

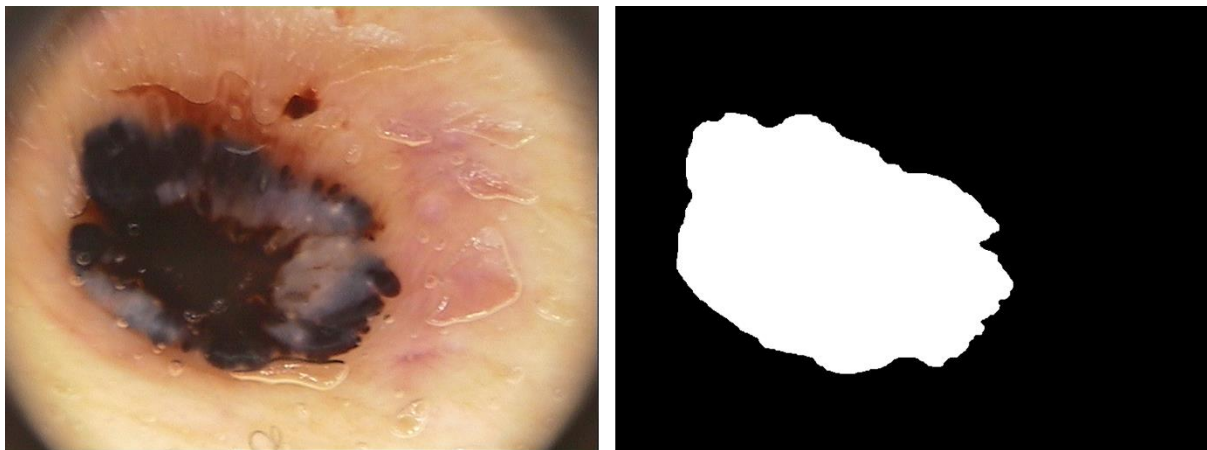
The increasing incidence of melanoma has recently promoted the development of computer-aided diagnosis systems for the classification of dermoscopic images. The PH² dataset has been developed for research and benchmarking purposes, in order to facilitate comparative studies on both segmentation and classification algorithms of dermoscopic images. PH² is a dermoscopic image database acquired at the Dermatology Service of Hospital Pedro Hispano, Matosinhos, Portugal.

The dermoscopic images were obtained at the Dermatology Service of Hospital Pedro Hispano (Matosinhos, Portugal) under the same conditions through Tuebinger Mole Analyzer system using a magnification of 20x. They are 8-bit RGB color images with a resolution of 768x560 pixels.











This image database contains a total of 200 dermoscopic images of melanocytic lesions, including 80 common nevi, 80 atypical nevi, and 40 melanomas. The PH² database includes medical annotation of all the images namely medical segmentation of the lesion, clinical and histological diagnosis and the assessment of several dermoscopic criteria (colors; pigment network; dots/globules; streaks; regression areas; blue-whitish veil).

You can access data using <https://drive.google.com/drive/folders/1fdwE1MDp7e7MKgtE71c4gQIPlg8fB6ck?usp=sharing>

In this data you are given manually segmented lesion masks along with each original colored image. The data also contains a detailed CSV file mentioning true label/class (common nevi, atypical nevi, melanoma) of each image.



Dermatologists generally use ABCD rule to differentiate between cancerous (Melanoma) and normal cases. Use your knowledge which you have gained so far and apply it to device the solution for differentiate between 3 groups of images.

NORMAL		CANCEROUS
	A: ASYMMETRY If you draw a line through the centre of the lesion, the two halves of a melanoma won't match.	
	B: BORDER IRREGULARITY The border of a melanoma is irregular, typically geographic: peninsulas, bays, islands.	
	C: COLOUR VARIEGATION Healthy moles are a uniform colour. A variety of different colours in the same lesion is suspicious.	
	D: DIAMETER > 6 MM Greater than 6 mm is suspicious, although melanomas can be smaller.	
	E: EVOLVING Recent change in size, shape or colour, or bleeding or scabbing are suspicious.	

You have to do following tasks in this assignment.

1. Divide data into 80-20 split. It means that you separate out last 20% data (images) from each group. Do the following steps on initial 80% data and keep remaining 20% data for last step (step-5)
2. Use CCA, transformations, histograms, spatial filtering, descriptors extraction to extract different attributes leading to a combined feature vector for each sample. (choose attributes wisely)
3. Use box plots and scatter plots to visually see the distribution of a feature against each class.
4. Do analysis on these plots and find out optimal values of threshold/condition which best segregate the cases.
5. Use your finalized scheme to see its effectiveness on remaining 20% data which was separated out in step-1. Calculate overall accuracy of the solution using this 20% data where accuracy is the ratio of truly detected image and total images.

Submission: You need to submit a report (word or pdf) against this assignment containing

1. In your report, Explain your working by clearly highlighting steps you followed for all above tasks.
2. Make a flow diagram for better explanation.
3. Copy complete editable code in a word file
4. Add tables to mention accuracy values per image and overall average accuracy values

Please note that your solution should contain concepts which we have covered till now. It should not use any other technique to solve the problem