# Team OCT-opus

Samin Saberi, Kathleen Wang,
Pei Lin Li, Kenneth Sinder, Céline O'Neil

4B Software Engineering - Research FYDP

[Presenter: Kenneth] Hi, we're team OCT-opus, and today we'd love to talk to you about the medical imaging problem we've been tackling for Prof. Dida Bizheva at UW Physics. Before we're able to jump into it, there's a little bit of background we need.

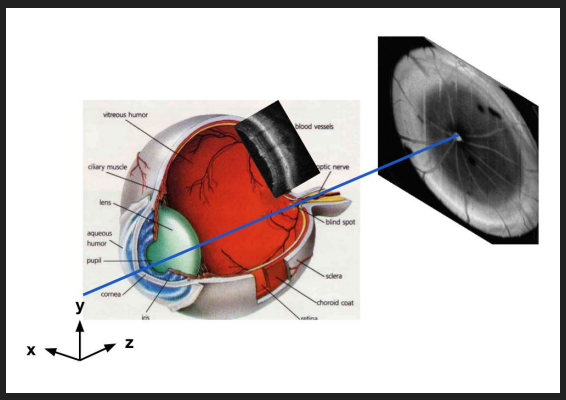
# Background

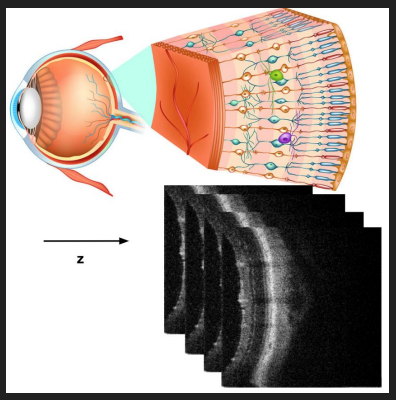
- **Optical coherence tomography (OCT):** non-invasive imaging technique used to capture cross-sections of retinas
- **OCT angiography (OCTA)**: extension of OCT, uses repeated acquisition of each cross-section & combines using software to image blood flow
  - One such algorithm: **optical microangiography (OMAG)**
  - However, due to multiple acquisition, might require sedation
  - In addition, the machine will need to be reconfigured

[Presenter: Kenneth]

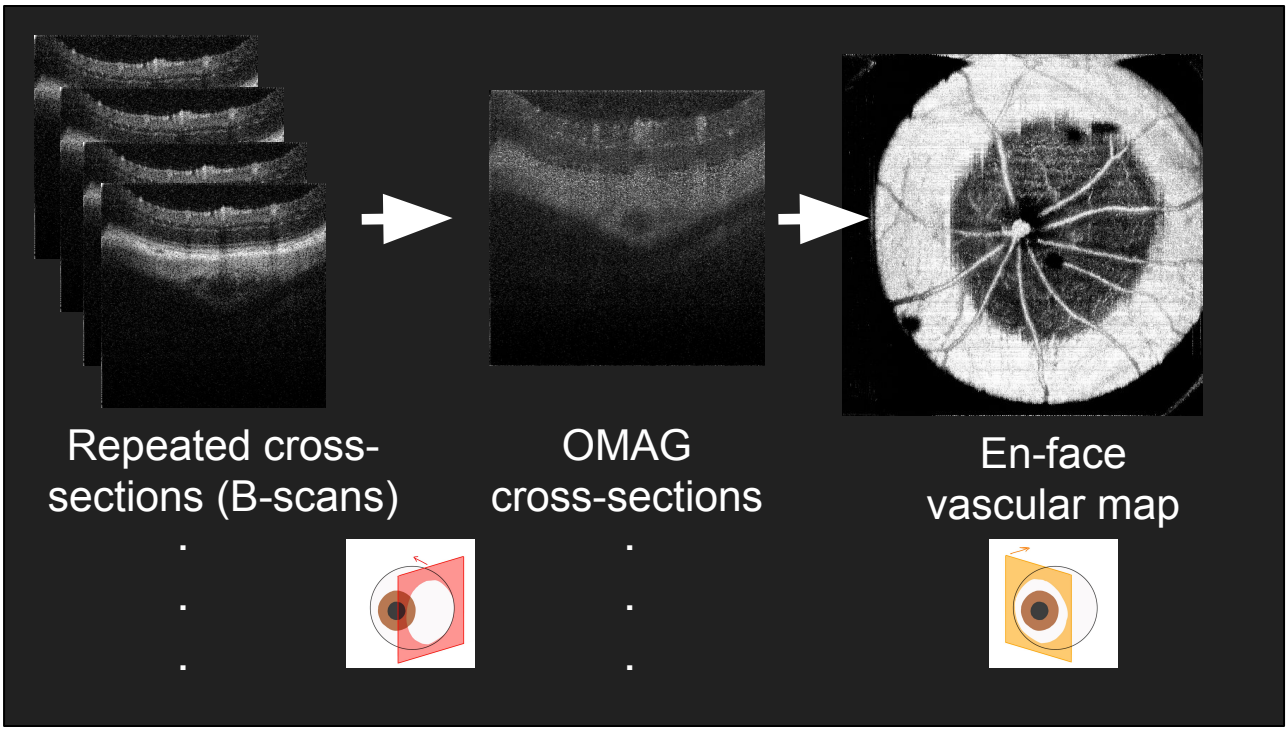
- OCT is a technique that's used for over two decades to capture high resolution images of living tissue up to 2mm deep non-invasively using interference of near-infrared light, and one of its applications is to capture a sequence of cross-section images representing a retina
- OCT angiography is an extension of OCT requiring, at minimum, additional software and hardware, which involves capturing repeated acquisitions of the same spot and combining them to effectively convert structural images into functional images capturing blood flow. One algorithm that combines each small chunk of images is called optical microangiography, or OMAG.

# Prerequisite Terminology




[Presenter: Pei Lin]




[Presenter: Pei Lin] When we use OCTA to capture cross-sections of retinas with the goal of imaging blood flow, we end up with a sequence of, say, 2048 B-scans, and then we run an algorithm like OMAG to collapse that down to 512 OMAGs, combining each set of 4 down to 1 subtracting off to reveal the blood flow. Then we take those 512 OMAGs and combine them to form a cube and take a slice of them to create a combined en-face vascular map of the retinal capillaries.

