Fovea Segmentation using Deep Learning

Project Report

Sahil Khose[1] Ankita Ghosh[2]

[1]4th year, ICT Department
Manipal Institute of Technology

[2]4th year, CSE Department
Manipal Institute of Technology

04 Sept 2021

Table of Contents

Dataset

We have trained the model on five datasets to perform binary segmentation. The datasets consist of coloured retina images and their corresponding masks:

1. Drive: **40** datapoints
2. Messidor: **180** datapoints
3. IDRiD: **58** datapoints
4. diaretdb0: **127** datapoints
5. diaretdb1: **79** datapoints

Preprocessing

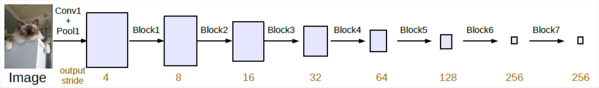
- For feature extraction for the deep learning model, we apply **bilateral filter** $\rightarrow$ two iterations of **dilation** $\rightarrow$ one iteration of **erosion**.
- For image augmentation we perform **shuffling, rotation, scaling, shifting and brightness contrast**.
- The images and masks are resized to 512×512 **dimensions** while training.
- We also added **cropping** to increase the number of datapoints by 20%.

Literature Review

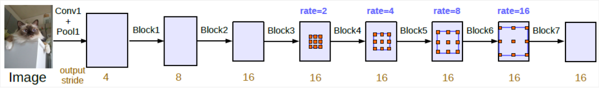
1. Paper 1: Tan et al
   - Converted RGB to **LUV** then processing on L to RGB
   - LReLU, softmax, Xavier initialization is used
   - **7-layer** custom network is trained
2. Paper 2: Sedai et al
   - 2 stage approach is used: **coarse network** followed by **fine network**
   - ImageNet pre-trained **VGG-16** model is used
   - **Class balanced cross entropy loss** is used to fix the imbalance problem

Architecture

- We trained a **DeepLabV3+** model with **EfficientNet-B3** as the backbone.
  - DeepLabV3+ is a refinement of DeepLabV3 which uses atrous convolution. Atrous convolution is a powerful tool to explicitly adjust the filter's field-of-view as well as control the resolution of feature responses computed by Deep Convolution Neural Network.
  - We use **encoder depth of 5** which refers to the number of stages used in encoder. The number of d**ecoder channels used is 256**.



(a) Going deeper without atrous convolution.



(b) Going deeper with atrous convolution. Atrous convolution with $rate > 1$ is applied after block3 when $output\_stride = 16$.

Figure: DeepLab V3+

Loss

- **Binary Cross Entropy Loss**
  $L_{BCE} = y \log(y_{hat}) + (1 - y) \log(1 - y_{hat})$
- **Dice Loss**
  $L_{Dice} = 1 - \frac{2TP}{2TP+FN+FP}$
- **Tversky Loss**
  $L_{Tversky} = 1 - \frac{TP}{TP+\alpha FN+\beta FP}$ where, $\alpha + \beta = 1$
- **Focal Tversky Loss**
  $L_{FTL} = (L_{Tversky})^{\gamma}$ where, γ controls the non-linearity of the loss.
- **Final Loss**
  $L_f = \lambda L_{BCE} + (1 - \lambda) L_{FTL}$ where, λ is the weight parameter.

Semi-supervised pipeline

Algorithm 1: Semi-supervised classification train loop

Input: Sample image
Output: Class of the given image

for $epoch \leftarrow 0$ **to** E **do**
  if $epoch < E_i^\alpha$ **then**
    $\alpha \leftarrow \alpha_i$
  else if $epoch < E_f^\alpha$ **then**
    $\alpha \leftarrow \frac{\alpha_f - \alpha_i}{E_f^\alpha - E_i^\alpha} * (epoch - E_i^\alpha) + \alpha_i$
  else
    $\alpha \leftarrow \alpha_f$
  end if
  Run the model on train set
  $loss \leftarrow BCE(l, \hat{l}) + \alpha * BCE(u_{epoch}, u_{epoch-1})$
  Generate the pseudo labels for unlabeled data
  Evaluate the model on validation set
end for

Figure: Semi-supervised algorithm

Evaluation and Results

Method	F1	MIoU	Sens	Spec	Acc
non-DL	0.804	0.688	0.816	**0.998**	0.996
Ours	**0.824**	**0.705**	**0.917**	0.997	**0.997**
Tan et al	-	-	0.885	0.991	-
Sedai et al	0.810	-	-	-	-

Table: Metrics Comparison

Image - Mask - Prediction

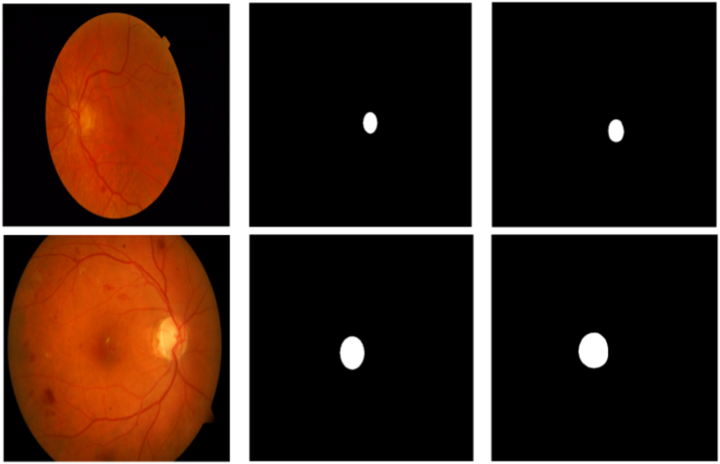


Figure: Image, Mask and Prediction (left to right)

Discussion

- We were able to get **state-of-the-art results** for fovea segmentation using our **semi-supervised segmentation method**.
- All our work for this **2.5 month** project has been summarised in our **5 bi-weekly reports** in detail.
- We are currently working on **AMD classification** and are aiming to perform equally well on that task.