Semantic Segmentation on Eye Images for Keratitis Detection

Jonathan Mak
Electrical Engineering
jmak@stanford.edu

Abstract

In this work, we explore semantic segmentation of OpenEDS images using various models based on a UNet Architecture. We first attempt binary segmentation of our problem, before moving into multi-class segmentation. The goal is to properly segment out various regions of the eye, including the iris, pupil, and sclera, and apply that to future optical diagnostics by testing our network on a novel dataset of images, provided by Sankara Nethralaya for those suffering either from viral, bacterial, or fungal keratitis. We preprocess our OpenEDS images on our training dataset through various augmentation techniques to try and mimic our real life keratitis images through random insertions of keratitis-like geometric objects, bilateral filtering methods to smooth out our image, and CLAHE normalization for improved contrast. While the mIoU on the OpenEDS validation set outperforms current SoTA at almost 98%, the mean IoU when evaluating on a real-life keratitis dataset reaches only 50%. However, these results are promising given that only 20 test images were used, and that the dataset is drastically different. Annotations for the keratitis were hand-done with the help of medical students from Stanford.

1 Introduction

The computer vision community has rapidly improved on object detection and segmentation results over a short period of time. More specifically, semantic segmentation, which is detection of correct masks of an image, and instance segmentation, which requires the correct detection of all objects in an image while also precisely segmenting each instance are now the new frontier of development. Developments on this frontier have allowed for more meaningful healthcare applications, and in this project, we will be investigating semantic segmentation of both healthy and unhealthy eyes in order to provide more meaningful clinical information on detecting keratitis. Our baseline model will be a traditional semantic segmentation model in UNet developed by Ronneberger et al. [2015] for multi-class semantic segmentation. We will be exploring a variety of image augmentation techniques in order to have our model perform well on our real life keratitis data set, which will include light normalization through Contrast Limited Adaptive Histogram Equalization (CLAHE), bilateral filtering, and geometric object augmentation. After testing our model on the provided OpenEDS images, we will apply our architecture to a novel dataset of unhealthy eye images, provided to us by a non-profit, AI For Eye, with images that were gathered from clinical practitioners and nurses from Sankara Nethralaya in Chennai, India. These images all show unhealthy and diseased eyes, which are all afflicted by various types of keratitis including bacterial, fungal, and viral. In this project, we hope that our model will be still be able to identify the 3 classes outlined above, and the unlabeled shapes in our image will therefore be the region affected by keratitis. The model is meant to perform the difficult task of semantic segmentation even when there is surrounding noise (keratitis), and our trained model will hopefully adapt given that some of the healthy eyes do not show the full circular shape of pupil, iris, and sclera.

CS271 (2020), Stanford.

We will evaluate our results using the labeled images from the dataset. Qualitatively, we expect to see each of the eye features (iris, pupil, and sclera) clearly segmented in each input image. Quantitatively, we will implement evaluation metrics such as precision, recall, and accuracy for analysis of the results (where IoU will be used to determine whether each instance is correctly segmented). Finally, we want to see real world generalizations and apply them to existing keratitis images to see if it would be able to detect abnormalities after being trained on normal images of eyes (correctly reject non-healthy images by only segmenting out healthy irises, etc). The latter step of identifying physically present diseases would be extremely helpful for clinicians in the future using this software in order to segment out abnormalities, and having an accurate segmentation model is extremely helpful for researchers working in the eye tracking space, which has gained momentum in the VR/AR space as well as for general diagnosis of mood and psychological well being. The segmentation can be refined later on for unhealthy images and more specifically, can later act as an important pre-processing step for identifying and marking out regions of the eye afflicted by keratitis, thereby providing a major leverage for real-life clinical use cases, as in AI For Eye.