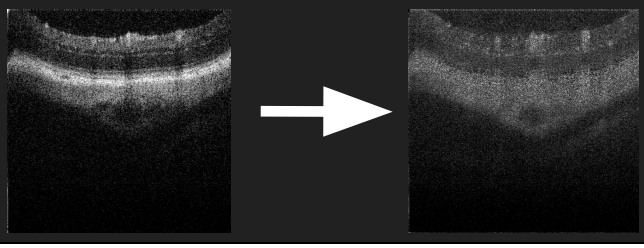
# Enface Fly-Through

[Presenter: Pei Lin]

## Problem

- We want to train a system that can **infer OMAG-like images from single (non-repeated) B-scans**
- ... and create a resulting en-face vascular map of comparable quality
- Image-to-image translation problem:

B-scan (Structural)



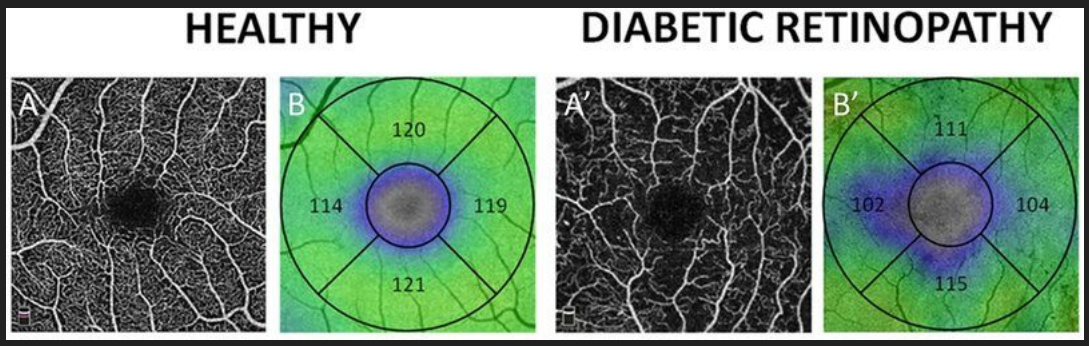
OMAG (Functional)

[Presenter: Kenneth]

- This leads us to the problem we're tasked with. We want to take some existing pool of OCTA data from Prof. Dida Bizheva's lab and use it to train a system that can predict OMAG-like blood flow cross sections that look reasonable in some sense from single OCT images.
- We can then run our script to assemble the en-face image from the inferred OMAG-like set of images.
- Fundamentally, this is an image-to-image translation problem, hinting at some possible traditional computer vision techniques or deep learning image processing approaches which might work here.

## Why?

- Vascular map used to diagnose and monitor diseases




- OCTA hardware is expensive
- Make use of existing OCT-only datasets

[Presenter: Kenneth]

- The core motivation for why we even care about these enface images in the first place is that quality enfaces can be indicators of diseases affecting the human eye. For instance, diabetic retinopathy is a condition that results as a complication of diabetes and produces often noticeable changes in the thickness and structure of retinal capillaries. Clinicians can carefully examine vascular maps and extract meaningful information.
- But why do we want to be able to get to this point from single acquisition OCT imaging?
- Repeated acquisition hardware and software can be very costly and time-consuming to execute and configure and few labs have the sort of OCTA equipment that Prof. Dida Bizheva's lab does
- Breathe new life into structural retinal B-scan datasets from previous years available on the internet and turn them into OMAG-like functional tissue imaging for further research
- Image Source: https://www.researchgate.net/figure/En-face-OCTA-angiograms-black-and-white-and-retinal-thickness-maps-color-coded-maps_fig3_334557094

## Why? (cont'd)

- Project originally proposed by **Prof. Kostadinka (Dida) Bizheva** from Physics & her grad student, **Zohreh Hosseinaee** out of multiple possible options
- Zohreh switched research groups but Dida still considers this problem worthy of solving & publishing on its own, as well as being useful in her lab with future students
- Original sedated rat data provided to us by Zohreh, captured by **Bingyao Tan**

[Presenter: Kenneth]

- This is one of three options that Prof. Dida Bizheva and her graduate student at the time, Zohreh, presented to us.
- We could have also worked on segmenting cross-section layers or processing corneal images, but Zohreh had the most data available for the retinal image translation project from sedated rat OCT that they'd done, and expressed that this was an important piece Prof. Dida Bizheva's lab is looking for.

# Stakeholders

Bingyao Tan
Postdoctoral Fellow
Dept. of Physics

Prof. Kostadinka Bizheva
Dept. of Physics

Zohreh Hosseinaee
Postdoctoral Fellow
Dept. of Physics

[Presenter: Kenneth]

- Bingyao Tan is another crucial stakeholder in this research because he was the individual who originally imaged the sedated rats, and there's Prof. Dida Bizheva and Zohreh from Physics.

# Additional Support

Prof. Alexander Wong
Dept. of Systems Design

Prof. Vasudevan "Vengu" Lakshminarayanan
Dept. of Physics

Eric Praetzel
Hardware Specialist
Dept. of ECE

Sourya Sengupta
M.Sc. Student
Dept. of System Design

[Presenter: Kenneth]

- We initially seeked technical machine learning mentorship from Prof. Alex Wong of Systems Design, but when he was busy, we pivoted to Prof. Vengu and his grad student Sourya, who both had plenty of time to mentor us and experience specifically working with very similar deep learning research and are continuing to communicate with us on this project
- Finally, we used Eric Praetzel's eceubuntu4 GPU machine initially but once that became a performance bottleneck for training our model, we talked to Vengu and Sourya and got access to Compute Canada's Sharcnet cluster to create GPU jobs from there which worked well from us.

## Why? (cont'd)

- Even with high freq imaging, human eyes have some motion which causes blur in the resulting OMAG
  - Only requiring a single capture per spot mitigates this

[Presenter: Kenneth]

- A final piece of motivation behind the importance of this project is that when Prof. Dida Bizheva's lab is able to gather human datasets and train and test on those instead of just rat datasets, we'll want to avoid problems caused by applying the OMAG algorithm to moving images, so our software will help us in that regard by requiring only single acquisitions and extrapolating everything we need from there.

