

Argus III: A Novel Image Optimization and Augmentation Framework to Enable an Improved Patient Experience for the Next Generation Epiretinal Prosthesis

William Huang

Palos Verdes Peninsula High School

Rolling Hills Estates, CA, USA

Introduction

I. Retinal Degenerative Diseases

Age-Related Macular Degeneration (AMD)

- 200 million people affected worldwide
- Dry AMD: center of retina deteriorates
- Wet AMD: leak in macula, impairs a patient from seeing clearly
- Leads to eventual blindness



Source: Artificial Retina

Retinitis Pigmentosa (RP)

- 1.5 million people affected worldwide
- Destroys rod cells in photoreceptors in the eye
- Patients cannot see in dimly lit environments
- Gradual vision loss
- Leads to eventual blindness



Source: Artificial Retina

II. Retinal Prostheses

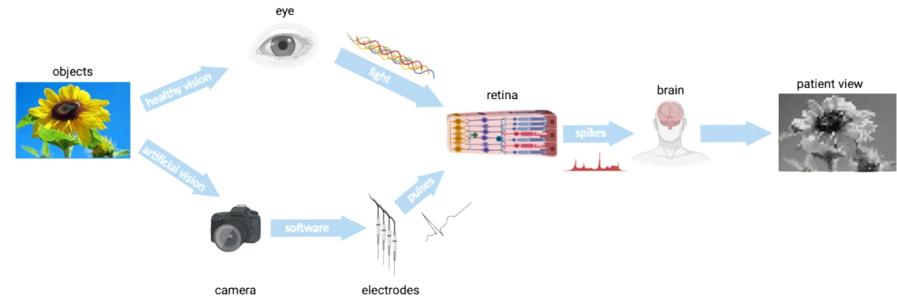


Figure 1: Current bionic eye strategy to restore vision. Source: author

- Retinal Prostheses → current most promising solution to allow AMD and RP patients to see through electrical stimulation of the retina
- It consists of a video-camera, virtual processing unit (VPU), and an external coil for transmission

1. Camera captures images that are converted into electrical parameters conveying spatial-temporal information
2. Image data is then sent to the chip through RF telemetry
3. Microelectrode array electrically stimulates the retinal neurons

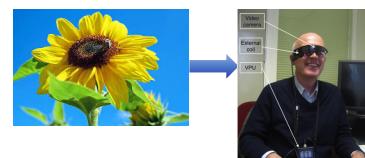


Figure 2. Low resolution grayscale images seen by patients. Source: Top right (Trinh et. al, 2020) & Bottom left (Farvardin et. al, 2018)

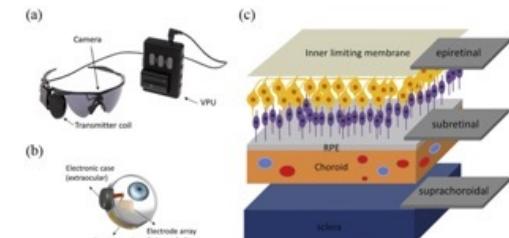


Figure 3. (a) External & (b) internal components of the Argus II system; (c) illustration of the implantation site of epiretinal prostheses. Source: (Yue et al., 2016)

Motivations and Contributions

Limitations of Retinal Prostheses

- **Low Resolution:** Compact environment for the prostheses on the retina, heat produced may cause damage, etc.; therefore, users can only see a grayscale (no color), pixelated image of surroundings (20-2000 pixels);
- **Lack of Adaptability:** Patients, especially first-time users, face a steep learning curve to identify objects
- **Current Improvement Strategies** do not address these shortcomings

Current Work	Novel Contributions
Edge detection and downsampling are the two main strategies to improve vision computationally ([7],[8]).	Magnified and improved the resolution of ROIs, encoded maximum spatial information, first study to consider color perception for patients
Salient object detection algorithms are inaccurate when images have multiple objects.	The virtual magnifier overcomes this by selecting only a certain area, therefore eliminating unnecessary information.
The existing conformal parametrization method for ROI enlargement may induce area shrinkage, which produces numerical instability.	The optimal transport map is area-preserving, thus the robustness is improved.
Traditional zoom in and segmentation algorithms remove information outside the ROI.	Magnifier maintained the information from the entire image, which is important for patients to gain a general sense of their surroundings.
Image libraries and datasets only contain high-resolution images. (not suitable for retinal prosthesis applications)	This research initiated the idea to provide patients with a low-resolution image training library via an AE-OT model while avoiding mode collapse.

