AI Diagnosis of HCQ Retinopathy

Long term treatments using hydroxychloroquine (HCQ), can lead to irreversible visual loss, due to its toxicity. OCT scans of the layers of the retina is a key tool in the diagnosis of HCQ retinopathy, to detect the disease in the early stages. The goal of this study is to explore the possibility of automating and improving the accuracy of OCT diagnosis using AI tools, particularly in the context of a small dataset. Throughout this work, we tried to evaluate the capacities of commonly used convolutional networks, such as ResNet or Vit at diagnosing retinopathy.

Dataset and Pre-processing

The dataset used for this study is comprised of 657 scans of the retina, coming from 308 patients. Each patient had both of their eyes scanned and sometimes, the same eye scanned several times at different moments, which shows the evolution of the retinopathy. Each scan was from any individual was treated as independent from the others. From the 657 scans of the retina, 122 of them show retinopathy.

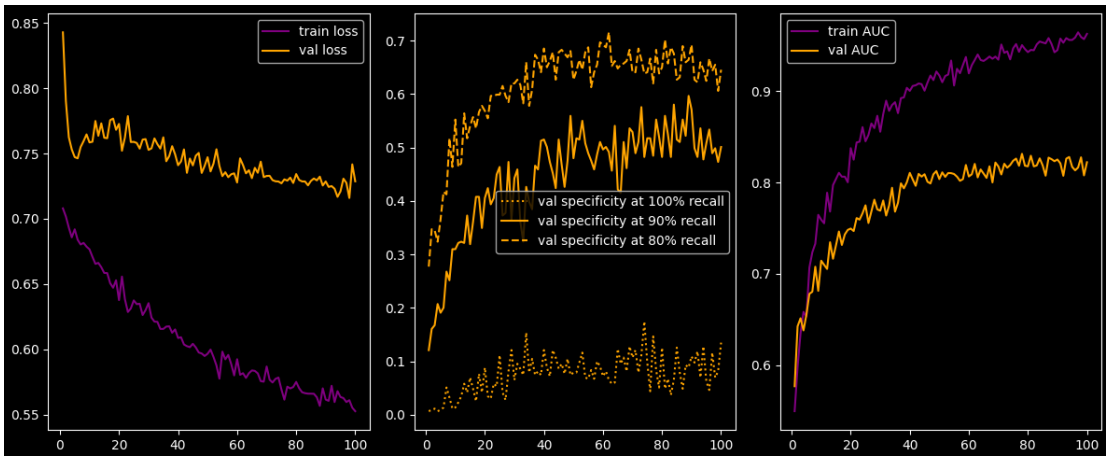
The test set is comprised of 20% of the dataset and the rest of the data is split in a 5-fold, each validation set containing 20% of the remaining data. The training set is also re-sampled to balance the frequencies of the classes. To have a good idea of a model's performance, the predictions of each fold are concatenated before calculating the metrics. The metrics therefore represent the performance of every fold put together. The main metric used to track the performance is the specificity at a fixed recall.

The data preprocessing involves image-wise standardization (z-score normalization) and resizing images that do not match the required format.

