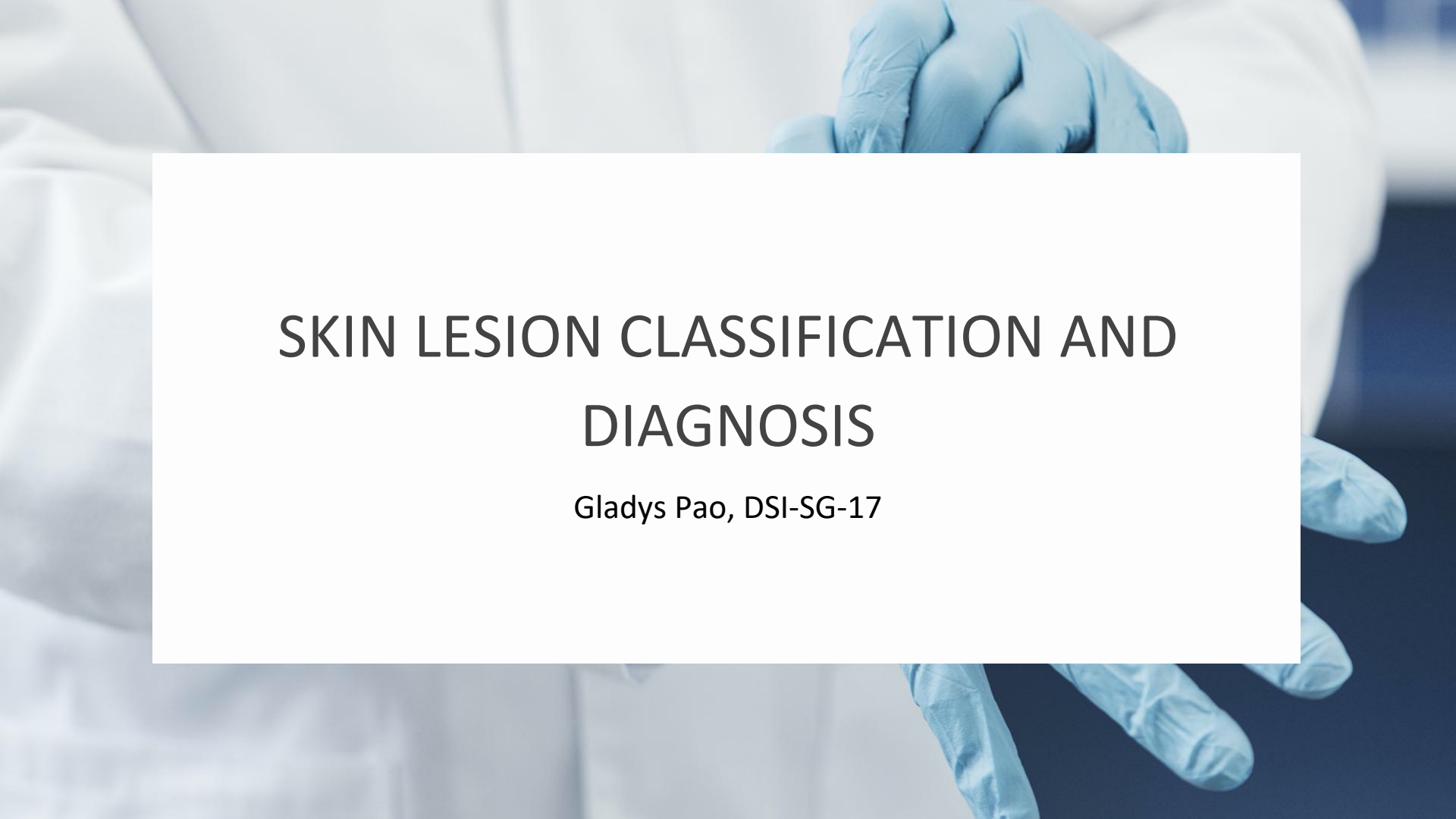


SKIN LESION CLASSIFICATION AND DIAGNOSIS

Gladys Pao, DSI-SG-17

PROBLEM STATEMENT

01 THE PROBLEM

- Skin cancer is the most common cancer globally
- Numerous lesions being misdiagnosed despite usage of dermoscopy
- $673 million in overall cost

02 THE STAKEHOLDERS

- Dermatologists
- Hospital and skin clinics in Singapore

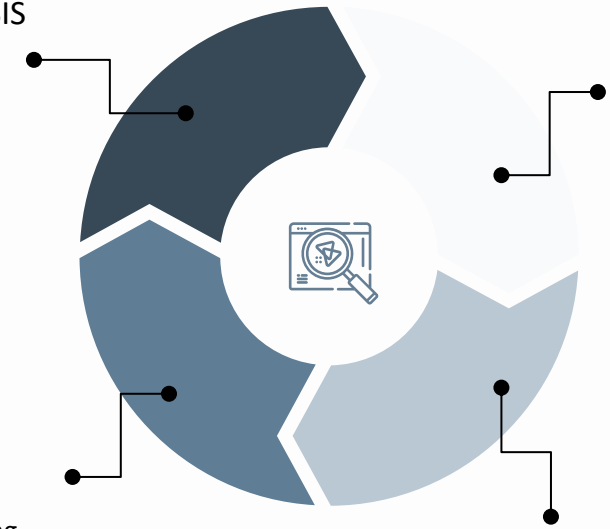
03 THE PURPOSE

- Improve diagnostic rate of skin cancer
- 2 classification tasks using Convolutional Neural Networks:
    - **Specific skin lesion diagnosis**
    - **Malignant, benign, or precancerous**
- Allow for appropriate treatments for patients with skin lesions/cancers