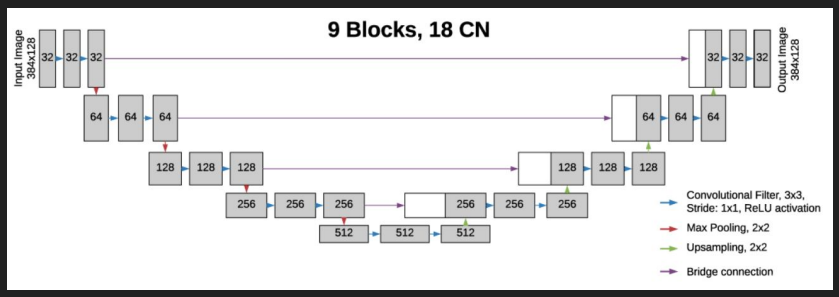
## Provided datasets

- **66** usable rat eyes' worth of data
- 512 (B-scan, OMAG) pairs per eye (4 B-scans used to construct OMAG, but 3 of every 4 B-scans discarded)
  - i.e. **~ 33 500** image pairs for training/testing
- Goal is to switch to using human eyes once they are imaged in Prof. Bizheva's lab

[Presenter: Kathleen]

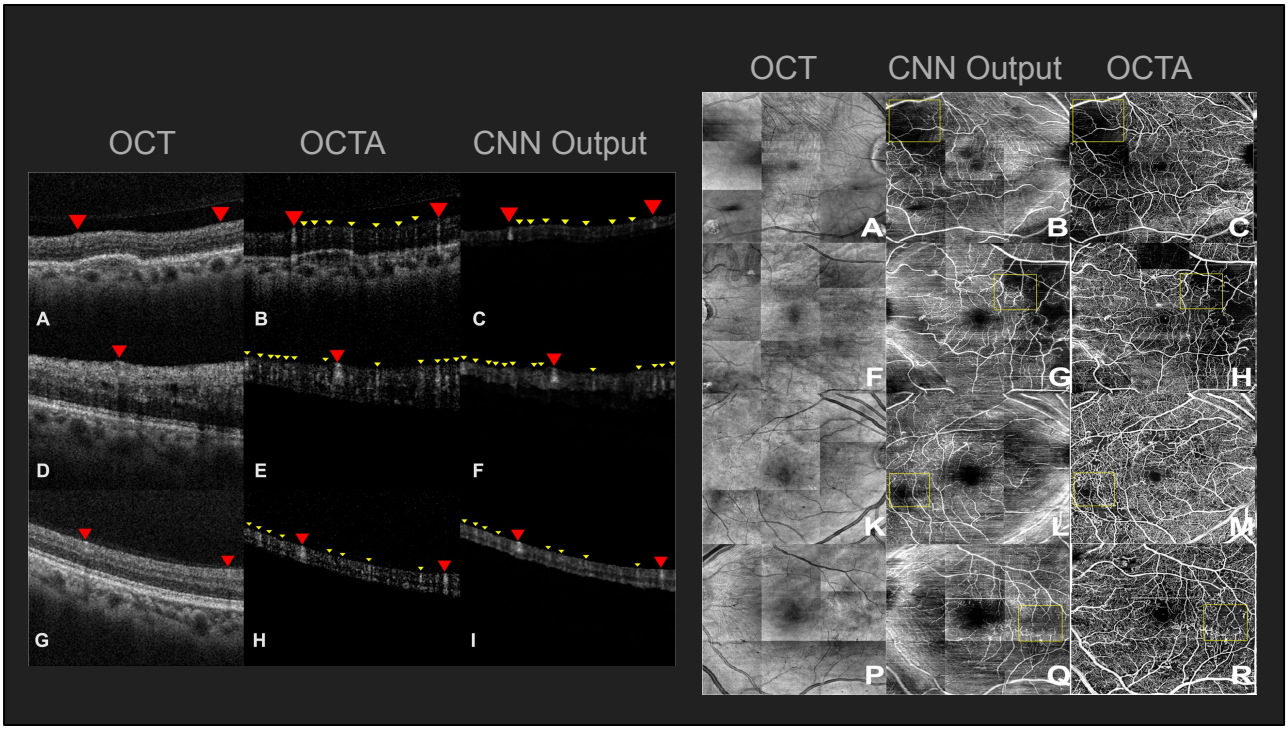
## Do existing solutions exist?

- Recent approach in research world uses CNN: “Generating retinal flow maps from structural optical coherence tomography with artificial intelligence” 2019 paper by Lee et. al




[Presenter: Kathleen]

- None that we could find that are nicely wrapped up with training and testing modes and a GUI that we can deliver to Dida’s lab right away.
- But this has started to crop up in medical imaging research, with a 2019 paper also taking the approach of using OCTA images as ground truth and applying a different deep learning model involving a convolutional neural network (CNN) to perform image-to-image translation.
    - Uses a CNN in the style of a u-net




[Presenter: Kathleen]

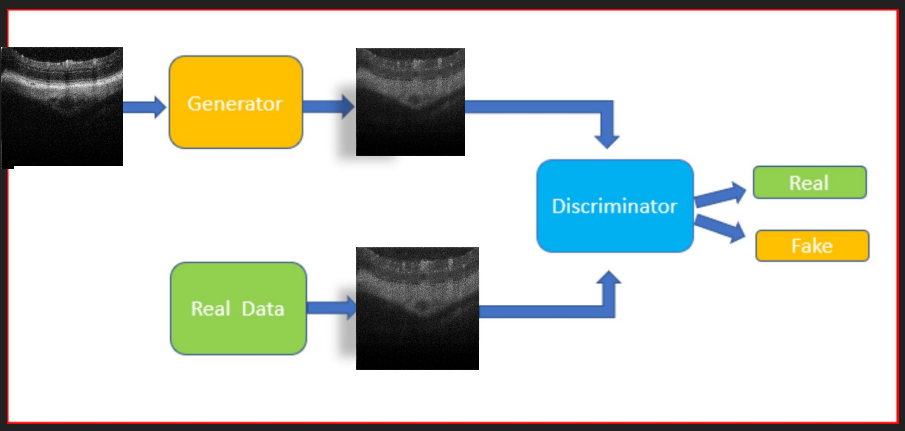
- 
- Working on implementing that baseline
- No code readily available online, reproducing it from the paper's model diagrams
- Idea is to compare the performance of state-of-the-art that was trained with our data with our own model

Another argument for the novelty of our work is once we hand it off, Dida can use it, retrain on human data when that becomes available, and more easily integrate it into her broader research in a paper she may want to write that documents her state of the art lab hardware, etc.

There's also value in applying an existing model to a new application, to enrich the academic field

# How?

- Generative adversarial network:

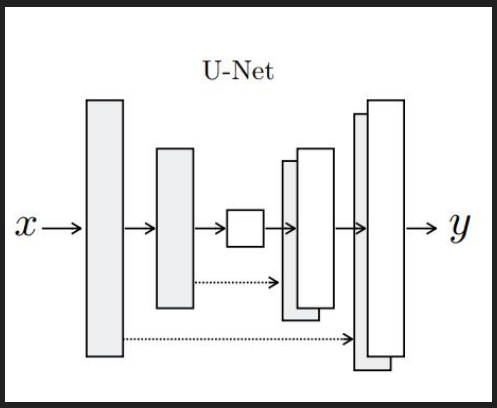



[Presenter: Celine]
Img src
Using midterm 490 slides for inspiration
https://docs.google.com/presentation/d/1K2UM_fVzbja74bkp1QK9WQ5OzqIA402T5HK3X6Ethqc/edit#slide=id.g5bd09486b1_6_47

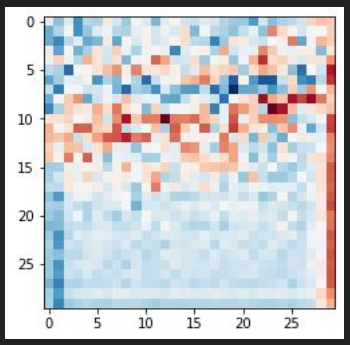
We first attempted using a Generative Adversarial Network, or GAN, which are good at image-to-image translation tasks. The general idea behind GANs is to play off two competing neural networks against each other. One, the generator, takes at BScan as input and attempts to produce an OMAG from it that will fool the second neural network, the discriminator, into thinking that it's the real OMAG. Both neural networks start out poorly, but quickly learn to better "outthink" each other.