Optimal Transport (OT)-Based Optimization Framework

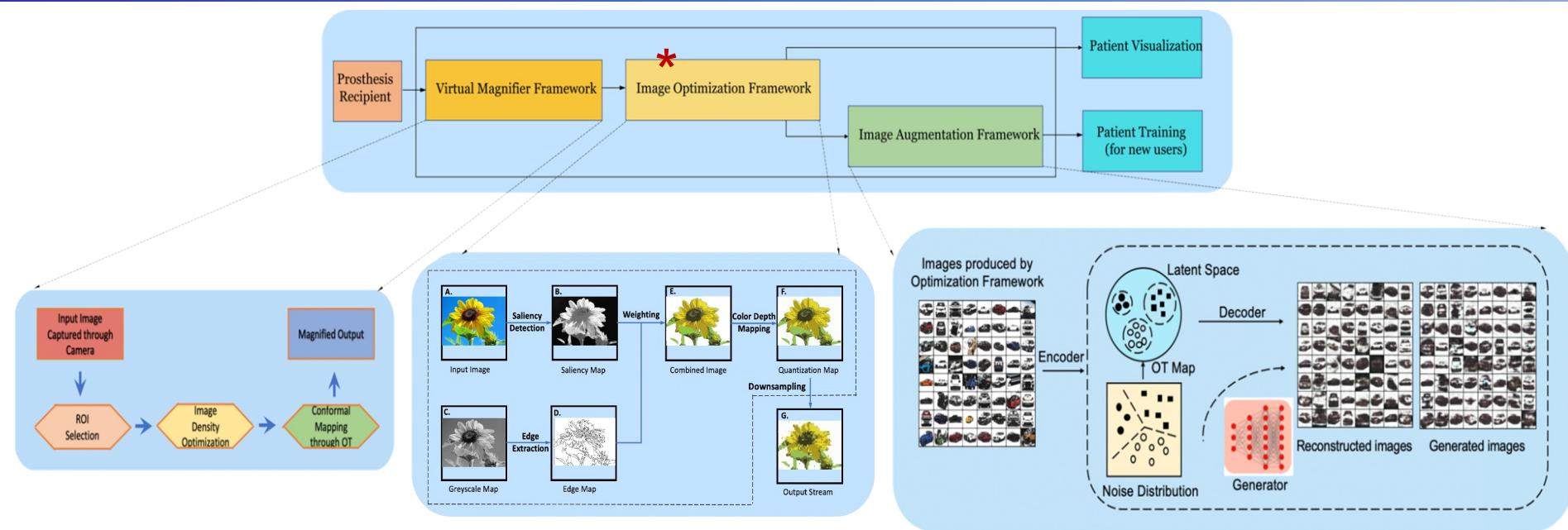


Figure 4: Overall methodological framework through a) Virtual Magnifier b) Optimization Framework c) Augmentation Framework. *Source: author*

* William Huang, "Enhancing the Bionic Eye: A Real-time Image Optimization Framework to Encode Color and Spatial Information Into Retinal Prostheses," 2021 IEEE MIT Undergraduate Research Technology Conference (URTC), 2021, pp. 1-5, doi: 10.1109/URTC54388.2021.9701618.