METHODOLOGY

STEP 1:
EXPLORATORY DATA ANALYSIS

- Analyse trends in labelled and image data



STEP 2: IMAGE PREPROCESSING

- Standardise images (size, colour etc.)
- Data augmentation

STEP 3: MODELLING AND EVALUATION

- Convolutional Neural Networks
- Evaluation metrics (Accuracy and Recall)

STEP 4: CONCLUSION

- Recommendations moving forward

THE DATASET

DATA SOURCES

TOTAL NUMBER OF IMAGES

10,276

MAIN DATASETS

International Skin Imaging Collaboration (ISIC):

- 2018, 2019 and 2020 datasets

EXTERNAL DATASETS

- DermNet NZ
- Dermoscopedia
- 7-point Criteria Evaluation Database

TARGET VARIABLES

7 SKIN LESIONS

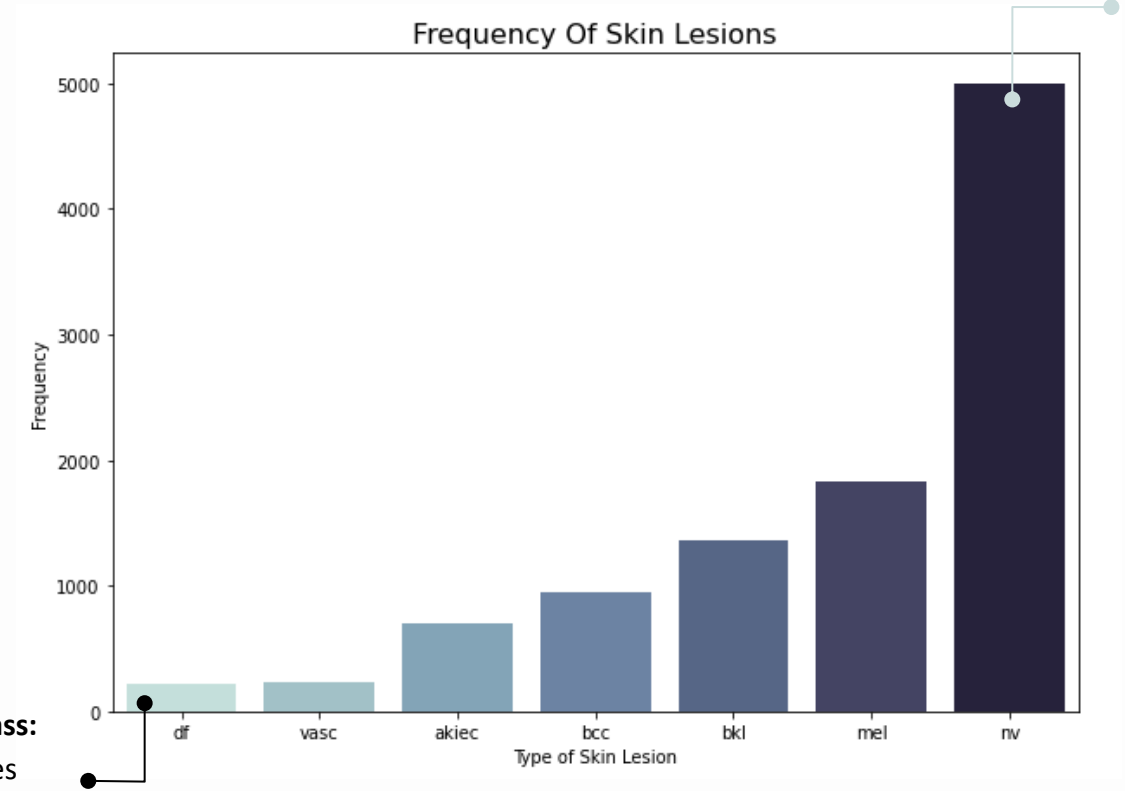
- Melanocytic Nevi (NV)
- Benign keratoses (BKL)
- Vascular lesions (VASC)
- Dermatofibroma (DF)
- Melanoma (MEL)
- Basal cell carcinoma (BCC)
- Actinic Keratoses (AKIEC)

3 TYPES OF DIAGNOSIS

- Benign
  - NV, BKL, VASC, DF
- Malignant
  - MEL, BCC
- Precancerous
  - AKIEC

EXPLORATORY DATA ANALYSIS

CLASS IMBALANCE



Majority class:
NV with
4992 images;

