AML - Challenge 2, Glaucoma diagnosis and segmentation

Context

Glaucoma is a **prevalent condition** related to an **abnormal fluid balance** in the **eye** that causes an **increase** of **internal ocular pressure**. The **increase** of **pressure** gradually **damages** the **eye optic nerve**. If undiagnosed, these damages can result in permanent vision loss. *Approximately 2.86% of the global population over 40 years old is affected by glaucoma* [1]. ***Patients affected by glaucoma usually do not present symptoms in the early stages of the disease, while an early diagnosis is critical to prevent irreversible damages***. It is thus important to develop inexpensive detection methods, in order to massively and systematically control patients, before the symptoms appear.

One way to **diagnose glaucoma** is to perform a **visual examination** of the **inside back surface** of the **eye** (**fundus**). The images are obtained by special cameras through a dilated pupil. The main advantage of this technique, called "**Retinal Fundus Imaging**", is to be **brief** (usually a minute long) and **painless** to the patient, making it suited for simple routine checks. However, **establishing** an **accurate diagnosis from these images** is **particularly difficult**. This task is currently performed by human experts, which represents a significant financial burden. Therefore, *automatic glaucoma detection methods from fundus imaging are both highly desirable and challenging*.

As most of the applications in medical imaging, this research topic essentially suffers from two limitations:

1. **The lack of data**: such data is costly to annotate, essentially due to the human experts involved.
2. **The disparity of data**: the data can vary a lot depending on the imaging device involved and its calibration. Therefore, one solution developed in one particular medical center may perform poorly in another medical center. It is thus crucial to design solutions that can achieve good generalization performance with few data.

Your objective

Your task in this challenge is to participate in the research of a **machine learning solution** to **automatically detect glaucoma from retinal fundus images**. *You will be provided with a baseline method based on both a Convolutional Neural Network and some useful human expert's knowledge, which you are free to use as a basis*. Previous works exist in this application domain [2], and may eventually inspire you. Your goal will be to improve our baseline. As usual, model performance will not be the only evaluation criterion, and serendipity will also be taken into account.

Metric

The metric for the evaluation of the results is the **Area Under Curve** (**AUC**) which is the area under the **Receiver Operating Characteristic** (**ROC**) curve.

Submission format

To evaluate your results on the leaderboard, you should submit a single csv file reporting the results of your model on the test images. Your csv submission file should report the image ID and the prediction of your model between `0` (healthy) and `1` (glaucoma) in the following format:

```
Id, Predicted
T0001, 0.9738
T0002, 0.1625
T0003, 0.7638
...
```

You can download an example submission file (sample_submission.csv) on the Data page.

Retinal Fundus images

On the ocular anatomy diagram shown in Figure 1, we can identify the different parts of the eye related to the retinal fundus images.

