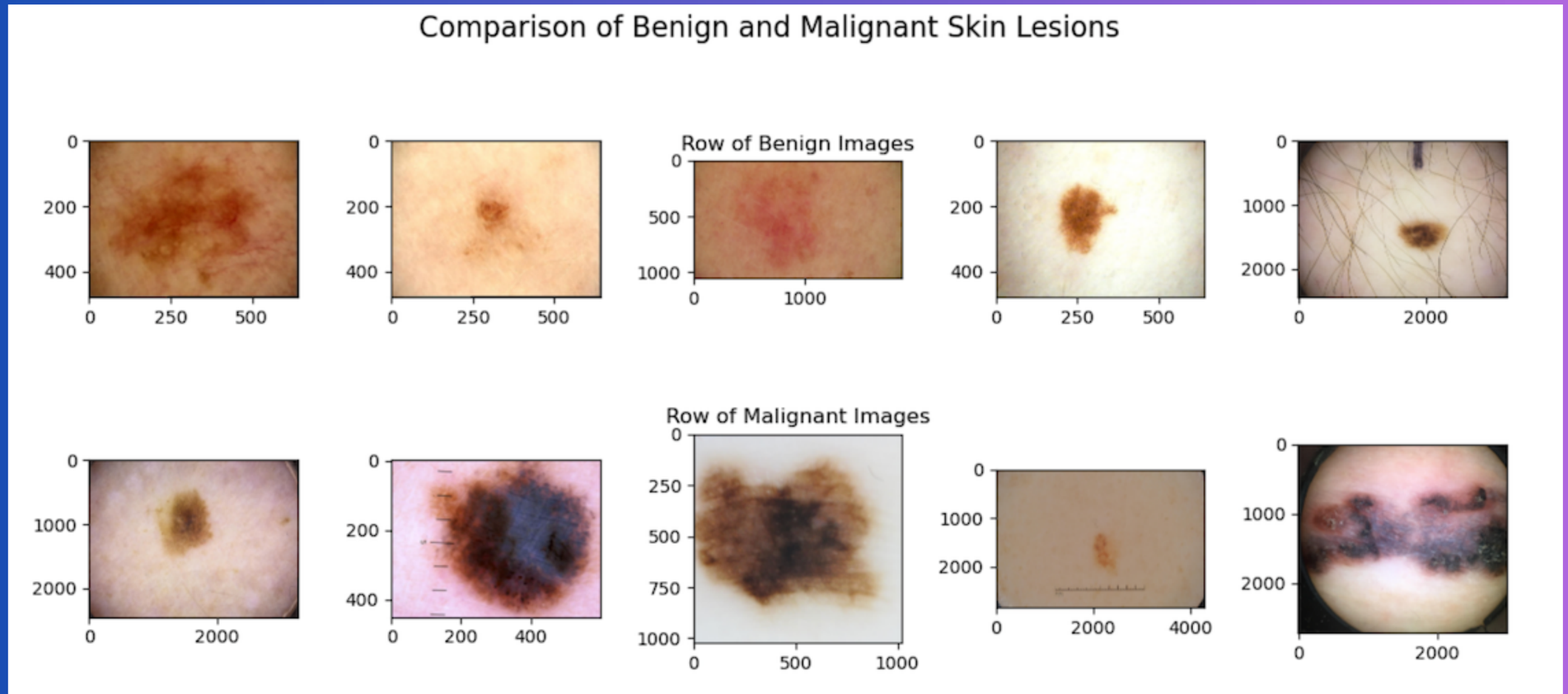
A symmetrical **B** order **C** olor **D** iameter **E** volving



EDA

- The public may have difficulties in identifying these differences.

