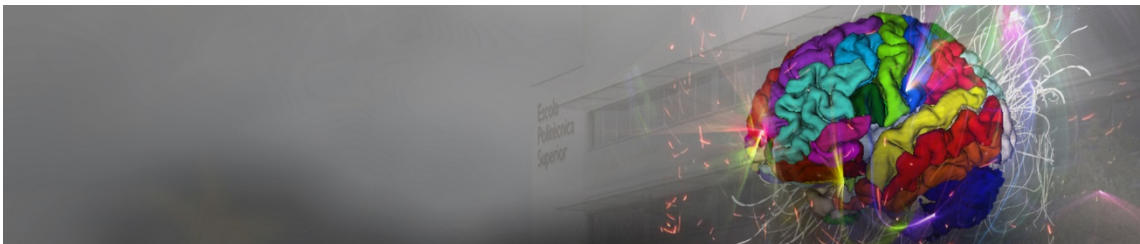




---

# Computer Aided Diagnosis

Final Project:  
Diagnosis in dermoscopic images

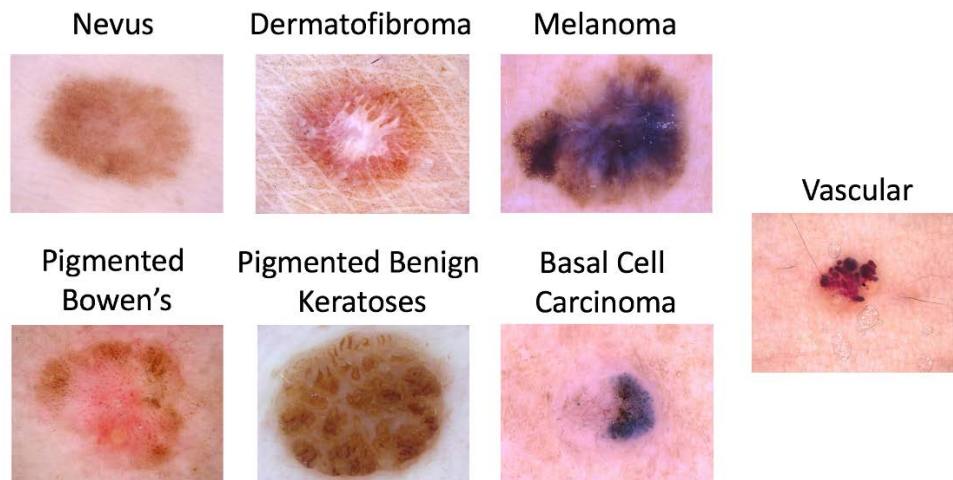


---

## Introduction

In this coursework, the goal is to develop a computer aided algorithm for the diagnosis in dermoscopic images. This final coursework is a mini-project where you can choose the strategy to solve the problem.

The input data are dermoscopic images in JPEG format. The distribution of disease states represents a modified real world setting whereby there are more benign lesions than malignant lesions, but an over-representation of malignancies. The lesion images come from the HAM10000 Dataset, and were acquired with a variety of dermatoscope types from different anatomic sites (excluding mucosa and nails). Images were acquired from a historical sample of patients that presented for skin cancer screening from several different institutions<sup>1</sup>.



We divided this project into two main challenges with a similar structure. Three different sets of images will be given: training, validation, and testing. For the training and validation images, you will have the corresponding ground truth, while for the testing set it will not be available. We will use the results in this testing set to independently evaluate your algorithms (this dataset will be given during the last week of the coursework).

For analysing and reporting the results of your approaches, you should use the validation set and the confusion matrix as a result. Note that, in this part, the validation images cannot be used for training. When evaluating the testing set, you can merge all the information (if you want).

As a suggestion, start with a simple approach and increase its complexity smoothly. You can use software already developed during all your MAIA studies or even from external sources (in this case, cite it properly). Report all the trials and experiments, analysing the improvement on each step.

- **Challenge 1:** The binary problem of classifying Nevus images vs all the others. We will give you 6000 images, 3000 being nevus, 3000 being a combination of the others to train the system. The test set will consist on 1015 images.
- **Challenge 2:** A three-class problem consisting on the classification of melanoma vs benign keratosis vs basal cell carcinoma. The training set will consist on 1000 images for the first two classes and 500 for the last one (imbalanced problem). The test set consist on 226 images.

---

<sup>1</sup> Images were collected with approval of the Ethics Review Committee of University of Queensland (Protocol-No. 2017001223) and Medical University of Vienna (Protocol-No. 1804/2017).

---

## Objectives

- A Information search. Teamwork.
- B To design, analyse and implement approaches for automatic diagnosis, i.e., to classify a given region of interest.
- C To test implementations at least with the provided images. To evaluate results using provided ground truth. To study problems and possible improvements. To submit a final result for the testing set (unknown ground-truth).
- D Documentation.

---

## Coursework:

- A Powerpoint with the following sections:
  - 1 Proposal analysis.
  - 2 Design and implementation of the proposed solutions.
  - 3 Experimental section and results analysis (qualitative/ quantitative analysis, speed, etc). Results should be provided for the validation set.
  - 4 Conclusions.
- B Code with comments.
- C The diagnosis of the images with unknown ground-truth.

---

## Coursework evaluation:

- A During the labs.
- B After the coursework.
- C Results of the unknown dataset will be also considered.

**DEADLINE:** It will be the one indicated in the moodle submission link. Late submission will be penalised.