48.6%

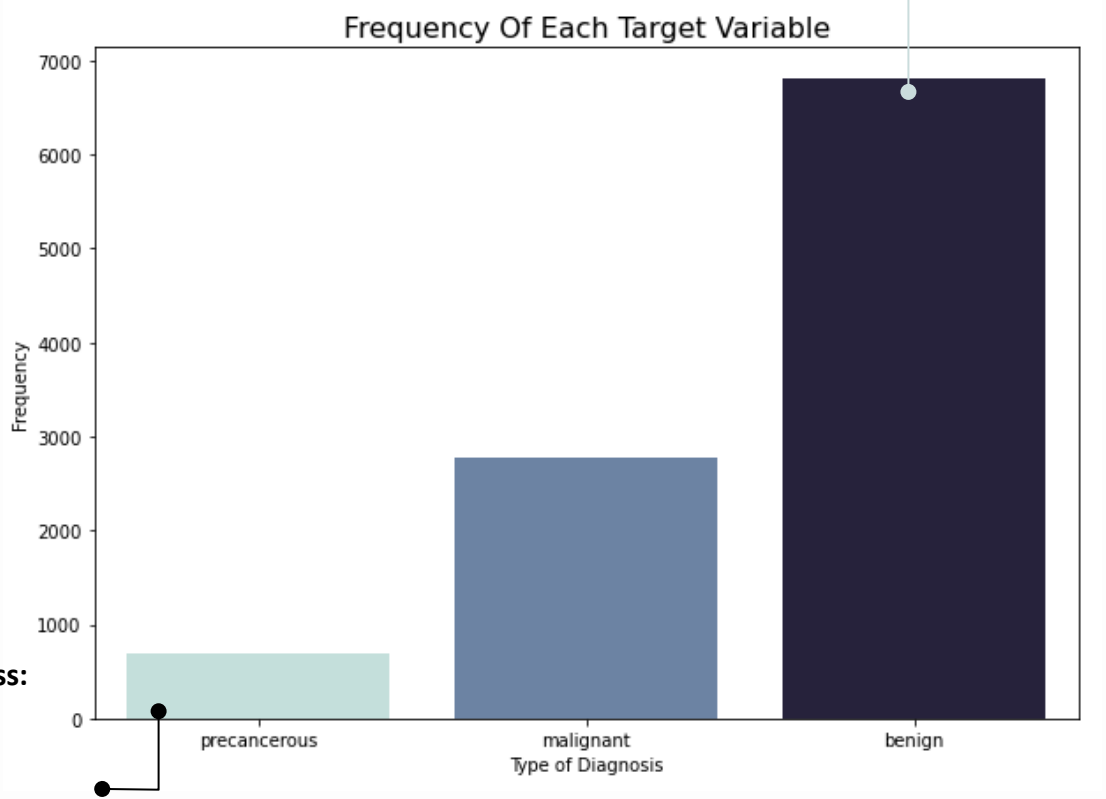
Lowest minority class:
DF with 217 images

2.11%

CLASS IMBALANCE

Majority class:
Benign with 6801 images

66.2%



Lowest minority class:
Precancerous with
696 images

6.8%

IMAGE SIZES

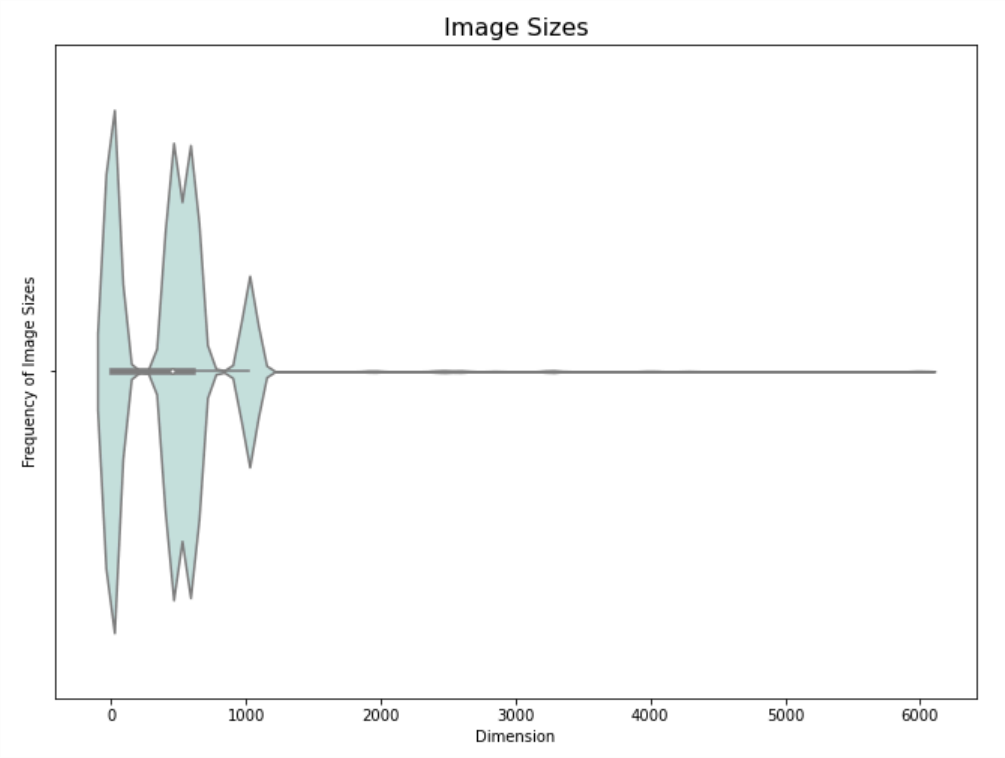
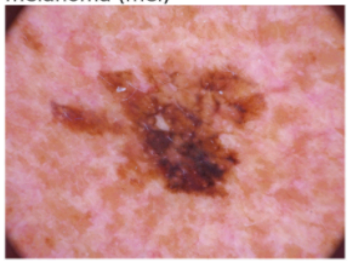