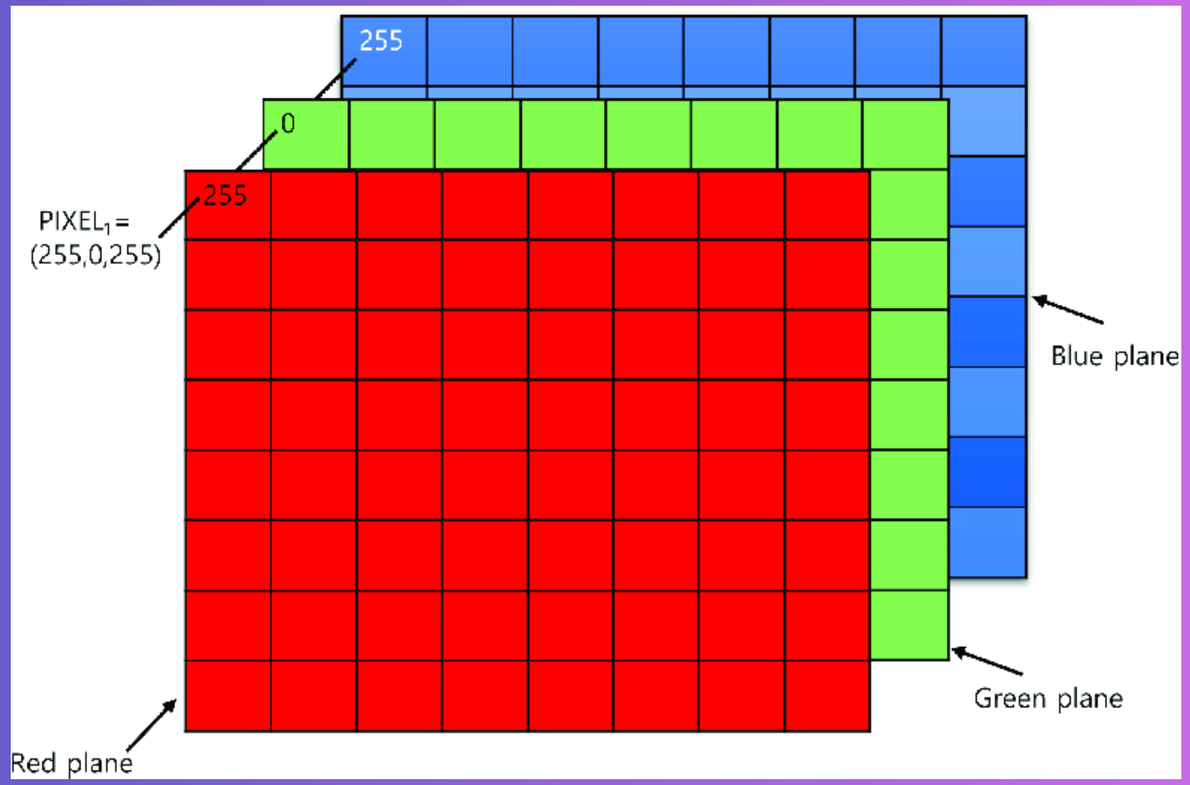
- A neural network model should be able to pick up on this hard to identify differences compared to an untrained eye.



EDA

Average Color Channel Intensity by Class

Each image in my dataset is a color image which means I can calculate the intensity of the RGB layers.