2 Related Work

Segmentation of the eye is primarily an interesting task to researchers because of its applications in VR tracking and being able to continuously follow the location of the focus of the eyes. An initial work by Thoma [2016] provided standardized references when dealing with evaluating pixel level segmentation models, which has been the foundation of the techniques we will use to evaluate our semantic segmentation models (IoU and Dice coefficient). Specific work regarding our dataset was done by Huynh et al. [2019]; they use a lightweight network (MobileNetV2) and tackle this task in as lightweight a fashion in order to actively deploy onto hardware. Their primary contributions come in three primary stages: get a grayscale image from the input, segment three distinct eye regions with a deep network, and remove incorrect areas with heuristic filters. Another model by Chaudhary et al. [2019] combined the DenseNet with UNet architecture alongside additional heuristic filtering on the OpenEDS dataset in order to get rid of the light scattering that is displayed by the camera.

In terms of effective segmentation, there have been several efforts done to identify parts of the eye, for example, starting first with iterative approaches to detect known eye features such as the iris and sclera, as noted in Adegoke et al. [2013] which used various signal processing techniques such as Wilde's and Daugman's method, and sclera biometric methods utilized by Das et al. [2013]. These papers gave way to eventually more advanced learned approaches to eye segmentation. A majority of the initial landscape focused on binary segmentation, where either the iris or sclera was detected. The work done by Sankowski et al. [2010] utilized learned methods for primarily boundary detection such as reflection localization and inpainting, iris boundary localization and eyelid boundaries localization. Radu et al. [2015] proposed a novel sclera segmentation algorithm for color images, operating at the pixel-level; it uses a two-stage classifier, where in the first stage a set of simple classifiers is employed and in the second stage a neural network classifier operates on the probabilities-space generated by the classifiers at the first stage. Das et al. [2017] hosted a similar competition to analyze the eye, which consisted of 2624 images taken from both eyes of 82 individuals. The winning team utilized heavy data augmentation, as well as used an encoder-decoder architecture that was supplemented with additional data. Finally, Lucio et al. [2018] utilized both a Fully Convolutional Network as well as Generative Adversarial Network, which was used to create the best possible segmentation. They ultimately achieved a dice coefficient/F1 score of 87.48%, which proved to be state of the art at the time.

Given the recent rise of segmentation problems, multi-class eye segmentation is still relatively new. These include Rot et al. [2018] and Luo et al. [2019]. Rot et al. [2018] trained a convolutional encoder-decoder neural network on a small data set of 120 images from 30 participants and primarily evaluated the multi-class SegNet architecture in comparisons with an ensemble of single-class techniques. Luo et al. [2019] trained a convolutional neural network coupled with conditional random field for post-processing, on a data set of 3161 low resolution images to segment two eye feature classes (iris and sclera). Hence, we will build off previous works done in semantic segmentation as well as newer methods of combining multiple architectures in order to create a more expressive model to capture all of the finer features in the eye.

3 Dataset

A large portion of our eye segmentation data will come from the Open Eye Dataset (OpenEDS) from Facebook's Research Labs, developed by Garbin et al. [2019]. This high resolution dataset consists of images of dimension 400 x 640 and is compiled from video capture of the eye-region collected from 152 individual participants and is divided into four subsets. The most important subsets we will use are the (i) 12,759 images with pixel-level annotations for key eye-regions (iris, pupil, and sclera), and (ii) 252,690 unlabelled eye-images. An example of the data is shown below in Figure 1, where points 1-3 are the pupil, and 4-6 are the iris.

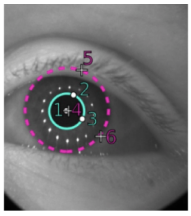


Figure 1: Example annotated image from OpenEDS

When split up, the un-annotated images and its respective ground truth mask are displayed below. Compared to Figure 1, Figure 2 shows an eye where the iris and sclera are obstructed and not fully visible, which will be helpful when applying the novel keratitis dataset and seeing whether the pupil, iris, and sclera can still be accurately detected.

(a) Unannotated Eye Image

(b) Ground Truth Mask

Figure 2: Eye Image and respective ground truth mask

The unhealthy eye data will be provided by a non-profit, AI For Eye, which currently consists of 20 standardized images of keratitis from various patients. The non-profit is currently partnered up with Sankara Nethralaya in Chennai, India, and the dataset will hopefully grow (pending partnerships with other large eye care hospitals). However, since the eye images are being collected in real time, it has been harder to gather a larger dataset. An example of an eye from this dataset is displayed below in Figure 3, where the pupil, iris, and sclera can still be identified, despite keratitis affecting the bottom right of the eye. Since the images are from a completely different, we pass it through an extensive preprocessing pipeline. We first resize our images to be the same aspect ratio, zero pad the top and bottom to make it the same as our testing images through bilateral filtering. Lastly, we normalize the zoom out to a relative perspective similar to the other images using CLAHE normalization, both of which are elaborated on in the subsequent section.

