AO-SLO Image Analysis Using Deep Learning

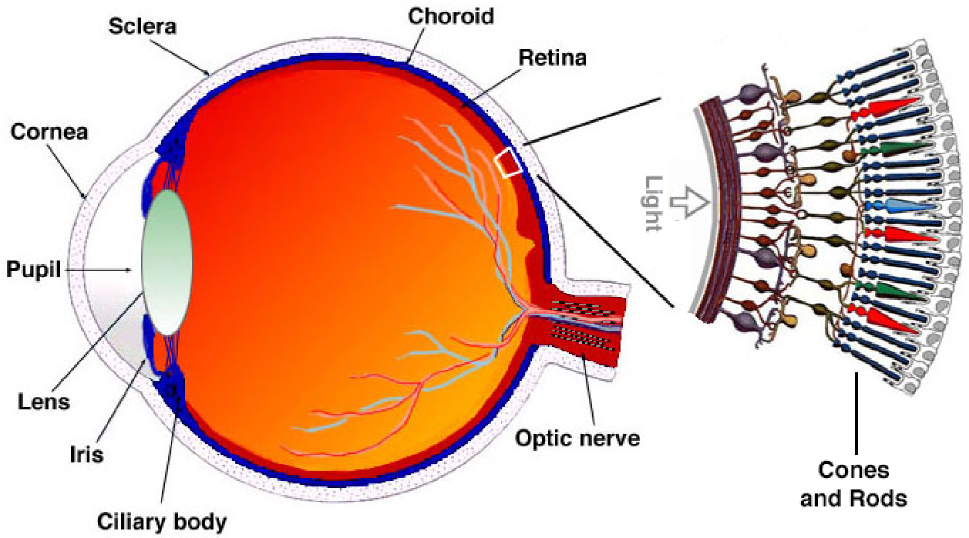
Cone/Rod Recognition

Mengxi Zhou
Sruthi Ammannagari
Fan Shen
Zilin Wang
The Ohio State University

Outline

- Background
- Problem Definition
- Dataset
- Proposed Method
- Experiments and Results
- Future Work

Background



In this project, we focus on the outermost layer of the retina where the layer of cones and rods resides. Cones and rods are the photoreceptors responsible for mediating the sense sight.

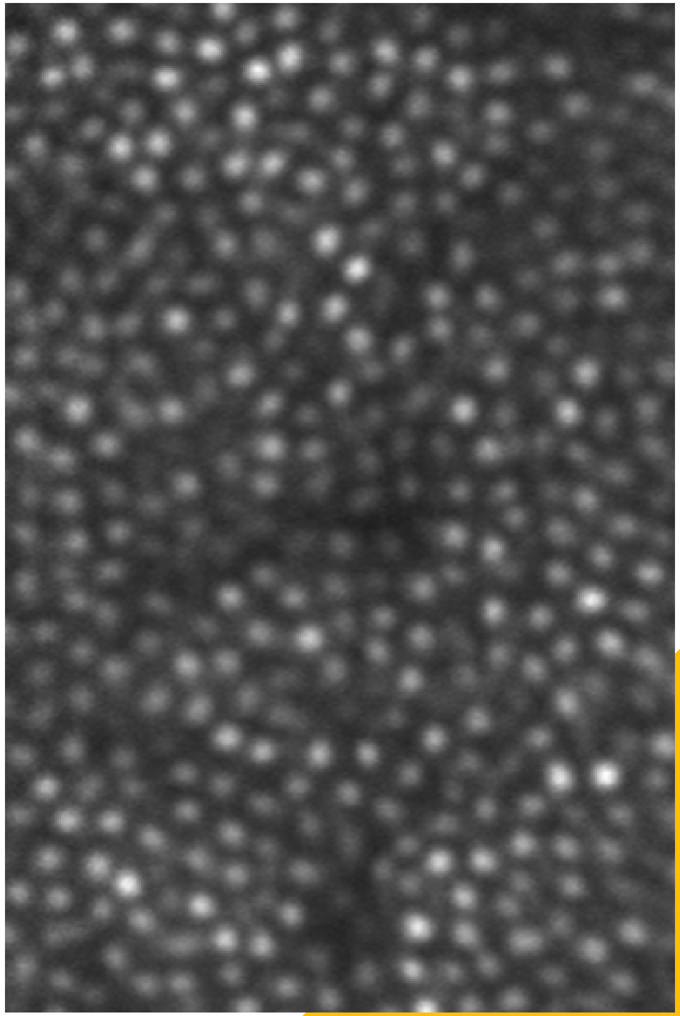
Cones: responsible for the visual acuity of the human eye (the ability of the eye to resolve and to pick up the minor details on an object) and for distinguishing colours.