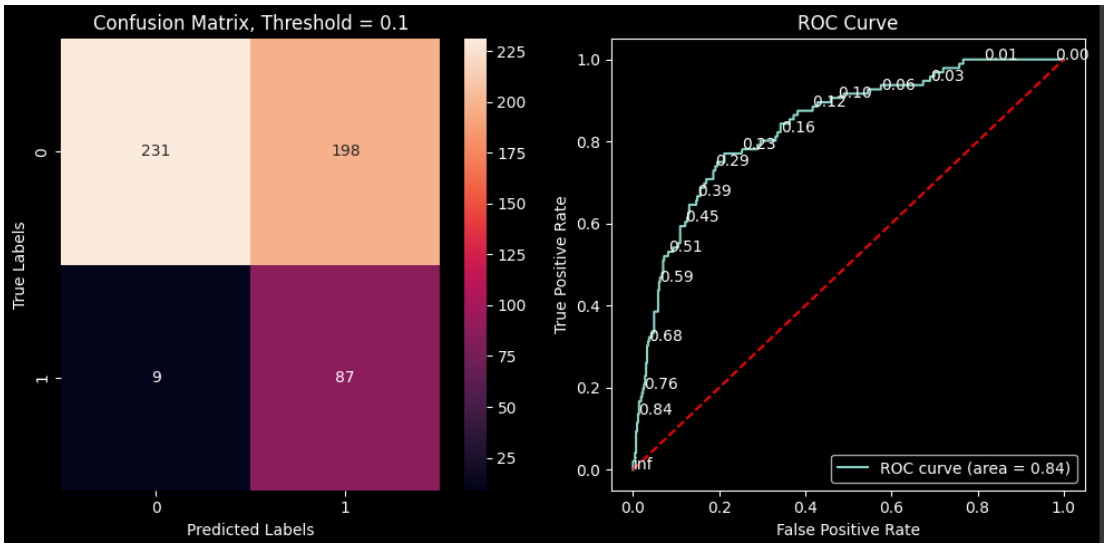
Models and results

The approach used is to reuse a pre-trained models and fine tune them, to avoid overfitting the data. Since these models usually take 3 channels images as an input, and each scan contains 6 OCT images of the same retina, we searched for the best way to use these images.

The first approach used was to only feed the 3 images that show best the retinopathy in the case of a diseased patient. The fully connected layer of the network was also changed to have a binary output. Only the new fully connected layer was trained until we reached a plateau (40 epochs), then the other layers of the model were also fine-tuned.



(left) Training and validation loss, (center) validation specificity at different recalls, (right) training and validation AUC, for the single model.



Confusion matrix and ROC curve, for the single model.

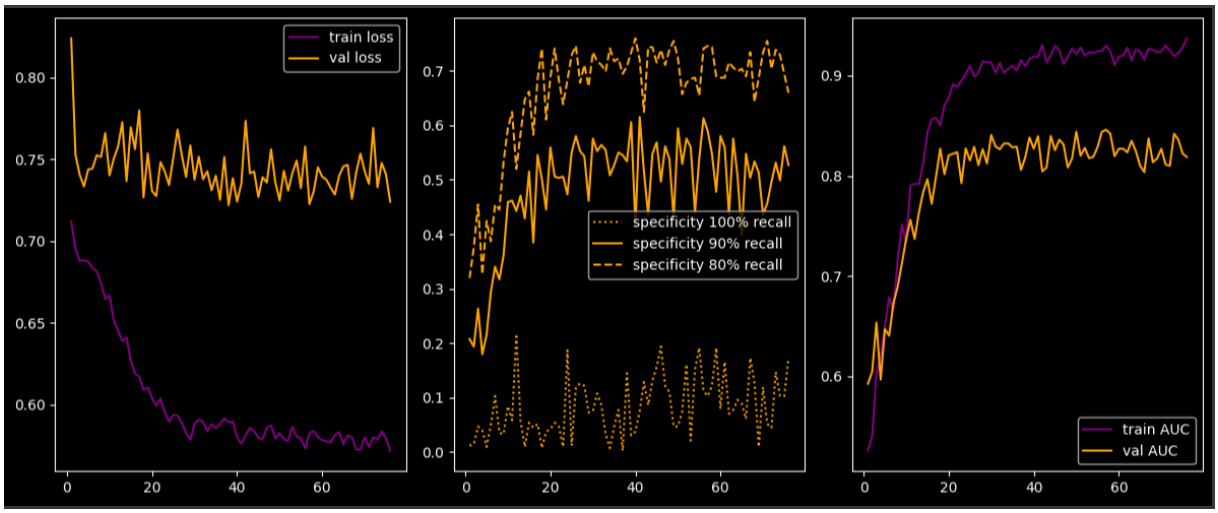
This model can achieve 54% specificity at 90% recall. The recall could be further increased by changing the thresholds, but as shows the training plots and the ROC curve, this would greatly impact the specificity.

