MACHINE LEARNING-BASED CLASSIFICATION OF SKIN LESIONS

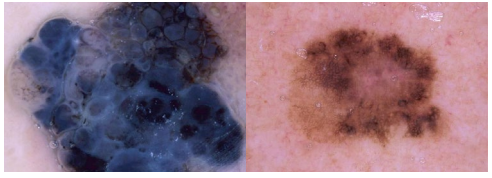
Machine Learning

Joan Carles Montero Jimenez

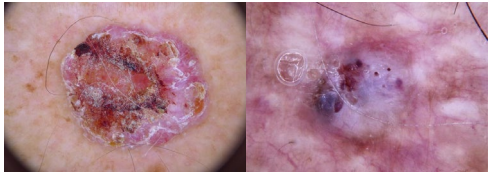
Curs 2023-24

INDEX

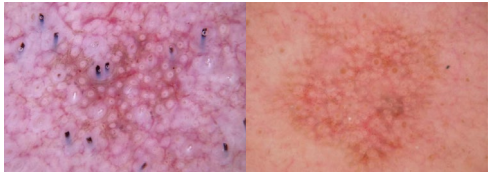
1. INTRODUCTION - Lesions classification



MEL

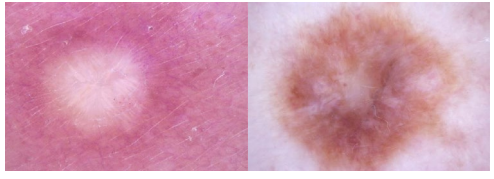


Basal Cell Carcinoma (BCC)

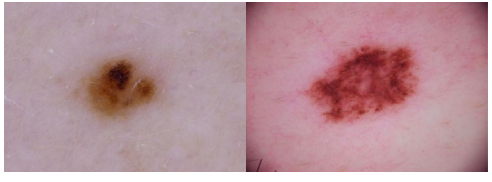


Bowen's Disease (AKIEC)

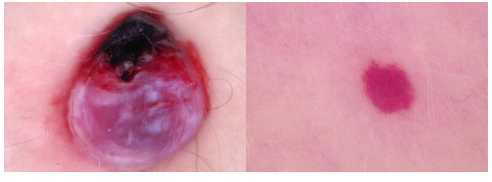
Bening Keratosis - like Lesions (BKL)



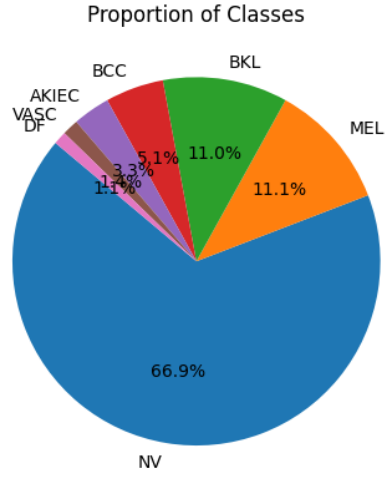
Dermatofibroma (DF)



Melanocytic nevi (NV)

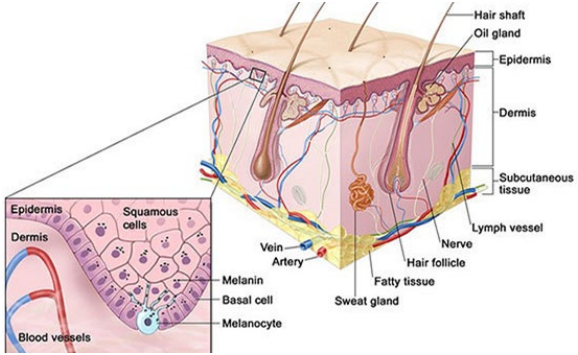


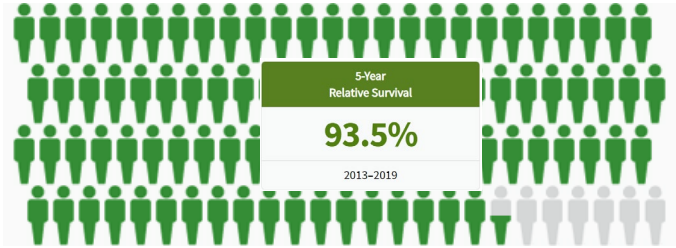
Vascular lesions (VASC)