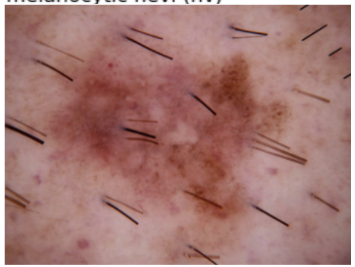
Image size	(1983, 1983)	(1775, 1775)	(4288, 2848)	(3024, 4032)	(2592, 3888)	(2317, 2317)	(511, 768)	(3456, 5184)	(2848, 4288)	(222, 294)	(4000, 6000)	(1936, 2592)	(2448, 3264)	(512, 768)	(1024, 1024)	(450, 600)
frequency	1	1	1	1	1	2	3	6	13	20	20	32	41	61	1770	8303

IMAGE ANALYSIS

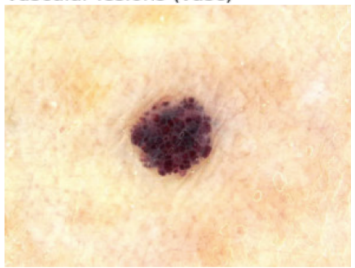
melanoma (mel)



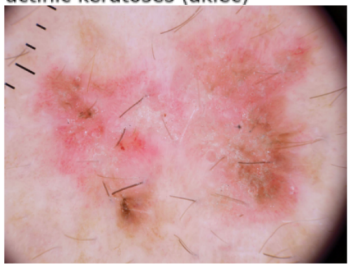
melanocytic nevi (nv)



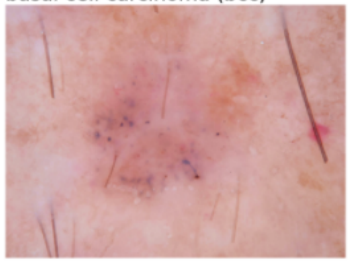
vascular lesions (vasc)



actinic keratoses (akiec)



basal cell carcinoma (bcc)



benign keratosis-like lesions (bkl)

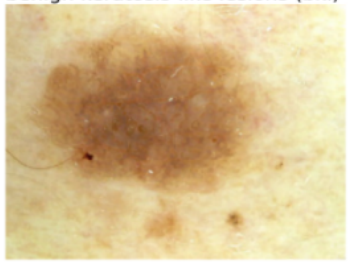


IMAGE PREPROCESSING

BLACK HAIR AND MARKING REMOVAL

ORIGINAL IMAGE

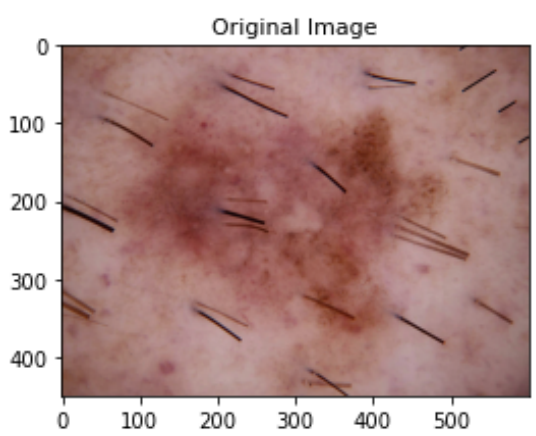
GREYSCALED IMAGE

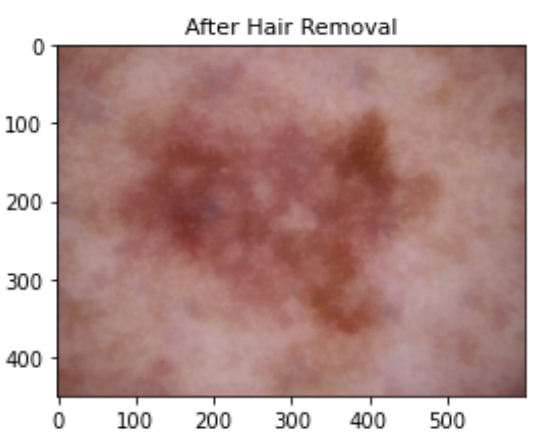
AFTER MORPHOLOGICAL FILTERING:
BLACK HAT FILTERING

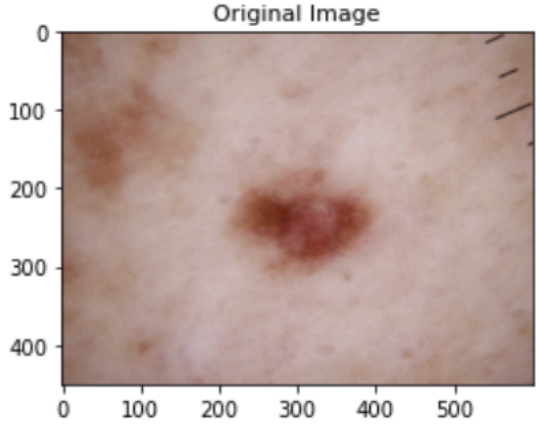
INCREASING THRESHOLD

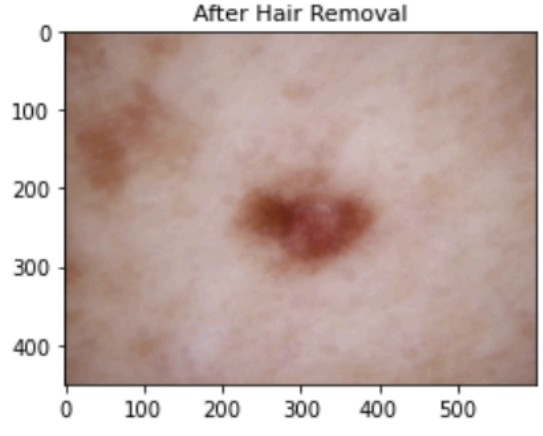
INPAINTING
(IMAGE INTERPOLATION)