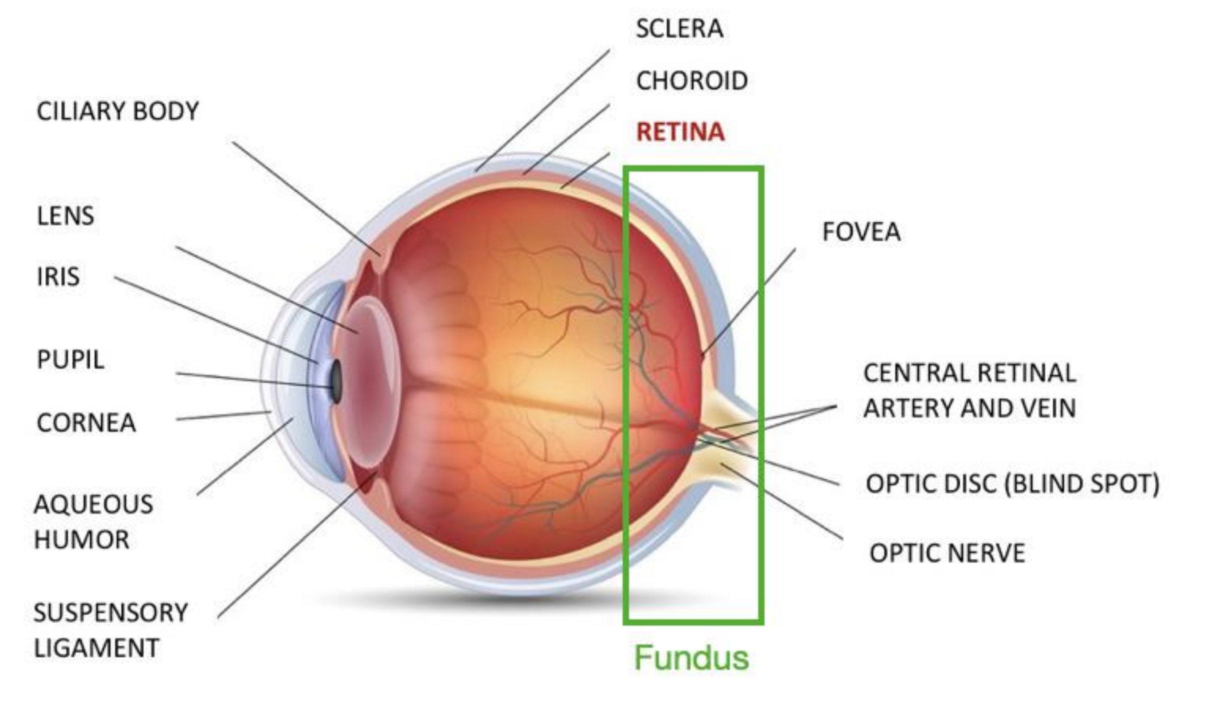

Figure 1: Eye anatomy with a focus on the fundus part of the eye

Damages on the optic nerve can be spotted by professional examination of the **Optic Disc** (**OD**) and the **Optic Cup** (**OC**) located inside the OD [3].

Figures 2 and 3 respectively show a healthy eye and a glaucomatous eye, taken from the dataset.