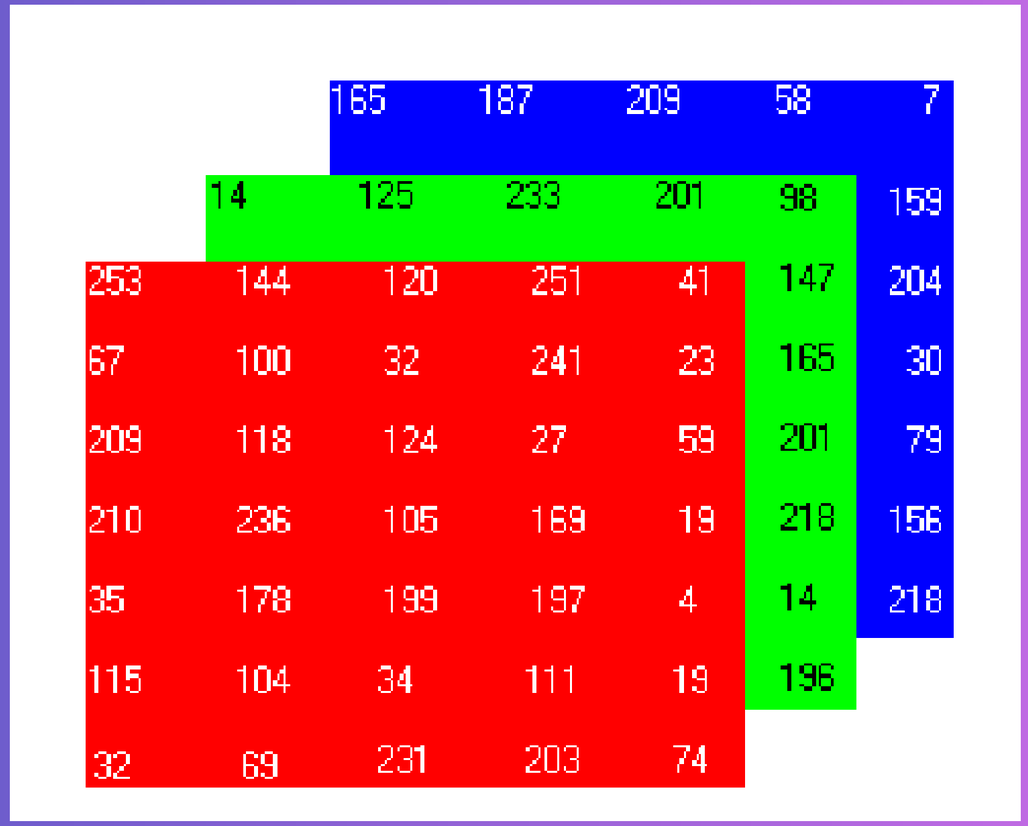


EDA

Average Color Channel Intensity by Class

For every image in each class:

1. Calculate mean value of each layer
2. Calculate the average of these mean values.

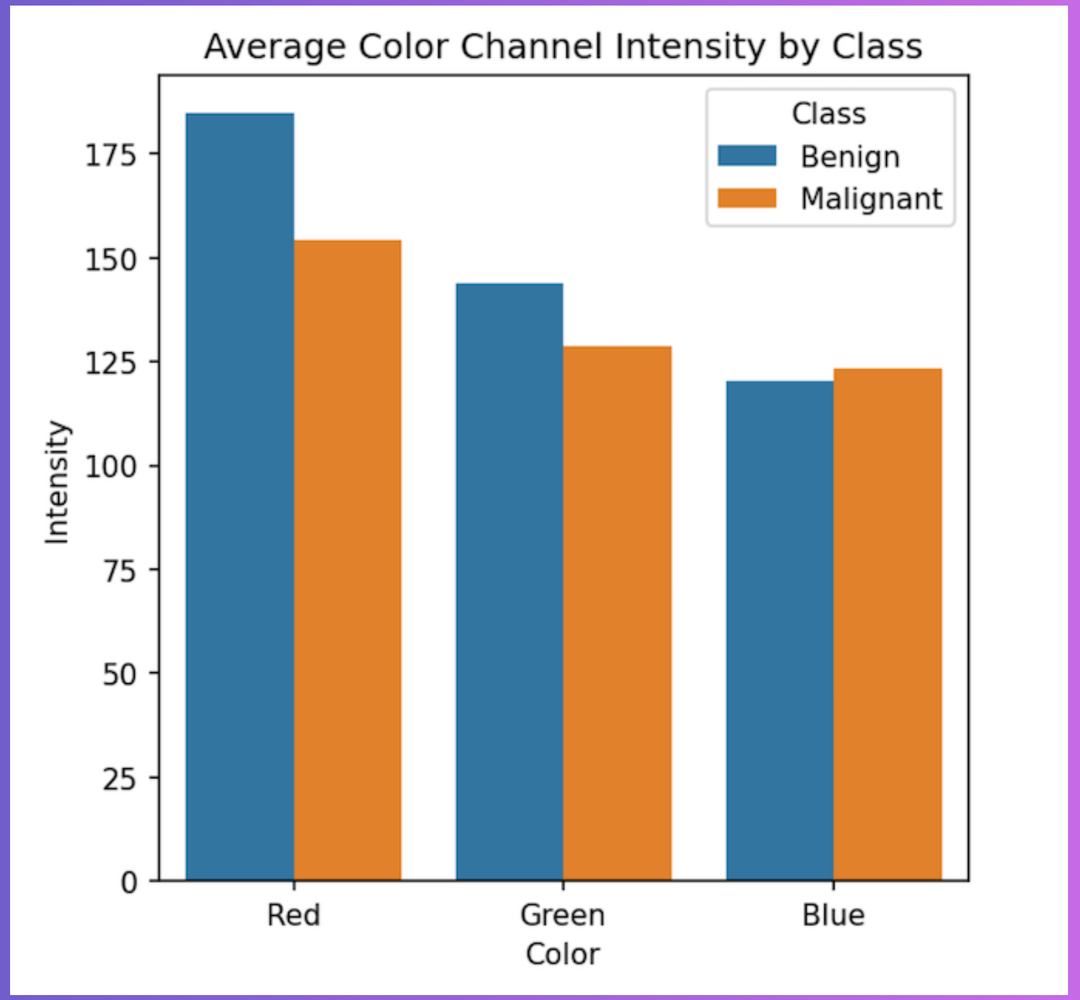


EDA

Average Color Channel Intensity by Class

There is indeed a difference between the color channel intensities of the two classes.

This is a good sign that a neural network will be able to pick up on these differences.



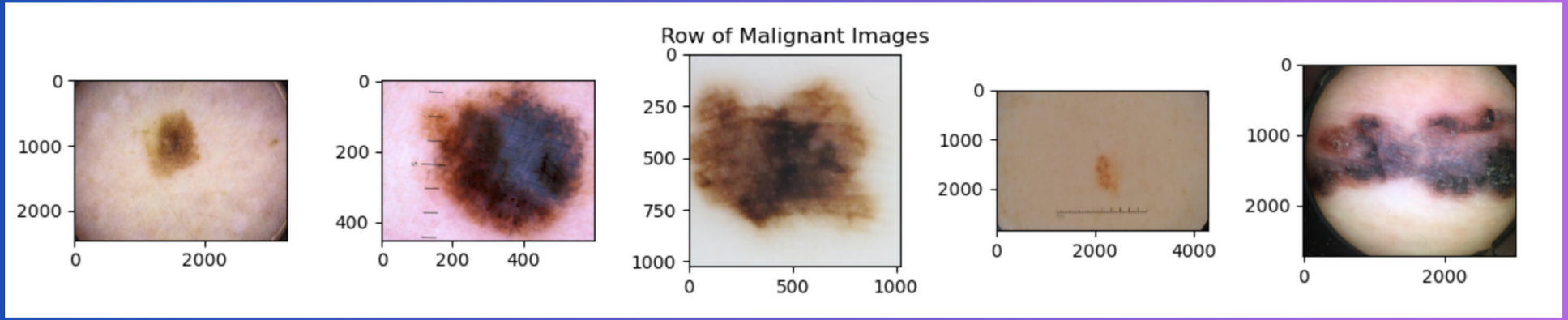
Modelling Process

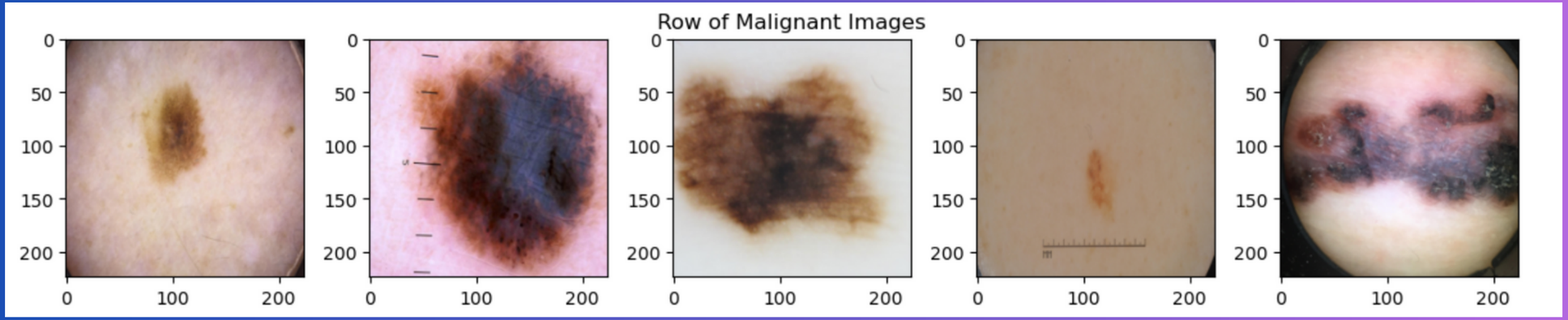
Data Pre-processing → Baseline Model → Model Tuning → Final Model → Model Deployment

Model selection will be based on its recall score.
(i.e. model's ability to correctly identify malignant skin lesions out of all the actual malignant skin lesions in the dataset.)