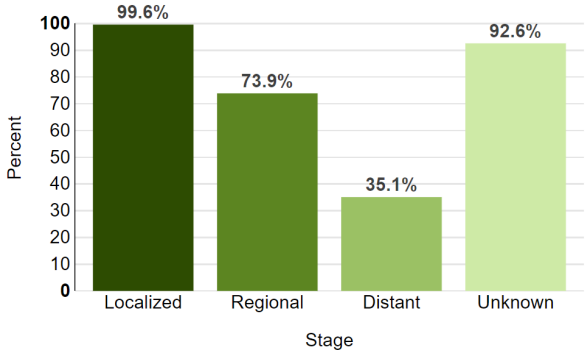
Dataset origin:
https://www.kaggle.com/datasets/volodymyrpivoshenko/skin-cancer-lesions-segmentation

1. INTRODUCTION - Melanomas





5-Year Relative Survival



	Common Types of Cancer	Estimated New Cases 2023	Estimated Deaths 2023
1.	Breast Cancer (Female)	297,790	43,170
2.	Prostate Cancer	288,300	34,700
3.	Lung and Bronchus Cancer	238,340	127,070
4.	Colorectal Cancer	153,020	52,550
5.	**Melanoma of the Skin**	**97,610**	**7,990**
6.	Bladder Cancer	82,290	16,710
7.	Kidney and Renal Pelvis Cancer	81,800	14,890
8.	Non-Hodgkin Lymphoma	80,550	20,180
9.	Uterine Cancer	66,200	13,030
10.	Pancreatic Cancer	64,050	50,550

Melanoma of the skin represents 5.0% of all new cancer cases in the U.S.

Source:
https://seer.cancer.gov/statfacts/html/melan.html

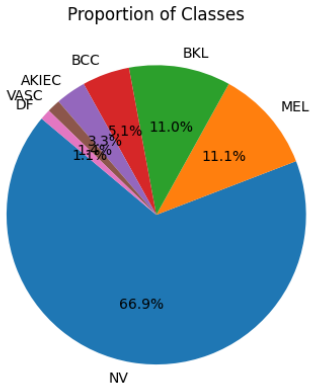
2. FIRST CONSIDERATIONS

Data reduction: 10015 images to 1001, 143 images for each class (14% each one)

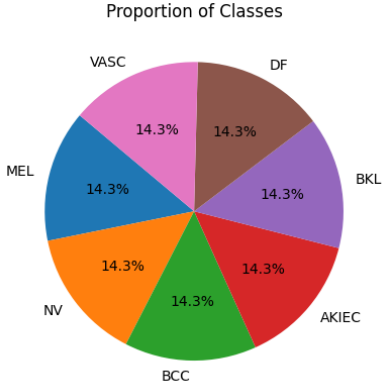
Size reduction: 600x450 pixels to 300x225 pixels

Use of masks to focus just on the lesion.

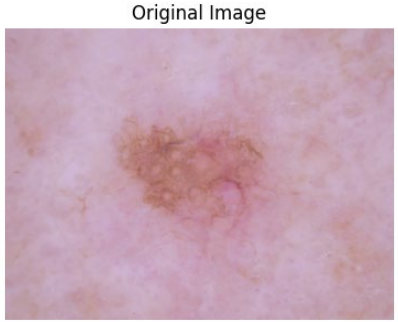
Main matric: **RECALL** as we pretend to minimize the number of FALSE NEGATIVES