Rods: are not able to distinguish colours. However, they are highly sensitive to low-intensity light, thus are mainly responsible for night vision.

Background

For certain diseases, such as glaucoma and macular degeneration, they can affect the layer of cones and rods even in the early phase of its formation. And the density of cones/rods can be a potential indicator of the disease. If an anomaly (of the density) is detected, disease can be spotted earlier thus proper treatments can be provided to the patient.

Adaptive Optics-Scanning Laser Ophthalmoscopy (AO-SLO) is utilized to visualize the layer of cones and rods. It is an emerging imaging technology which provides more detailed (cell-leveled) information.



A Sample AO-SLO Image

Problem Definition

In order to compute the cones/rods density, it is equivalent to count all cones/rods in the AO-SLO image (and then divided by the actual area). Further, it is equivalent to identify all centers of cones/rods and count the centers. Thus, the problem is defined as the following:

Given an input AO-SLO image I, the problem is to classify each pixel in I as either 0 (not at the center of a cone) or 1 (at the center of a cone).

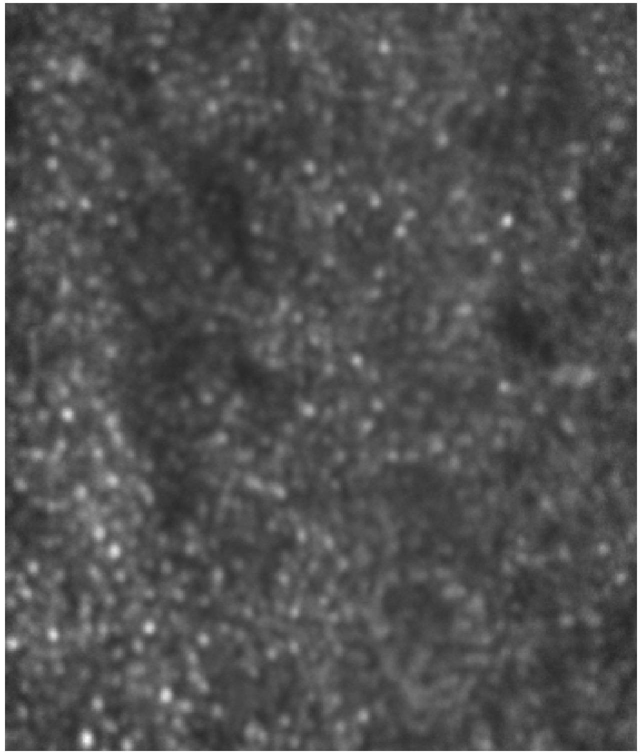
Dataset

The dataset is generated and provided by the college of optometry, OSU. Currently the dataset contains:

- 91 images (from 12 subjects)
- x, y coordinates of the cone centers on each image.

The quality of the images varies (in terms of contrast and clarity). The labels are generated by first finding the intensity peaks using a rule-based algorithm, then manually modified by the domain experts.

A sample AO-SLO image along with the x, y coordinates for cone centers is shown on the right.



Cone Index	x_Coord_Ref	y_Coord_Ref
1	590	458
2	587	66
3	526	270
...	...	...

Model 1: CNNs

Cunefare, David, et al. [3] proposed a method using Convolutional Neural Networks (CNNs) to recognize the cones/rods. Essentially, the method tried to divide a whole AO-SLO image into small patches and tried to predict each small patch as either having a cone or not having a cone. During training, the method extracted both positive and negative patches according to the ground-truth. When doing inference, a small sliding window would go through the image and made prediction at each position.