Data Pre-processing

Resized images to a standardised size of 224 pixels by 224 pixels to match pre-trained models input.





Baseline Model Selection

Shortlisted Pre-trained Models

MobileNetV3	EfficientNetB3

Both produce:

1. SOTA results in computer vision problems.
2. Balance model complexity and accuracy compared to other pre-trained models.

Selected Pre-trained Model

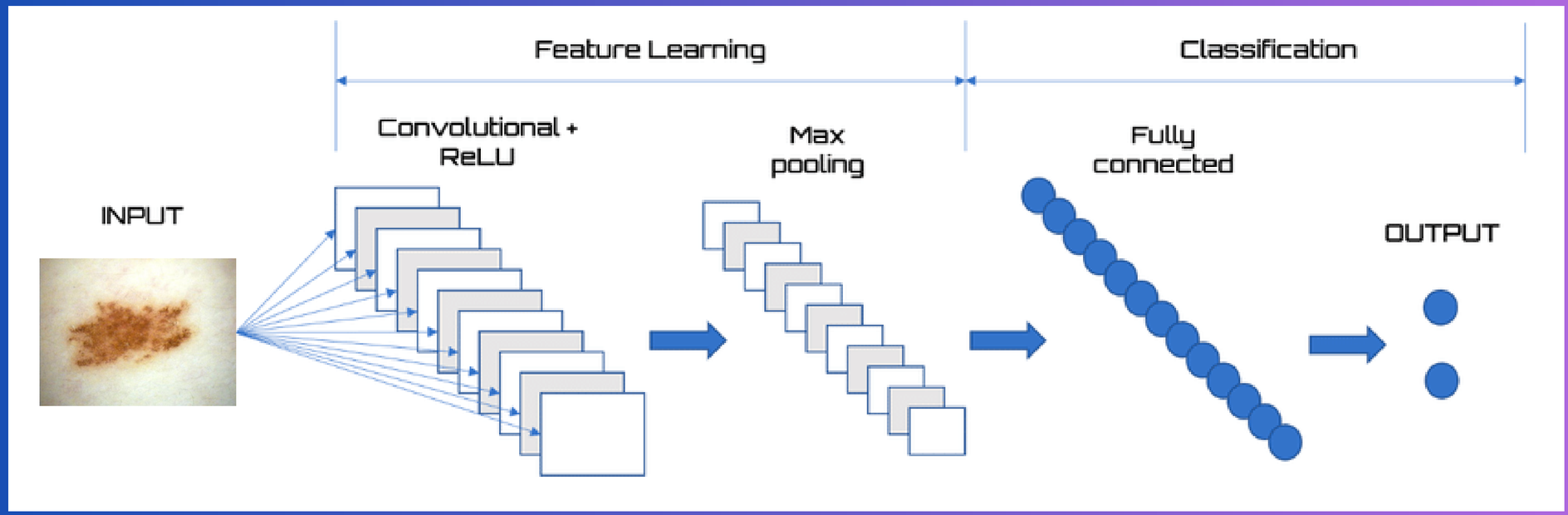
MobileNetV3

1. Better mean recall score (0.81 vs 0.77)
2. Quicker to train
3. Less computationally intensive
4. Better for model deployment