Other variations of this approach, which gave similar or worse results:

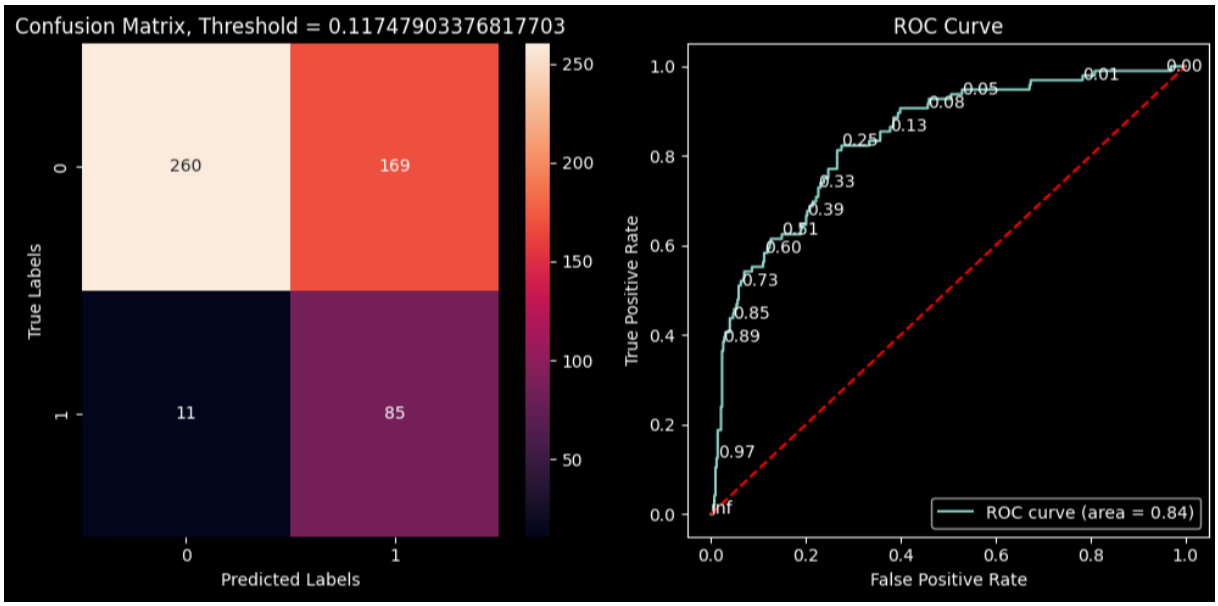
- Use of other networks (ResNext, EfficientNet, VIT, GoogleNet)
- Data Augmentation (image rotation, contrast). The choices of augmentation are limited as the images are quite uniform throughout the dataset.
- The same network, with other images inputted. The images that seem to give the most information about retinopathies are images 2 to 4, while images 0, and 5 give worse performance.
- Training the entire network without pre-trained weights.

The second approach that was explored was to create an ensemble mode. The base model consists of the same modified ResNet50 described above. The fully connected layer was modified to output a

feature map of 256 values. All 6 images are propagated through the model, by inputting the same image on all 3 channels, one by one. The 6 feature maps are concatenated together, then passed through the fc layer of the ensemble model, consisting of linear layers, Relu, Dropout and sigmoid.



(left) Training and validation loss, (center) validation specificity at different recalls, (right) training and validation AUC, for the ensemble model.



Confusion matrix and ROC curve, for the ensemble model.

This model can achieve 58% specificity at 90% recall.

Other variations of this approach, which gave similar or worse results:

- Modified base model with single channel input.
- Inputting 3 different images at a time on the 3 channels twice (0 to 2 and 3 to 5).
- Inputting 3 different images at a time 6 times (increment by 1 each time)
- Using the entire base model with a single output. Then use pooling (average or maxpooling) to process the 6 outputs.

- Using other models as base models (VIT, ResNext, GoogleNet, EfficientNet).

Conclusion:

The results obtained from the models demonstrate the potential of AI-based methods in diagnosing hydroxychloroquine retinopathy. These models can be used to facilitate and accelerate diagnosis, acting as an assistive tool for practitioners. However, the performance of the models developed in this project is not sufficient for real world application. The main limitation lies in the size of the dataset, particularly the limited number of positive cases available for training.

Developing a more robust model would require to use a larger dataset or different ways to extract information from the scans, such as leveraging retinal layer segmentation could be explored. The latter option might offer a promising improvement and will be considered in future work.