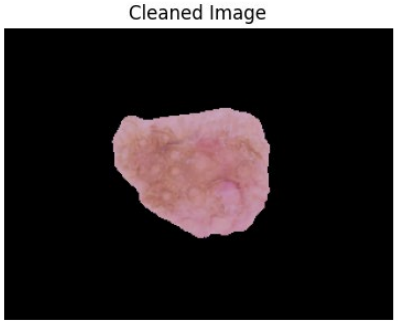


10015 images



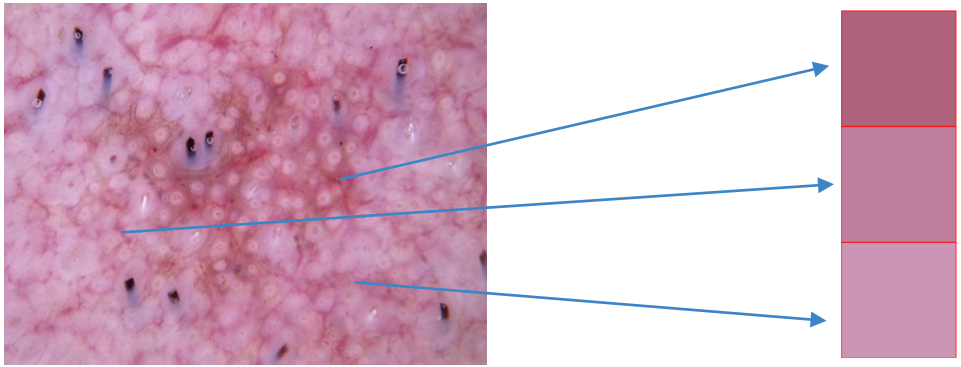
1001 images

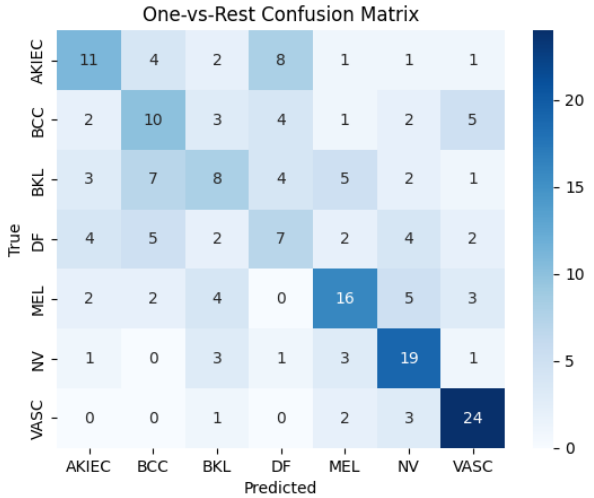




3.1 Approach with reduced dimensionality

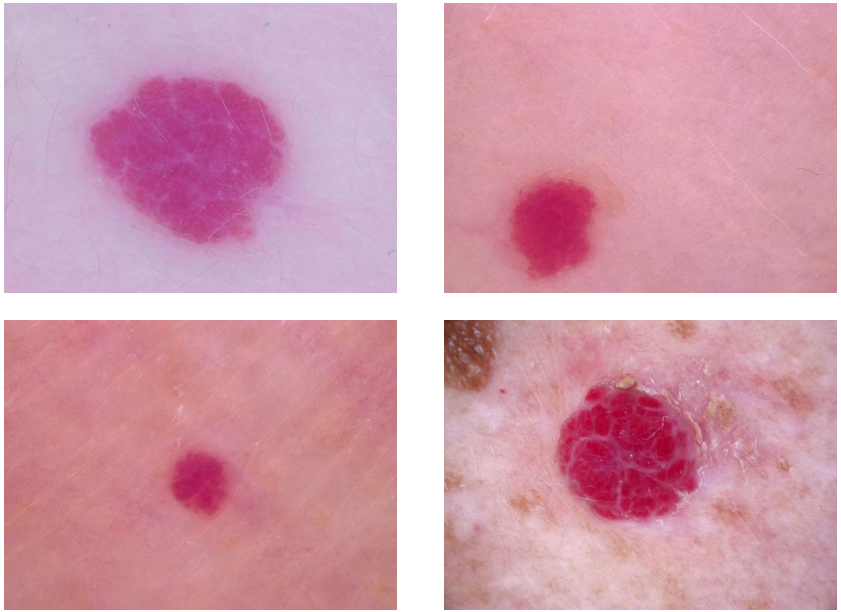
- **K** Means feature extraction with k=3, getting the colours of the 3 centroids and the colours standard deviation, in total 12 values per image (reduced from the current 300x225x3 values).
- Best classifier: One-vs-Rest with MEL recall of 50% and precision and recall global values of 45 and 47%.

