Neur Project Description

The dataset contains a total of 2355 images of 785 subjects taken by a holographic microscope. Each subject has 3 images belonging to it starting with the same filename_id but with different suffixes:

- an amplitude image (ends with "_amp.png")

- a phase image (ends with "_phase.png")

- a computer generated segmentation mask (ends with "_mask.png")

There are 5 different classes of subjects, grouped into their respective folder:

- chlorella

- debris

- haematococcus

- small haematococcus

- small particles

3 images are provided for each subject, but you don't have to use them all. You can read more about holographic microscopy at: https://www.holodetect.com/ In the train dataset, classes are provided (folders), but the test dataset contains only a single folder with a mix of images in it.

Your task is to have at least 0.5 recall and a high precision for the chlorella class! Submission File:

For each ID in the test set, you must predict a class number for the TARGET variable. The ID is the name of each image in the test set (e.g. 34.png has the ID 34). The file should contain a header and have the following format:

ID,TARGET

1,0

2,3

3,2

etc.

Target class numbers: 0 - chlorella 1 - debris 2 - haematococcus 3 - small haematococcus 4 - small particles

Files

- train - the training set, amplite/phase/mask images in .png

- test - the test set, amplite/phase/mask images in .png

- example_solutions.csv - an example solutions file in the correct format (solutions are just random numbers!)

Columns of example_solutions.csv

- "ID" - id of a single subject without suffix

- "TARGET" - class label 0-4