Model Tuning

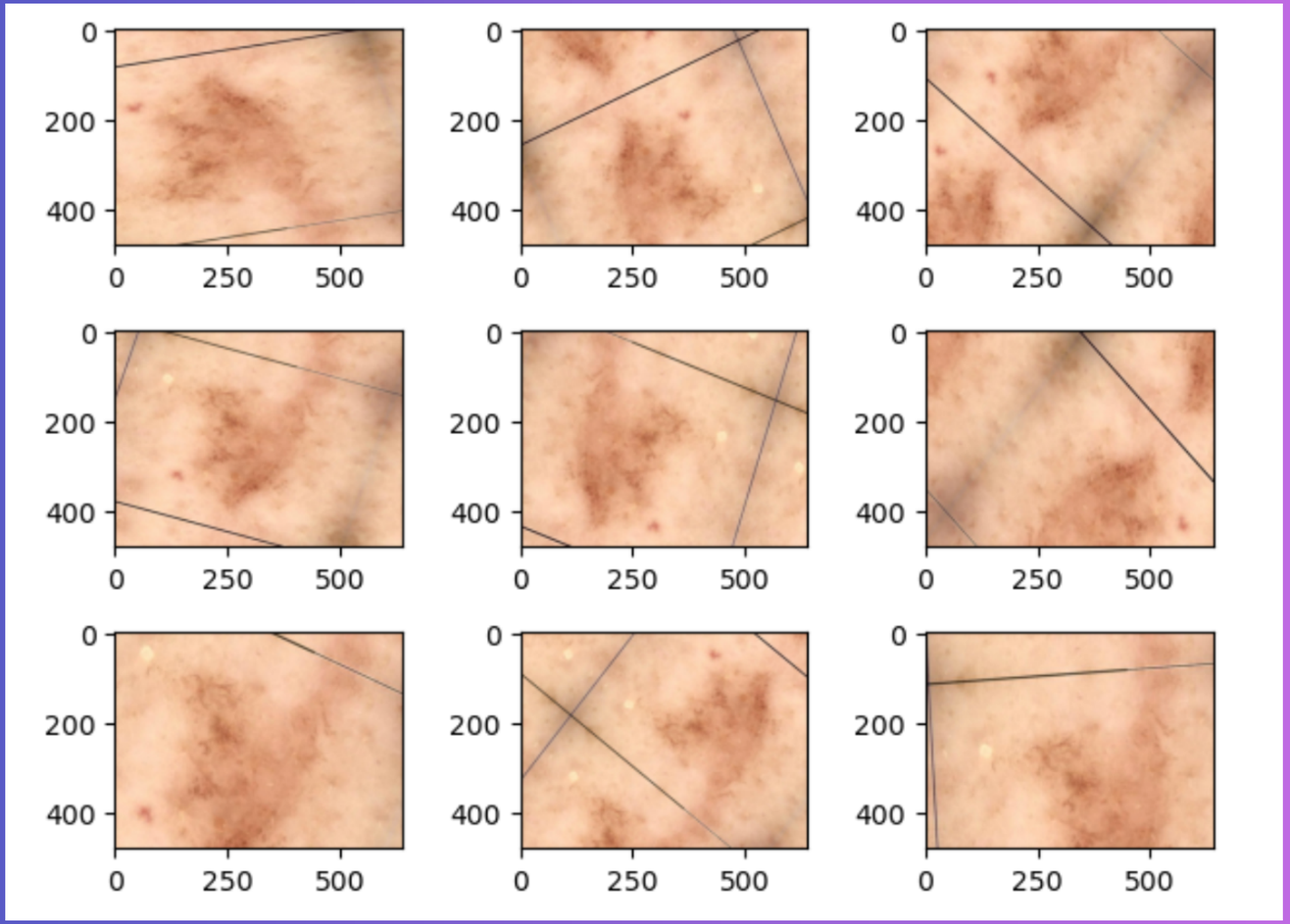
Following methods were implemented:

1. Changing global pooling Layer
2. Increasing hidden layers
3. Data augmentation



Model Tuning - Data Augmentation

Image data was augmented (e.g. zoomed in, rotated, shifted, etc.)



Model Summary

Note: For a full breakdown on the hyperparameters for each model (number of epochs, batch size, type of global pooling layer, etc.), as well as the specific differences between each model, please view the code notebook.

Model	Baseline Model - MobileNetV3	Improved Model #1	Improved Model #2	Improved Model #3	Improved Model #4	Final Model
Main Difference from previous model	Pre-trained model straight to output layer	Global Max Pooling instead of Global Average Pooling	Increasing number of hidden layers	Even more hidden layers	Augmentation of Images	Pre-trained model with dropout layer included
Train Recall Score	0.87	0.87	0.85	0.85	0.83	0.85
Validation Recall Score	0.81	0.73	0.77	0.76	0.69	0.84
Holdout Recall Score	n.a	n.a	n.a	n.a	n.a	0.96

After model tuning:

1. The best model so far is the baseline model.
2. Pre-trained model may already have sufficient complexity and capacity.