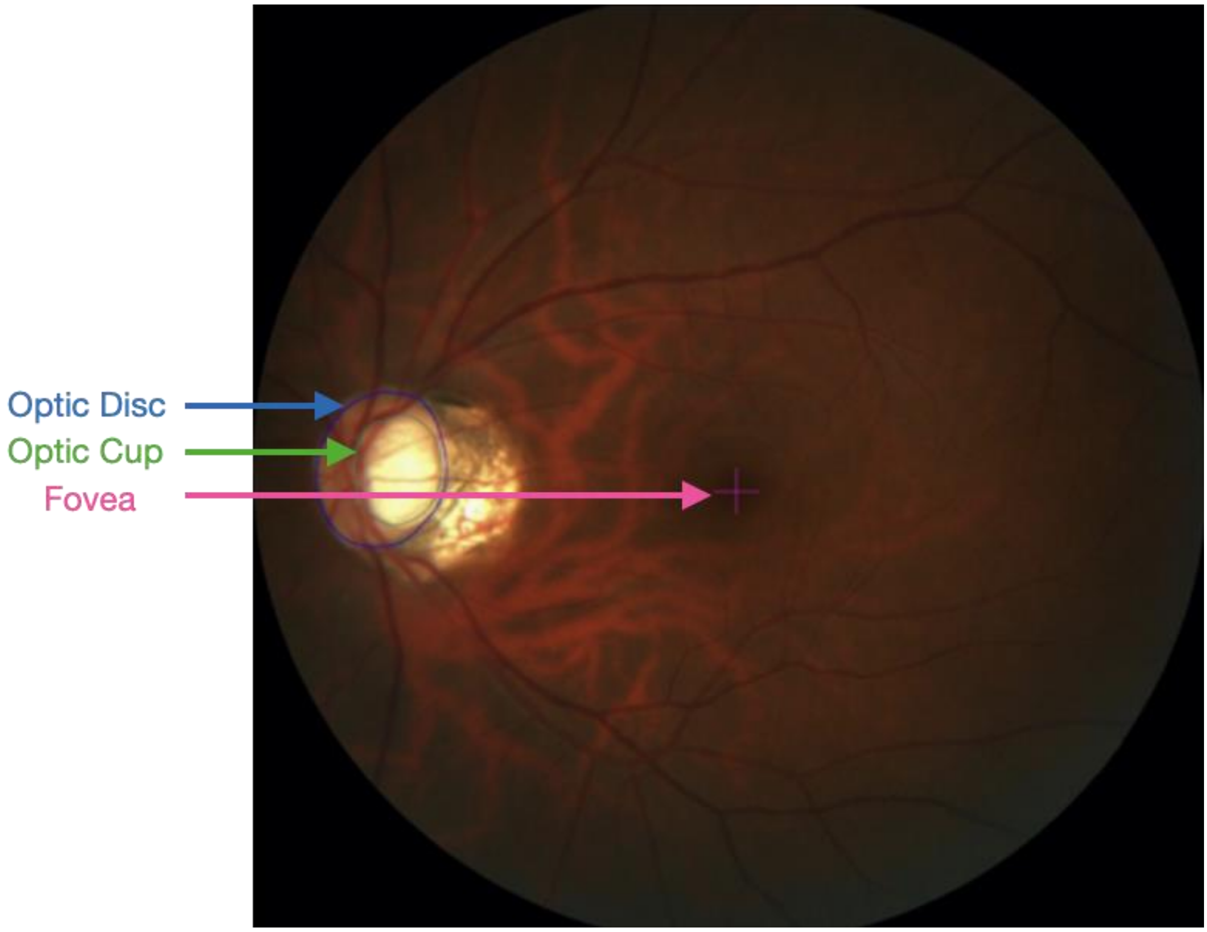

Figure 2: Optic Disc, Optic Cup and Fovea in fundus image

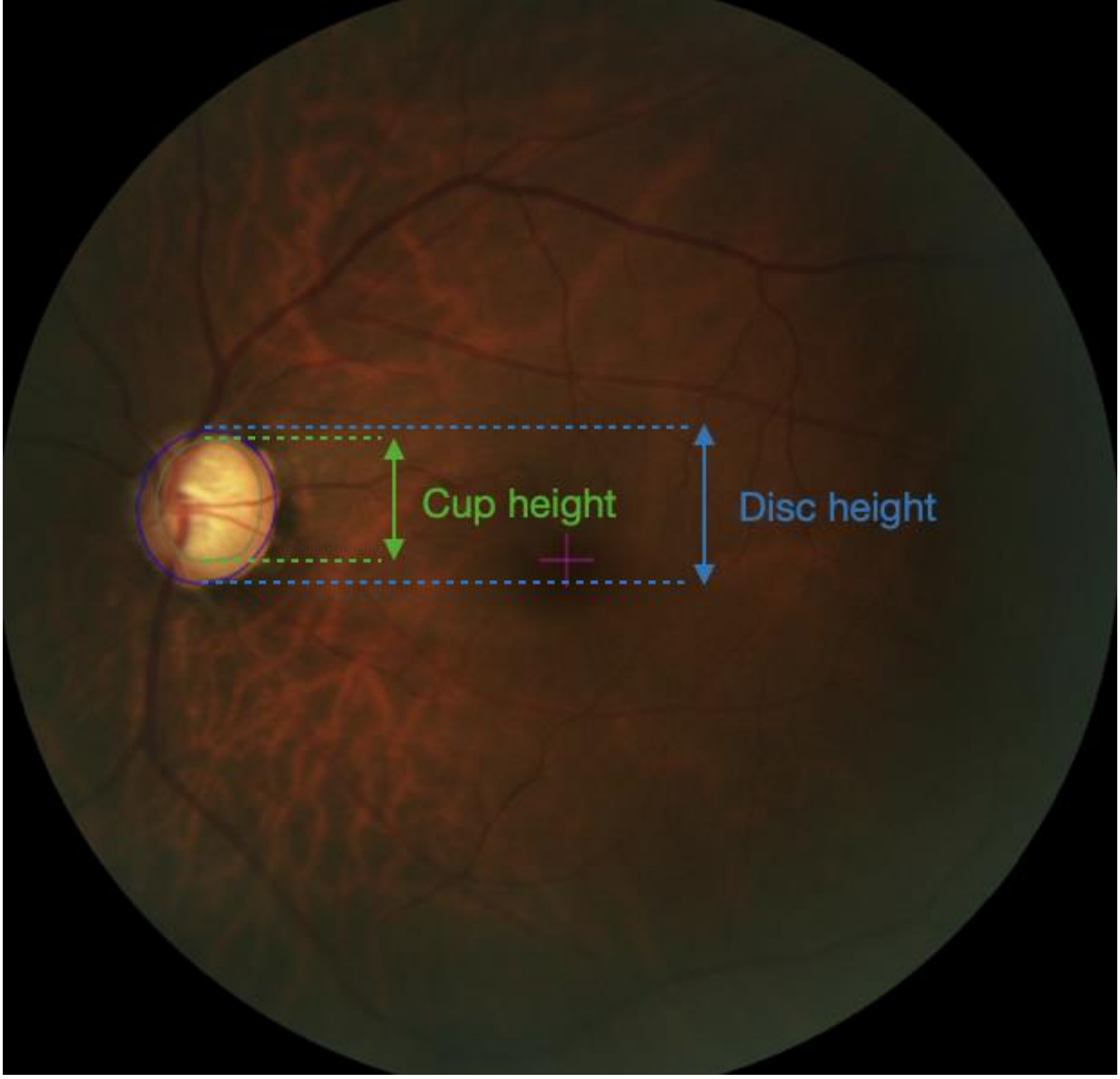

Figure 3: Disc height (OD height) and Cup height (OC height) used in vCDR are represented on a glaucomatous eye fundus image.