# Architectural Diagram

U-Net (Generator)




PatchGAN (Discriminator)




[Presenter: Celine]

The generator takes in a BScan and must output its best approximation of the OMAG corresponding to that BScan. In order to do this, the generator needs to learn how to recognize structural data in the BScan image. To force this, we use a U-Net, a neural network that successively downsamples the input image and then upsamples it again. This forces the neural network to choose what information is most important to pass through the bottleneck; ideally this is the structural information that it needs to infer a realistic OMAG.

Rather than get the discriminator to classify the entire image as “fake” or “real”, we add a little bit of nuance by making it a PatchGAN; dividing the images into a set number of “patches” and predicting the likelihood of “realness” for each patch. This forces the generator to accurately represent details of the OMAG of the image in order to fool the discriminator; simply getting the general image structure right would not be enough.

## Chosen architecture: “pix2pix” cGAN

- Conditional GAN: generation of the output is conditional on an input
- pix2pix is a cGAN framework for image-to-image translation problems
- Uses modified U-Net for the generator and a PatchGAN for the discriminator
- Augmenting images using random crops, rotations up to 30 degrees, and flips

[Presenter: Celine]
“Image-to-Image Translation with Conditional Adversarial Networks”
Berkeley AI Research Laboratory, 2018

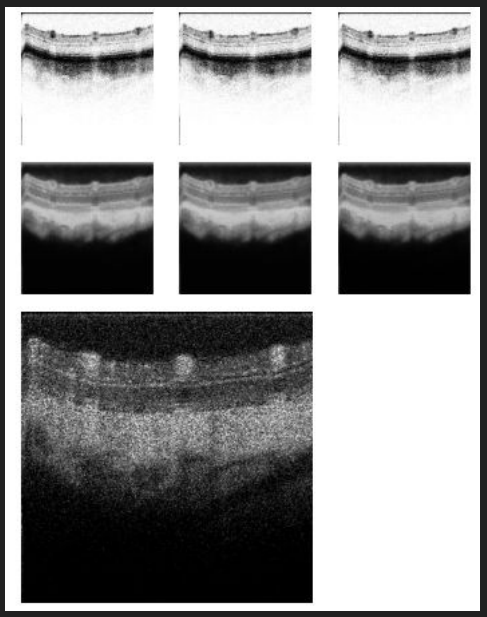
## Why pix2pix?

- Natural choice for image-to-image translation framework
  - Already set up to train on paired images, etc.
- Working on comparison with convolutional neural network (CNN), but pix2pix provides solid baseline to start
  - Would have pivoted to entirely different architecture here if it was necessary
- Fairly recent framework yet has demonstrated success

[Presenter: Celine]

## Modifying pix2pix

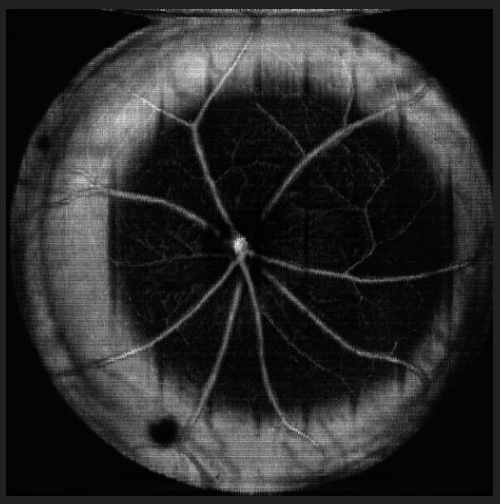
- Capillaries have continuity; the model might be able to use information from adjacent BScans to augment the current BScan
- Experimented with generator taking 3 adjacent BScans and discriminator evaluating the central one
- Inspired by a similar approach in “Image Synthesis in Multi-Contrast MRI with Conditional Generative Adversarial Networks”, Dar et. al.

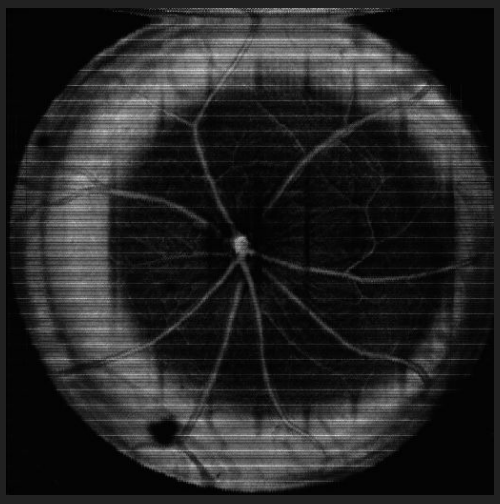


Our main goal is to get a better vascular map of the retina with fewer BScan acquisitions, and the purpose of multiple acquisitions is to get a more accurate idea of the location and size of capillaries across images that might capture slight variations in bloodflow and eye movement. Since capillaries are connected across scans, we theorized that we might be able to get a lot of that information augmentation from scans of adjacent locations, rather than multiple acquisitions of the same location.