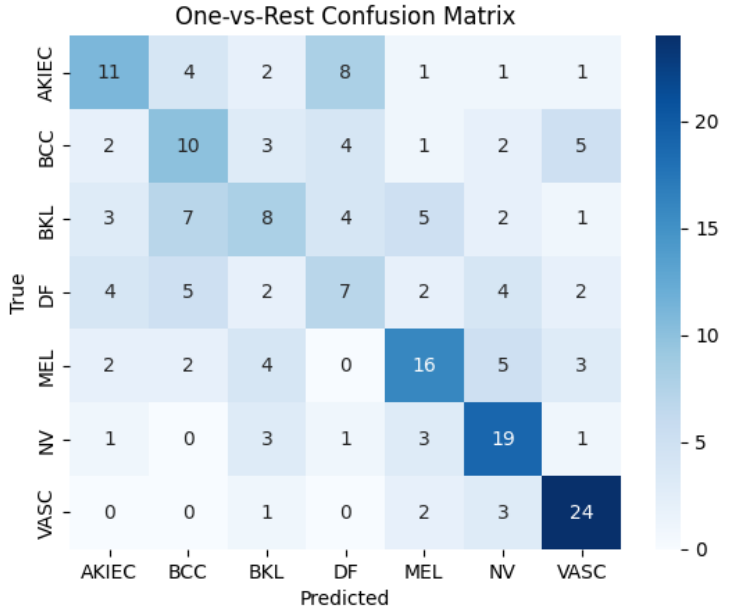




3.1 Approach with reduced dimensionality

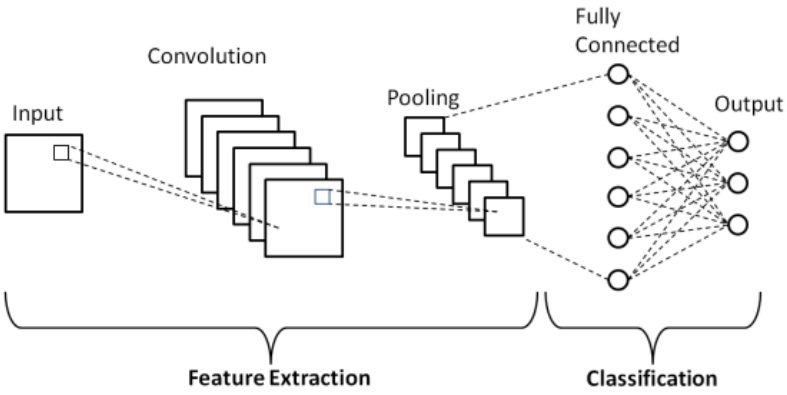
- **Observation:** great results when predicting vascular lesions, as it is a characteristic lesion with red/pink colours and circular shape.