BLACK HAIR AND MARKING REMOVAL

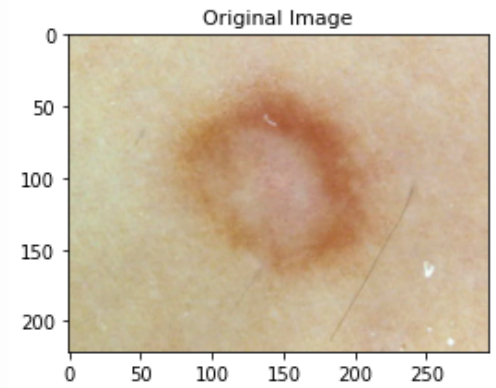


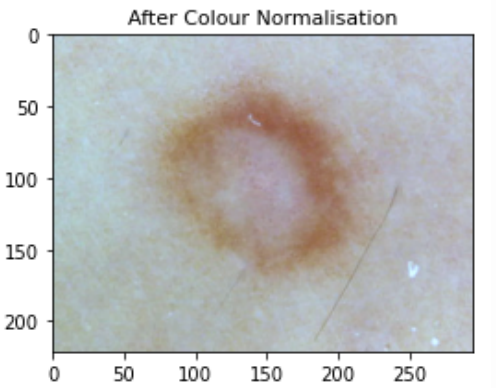


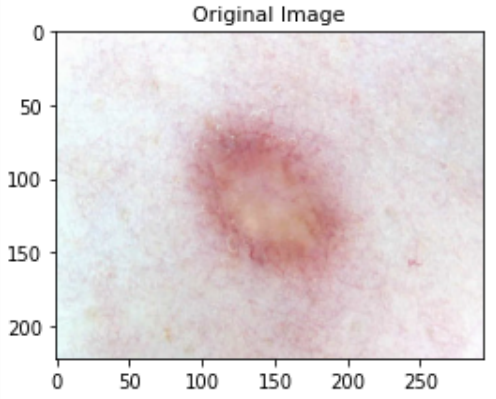


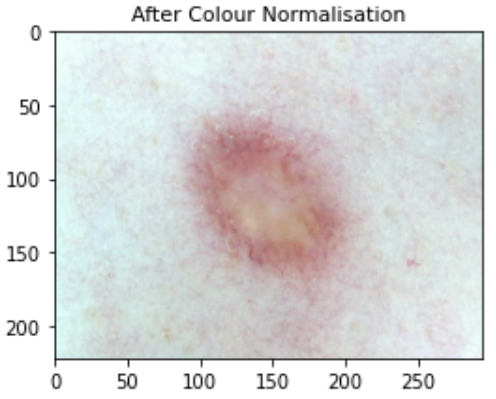


COLOUR NORMALISATION

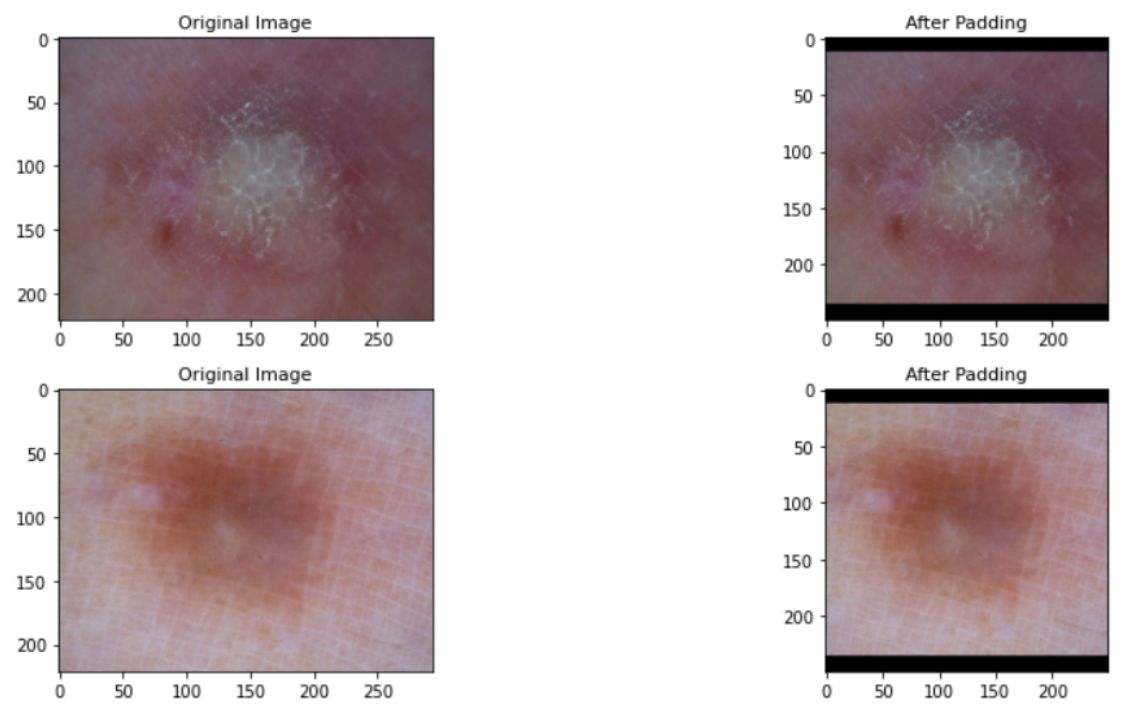




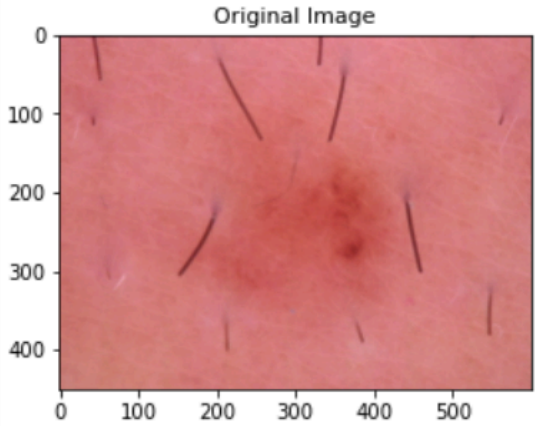


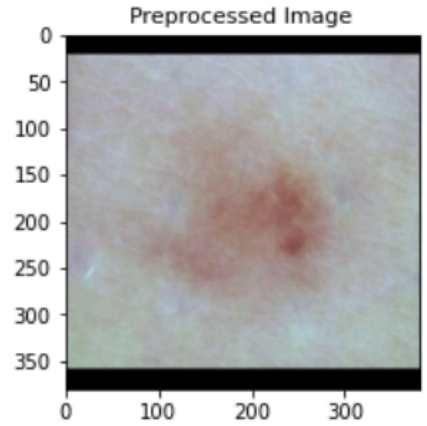


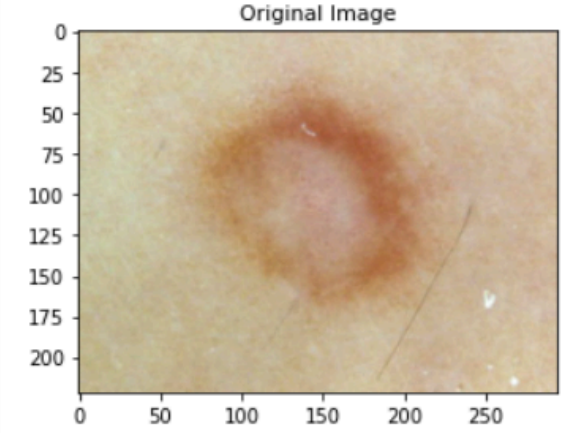
PADDING, CROPPING AND RESIZING

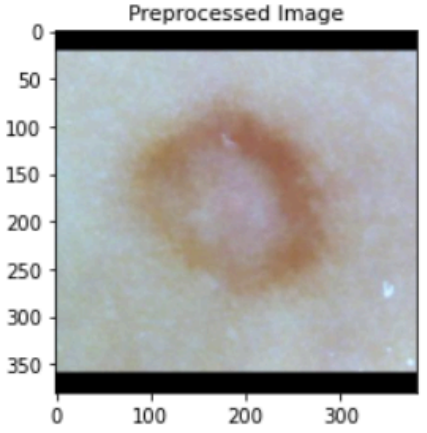


ORIGINAL VERSUS PREPROCESSED









SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE (SMOTE)

- Data augmentation for minority classes
- Generates synthetic samples for the minority classes
- Obtain a synthetically class-balanced training set, which is then used to train the classifier

TASK 1:
SKIN LESION CLASSIFICATION

20,965

images

TASK 2:
DIAGNOSIS CLASSIFICATION

12,240

images

MODELLING AND EVALUATION

CONVOLUTIONAL NEURAL NETWORKS

- Similar model architecture for both tasks
- Tuned parameters for best performance
- The usage of Batch Normalization layer right before activation layer
    - Standardise inputs
    - Model trained faster
    - Better model performance
    - Better regularization effect (compared to Dropouts)

- ReduceLROnPlateau
    - Keras callback API
    - Reduce learning rate by a factor of 0.1 and minimum learning rate of 0.00001 (10e-6)
    - Combination of this feature with initial learning rate of 0.0001 (10e-5) helped to improve model performance

Task 01

SKIN LESION CLASSIFICATION

7-way Classification

BASELINE MODEL:

Artificial Neural Networks

56.8%

ACCURACY RATE

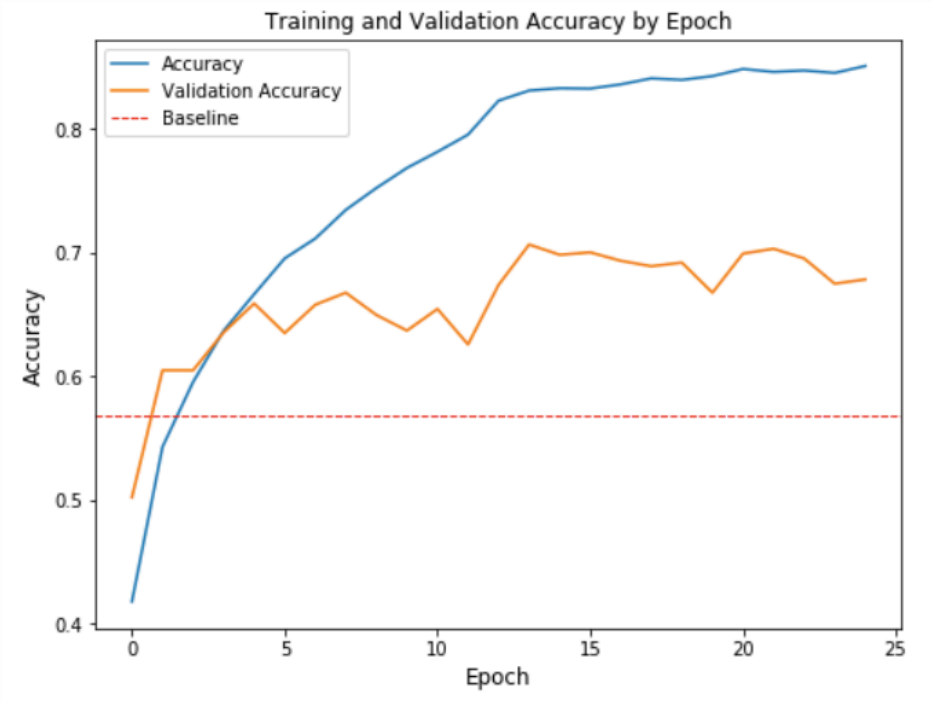
75.6%

TOP 2
ACCURACY RATE

38.4%

AVERAGE
RECALL RATE

CONVOLUTIONAL NEURAL NETWORKS



70.7%

ACCURACY RATE

86.2%

TOP 2
ACCURACY RATE

57.7%

AVERAGE
RECALL RATE

MODEL EVALUATION

70.5%

TEST
ACCURACY RATE

0.1%

DIFFERENCE

86.2%

TEST TOP 2
ACCURACY RATE

0%

DIFFERENCE

60.7%

TEST AVERAGE
RECALL RATE

3%

DIFFERENCE

Task 02

DIAGNOSIS CLASSIFICATION

3-way Classification: Benign, Malignant or Precancerous

CONVOLUTIONAL NEURAL NETWORKS

75.8%

ACCURACY
RATE

>

65.8%

DERMATOLOGISTS'
ACCURACY
RATE

63.6%

AVERAGE
RECALL RATE

>

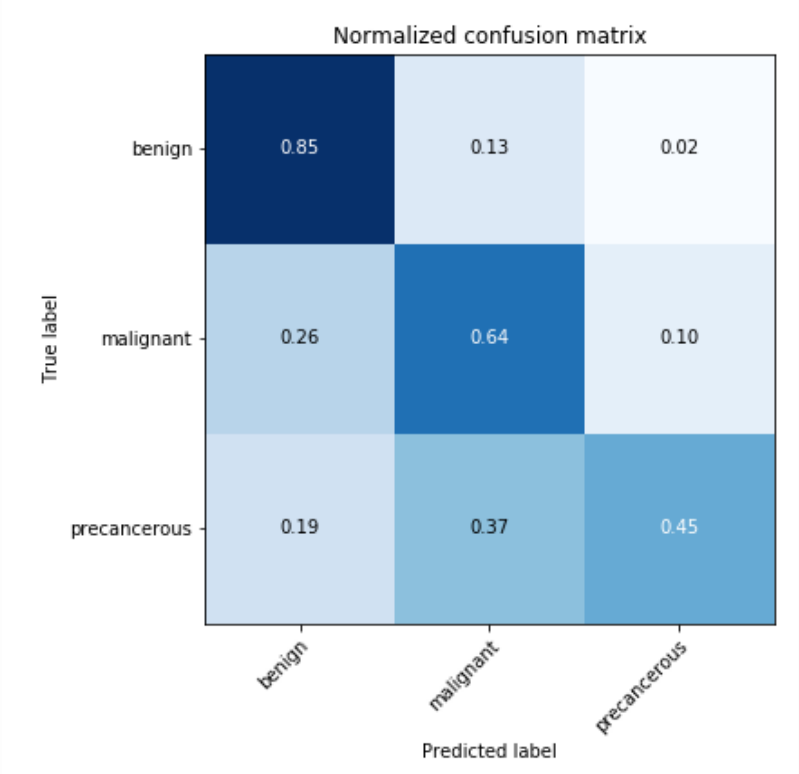
49.2%

DERMATOLOGISTS'
AVERAGE
RECALL RATE

MODEL EVALUATION

	Accuracy	Average Recall Rate
Validation	75.8%	63.6%
Test	76.4%	64.4%
Difference	**0.6%**	**0.8%**