❖ Goals:

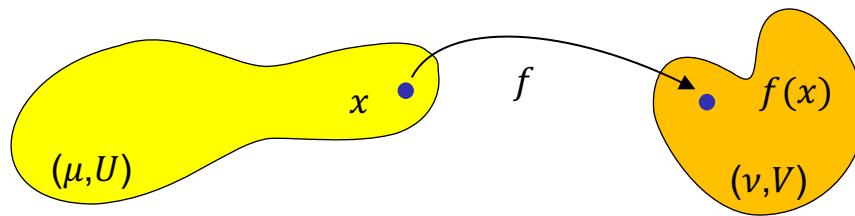
- (1) localize and “magnify” ROIs in an image frame while preserving important features and areas of objects
- (2) optimize the magnified segments to encode the maximum amount of spatial and color information to the patients through attention mechanisms as well as color scheme comparisons
- (3) augment the optimized images to enable new patients to quickly adapt to the prosthesis.

❖ Strategy:

- (1) a virtual magnifier based on the OT theory
- (2) an image optimization model through region-contrast based saliency maps, color depth mapping, & color space selection
- (3) an image augmentation model based on an autoencoder-OT model for new patient training.

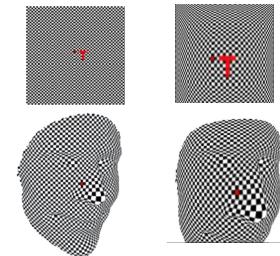
Procedure

1. Virtual Magnifier Based on Optimal Transportation Theory



Transport Mapping

For a transport scheme s (a mapping from U to V)
 $s : \mathbf{x} \in U \rightarrow \mathbf{y} \in V$,
the total cost is
 $C(s) = \int_U \mu(\mathbf{x}) c(\mathbf{x}, s(\mathbf{x})) d\mathbf{x}$
where $c(\mathbf{x}, \mathbf{y})$ is the cost function.



The virtual magnifier quantitatively specifies the density of a ROI and maps an initial density to the desired density to best preserve the critical features of objects in the segments or ROIs before resolution reduction. This strategy is based on the OT map. The OT map is the measure-preserving map T that minimizes the total transport cost:

$$\min_{T \# \mu = \nu} \int_{\Omega} c(x, T(x)) d\mu(x)$$

Continuous category \rightarrow Brenier's theorem shows that for a quadratic cost function $c(x,y)$, the OT map is the gradient map or the Brenier potential that satisfies the Monge-Ampere equation

$$\det(D^2)(x) = \frac{f(x)}{g \circ \nabla u(x)}$$

In 2D cases (e.g., images), the Monge-Ampere equation can be written in term of the Poisson equation:

$$u_{xx}u_{yy} - u_{xy}^2 = \frac{f}{g \circ Du},$$

By sampling regular grid points on an image, we can define finite difference operators to solve the discrete Poisson equation and find the OT map.

Transport Mapping. Source:
Prof. Gu

Algorithm 1 Image segmentation through optimal mass transport map.

Require: ROI \mathcal{A} , source density $f(i, j)$ and target density $g(i, j)$ of \mathcal{A} , feature threshold ϵ
Ensure: The optimal mass transport map $T : (\Omega, f) \rightarrow (\Omega^*, g)$
 $g(i, j) \leftarrow$ initial density
 $\epsilon \leftarrow \epsilon^*$
while $\epsilon < \epsilon^*$ **do**
 Find T by solving Eq.7: $\Delta u = \sqrt{u_{xx}^2 + u_{yy}^2 + 2u_{xy}^2 + 2\frac{f}{g \circ Du}}$
 Magnify \mathcal{A} through $T : (\Omega, f) \rightarrow (\Omega^*, g)$
 Update ϵ
 Increase density $g(i, j)$
end while

Algorithm 1: Pseudo-code of OT magnifier. *Source: author*

Define an operator: $\mathcal{T} : H^2(\Omega) \rightarrow H^2(\Omega)$,
 $\mathcal{T}[u] = \Delta^{-1} \left\{ \sqrt{u_{xx}^2 + u_{yy}^2 + 2u_{xy}^2 + 2f/g \circ Du} \right\}$, Brenier potential is
fixed point of \mathcal{T} ,
convert \mathcal{T} to \mathcal{P}

$$\mathcal{P}(\varphi) := \Delta^{-1} \left\{ \sqrt{(\varphi_{xx} + 1)^2 + (\varphi_{yy} + 1)^2 + 2\varphi_{xy}^2 + 2f/g \circ (Id + D\varphi) - 2} \right\}$$

