Automatic Optic Disk and Cup Segmentation of Fundus Images Using Deep Learning

Venkata Gopal Edupuganti*, Akshay Chawla$ and Amit Kale*

*Bosch Corporate Research, $Carnegie Mellon University

Objectives

- To automatically segment optic disk (OD) and cup regions in fundus images to derive clinical parameters, such as, cup-to-disk diameter ratio (CDR), to assist glaucoma diagnosis. An eye fundus camera is non-invasive and low-cost, enabling the screening of a large number of patients quickly.
- Discuss various strategies on how to leverage multiple doctor annotations and prioritize pixels belonging to different regions during network optimization.
- Evaluate proposed approaches on the Drishti-GS dataset.

Motivation

Due to a lack of depth information, an expert takes **eight minutes** to annotate each fundus image. Hence, an automated system that marks the boundaries will be extremely helpful. Although several segmentation methods exist, generation of a reliable cup boundary from the CFI is still a challenging task due to the lack of a clear visual demarcation between the cup and the disk as shown in figure below.

Figure 1: Example fundus image

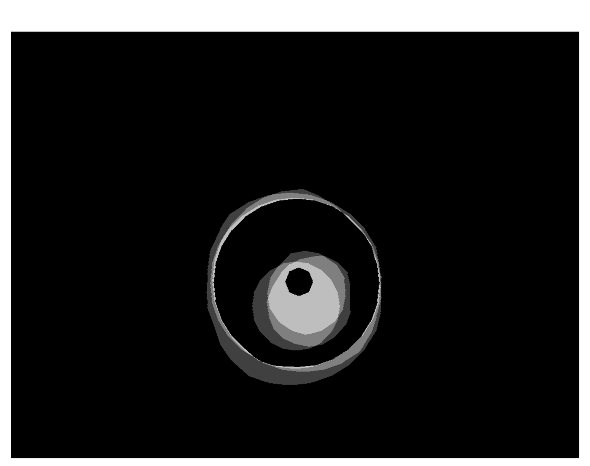

Figure 2: Disagreement softmask - cup(inner) and disk(outer)

Existing Approaches

- **Classical Approaches:** A combination of image processing methods, such as, edge detection, wavelets, hough transforms, active contours based approach, to parameterize disk and cup boundaries needing significant hand tuning of parameters
- **DL Based Approaches:** Separate cup and disk segmentation pipelines requiring special pre-processing and significant image resizing. Features from imagenet pre-trained models with layers for disk and blood vessel segmentation but not cup segmentation

Proposed Approach

- Present a system using FCN8s architecture that generates cup and disk segmentation in a single shot using one deep neural network on full resolution images
- Propose various strategies of utilizing multiple expert annotations and prioritizing certain regions during training for optimal boundary retrieval.

1. **Exp 1:** Each pixel contributes equally to the loss
2. **Exp 2:** Cup/Disk pixels contribute 10x more loss than background pixels
3. **Exp 3:** Cup/Disk boundary pixels contribute 10x more loss than other pixels
4. **Exp 4:** Ignore the regions of contention between annotators for loss computation
5. **Exp 5:** Pixels with higher disagreement contribute less to the loss computation

Network training

Training of FCN involves minimization of cross entropy loss between groundtruth and the network output through backpropagation. Inputs (shown in green) are the RGB image, groundtruth segmentation and an optional weight mask (dotted box).