As shown in Figures 4 and 5, one relevant feature for glaucoma detection is the **vertical Cup-to-Disc Ratio** (**vCDR**) which is the **ratio between OC and OD heights**, both represented in Figure 3.

$$vCDR = \frac{OC_{height}}{OD_{height}}$$

Although other relevant markers exist (Figure 5), the vCDR is the easiest to exploit. The baseline method presented in the following section is thus based on the computation of the vCDR.

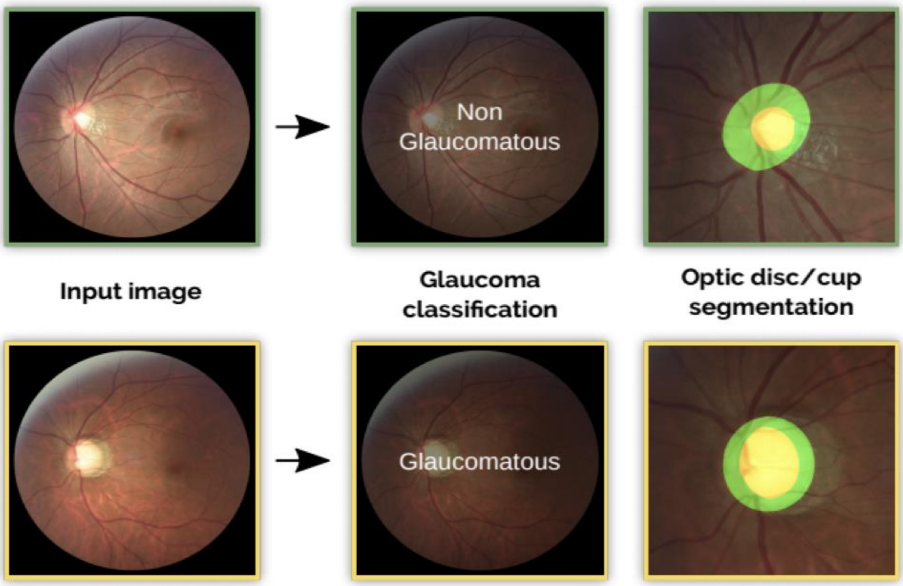

Figure 4: Comparison of Optic Disc (OD) and Cup (OC) for healthy and Glaucomatous eyes. A cupping effect is observed in the second case.

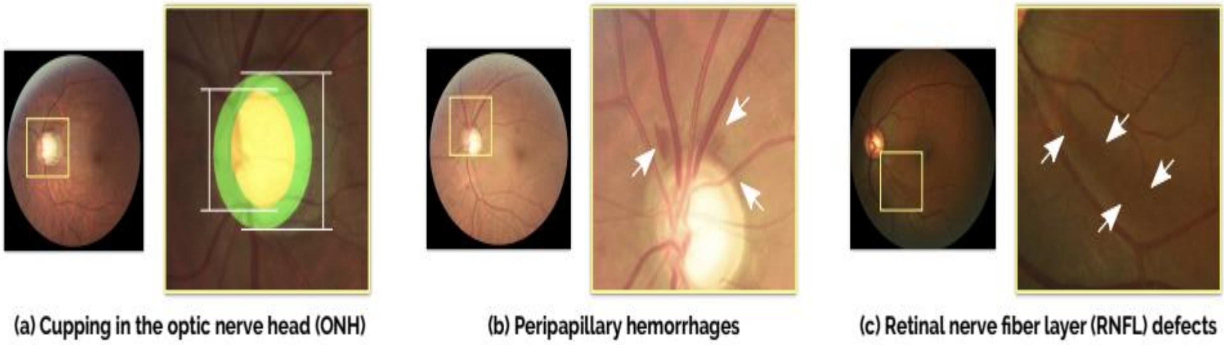

Figure 5: Three relevant markers for glaucoma.