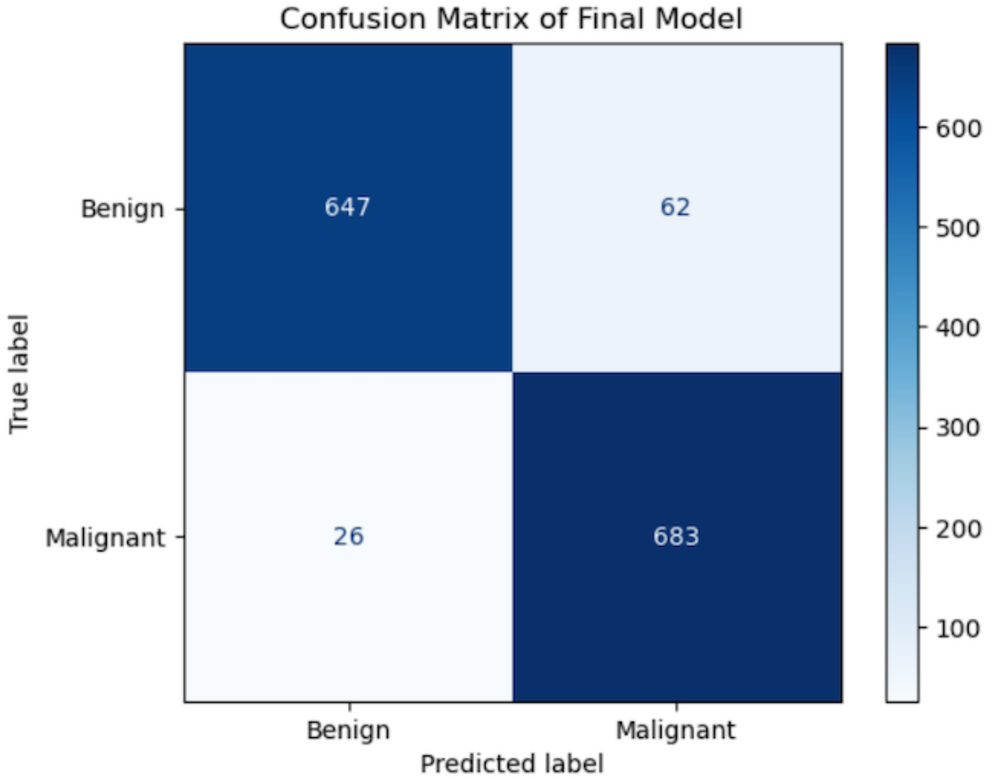
The final model, was the baseline model but with one dropout layer before the output layer to reduce overfitting.

Final Model Performance

Train Recall Score	0.85
Validation Recall Score	0.84
Holdout Recall Score	0.96

Holdout Set Confusion Matrix and Classification Report

	precision	recall	f1-score
Benign	0.96	0.91	0.94
Malignant	0.92	0.96	0.94



Model Deployment - Streamlit

https://btanyh-streamlit-apps-skin-lesion-classifier-f4yhts.streamlit.app/

App Showcase

Skin Lesion Classifier

Please upload an image of a skin lesion in either jpg or png extension.

Drag and drop file here
Limit 200MB per file • JPG, PNG
Browse files

Please upload an image file to be classified as benign or malignant.

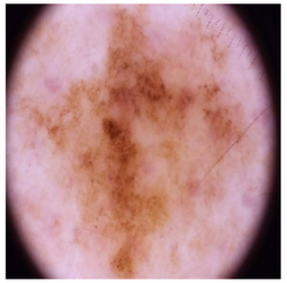
After uploading the image, the program will automatically classify the image as either malignant melanoma or benign with recommended next steps.

Skin Lesion Classifier

Please upload an image of a skin lesion in either jpg or png extension.

Drag and drop file here
Limit 200MB per file • JPG, PNG
Browse files

ISIC_0026647.JPG 27.8KB



Your skin lesion is suspected to be malignant melanoma. Please visit a doctor IMMEDIATELY!

Conclusion

Task	Result	Outcome
Develop an accurate neural network model.	Recall score of 0.96	
Deploy the model to an app which the public can use.	Functioning app on Streamlit	
The goals of the project have been achieved		

Cost Benefit Analysis

Each time the app is used successfully in early detection, massive financial savings will be achieved by both the healthcare sector and the patient.

Treatment Cost per Year	
Stage I	Stage IV
USD 15,000	USD 152,000

Recommendations for Future Works

1. More Image Pre-processing can be done

2. More advanced pre-trained models

3. Incorporating a diverse dataset.


Thank
You