Finite difference operators when sampling Ω by regular grid points with horizontal and vertical steps

$$\mathcal{D}_{xx}^2 u_{ij} = \frac{1}{h_x^2} (u_{i+1,j} + u_{i-1,j} - 2u_{i,j})$$

$$\mathcal{D}_{yy}^2 u_{ij} = \frac{1}{h_y^2} (u_{i,j+1} + u_{i,j-1} - 2u_{i,j})$$

$$\mathcal{D}_{xy}^2 u_{ij} = \frac{1}{4h_x h_y} (u_{i+1,j+1} + u_{i-1,j-1} - u_{i-1,j+1} - u_{i+1,j-1})$$

Obtain discrete Poisson equation

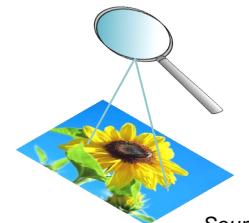
$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j} = \rho_{i,j},$$

Procedure (Continued)

2. Image Optimization

1. Edge detection for object-background identification:

- A grayscale map of the image is generated because most edge detection algorithms *rely on varying pixel intensities*
 - $I = 0.299R + 0.578G + 0.144B$, where I is the intensity value of a pixel
- A Gaussian filter is passed over the image to *reduce noise*
$$H_{ij} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(i-(k+1))^2 + (j-(k+1))^2}{2\sigma^2}\right); 1 \leq i, j \leq (2k+1)$$
- Canny edge operator applied to grayscale image by measuring contrasting pixel intensities



Source:
author

2. Region-contrast based salient object detection:

- Given a saliency value, r_k , the *color contrast of the region to other regions is found through*

$$S(r_k) = \sum_{r_k \neq r_i} w(r_i) D_r(r_k, r_i)$$

- To find $D_r(\cdot, \cdot)$, where D_r is *the color distance between two regions*,

$$D_r(r_1, r_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f(c_1, i) f(c_2, j) D(c_1, i, c_2, j)$$

- To prioritize closer regions, a spatial weighting term is introduced

3. Color Quantization:

$$S(r_k) = w_s(r_k) \sum_{r_k \neq r_i} e^{-\frac{D_S(r_k, r_i)}{\sigma_S^2}} w(r_i) D_r(r_k, r_i)$$

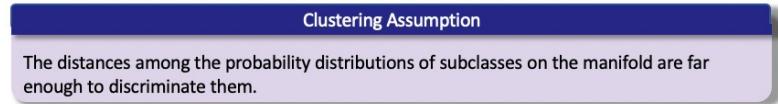
- A grayscale map of the image is generated because most edge detection algorithms *rely on varying pixel intensities*
- *Image is partitioned into n-distinct clusters* with each cluster representing a unique color using MiniBatchKmeans
- *Distance of each pixel value from the centroids* is calculated for successful mapping
- Pixel values are assigned to $S_i^{(t)}$

$$S_i^{(t)} = \left\{ x_p : \|x_p - m_i^{(t)}\|^2 \leq \|x_p - m_j^{(t)}\|^2 \forall j, 1 \leq j \leq k \right\},$$

Procedure (Continued)

3. Autoencoder-OT for Augmentation

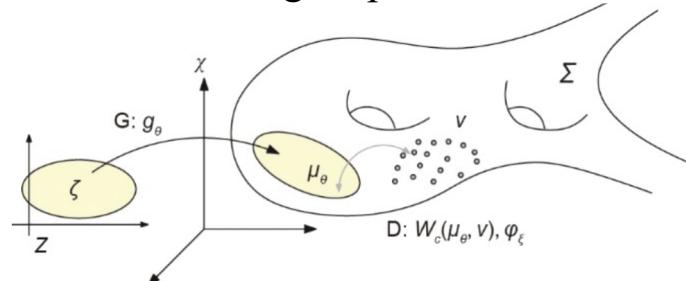
Generative adversarial networks (GANs) mainly accomplish:



1. Manifold learning
2. Probability distribution transformation. The generator map is decomposed as

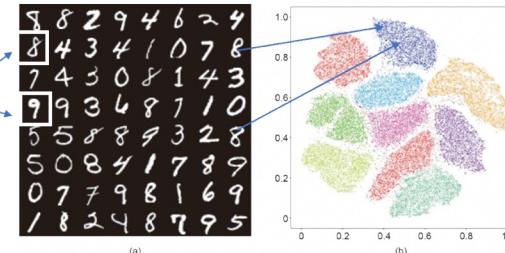
$$g_\theta = h \circ T; \quad h : \mathcal{Z} \rightarrow \Sigma; \quad T : Z \rightarrow Z$$

where the decoding map h is for manifold learning, and the map T is for measure transportation.



A manifold structure in data. Source: (Lei et al., 2020)

Each handwritten digit image has a dimension of 28×28 , and is treated as one point in the image space \mathbb{R}^{784}



Each 28×28 image is mapped onto a point on 2D surface

OT theory can be utilized to compute the map T and the minimal total transportation cost of an OT map is called the Wasserstein distance:

$$W_c(\mu, \nu) = \min_{T: \mu \rightarrow \nu} \int_{\Omega} c(x, T(x)) d\mu(x)$$

OT interpretation of GANs makes part of the black box transparent and OT-based GANs avoid mode collapse. This research improves the performance of the AE-OT model by pre-processing the input images by increasing the density of critical objects through the virtual magnifier.

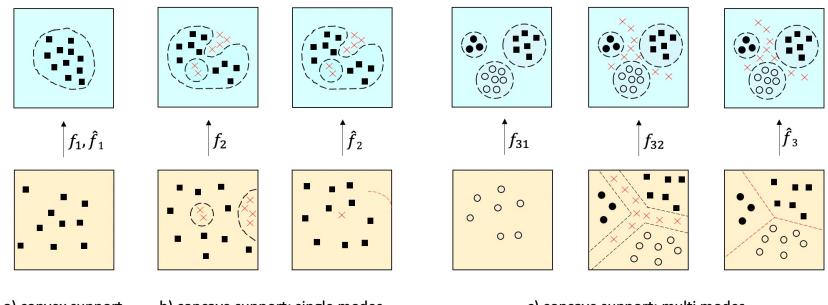


Figure 5. Mode collapse. Source: (An, 2020)

Results

1. Virtual Magnifier: Localization and Distortion of Regions of Interest

Helping patients focus on ROIs in a scene, the virtual magnifier warped the tested images to enlarge specified objects while maintaining local details and angles. The magnifier was evaluated on 5 different tests: visual, accuracy, scalability, user preference, and comparison.