Data and System Overview

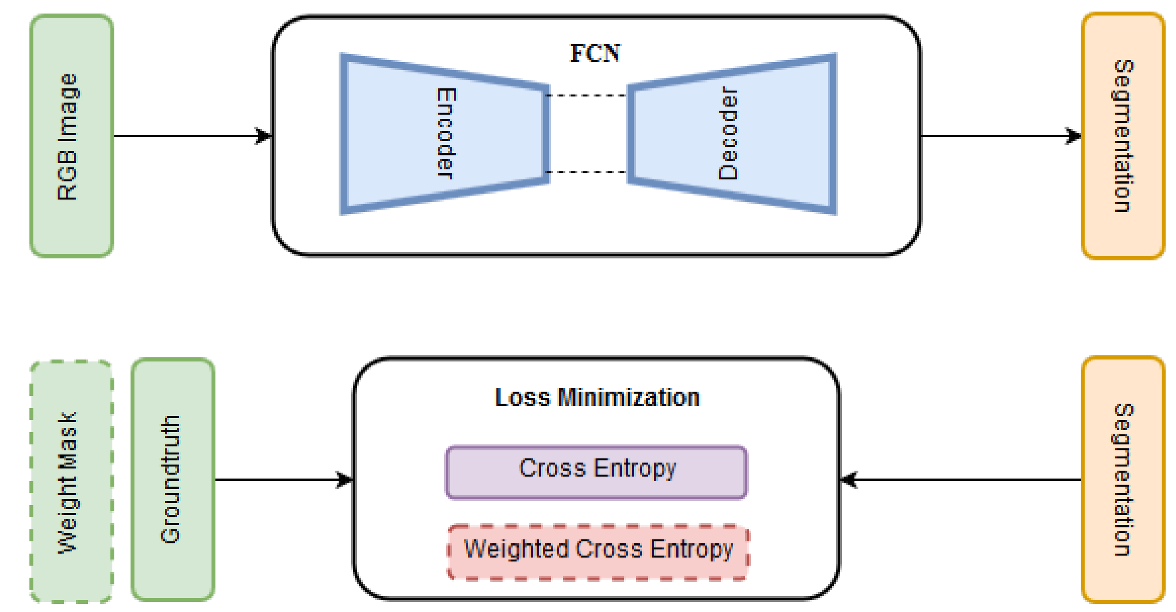

Figure 3: System overview

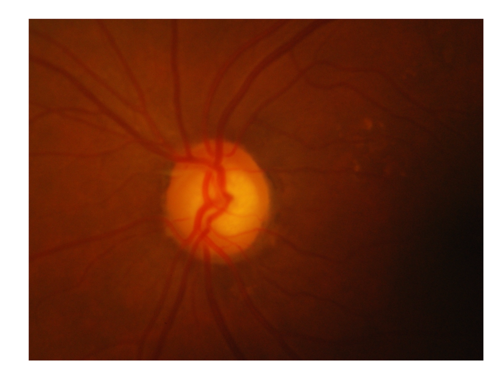

Figure 4: Input image

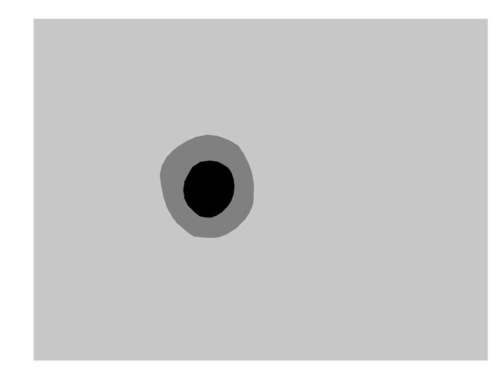

Figure 5: Ground truth

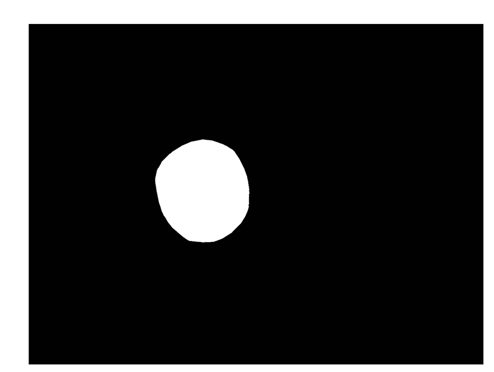

Figure 6: Mask - Exp 2

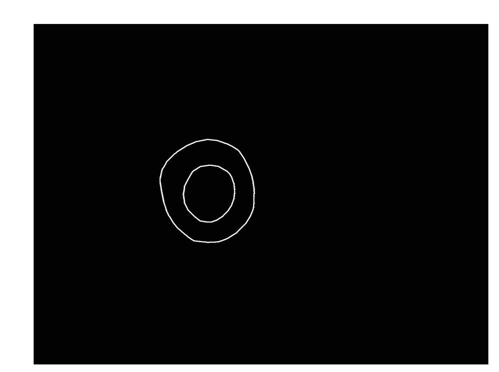

Figure 7: Mask - Exp 3

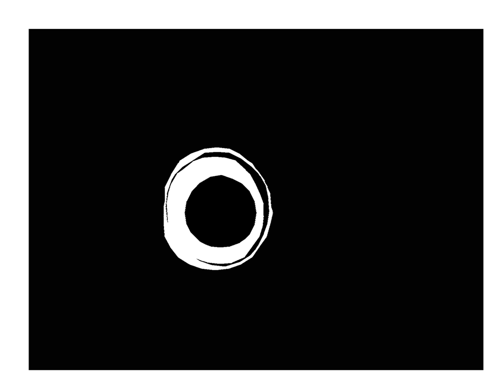

Figure 8: Mask - Exp 4

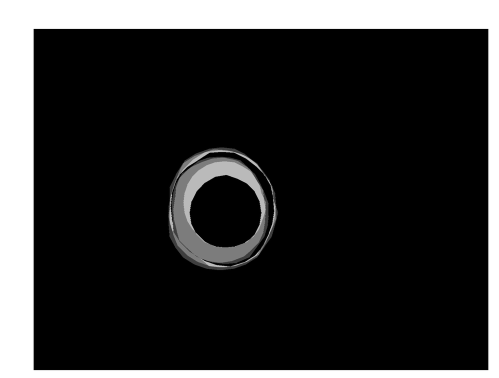

Figure 9: Mask - Exp 5

Evaluation metrics

$$IoU = \frac{TP}{TP + FP + FN} \qquad (1)$$

$$F1 = \frac{2 * TP}{(2 * TP) + FP + FN} \qquad (2)$$

Results

Table 1: Best validation set checkpoint

	Segmentation Strategies				
IoU	Exp1	Exp2	Exp3	Exp 4	Exp5
Mean	0.8509	0.8559	0.8660	0.8823	0.8055
Cup	0.8055	0.8240	0.8284	0.8450	0.7640
Disk	0.7507	0.7482	0.7743	0.8030	0.6591
Void	0.9964	0.9957	0.9962	0.9970	0.9926

Table 2: Test Set

	Segmentation Strategies				
IoU	Exp1	Exp2	Exp3	Exp4	Exp5
Mean	0.8346	0.8341	0.8266	0.7940	0.7795
Cup	0.8122	0.8139	0.7920	0.7720	0.7660
Disk	0.6958	0.6931	0.6911	0.6160	0.5802
Void	0.9959	0.9952	0.9959	0.9940	0.9917

Table 3: F-score comparison with prior state of the art

	cup	disk
G. D. Joshi et al.	0.84	0.97
J. Sivaswamy et al.	0.79	0.96
J. Zilly et al.	0.871	0.973
A. Sevastopolsky	0.85	–
Proposed Exp1	0.897	0.967

Inconsistencies of network performance between validation and test set are attributed to difference in distribution of training and test data. However, our model still outperforms prior methods on the test set.

Conclusion

Experimental evaluations on Drishti-GS dataset have shown comparable and superior F-score to prior work on optic disk and cup segmentation, respectively. Due to the complexity of the network, a focus of future work could be on using network pruning techniques for parameter reduction and inference acceleration, and coming up with a compact architecture.