Figure 3: An eye suffering from fungal keratitis

4 Technical Approach

All models utilized a LR of 0.001 with a progressive learning rate scheduler for initial experimentation, and given the data resolution and the resources we have, the models were trained for 20 epochs (which still took 12+ hours each).

4.1 Baseline Model

The beginning baseline was a simple baseline of a UNet Model with a cross entropy loss, first experimenting with a simply binary problem (sclera vs background). The purpose of this was to examine the efficacy of the UNet and how it applied to a binary semantic segmentation task, which it was built for despite having a novel dataset, since the dataset was recently released in 2019. Then, the UNet model was applied into our multiclass problem by refactoring it to allow for all 4 classes to be segmented. All the images were preprocessed to be into the 400×640 dimension sizing, and one channel of the gray image was extracted for training.

4.2 UNet + DenseNet

In order to expand more upon the expressiveness of our model, we needed to increase the complexity. Instead of using simply a UNet, the architecture was also connected to a DenseNet. In addition, we augmented our loss term for our baseline semantic segmentation model with additional Surface Loss and Dice Loss terms (also shown below), and increasing the dropout percentage to 40% in order to prevent our model from overfitting. We chose these loss functions primarily because surface loss (a version of L_2) can help us regularize our model and recover important features in unbalanced segmentation tasks as shown in Kervadec et al. [2018], and Dice Loss is commonly used in semantic segmentation tasks. Our final architecture consisted of a UNet and a DenseNet, which was developed by Huang et al. [2016] and widely used in similar tasks in the past. Lastly, both early stopping and learning rate schedulers were used, starting after epoch 10. Our final loss function for our final semantic model was a weighted linear combination of our categorical cross entropy loss, dice loss, and surface loss. The hope here was to make the model as expressive as possible.

$$\text{Categorical Cross Entropy Loss} = -\sum_{i=1}^{C} t_i \ln(f_i(s)); f_i(s) = \frac{e_i^s}{\sum j^C e_j^s} \tag{1}$$

$$\text{Dice Loss}(y, \hat{y}) = 1 - \frac{2\sum y_{h,w}\hat{y}_{h,w}}{\sum y_{h,w} + \sum \hat{y}_{h,w}} \tag{2}$$

$$\text{Surface Loss} = \int_{h,w} \phi_g(q) f_i(q) dq; f_i(s) = \frac{e_i^s}{\sum j^C e_j^s} \tag{3}$$

4.3 Preprocessing Data Augmentation

Besides deepening our model architecture to incorporate better latent space representations, the fine tune our semantic segmentation model to be more predictive, so we decided to take inspiration from previous works by adding in additional filtering and data augmentation techniques. The goal behind this was to generalize to both the OpenEDS testing data and also our real life keratitis dataset. Since the keratitis dataset is so small (20 images), it cannot be properly used as a training set, but rather an evaluation metric to see how well expressive and generalizable our network can become. To fit it better to real life images, a series of augmentation transformations was done in order to make them as similar as possible. For our preprocessing, we experimented with three linear combinations of 1) geometric masks, 2) bilateral filtering, 3) CLAHE Normalization.

4.3.1 Geometric Mask Augmentation

Oftentimes, data augmentation is simply done by traditionally, such as rotating, flipping, zooming and cropping. However, an alternative method that is used in machine learning, especially in generative modeling nowadays is the idea of masks. The neural network, in a generative modeling case, should learn to fill in the mask properly given the surrounding information, and is now widely used in

generative flow models. However, masking as data augmentation was used recently by Chen [2020], who injected squares across each image, as augmentation through information removal. They were able to achieve state-of-the-art results in a variety of computer vision tasks. In this case, we will be using similar masks to improve accuracy.