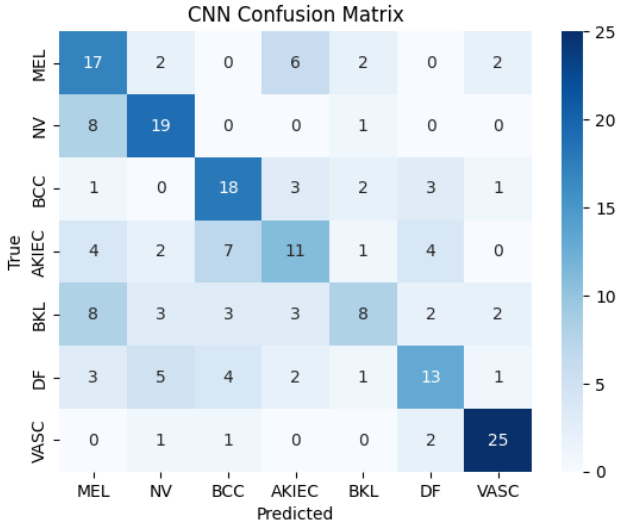




3.2 CNNs for multi class classification

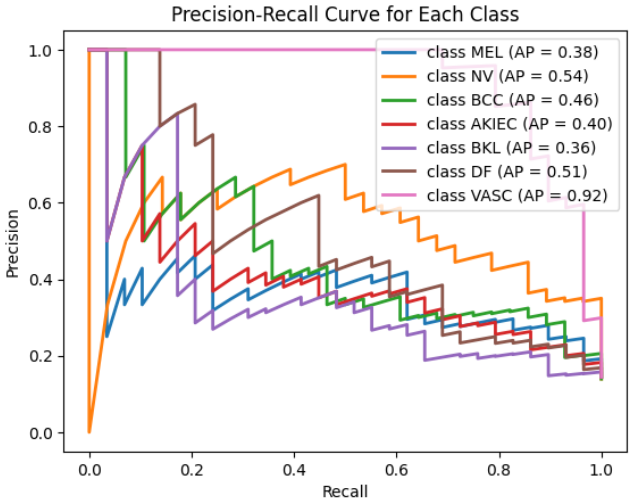
- Improvement of the global precision and recall to 55%. Recall value of 55% with melanomas and 59% with basal cell carcinomas.
- Multi class classification done by doing np.argmax(), not by a threshold (as happened with the One-vs-Rest classifier)

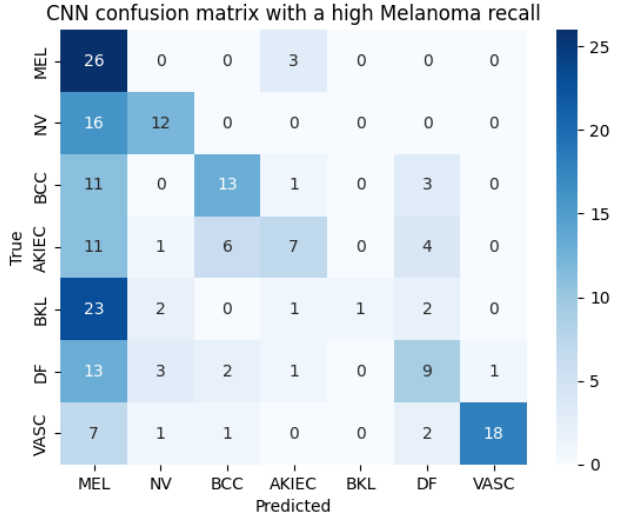




3.2 CNNs for multi class classification

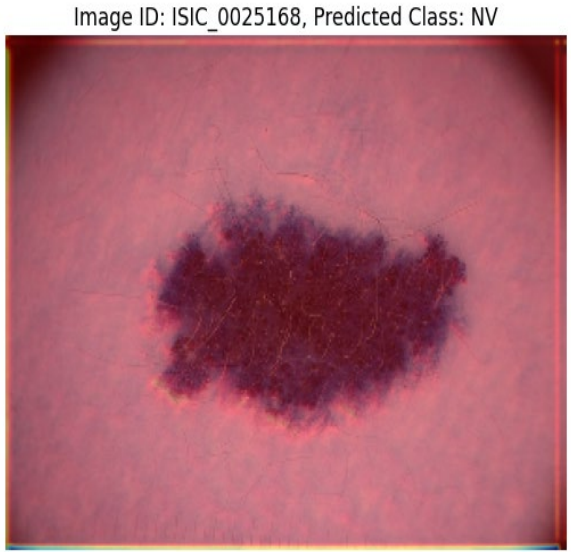
- Modifying the decision function to ensure minimum recall for melanoma of 90% by checking the individual PR curve.
- The new function will first check the assigned probability for melanoma and if it does not exceed it, it will proceed normally with np.argmax().