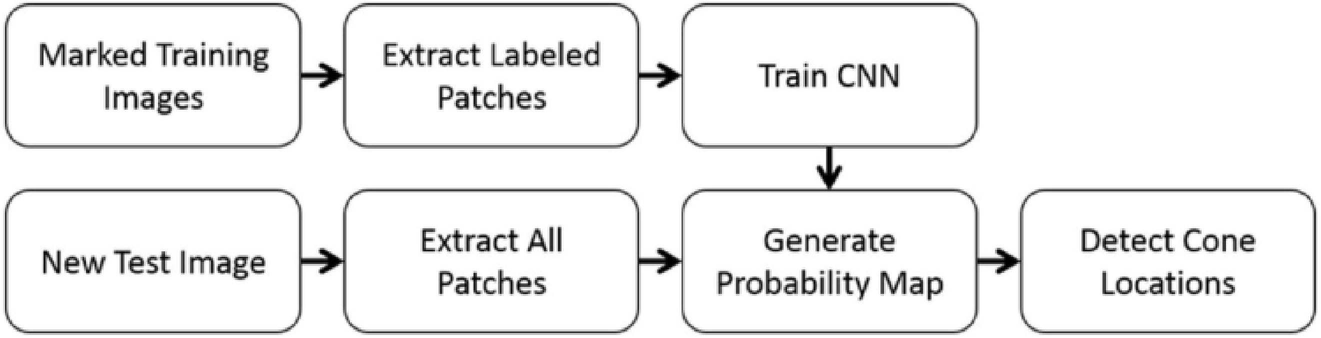
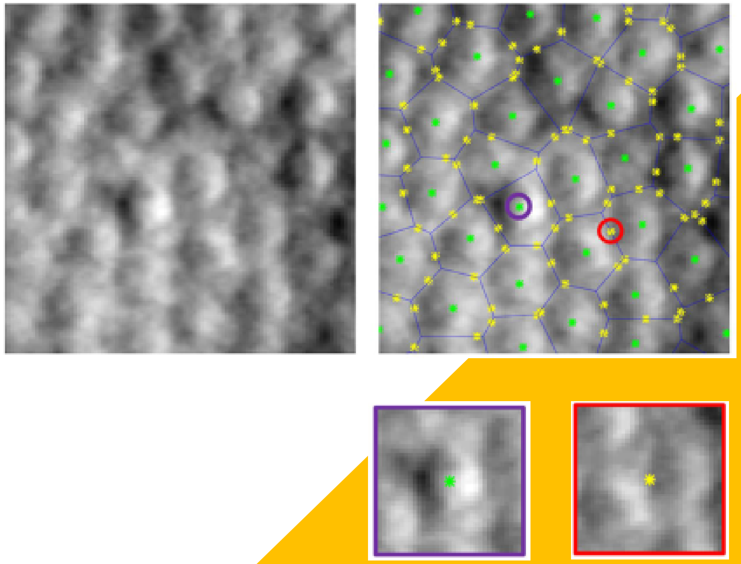