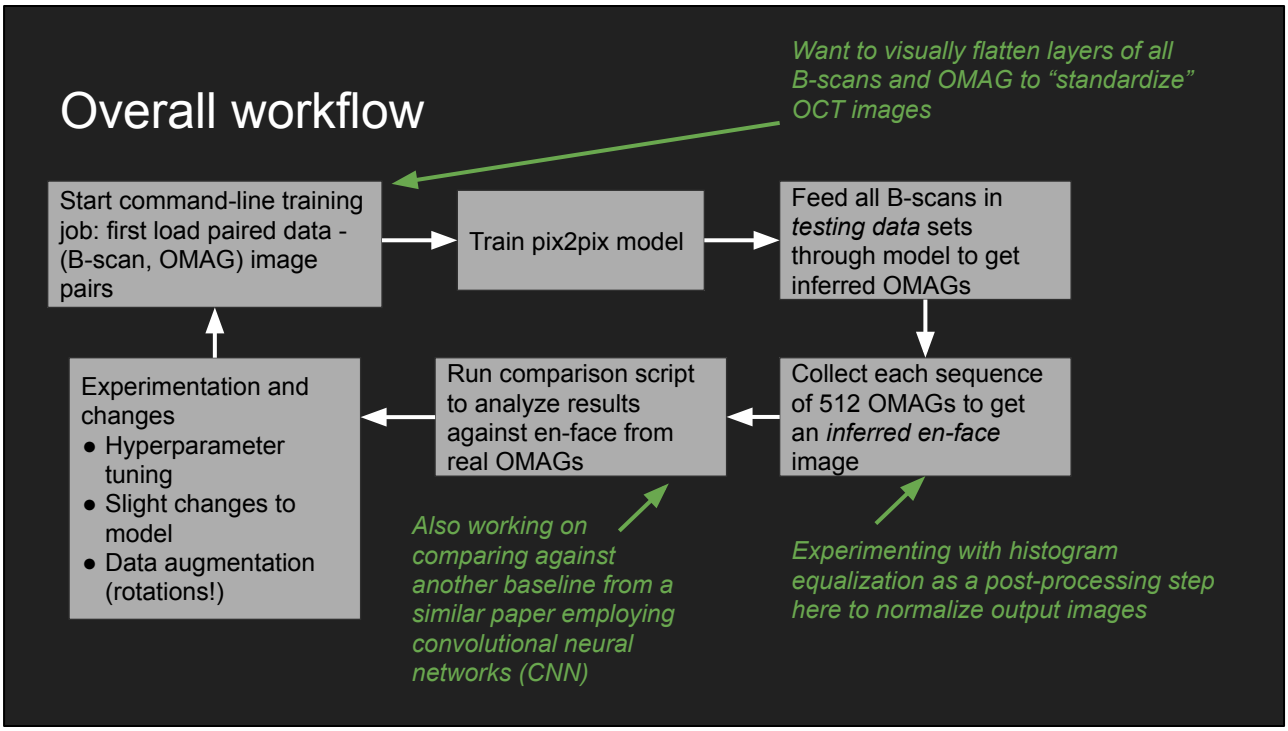
We looked to see if anyone else had done something similar in medical imaging, and based on an approach used in image-to-image translation of MRI cross-sections, we experimented with a generator that takes in 3 adjacent BScans, with the discriminator judging the accuracy of the central scan.

## Modifying pix2pix



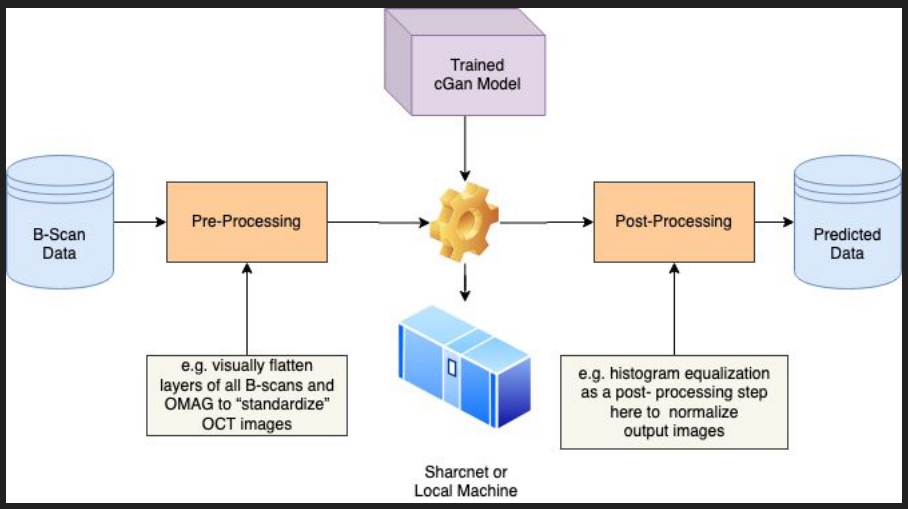


Unfortunately, the added complexity of this approach hasn't yet yielded improved results over the simpler single-BScan model; it currently leads to more artifacts and discontinuities. We're currently sticking to the basic pix2pix pipeline as it's much more reliable and has been producing sufficiently good results.

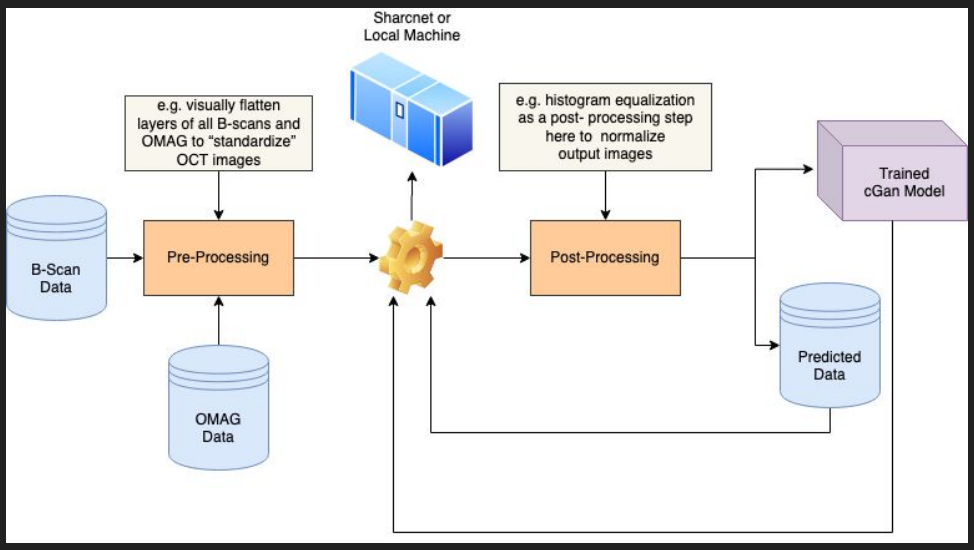



[Presenter: Pei Lin]