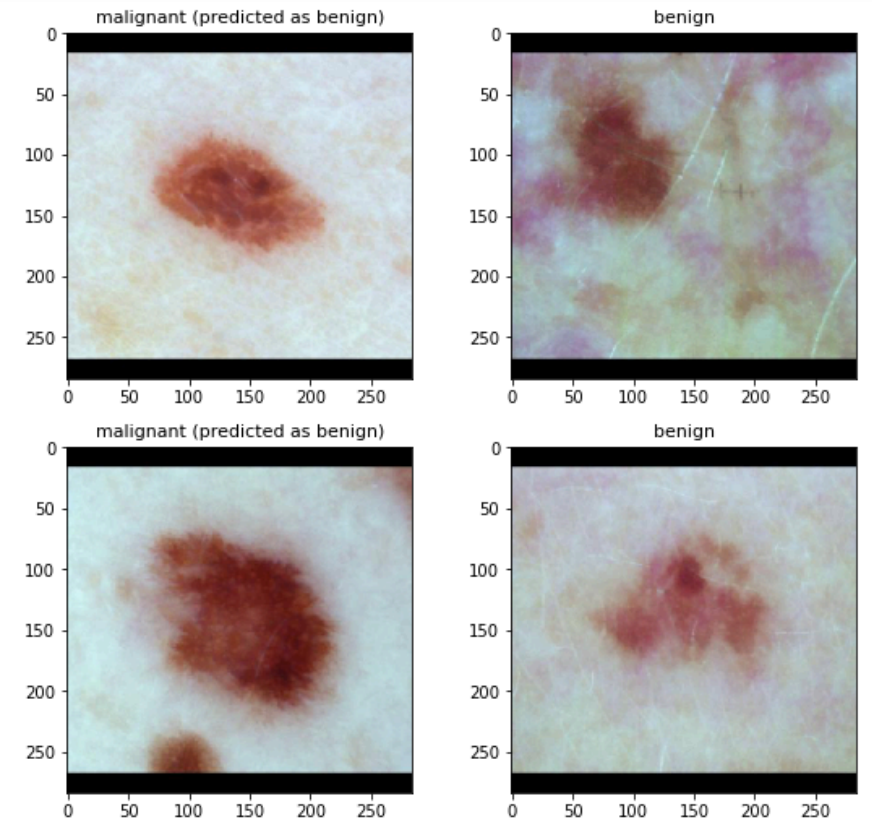
MODEL EVALUATION



MODEL EVALUATION

	Recall Rate
Benign	84.6%
Malignant	64.0%
Precancerous	44.6%
Weighted Recall Rate	76.4%

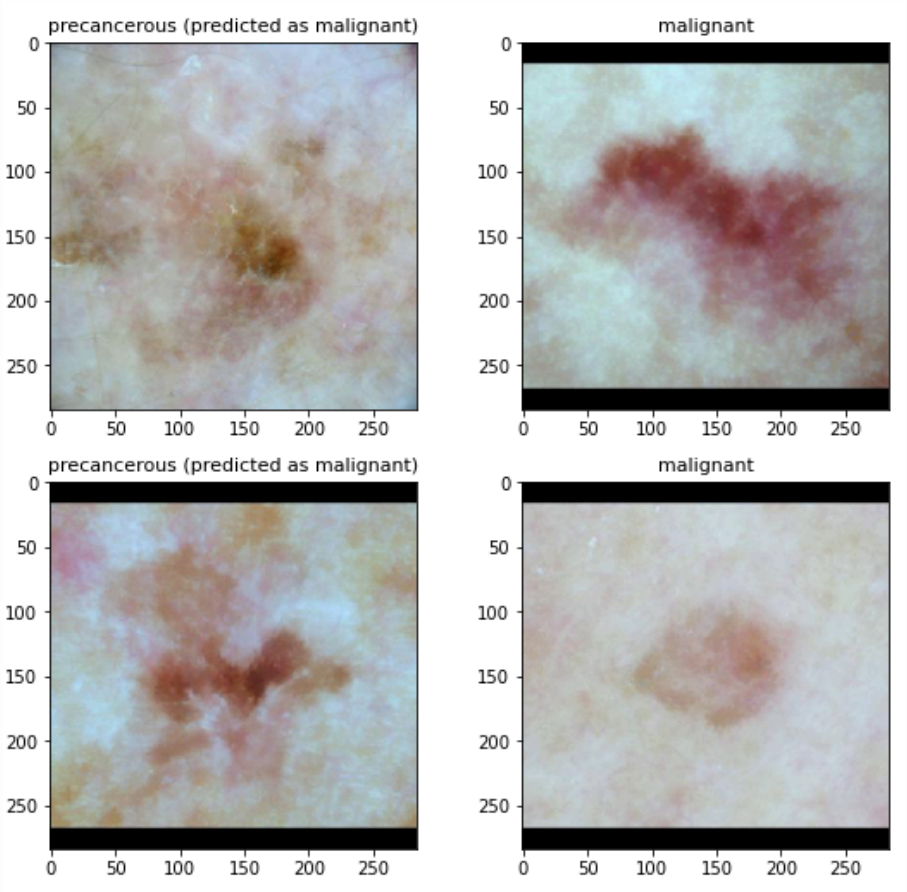
WRONG PREDICTIONS (MALIGNANT)



Lichen-planus like keratoses (LPLK)

- Show features that are similar to MEL

WRONG PREDICTIONS (PRECANCEROUS)



RECOMMENDATIONS AND CONCLUSION

RECOMMENDATIONS

1. Collect more data for minority classes
2. Build a web application for secure in-house usage at hospitals and skin clinics
3. Add additional medical data (metadata of patients, and characteristics of lesions)
4. Expand classification task to include other type of lesions
    - Non-neoplastic lesions (diagnosis classification)
    - Cystic lesions (skin lesion classification)

CONCLUSION

- Moderately high accuracy at 70.5% for the skin classification task, and 76.4% for the diagnosis classification task
- Recall rates surpassed baseline scores
- Performed very well for most benign skin lesions (especially for diagnosis classification task)
    - Serve as great models specifically for benign skin lesions in the future
- Calibrate model with more data and train on other type of lesions
    - Better guide for dermatologists
    - Improve clinical diagnostic accuracy

THANK YOU