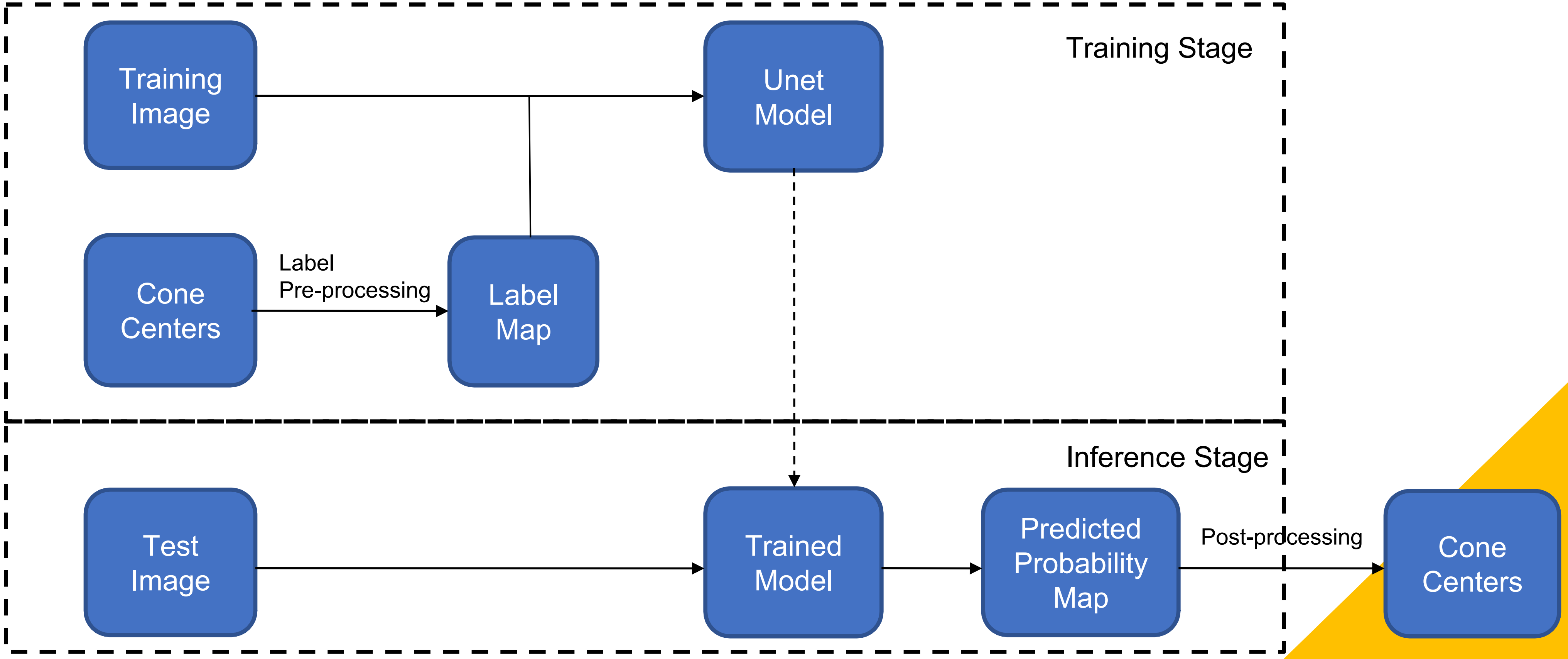