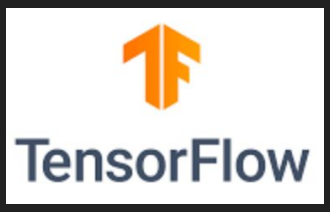
# Prediction Workflow



# Training Workflow



## Software and Hardware



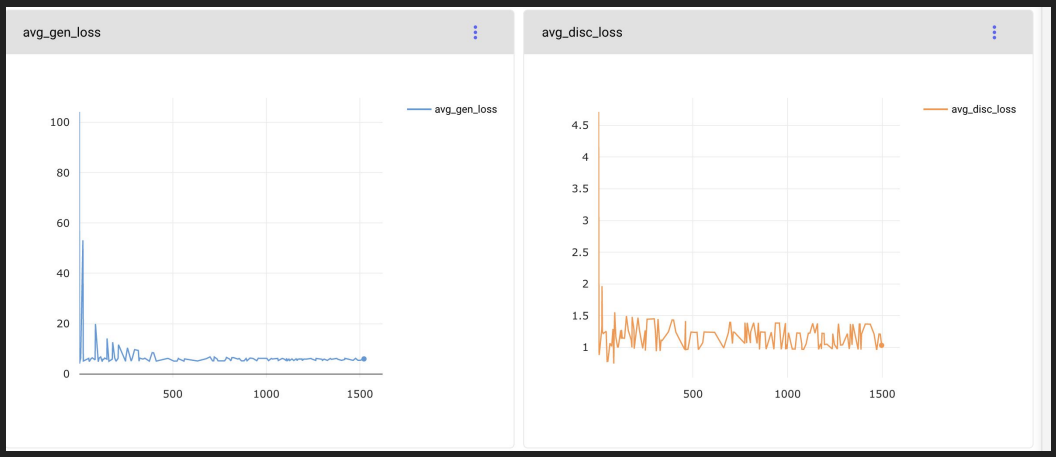




- TensorFlow 2.0
- Python 3
- Compute Canada's Sharcnet GPU Cluster
- Eric Praetzel's ECE Ubuntu 4 (GTX 1060)

[Presenter: Pei Lin]

# Generator & discriminator loss during training




[Presenter: Celine]

# k-folds cross-validation results (k=5)

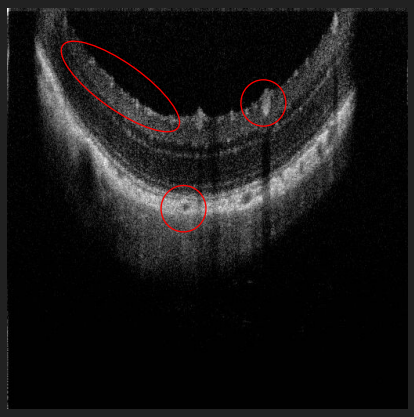
| | Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 | Average |
|---|---|---|---|---|---|---|
| **SSIM** [-1.0, +1.0] Structural Similarity | **0.667** | **0.591** | **0.637** | **0.699** | **0.561** | **0.631** |
| **PSNR** (dB) [0.0, +INF) Peak Signal to Noise Ratio | 22.468 | 20.111 | 21.288 | 24.189 | 17.575 | 21.126 |
| **MSE** [0.0, 65025.0] Mean Squared Error | 634.600 | 1144.652 | 672.697 | 337.727 | 1579.686 | 873.872 |
| **NRMSE** [0.0, 1.0] Normalized Root Mean Squared Error | 0.171 | 0.220 | 0.181 | 0.129 | 0.284 | 0.197 |
| **MAE** [0.0, 255.0] Mean Absolute Error | 52.532 | 69.538 | 57.817 | 41.074 | 91.609 | 62.514 |

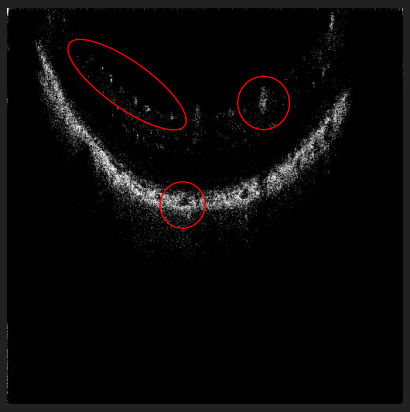
Ranges are for 256-pixel images. Green indicates value for identical images.

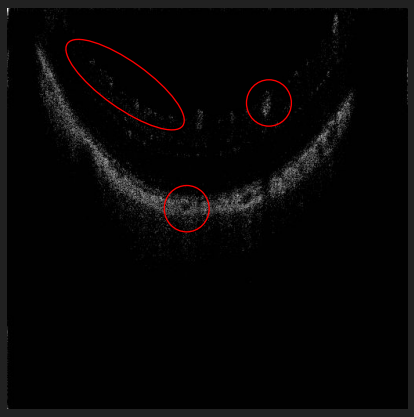
[Presenter: Pei Lin]
SSIM: structural similarity (measure); Range [-1, +1] where +1 is identical
PSNR: peak signal to noise ratio; Units are in dB (higher is better)
MSE: mean squared error; Range [0, +INF) where 0 is identical
NRMSE: normalized root mean squared error; 0 is identical
MAE: mean absolute error; Range [0, +INF) where 0 is identical

Our metric of focus is the structural similarity index, which goes from -1 to 1, where 1 is the most similar, and we see that comparing enfaces generated from the real OMAGs compared to our inferred OMAGs produces an SSIM of 0.631.

# Side-by-Side Comparison (B-Scans)





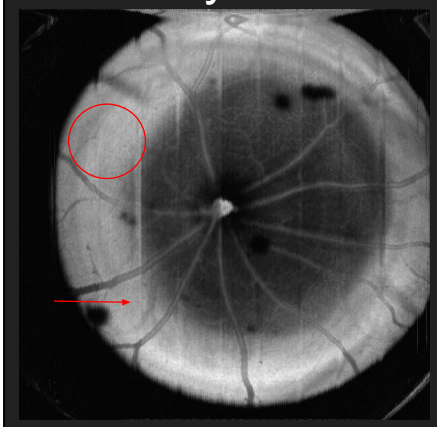


From B-scans | Predicted OMAG (ML) | Real OMAG/OCTA

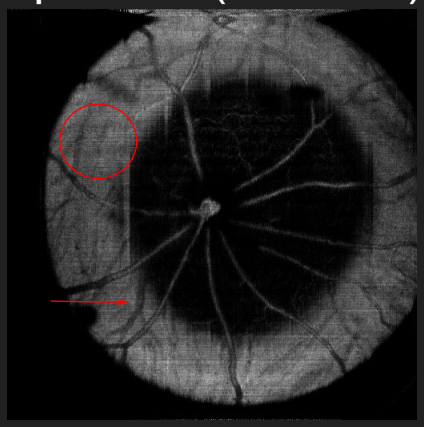
**Data**: 2015-10-21___512_2048_Horizontal_Images100, cross-section #140 out of 512, inverted colors

[Presenter: Pei Lin]