Since our keratitis images have geometric circular shapes blocking out the iris, pupil, and sclera as seen in 3, we incorporate some initial augmentation to our dataset where we inject / block off our eyes with similar looking geometric shapes, so that hopefully evaluation on our keratitis dataset would improve. Figure 4 shows an example mask that is randomly overlay on top of images with a probability p = 0.2, in order to similarly mimic the fungal keratitis in 3 where the previous eye in Figure 3 can be seen to beaffected by fungal keratitis (in the bottom right). The goal of this mask was to replicate the region of keratitis by applying these concentric circles mask on the image. The rest of the 20 images was analyzed for the most common locations of the keratitis, and rough approximations of the geometric object were created as masks to create a total of 3 different masks. The masks were chosen randomly, and as stated above, applied to our training images with probability p = 0.2.

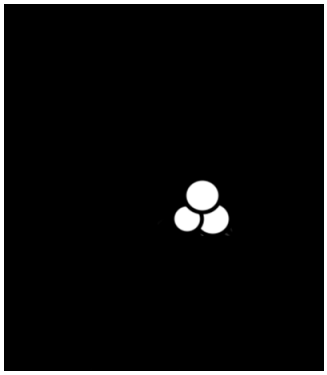


Figure 4: An example of inserting 3 geometric circles mimicking fungal keratitis

4.3.2 Bilateral Filtering

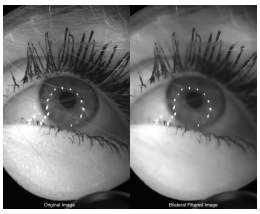
In addition, we apply advanced denoising/filtering techniques such as the bilateral filter, which is a non-linear, edge-preserving, and noise-reducing smoothing filter for images. The bilateral filter, defined below, helps by sharpening edges in the image in the pre-processing stage so that features are more distinct and any residual noise in the image/non-significant variance is reduced. It is hypothesized that this should aid in improving the predictions made by the semantic segmentation model (UNet + DenseNet).

$$I(x) = \frac{1}{W_p} \sum_{x_i} I(x_i) f_r(||I(x_i) - I(x)||) \mathcal{N}(||x_i - x||); W_p = \sum_{x_i} f_r(||I(x_i) - I(x)||) \mathcal{N}(||x_i - x||) \tag{4}$$

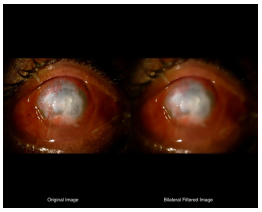
Essentially, the process behinds this first starts with linear Gaussian smoothing, similar to a Gaussian blur. Then, a weighting multiplied by the normalized tonal distance ($||x_i - x||$) is applied, so that features are enhanced more. This technique was applied in order to sharpen the edges between the pupil, iris, and sclera, so that our classifier would learn to identify both rough and smooth outlines of this. This filtering was applied to both the OpenEDS images and the keratitis dataset in order to have the model learn representations of outlines in the same manner. Figure 5 shows the effect of bilateral filtering on an OpenEDS Image and also on the Keratitis Images.

4.3.3 CLAHE Normalization

Lastly, we apply Contrast Limited Adaptive Histogram Equalization (CLAHE) which equalizes out the amount of light contrast seen in the image. It is a variant of adaptive histogram equalization where the contrast amplification is limited, since normal adaptive histogram equalization tends to overamplify regions in constant regions. We use this light normalization in order to make sure that all of the images are lighted in the same way, and normalized to have the same relative lighting



(a) OpenEDS vs Bilateral Filtered



(b) Keratitis vs Bilateral Filtered

Figure 5: Images before and after using Bilateral Filtering

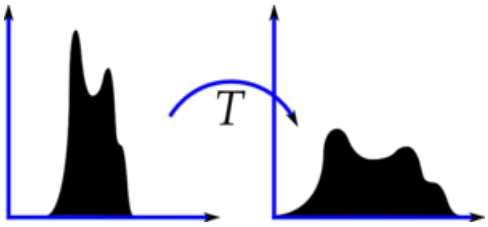


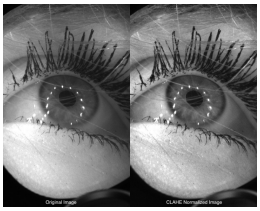
Figure 6: Example of Histogram Equalization on Images

conditions. This can be especially important when generalizing towards our keratitis dataset, since photos are taken in various lighting conditions in the clinic. In Figure 6, you can see what happens to the values of the pixels in the image before and after histogram equalization. In our case, the OpenEDS is a brighter image (since its primarily white), and has pixels confined to higher values. However, you usually want to have an image span the pixel range, so here the histogram was stretched out to include pixels from all regions in order to improve the contrast. Figure 7 shows the effect of histogram equalization on an OpenEDS Image and also on the Keratitis Images.