Figure 2. CNN based cone detection algorithm schematic.



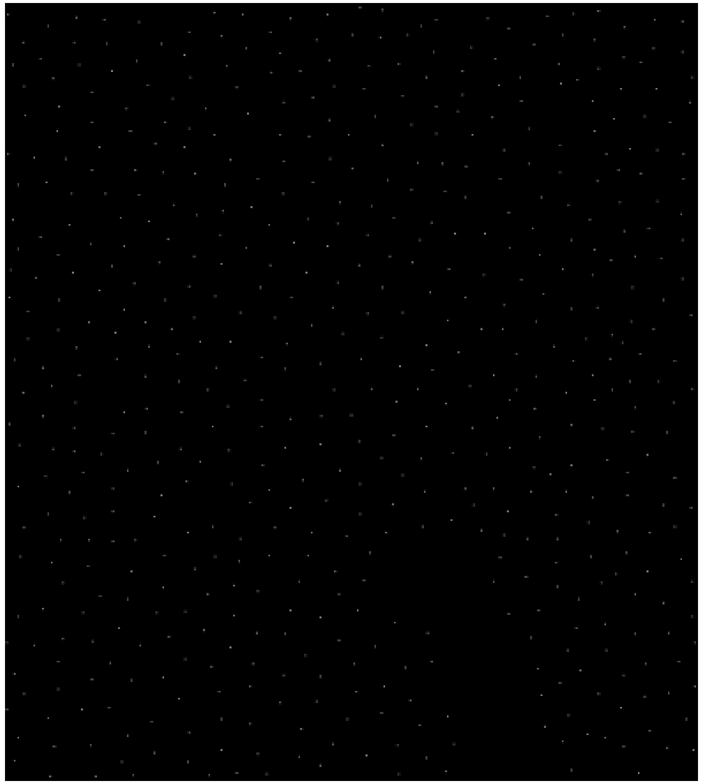
Proposed Method



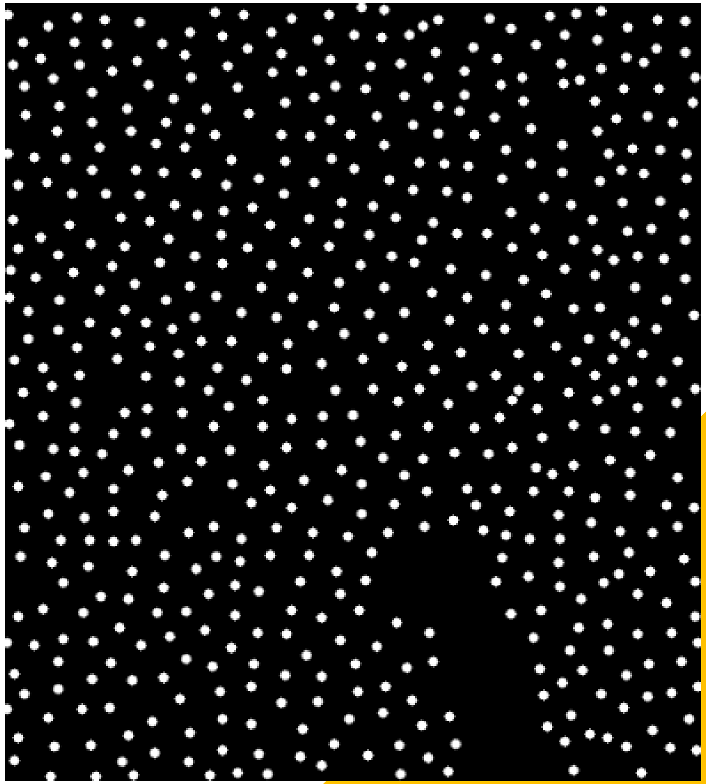
Label Pre-processing

Cone Index	x_Coord_Ref	y_Coord_Ref
1	590	458
2	587	66
3	526	270
...	...	...

Original Label Sheet



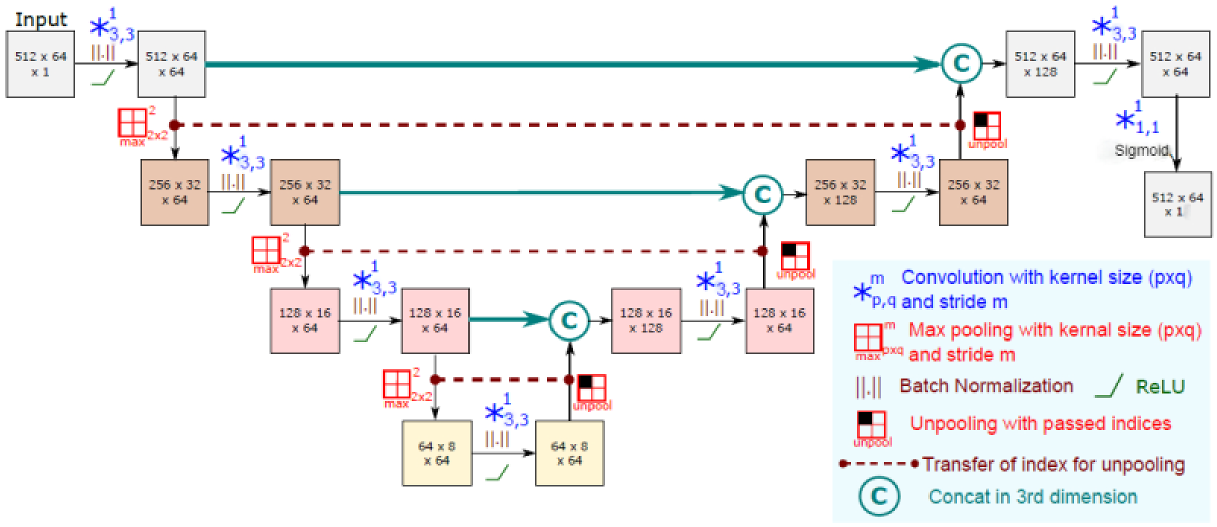
(1) 1-pixel Label Map



(2) Expanded Label Map

Model 2: Unet

Unet is a convolutional neural network that is largely used for image segmentation tasks. It consists of an encoder which down samples the input image to dense feature maps, and a decoder that up samples these feature maps back to original dimensionality retaining only the desired representations, in our case:

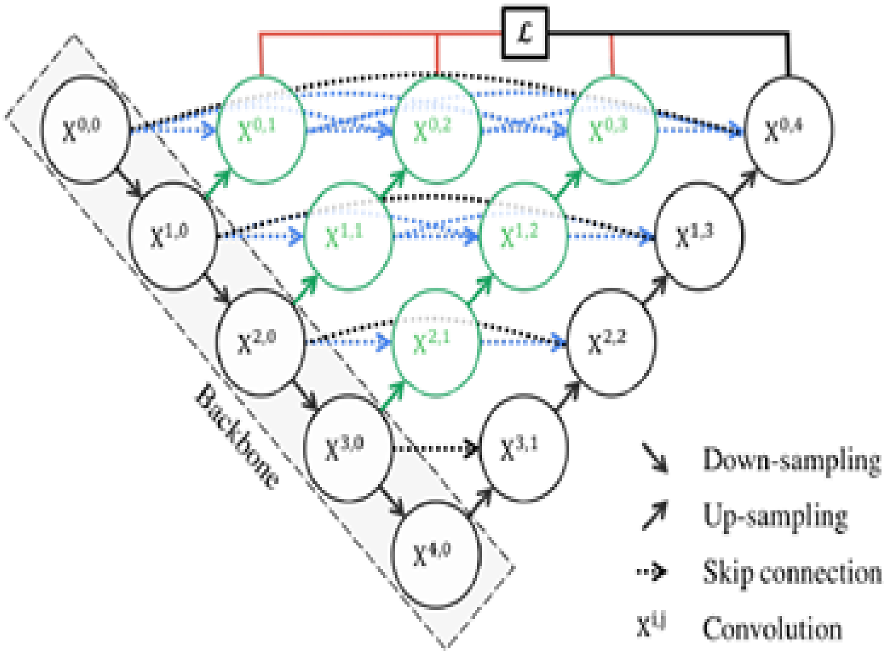


- 0: the pixel is not at the center of a cone
- 1: the pixel is at the center of a cone

Model 3: Unet++

Unet++ is the modified version of Unet, developed by researchers in Arizona State University. It shares the similar encoder/decoder architecture except the following aspects:

- Convolutional layer in the skip pathways
- Nested and dense skip connections in the skip pathways
- Deep supervision: accurate mode and fast mode



Post-processing

1.Non-maximum suppression: This step aims to only retain the strongest prediction/pixel within a small area (α * α), since each cone should only have one center. Other predictions/pixels will be set as 0s.

2.Thresholding: For the prediction at each pixel, we did a thresholding (>=β) to binarize the decision (to 0/1).

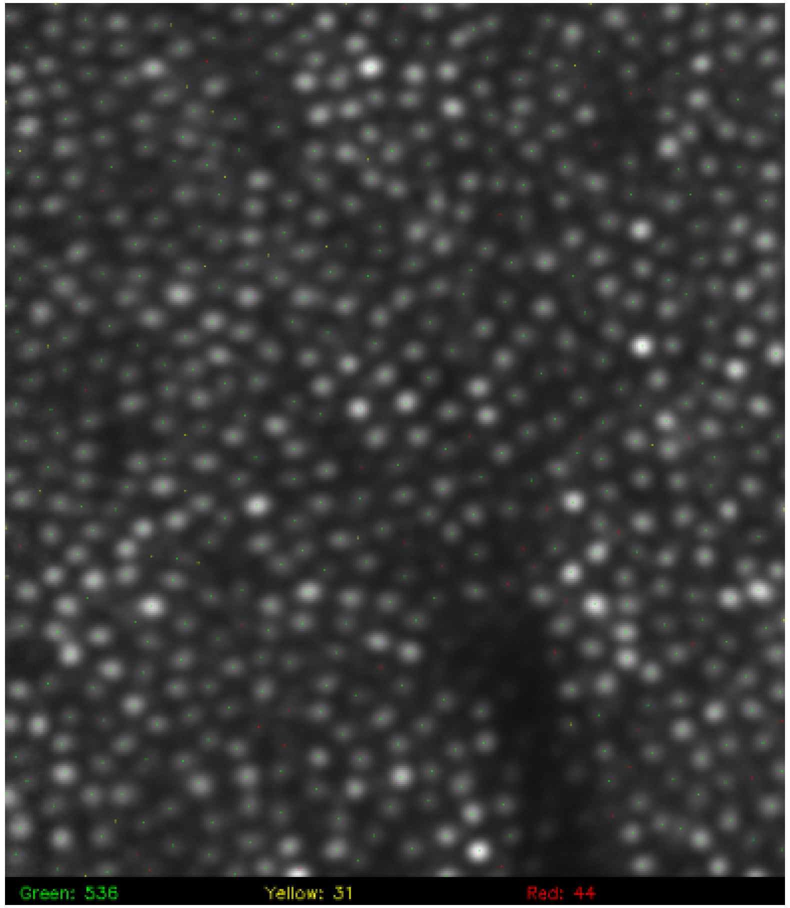
3.Ground-truth Matching: Considering that a predicted pixel (potential cone center) may not be at the exact location where the ground truth marked, we give a small allowance (10 pixels in the current experiments) for a positively predicted pixel to be matched to a truly labeled pixel.

