BMIA Practical Exercise 2:

Catuxa Seoane Botana
University of Santiago de Compostela
catuxaseoanebotana@gmail.com

I. Introduction

Macular Edema (ME) constitutes serious eye conditions that are caused mainly by theabnormal formation of the retinal microvasculature that is associated with a prolongedhyperglycemia of the diabetes mellitus. They represent the leading causes of preventableblindness in the working-age population of developed countries, representing a seriouspublic health problem. In particular, ME is defined as a focal or diffuse retinal thickeningderived by the intraretinal fluid accumulation in the macula, which is the central region of theretina where the vision is most acute.

Currently, the diagnosis and monitoring of the ME disease is mostly performed through thevisual analysis of Optical Coherence Tomography (OCT) scans. OCT is a non-invasiveimaging modality that is capable of providing high-resolution cross-sectional images of theneurosensorial retina in real time with micron-level resolution. In daily clinical practice, thediagnosis and monitoring processes of patients with ME disease are performed through thevisual inspection of several OCT scans by the clinicians, a process which is extremelytedious and time-consuming. For that reason, a fully automated system for the segmentationof pathological fluid regions in OCT images is crucial for the accurate diagnosis andmonitoring of this relevant ocular pathology.

The main goal of this practical exercise is to implement a complete methodology for the automatic segmentation of pathological fluid regions using OCT scans. To carry out the work, a set of OCT images is provided to support and validate the methodology developed.

A schematic representation of the proposed methodology can be seen in this figure:

Fig. 1: A schematic representation of the proposed methodology

II. Dataset

The dataset used for this work consists of 100 OCT scans, which are divided into two folders: one of them consists of the OCT scans with 50 images and the other consists of another 50 images representing the Ground Truth of the mentioned OCTs. At first, not all the images have the same size, so later on they will be transformed. In this figure it is shown an example of this dataset:

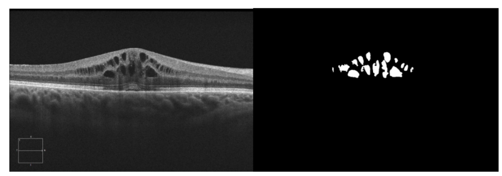


Fig. 2: Right: OCT image with pathological fluid regions. Left: Ground Truth image with annotations by clinical experts

III. Methods

To carry out this project, a code was made in Python. The main Python library used is Pytorch.

A. Data Preparation

The dataset images need to be previously processed and transformed in order to apply them to the model. First, using the split-folders function, the dataset is divided (both the GT and the images) into train(60%), test(20%) and validation(20%). Once the separation is done, the images are read and sent to a class called Dataset(). When an object of this class is created, the first thing that is done is to modify the images and masks(GT). A reshape is performed so that all the data have the same size, taking into account that each dimension is a multiple of 32, due to the different encoders that could be used later in the job. Another transformation that the dataset must undergo is to become a torch type. Finally, it is necessary to create DataLoaders, so using torch.utils.data import DataLoader() from the torch.utils.data library, these are built with the train dataset and validation dataset, with a batch size of 1 and a number of workers of 0. From here a model is chosen and trained.

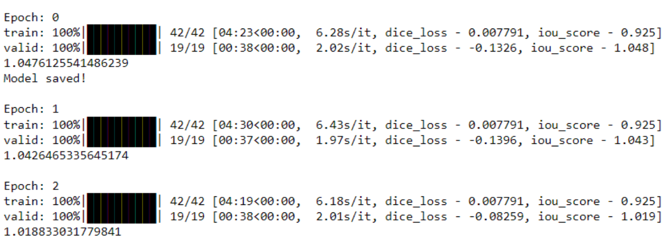
B. Learning Algorithm

Two models were used to train the algorithm and to be able to compare the results.

