Practical block I: Dermatological moles segmentation

Practicum and System Requirements

Please note that the session will be performed in python using Keras with TensorFlow Backend. Both Tensorflow 1.x or 2.x versions can be used (TF 2.x is recommended). Code using other programming languages will not be evaluated. The use of Google Colab is allowed.

A basic guide for installing Keras 2.3.1 and Anaconda can be found in a separate document in the campus virtual.

Practical session will be performed individually. Please code using python jupyter, in order to facilitate the results comparison.

1. Introduction

Skin cancer is a major public health problem, with over 5,000,000 newly diagnosed cases in the United States every year. Melanoma is the deadliest form of skin cancer, responsible for an overwhelming majority of skin cancer deaths. In 2015, the global incidence of melanoma was estimated to be over 350,000 cases, with almost 60,000 deaths. Although the mortality is significant, when detected early, melanoma survival exceeds 95%.

A popular method for remembering the signs and symptoms of melanoma is the evaluation of the "ABCDE" features:

- **A**symmetrical skin lesion.
- **B**order of the lesion is irregular.
- **C**olor: melanomas usually have multiple colors.
- **D**iameter: moles greater than 6 mm are more likely to be melanomas than smaller moles.
- **E**volution: The evolution (i.e. change) of a mole or lesion may be a hint that the lesion is becoming malignant.

To identify some of these features the segmentation of the mole is a critical task.

2. Dataset

Since 2016 the ISIC community provide datasets of dermatoscopic images for Machine learning training.

https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main

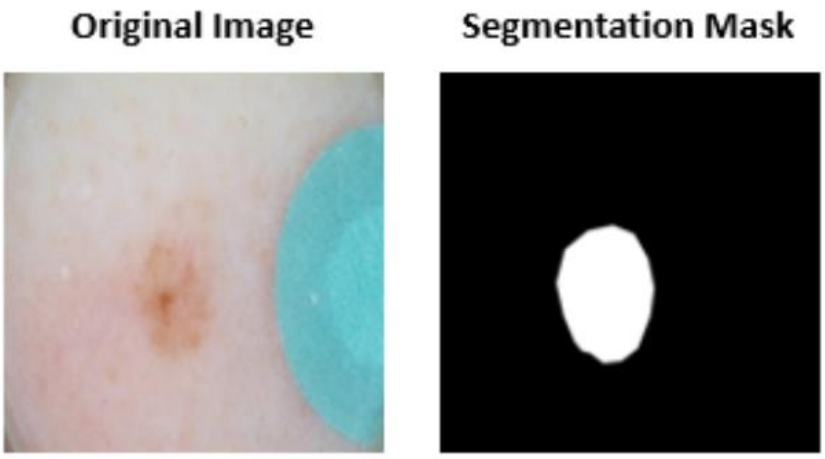
In this session the 2016 challenge will be considered. the dataset provided includes several tasks

- Part 1: Lesion Segmentation
- Part 2: Detection and Localization of Visual Dermoscopic Features/Patterns
- Part 3: Disease Classification

In this practical session we will focus on Part 1.

https://challenge.isic-archive.com/landing/2016/37

The **Training Data** file is a ZIP file, containing 900 dermoscopic lesion images in JPEG format, along with the corresponding segmentation mask.



Evaluation

Segmentations should be compared using a variety of metrics, all computed at the level of single pixels, including: sensitivity, specificity, accuracy, Jaccard index, Dice

coefficient, Roc curve. Please see the 2016 page for additional resources https://challenge.isic-archive.com/landing/2016/37

3. Practical session

The practical session is designed for three weeks of work (not more!!!). No need to do extra implementations.

The aim of this session is to implement a segmentation method based on a U-net.Then adapt the example code in order to receive the ISIC 2016 dataset.

In order to limit the scope of the session, the following goals are recommended:

1) FIRST WEEK. Adapt a U-net from keras (or pytorch if you are more familiar with that library).
    a. U-net is a network that have been described in other master courses. Instead of using code provided in other courses, students are recommended to adapt a basic U-net implementation provided in: https://keras.io/examples/vision/oxford_pets_image_segmentation/

       In case you have an equivalent implementation with PyTorch (and you are familiar with such code) you are allowed to use pytorch instead.

    b. Adapt the input layer of your neural network in order to receive the data of the challenge
    c. Resize the images in order to complete a training in LESS THAN A DAY of computation
    d. Compute the performances of this BASELINE approach
2) SECOND WEEK. fine-tune the network,
    a. improve the segmentation by including data-augmentation, parameter optimization.
    b. Remember that the main drawback of the dataset is the limited amount of data. Propose ideas how to mitigate this problem.
    c. Analyze and compare the improvements with respect to the Baseline, and motivate your choices
3) THIRD WEEK. implement the metrics described in https://challenge.isic-archive.com/landing/2016/37
    a. Present the data as table comparing the BASELINE with the improvements. Use plots if needed to show performance increase.
    b. Compare in the table the results with the challenge winner

What should NOT be done and it is out of the scope of this session

1) Compare different neural networks
2) Comparing 3D vs 2D approaches
3) Spend days of training
4) Attempt to reach the performance of the challenge winner.

Please discuss your approach with the teacher during the practical sessions before implementing the code

In case of doubts, do not hesitate to ask to the teacher.

4. Deliverable

Submit a **SHORT** report using jupyter Notebook, commenting your results using markdown of python. No need to include the original dataset images.

IMPORTANT, only files delivered using the campus virtual will be evaluated. Please upload in the corresponding task your file before the dead line
IN CASE YOU HAVE NO ACCCESS TO THE CAMPUS VIRTUAL please contact the professor of the practical course to solve the issue.