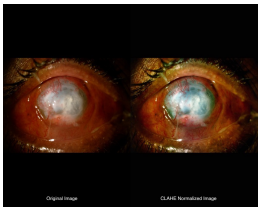
5 Results and Experiments

Since we use different loss functions in each model (with each model getting consistently more expressive), we standardized our experiments by evaluating our model both by using the Jaccard Index and calculating the mean IoU (mIoU) over each batch on the OpenEDS dataset, rather than keeping track of accuracy. We note that the Binary Model seems to perform better than Multiclass since it's prediction task is easier. The original UNet suffers from a drop in performance but adding on additional network architectures and having a more expressive loss function is able to account for the original drop in performance.

$$\text{Jaccard Index} = \text{IoU} = \frac{TP}{TP + FP + FN} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \tag{5}$$



(a) OpenEDS vs CLAHE Normalized



(b) Keratitis vs CLAHE Normalized

Figure 7: Images before and after using CLAHE Normalization

5.1 EDS Validation Evaluation

The results of testing our different experiments with varying levels of augmentation and model architectures are presented in Table 1. We can see that as the the model architecture gets more complex, the mIoU continues to rise on the validation set. Surprisingly, adding in geometric mask helps the model improve the most, but bilateral filtering and CLAHE normalization do not provide increase our mIoU as much. This is in alignment with our prediction that augmenting our data slightly can help our model generalize better. The linear combination of all 3 data augmentation techniques seems to fare only slightly better than our UNet + DenseNet Baseline, and once we start choosing random keratitis geometric masks to be injected, the mIoU goes down, which makes sense since we start significantly modifying our training data in order to better generalize towards our keratitis dataset rather than our OpenEDS dataset. Figure 8 shows the mIoU per epoch. We notice that with significant augmentation, the mIoU ascends more slowly, but eventually reaches a similar, albeit lower, mIoU compared to the other models.

Model	EDS mIoU
Baseline UNet (Binary)	0.9942
Baseline UNet (Multi-Class)	0.6190
UNet + DenseNet	0.9722
UNet + DenseNet + Geometric	**0.9755**
UNet + DenseNet + Bilateral	0.9733
UNet + DenseNet + CLAHE Norm	0.9747
UNet + DenseNet + all 3	0.9730
UNet + DenseNet + all 3 + Random Masks	0.9674

Table 1: Mean IoU comparison on OpenEDS Validation Set

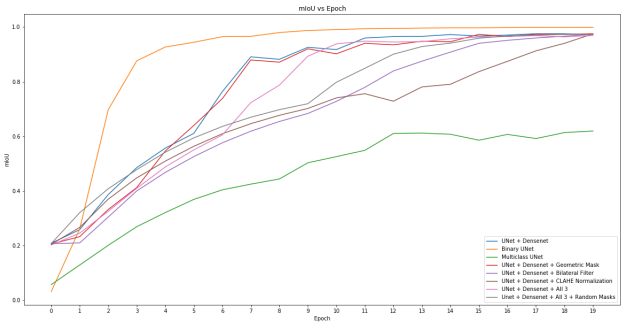


Figure 8: mIoU over batch per epoch

5.2 Keratitis Evaluation

We then transferred our trained OpenEDS validation models and tested them on our novel keratitis dataset. The results of testing our different experiments on our novel keratitis data set of 20 images with varying levels of augmentation and model architectures are presented in Table 2 after training for 250 epochs. The keratitis images were hand labeled with the help of medical students and faculty, and the mIoU was then calculated across the 20 images. We note that the mIoU is significantly lower, partially because there are not that many test images, but also because the images are inherently different. The problem of having a lower amount of test images will be resolved as the non profit, AI For Eye, presents us with more data, but we can see that with our data augmentation techniques, we are able to gain increasingly higher mIoU. This reiterates the fact that oftentimes real-life data can indeed be very messy, despite performing extremely well on our test OpenEDS dataset. In addition, the keratitis can affect and obscure regions of the eye, more so than was present in the OpenEDS dataset. For example, as shown in Figure 9, the affected keratitis area is on the cornea, which obscures both the iris and pupil in this image. AS a result, the ground truth segmentation mask