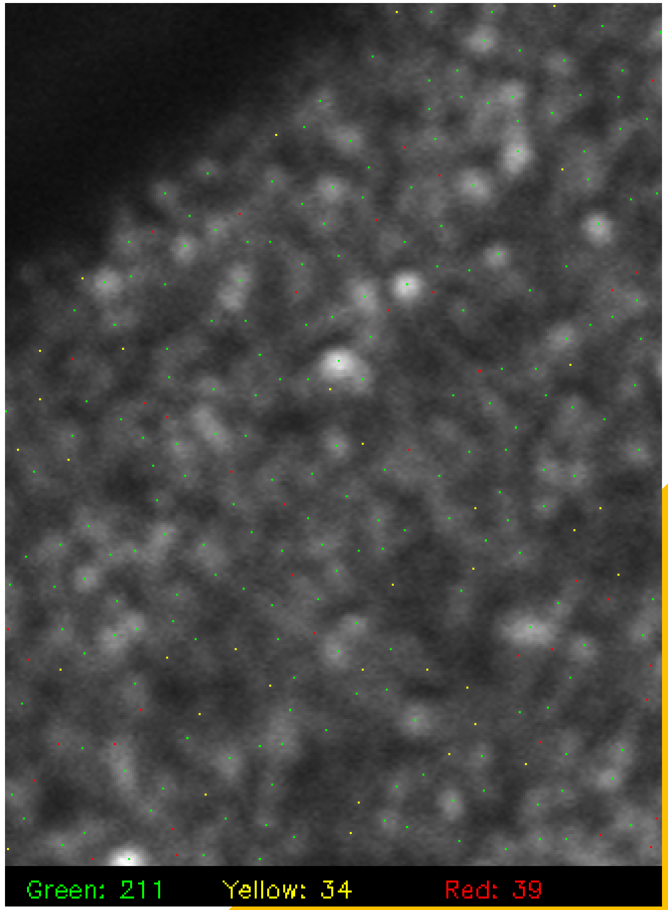
Experiments and Results

Green: True Positive
Yellow: False Positive
Red: False Negative

Left: good imaging
Right: poor imaging

(Ground-truth refinement may be needed.)





Experiments and Results

model_name	precision	recall	f1	loss	Pre-processing
Unet_v0	0.7884	0.8557	0.8207	bce	1-pixel
Unet_v1	0.8182	0.8178	0.818	bce	1-pixel
Unet_v1	0.8084	0.8353	0.8216	bce	expanded
Unet++_v0	0.7916	0.8755	0.8314	bce	1-pixel
Unet++_v1	0.7739	0.9034	0.8337	bce	1-pixel
Unet++_v1	0.734	0.8861	0.8029	focal	1-pixel
Unet++_v1	0.8482	0.8241	0.836	bce	expanded
CNN(David Cunefare)	0.816	0.9374	**0.8725**	bce	-

Another Approach: Object Detection

```
import rioxarray

rds = rioxarray.open_rasterio("cropped_im.tif")
df = rds.to_dataframe(name = "value")
```

```
df = df.reset_index()
df = df.drop(['spatial_ref','band'], axis = 1)
df
```

	y	x	value
0	0.5	0.5	177
1	0.5	1.5	173
2	0.5	2.5	173
3	0.5	3.5	173
4	0.5	4.5	171
...	...	...	...
289333	580.5	493.5	169
289334	580.5	494.5	154
289335	580.5	495.5	143
289336	580.5	496.5	131
289337	580.5	497.5	120

Rioxarray:
Open Rasterio Adaptation that preserves the location data of the pixels

Another Approach: Object Detection

Out[170]:

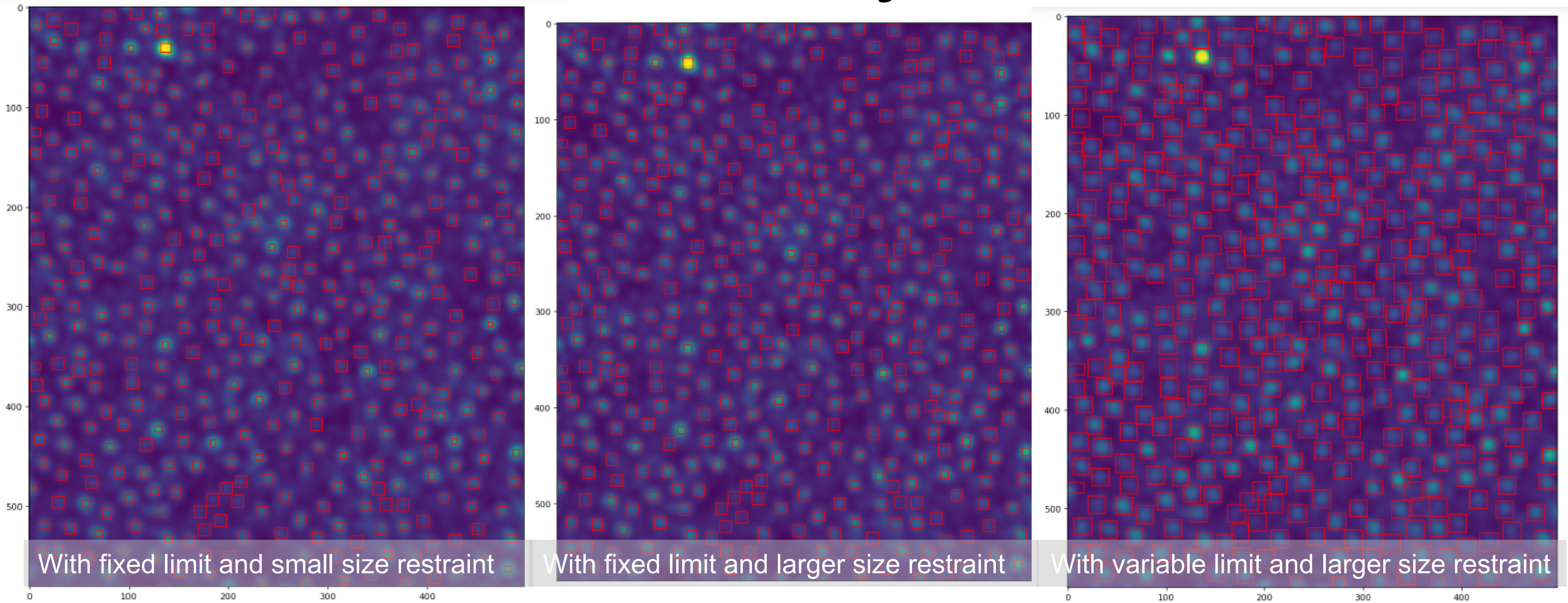
	0	1	2
0	[132.0, 36.0]	9.0	10.0
1	[99.0, 38.0]	6.0	5.0
2	[486.0, 443.0]	7.0	6.0
3	[337.0, 362.0]	6.0	5.0
4	[126.0, 420.0]	6.0	7.0
...	...	...	...
429	[103.0, 72.0]	10.0	10.0
430	[117.0, 74.0]	10.0	10.0
431	[443.0, 84.0]	12.0	11.0
432	[238.0, 134.0]	12.0	12.0
433	[346.0, 299.0]	11.0	12.0

Center coordinate available

Extends the border of the bounding box when extreme drop off in pixel intensity is presented

Calculate the upper left corner of the bounding box and size

Another Approach: Object Detection



Results with different parameter tuning

Future Work

- Ground-truth refinement
- Complete and examine object detection models
- Model exploration for superior performance
- Distinguish cones and rods
- User study on applying the model to the current analysis loop for domain experts

References

- [1] Roy, Abhijit Guha, et al. "ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks." Biomedical optics express8.8 (2017): 3627-3642.
- [2] Liang, Jiongqian, et al. "Semi-supervised embedding in attributed networks with outliers." Proceedings of the 2018 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2018.
- [3] Cunefare, David, et al. "Open-source software for automatic detection of cone photoreceptors in adaptive optics ophthalmoscopy using convolutional neural networks." Scientific reports 7.1 (2017): 1-11.
- [4] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenetclassification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
- [5] Rawat, Waseem, and Zenghui Wang. "Deep convolutional neural networks for image classification: A comprehensive review." Neural computation29.9 (2017): 2352-2449.

References

- [6] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
- [7] Wells-Gray, Elaine MM, et al. "Volumetric imaging of rod and cone photoreceptor structure with a combined adaptive optics-optical coherence tomography-scanning laser ophthalmoscope." Journal of biomedical optics 23.3 (2018): 036003.
- [8] Girshick, Ross. "Fast r-cnn." Proceedings of the IEEE international conference on computer vision. 2015.
- [9] Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems. 1097–1105 (Springer, 2012).