The first of these is the Uned model, which is trained with encoder_name='resnet50', encoder_weights='imagenet', in_channels=3, classes=1.The u-net is convolutional network architecture for fast and precise segmentation of images. Up to now it has outperformed the prior best method (a sliding-window convolutional network) on the ISBI challenge for segmentation of neuronal structures in electron microscopic stacks. It has won the Grand Challenge for Computer-Automated Detection of Caries in Bitewing Radiography at ISBI 2015, and it has won the Cell Tracking Challenge at ISBI 2015 on the two most challenging transmitted light microscopy categories

The othe one model is FPN, which is trained with 'resnet34'and in_channels=3. Feature Pyramid Network (FPN) is a feature extractor designed for such pyramid concept with accuracy and speed in mind. It replaces the feature extractor of detectors like Faster R-CNN and generates multiple feature map layers (multi-scale feature maps) with better quality information than the regular feature pyramid for object detection.

The epochs used in training are 10. Epoch is the number of times the forwardpropagation and backpropagation algorithms will be executed. In each cycle (epoch) all the training data passes through the neural network so that it learns about them, if there are 10 cycles and 1000 data, each cycle the 1000 data will pass through the neural network.To validate the model it is used the IoU metric.Intersection over Union (IoU) is used when calculating mAP. It is a number from 0 to 1 that specifies the amount of overlap between the predicted and ground truth bounding box.



```
Epoch: 0
train: 100%|██████████| 42/42 [04:23<00:00,  6.28s/it, dice_loss - 0.007791, iou_score - 0.925]
valid: 100%|██████████| 19/19 [00:38<00:00,  2.02s/it, dice_loss - -0.1326, iou_score - 1.048]
1.0476125541486239
Model saved!

Epoch: 1
train: 100%|██████████| 42/42 [04:30<00:00,  6.43s/it, dice_loss - 0.007791, iou_score - 0.925]
valid: 100%|██████████| 19/19 [00:37<00:00,  1.97s/it, dice_loss - -0.1396, iou_score - 1.043]
1.0426465335645174

Epoch: 2
train: 100%|██████████| 42/42 [04:19<00:00,  6.18s/it, dice_loss - 0.007791, iou_score - 0.925]
valid: 100%|██████████| 19/19 [00:38<00:00,  2.01s/it, dice_loss - -0.08259, iou_score - 1.019]
1.018833031779841
```

Fig. 3: Output's example of the model

IV. Results

The following table shows the IoU values for each model and each epoch. The closer this value is to the GT threshold, the better results it offers.

Model	Epoch 0	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5	Epoch 6	Epoch 7	Epoch 8	Epoch 9
Unet, 'resnet5'	1.0476	1.0426	1.0188	1.0070	1.02228	1.02609	0.98659	1.02651	0.99828	1.00143
FPN,'resnet34'	4.1713	0.01954	14.8299	0.06025	0.04078	4.0749	0.19641	36.82	26.1877	1.21266

TABLE I: Table of IoU score for every Epoch fo every model

The following two figures show examples of the code's outputs; in which examples of both the original images, the GT and the final predicted skins are shown.

Fig. 4: Examples of the output of the program: image, GT mask and predicted mask. Model: 'Unet' Encoder: 'resnet50'

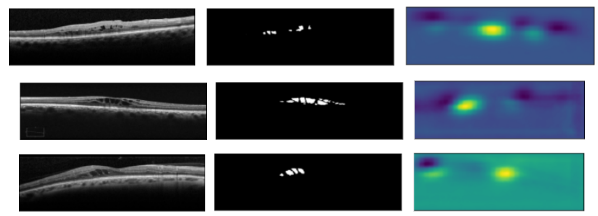


Fig. 5: Examples of the output of the program: image, GT mask and predicted mask. Model: 'FPN' Encoder: 'resnet34'

References

[1] https://pypi.org/project/pyfeats/#description
[2] http://rasbt.github.io/mlxtend/user_guide/feature_selection/SequentialFeatureSelector/
[3] https://github.com/qubvel/segmentation_models.pytorch/blob/master/examples/cars%20