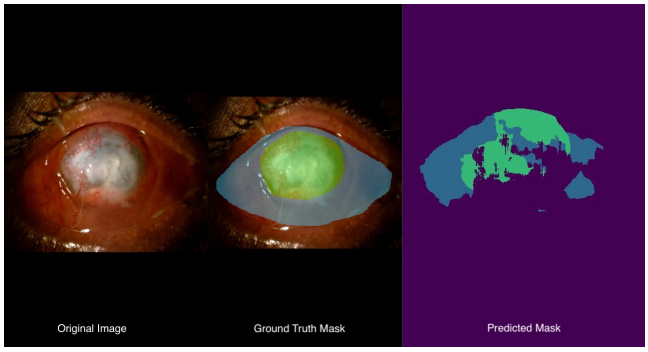


Figure 9: Obscured Eye from Keratitis

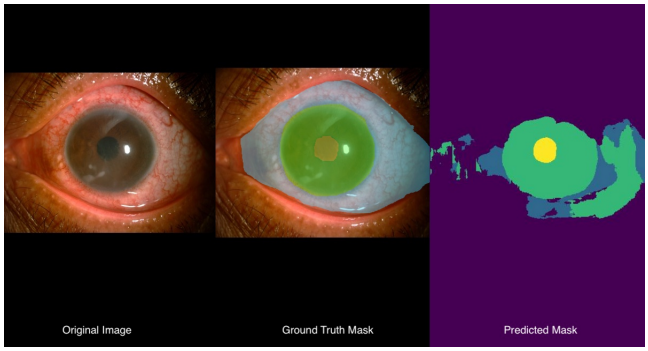


Figure 10: Segmented Eye with Keratitis

is simply labeled as the iris. During these tougher situations, the model doesn't perform as well since the OpenEDS data was much cleaner and had healthy eyes. However, in situations where the pupil, iris, and sclera are clearly visible alongside with the keratitis, our segmentation model performs much better, as shown in Figure 10. Our hypothesis regarding augmentation also holds, as we see increasing the amount of segmentation helps increase our mIoU on our test set, with the random masks providing the highest mIoU. Given that the dataset is completely different, these results still seem promising, and once more keratitis images are garnered, the model should be able to improve.

Model	Keratitis mIoU
Baseline UNet (Binary)	0.2034
Baseline UNet (Multi-Class)	0.3451
UNet + DenseNet	0.3847
UNet + DenseNet + Geometric	0.4256
UNet + DenseNet + Bilateral	0.4034
UNet + DenseNet + CLAHE Norm	0.4125
UNet + DenseNet + all 3	0.4491
UNet + DenseNet + all 3 + Random Masks	**0.4729**

Table 2: Mean IoU comparison on Keratitis Images

6 Conclusion

The field of semantic segmentation has recently received a lot of attention because of its high applicability in the computer vision. With the healthcare industry continuing to move towards using more technologically dirven tools for diagnostics, semantic segmentation models can be extremely useful in helping clinical practitioners make more informed decisions. In this paper, we presented a UNet + DenseNet architecture similar to other papers in the past, but utilized extensive preprocessing pipeline in order to prepare for our real world task of segmenting against a real life keratitis dataset

provided for us by AI For Eye through Sankara Nethralaya in Chennai, India. Geometric masking was utilized in order to mimic keratitis-like objects, bilateral filtering was used to increase smoothing of the edges but to identify outlines in a more distinct manner, and finally CLAHE normalization was used to normalize lighting conditions through equalized contrast and pixel values. We see that the augmentation improves the performance against the validation OpenEDS image set, but as more extensive augmentation is used, performance goes slightly down (albeit better than the original baseline). We notice that extensive augmentation on the OpenEDS set leads to improved performance on the keratitis imageset, which is in alignment with our predictions since the augmentations during preprocessing were to help the model generalize better to this dataset. Semantic segmentation is useful in this case because it can highlight to practitioners exactly where keratitis is located, and which regions of the eye it affects, and can prove to be a very powerful tool in the future.