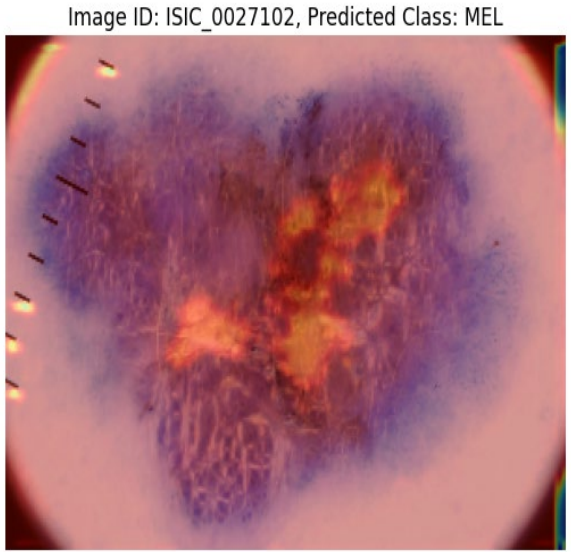


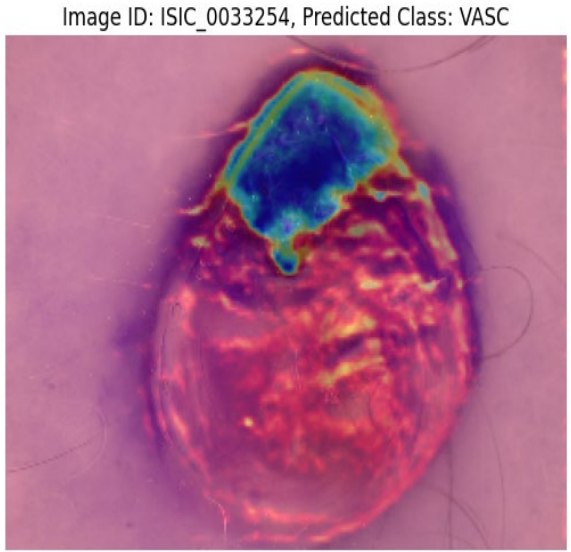


3.3 CNNs Grad - CAM

Heatmap that focuses on the **last convolutional layer** of the CNN, that captures the most abstract and detailed features of the skin lesions in the input images