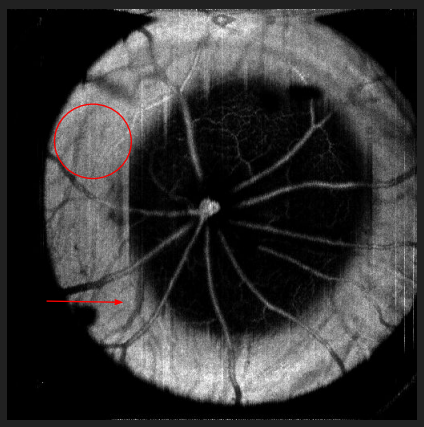
# Side-by-Side Comparison (Enfaces)



From B-scans



From predicted OMAGs (ML)

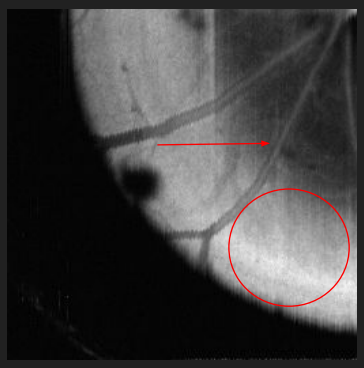


From real OMAGs/OCTA

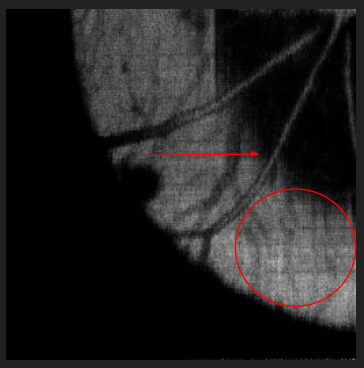
**Data**: 2015-10-21___512_2048_Horizontal_Images100, all 512 cross-sections

[Presenter: Pei Lin]

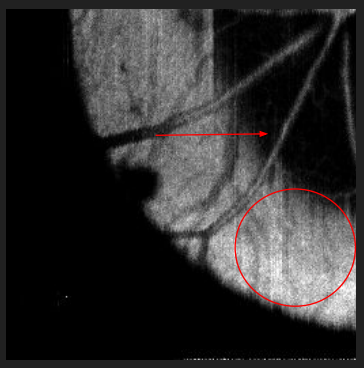
# Side-by-Side Comparison (Enfaces) - Vessel Walls



From B-scans



From predicted OMAGs (ML)



From real OMAGs/OCTA

**Data**: 2015-10-21___512_2048_Horizontal_Images100, all 512 cross-sections

[Presenter: Pei Lin]
(Discuss how the area of interest is the contrast difference of the blood vessel walls)

## Achieving Goals

- Thumbs-up from Prof. Bizheva and Zohreh, particularly impressed with:
  - Contrast of blood flow in vessels vs. vessel wall
  - Ability to infer some of the smaller capillaries
- SSIM of 0.63 (range -1 to +1, closer to 1 better)
- Enhancements include human datasets, get peer review, bring in clinicians to evaluate suitability for diagnosis, etc.

[Presenter: Pei Lin]

## Handoff to Prof. Bizheva: GUI

- Dida's future students will want a graphical user interface - we will provide 2 modes:
- **Training mode**
  - When lab acquires new images, retrain on those
  - Will help transition from rat data to human data
- **Testing mode**
  - Load checkpoints for already trained model & map B-scans to inferred OMAG & en-face
  - Output images & quality metrics for inferred enface

[Presenter: Samin]

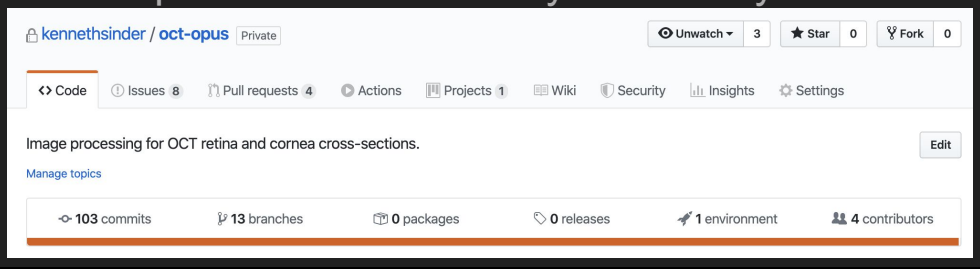
While the main aim of our project was to complete research within a cross-domain context, when we presented our results to our stakeholders in a final check-in meeting, they were pleased with our work and professor bizheva expressed to us that she would like a GUI that students in her lab could use to interface our work. As a result, some of our next steps are to wrap up our work to be used practically by our stakeholders in a way that will be accessible to non-developers.

Our gui will include a training mode that will allow our model to be retrained as more new data becomes available. This would allow for the possibility of training on human data which we did not have access to at the time of development.

A testing mode will allow the trained model to be used as is by loading existing checkpoints and loading the resulting inferred images and a set of quantitative quality metrics such as those that were discussed earlier by Pei Lin in his k-folds quantitative explanation.

## Some codebase details

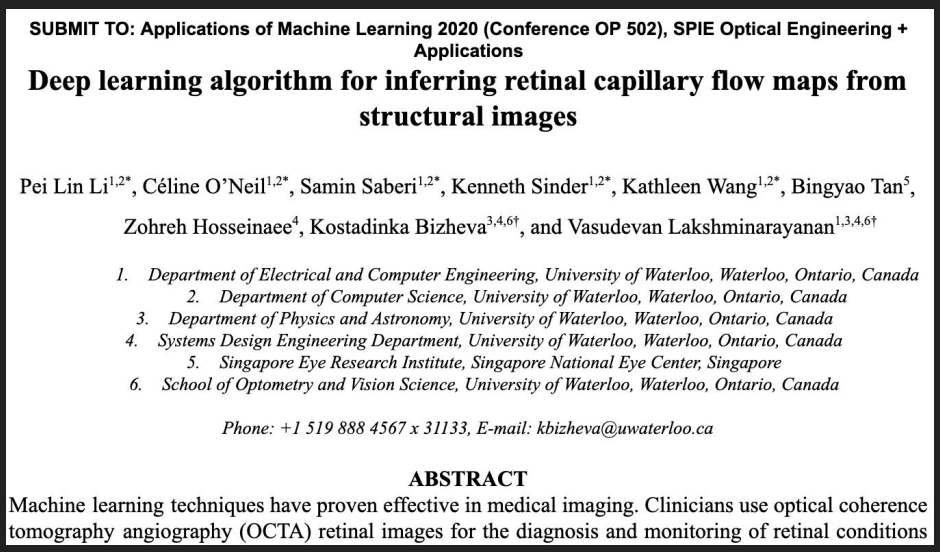
- Private GitHub repo
  - Using GitHub issue tracking & PRs + Dropbox Paper docs for organization and team meeting notes
- Thousands of LOC including related scripts & split out over multiple files for readability/modularity




[Presenter: Samin]

Our codebase is currently stored in a private github repository. Prior to handing off our work, we will also create documentation for  the use of our code and GUI. Our codebase includes thousands of lines of code which include related pre-processing and post-processing scripts,(such as histogram equalization and flattening as described by pei lin) as well as our modularized model.
In terms of the workflow we used for our project, we used a combination of Github issues and Dropbox paper documents to keep track of tasks within our team. We created detailed notes for each of our team meetings from which we produced action items worked off from.