6.1 Appendix

A YouTube video highlighting this paper is available at: `https://youtu.be/k7G7f-53FVk`
Code is available at: `https://github.com/jonathanjmak/cs271proj`

Thank you to all medical school students and faculty as well as Serena Yeung and TAs, especially Joy Hsu, who provided unwavering support and guidance during this project.

Bibliography

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. URL `http://arxiv.org/abs/1505.04597`.

Martin Thoma. A survey of semantic segmentation. *arXiv preprint arXiv:1602.06541*, 2016.

Van Thong Huynh, Soo-Hyung Kim, Guee-Sang Lee, and Hyung-Jeong Yang. Eye semantic segmentation with a lightweight model. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pages 3704–3707. IEEE, 2019. doi: 10.1109/ICCVW.2019.00457.

Aayush K. Chaudhary, Rakshit Kothari, Manoj Acharya, Shusil Dangi, Nitinraj Nair, Reynold Bailey, Christopher Kanan, Gabriel Diaz, and Jeff B. Pelz. Ritnet: Real-time semantic segmentation of the eye for gaze tracking, 2019.

BO Adegoke, EO Omidiora, SA Falohun, and JA Ojo. Iris segmentation: a survey. *International Journal of Modern Engineering Research (IJMER)*, 3(4):1885–1889, 2013.

Abhijit Das, Umapada Pal, Michael Blumenstein, and Miguel Angel Ferrer Ballester. Sclera recognition-a survey. In *2013 2nd IAPR Asian Conference on Pattern Recognition*, pages 917–921. IEEE, 2013.

Wojciech Sankowski, Kamil Grabowski, Małgorzata Napieralska, Mariusz Zubert, and Andrzej Napieralski. Reliable algorithm for iris segmentation in eye image. *Image and vision computing*, 28(2):231–237, 2010.

Petru Radu, James Ferryman, and Peter Wild. A robust sclera segmentation algorithm. In *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–6. IEEE, 2015.

Abhijit Das, Umapada Pal, Miguel A Ferrer, Michael Blumenstein, Dejan Štepec, Peter Rot, Žiga Emeršič, Peter Peer, Vitomir Štruc, SV Aruna Kumar, et al. Sserbc 2017: Sclera segmentation and eye recognition benchmarking competition. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 742–747. IEEE, 2017.

Diego R Lucio, Rayson Laroca, Evair Severo, Alceu S Britto, and David Menotti. Fully convolutional networks and generative adversarial networks applied to sclera segmentation. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–7. IEEE, 2018.

Peter Rot, Žiga Emeršič, Vitomir Struc, and Peter Peer. Deep multi-class eye segmentation for ocular biometrics. In *2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI)*, pages 1–8. IEEE, 2018.

Bingnan Luo, Jie Shen, Yujiang Wang, and Maja Pantic. The ibug eye segmentation dataset. In *2018 Imperial College Computing Student Workshop (ICCSW 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

Stephan J Garbin, Yiru Shen, Immo Schuetz, Robert Cavin, Gregory Hughes, and Sachin S Talathi. Openeds: Open eye dataset. *arXiv preprint arXiv:1905.03702*, 2019.

Hoel Kervadec, Jihene Bouchtiba, Christian Desrosiers, Éric Granger, Jose Dolz, and Ismail Ben Ayed. Boundary loss for highly unbalanced segmentation, 2018.

Gao Huang, Zhuang Liu, and Kilian Q. Weinberger. Densely connected convolutional networks. *CoRR*, abs/1608.06993, 2016. URL `http://arxiv.org/abs/1608.06993`.

Pengguang Chen. Gridmask data augmentation. *arXiv preprint arXiv:2001.04086*, 2020.