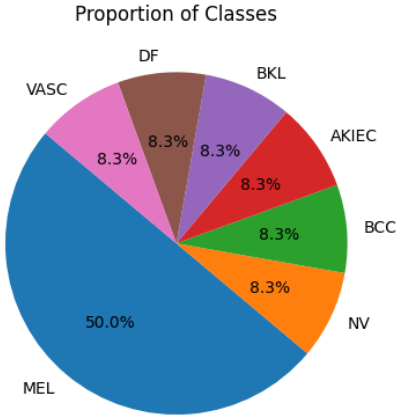


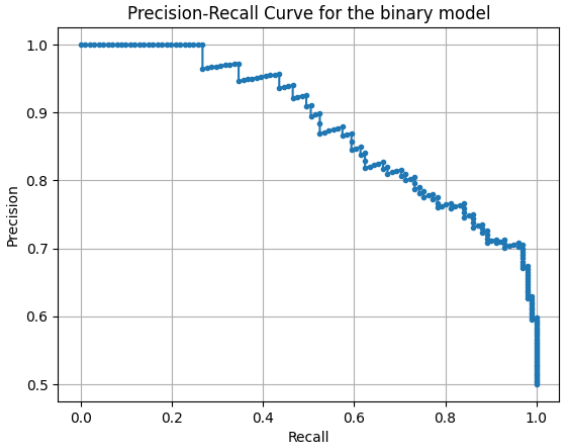


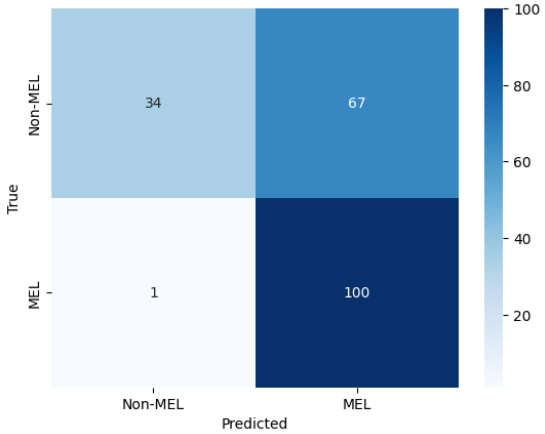
Source:
https://keras.io/examples/vision/grad_cam/

3.4 CNNs for binary classification

Required splitting the data into 50% melanomas and 50% non-melanomas.
Now, “naturally” moving the threshold to achieve 99% of recall with great results.



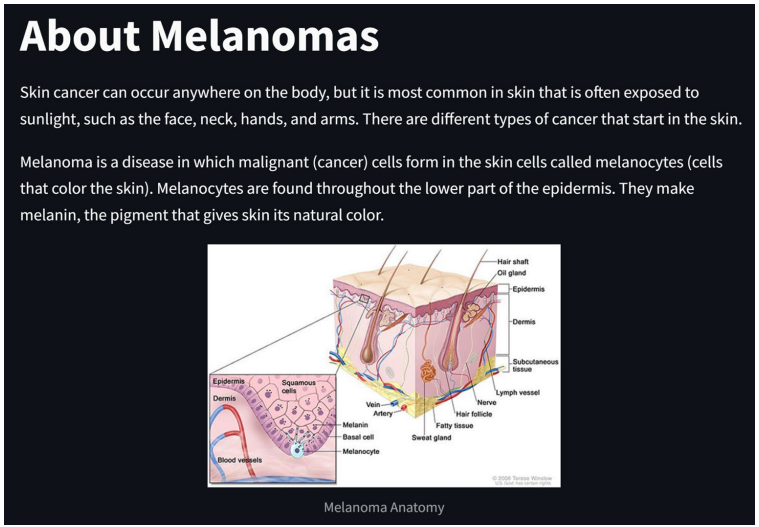




3.5 Streamlit app implementation

Two screens. A melanoma information screen and the home app, to make predictions about uploaded images and instantly get all probabilities as well as GRAD-MAPs.

Built with docker, with the chance to download the image and test it out.



Melanoma Anatomy

docker pull mjoancarles/skincancer:latest

Skin Lesions Detection App

Melanoma is a serious form of skin cancer that originates in the pigment-producing melanocytes. Early detection and treatment are crucial for a positive outcome. If you're interested in learning more about melanomas, their characteristics, and their treatment, please visit the 'About Melanomas' tab.

This app pretends to classify skin lesions into 7 different categories in which melanoma is one of them. When uploading a picture, if the model predicts it as a melanoma, please consult a dermatologist.

Choose an image...

Drag and drop file here
Limit 200MB per file • JPG, JPEG

Browse files

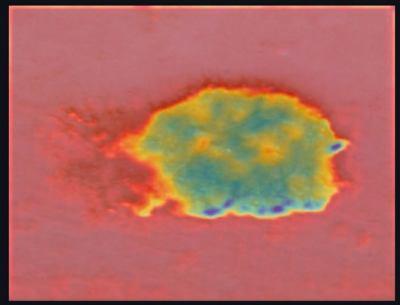
ISIC_0026081.jpg 257.6KB

Uploaded Image

The lesion is likely to be MEL with confidence:

	Class	Probability
0	MEL	0.40
1	NV	0.39
2	BCC	0.00
3	AKIEC	0.01
4	BKL	0.20

The highlighted areas in the following image show where the model focused its attention while making the prediction.



Grad-CAM Heatmap Overlay

CONCLUSIONS

THANKS FOR YOU ATTENTION :)

Joan Carles Montero Jimenez

Curs 2023-24