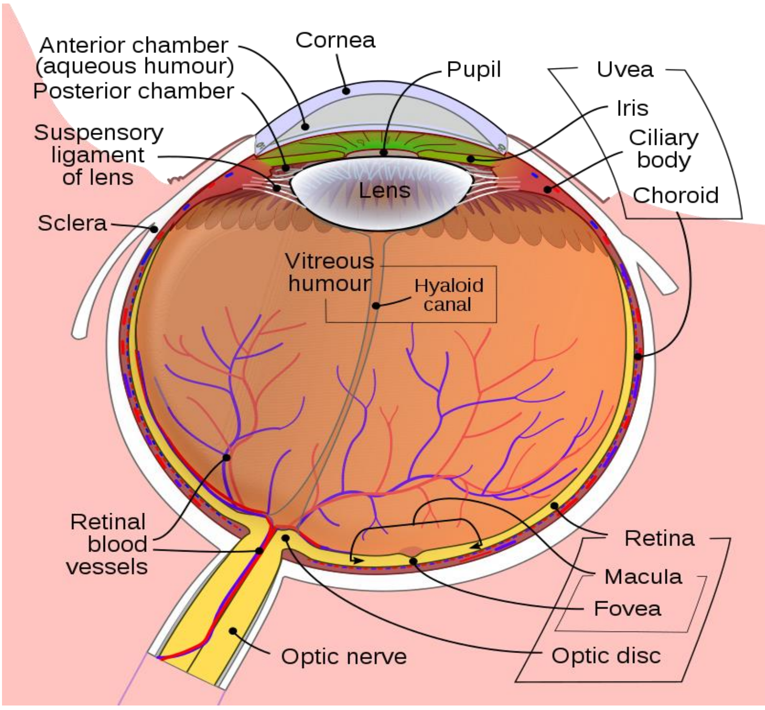

Diagram of the human eye.

It shows the lower part of the right eye after a central and horizontal section.

Baseline

Our pipeline (Figure 6) is made of four main steps, and aims at diagnosing glaucoma based on the vCDR of the eye, which is obtained from OD and OC segmentations.

The four steps are described below.

Pre-processing: Input images are of high resolution, which would considerably increase the processing time of a deep convolutional network. We thus systematically **resize** (**without crop**) **the input data to 256*256**.

Deep CNN: We use a **UNet** [4] , which is a **CNN encoder-decoder designed for image segmentation**, and well known in medical imaging. Its main specificity is to **combine** the **encoder feature maps with those of the same scale in** the **decoder**, in order to **restitute more accurate spatial information** (it creates a sort of "U" shape of the network). The network outputs a prediction maps for both OD and OC segmentation, of the size of its input (256*256 in our case), and is **trained with the segmentation ground truth**. We monitor the Dice Coefficient during training to monitor the segmentation network performances.

Post-processing: In order to accurately determine the OD and OC heights, we only retain the largest connected component of the prediction. This will filter potential noise in the prediction.

Linear classifier: We use a simple logistic regression model, which predicts the final classification probability from the vertical Cup-to-Disk Ratio (vCDR). The classifier is trained with the classification labels.

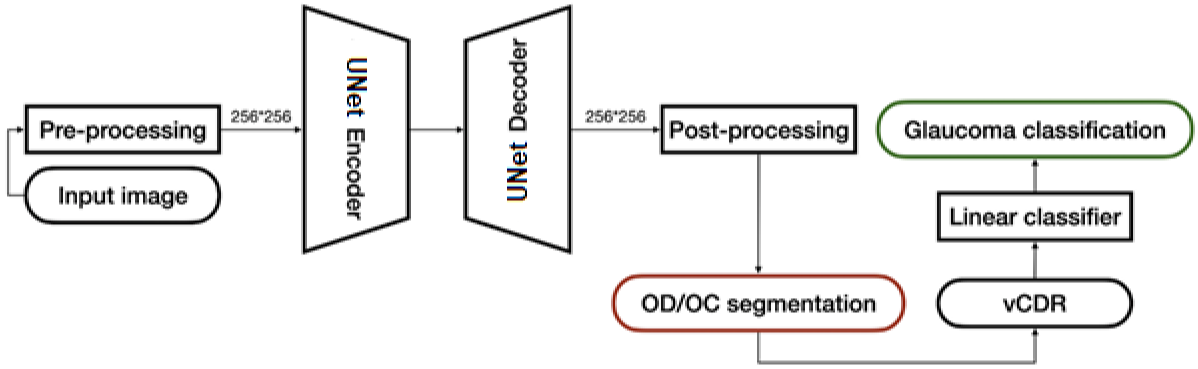

Figure 6: Baseline solution pipeline

Although this baseline produces good performance, there are many ways you could think of to improve it, either by modifying some of the four processing steps, or by improving/replacing the pipeline itself.