# O + P Paper: in progress, due July



**SUBMIT TO: Applications of Machine Learning 2020 (Conference OP 502), SPIE Optical Engineering + Applications**

**Deep learning algorithm for inferring retinal capillary flow maps from structural images**


Pei Lin Li[1,2*], Céline O'Neil[1,2*], Samin Saberi[1,2*], Kenneth Sinder[1,2*], Kathleen Wang[1,2*], Bingyao Tan[5], Zohreh Hosseinaee[4], Kostadinka Bizheva[3,4,6†], and Vasudevan Lakshminarayanan[1,3,4,6†]

1. *Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada*
2. *Department of Computer Science, University of Waterloo, Waterloo, Ontario, Canada*
3. *Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, Canada*
4. *Systems Design Engineering Department, University of Waterloo, Waterloo, Ontario, Canada*
5. *Singapore Eye Research Institute, Singapore National Eye Center, Singapore*
6. *School of Optometry and Vision Science, University of Waterloo, Waterloo, Ontario, Canada*

*Phone: +1 519 888 4567 x 31133, E-mail: kbizheva@uwaterloo.ca*



**ABSTRACT**

Machine learning techniques have proven effective in medical imaging. Clinicians use optical coherence tomography angiography (OCTA) retinal images for the diagnosis and monitoring of retinal conditions


[Presenter: Samin]
Abstract Due:
12 February 2020

Author Notification:
20 April 2020

Manuscript Due Date:
29 July 2020

San Diego Convention Center
San Diego, California, United States
## 23 - 27 August 2020

While working on our model, we had several conferences in mind to aim towards. We ended up applying to SPIE's O+P (optics and photonics) conference which is tentatively occuring in august in san diego. We submitted out abstract in february and expect to be notified of our abstract results mid april.

**ABSTRACT**

Machine learning techniques have proven effective in medical imaging. Clinicians use optical coherence tomography angiography (OCTA) retinal images for the diagnosis and monitoring of retinal conditions such as diabetic retinopathy. The optical microangiography (OMAG) algorithm used to construct OCTA images requires multiple structural OCT acquisitions per location. We created a system to extrapolate microvasculature from single acquisitions with the aim of lowering the resource costs to obtain improved functional imaging. This model can be used to enhance existing OCT datasets where one might not be able to easily acquire new OCTA images.

We used a paired, unsupervised deep learning approach through implementation of a conditional generative adversarial network (cGAN) for inferring detailed retinal capillary flow maps from standard OCT images. Our cGAN architecture is based on the pix2pix image-to-image translation framework, which uses a modified U-Net for the generator and a PatchGAN for the discriminator. OCTA images generated by the OMAG algorithm applied to cross-sections captured in-vivo, non-invasively from sedated rats were used as the ground truth: 66 rat eye data sets (51 for training, 15 for testing) with a sequence of 512 paired cross-sections each. Cross-sections were augmented with random jitter and rotations to provide the neural networks with a richer training set. The trained system generates en-face blood flow maps of comparable quality to those generated from OMAG. Applying k-folds cross-validation (k=5), we observed a mean Structural Similarity Index Measure (SSIM) of 0.63. We will compare the visibility of small capillaries to a state-of-the-art convolutional neural network baseline.

[Presenter: Samin]

Here is a glimpse at the abstract that we've submitted to the conference, which include information about the motivation of our work, the collection of the data, our model's architecture and the results.

## Summary

- Presented with (1) an image-to-image translation problem for retinal cross-sections, & (2) sedated rat data
- Success with applying image-to-image cGAN framework to this domain
  - Able to infer capillaries well compared to ground truth
- Documented codebase to hand-off to Physics lab & planned paper are project artifacts we're proud of

[Presenter: Samin]

To summarize our work, we were originally presented with a set of problems by professor bizheva, including this image to image translation problem for retinal cross-sections and a large amount of sedated rate data to accompany the problem. As can be concluded from our results and the responses from our stakeholders, we believe to have found success in applying an image to image cGAN framework to this domain, and compared to the ground truth of OMAG images and the CNN baseline which we have replicated to the best of our abilities, our model is of comparable quality in being able to infer capillaries.

As a result of our success in this application, we are working towards our newly presented target of handing off our work to the Physics lab, and will be working towards the creation of our full manuscript.

## Possible Future Directions

- Automated detection/classification of eye diseases
  - Requires labelled training data
- Try improving quality of predictions by:
  - Taking previous and next B-scan as input
    - (We started to investigate this)
    - Motivation: we have constraint of blood vessel continuity
  - Further hyperparameter tuning

[Presenter: Samin]

In terms of the future work that can be built from our research, some of the potential outcomes of our work can be the automated detection/classification of retinal conditions such as diabetic retinopathy as discussed earlier. To achieve this goal, we would need access to an extensive dataset which would include human data, with diseased eyes, all labelled to be trained on. This is data that would could be generated in Professor Bizheva's lab.

Additionally, we can always try to further the quality of our results by tuning hyperparameters further and trying methods such as taking into account data between adjacent b-scans in the cross section which could potentially result in better continuity of the blood vessels in our image. (+ vengu thinks is good?/conference acceptance likely)

# References


- Slide 7 Image source https://www.researchgate.net/figure/En-face-OCTA-angiograms-black-and-white-and-retinal-thickness-maps-color-coded-maps_fig3_334557094
- Slide 13 cite “Generating retinal flow maps from structural optical coherence tomography with artificial intelligence” 2019 paper by Lee et. al
  - Image source https://www.nature.com/articles/s41598-019-42042-y/figures/1
- Slide 14 Image sources
  - https://www.nature.com/articles/s41598-019-42042-y/figures/2
  - https://www.nature.com/articles/s41598-019-42042-y/figures/3
- Slide 15 Image source https://medium.com/datadriveninvestor/generative-adversarial-network-gan-using-keras-ce1c05cfdfd3
- Slide 16 UNet Image source: “Image-to-Image Translation with Conditional Adversarial Networks” by Isola et. al
  - https://arxiv.org/pdf/1611.07004.pdf
- Slide 19 “Image Synthesis in Multi-Contrast MRI with Conditional Generative Adversarial Networks” by Dar et. al
  - https://arxiv.org/pdf/1802.01221.pdf

# Thank you!

Questions?