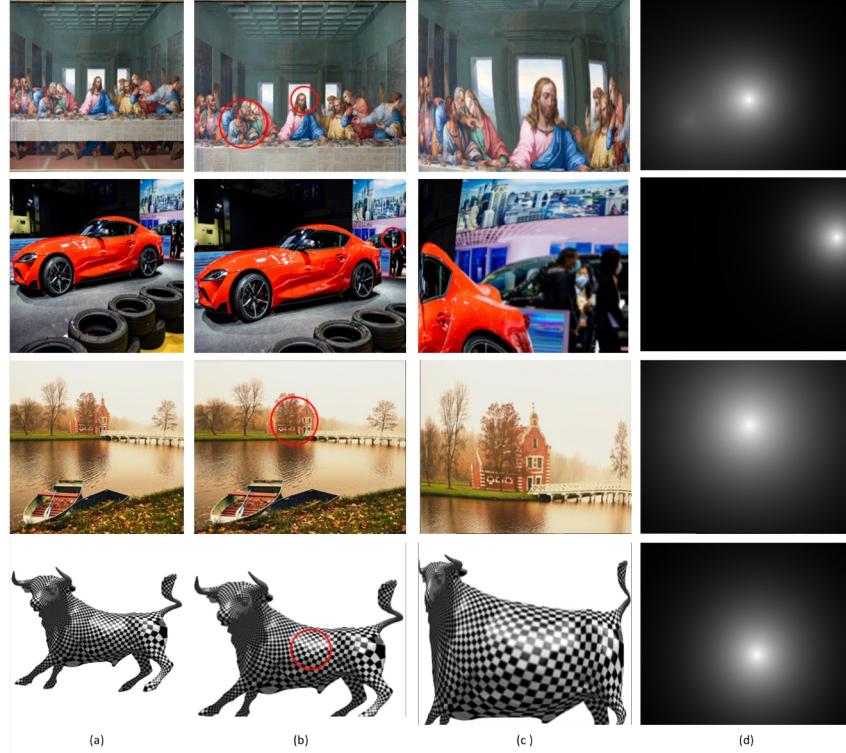


Figure 6. Visual Test: (a) Input image, (b) Means & standard deviations, (c) OT map, (d) Gaussian mixture. *Source: author*



Figure 9. Magnified ROI with increasing density.
Original face scan courtesy to Prof. Gu

Table 1. Scalability Test *Source: author*

Image	256	512	1024
Car	5.702	6.886	8.772
The Last Supper	24.121	26.894	34.796
Female	3.95	4.008	4.487
Male	3.628	4.036	5.502
Sunflower	9.21	9.617	10.228

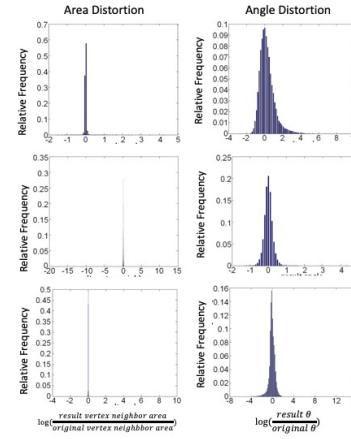


Figure 7. Accuracy Test:
Histogram of area and angle
distortion factors on different
images. *Source: author*

Table 2. User Preference Test *Source: author*

Magnifier/Score	1	2	3	4
Zooming	63.49	36.51	0	0
Bifocal	52.38	39.68	7.93	0
Perspective Wall	50.79	46.03	3.17	0
Fisheye	0	60.32	28.57	11.11
Poly-focus	0	26.98	50.79	22.22
Virtual Magnifier	0	0	18.512	81.481

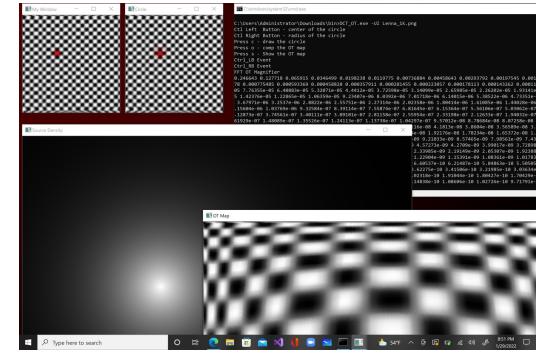


Figure 8. ‘Digital knob’ user interface.
Source: author

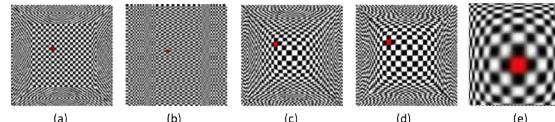


Figure 10. Comparison Test: (a) Bifocal; (b) Perspective Wall;
(c) Fisheye; (d) Polyfocal; (e) virtual magnifier. *Source: author*

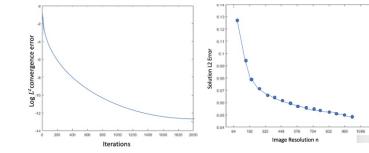


Figure 11.
 L^2 error.
Source: author

Results (Contin.)

2. Image Optimization: Salient Object Detection and Color Quantization

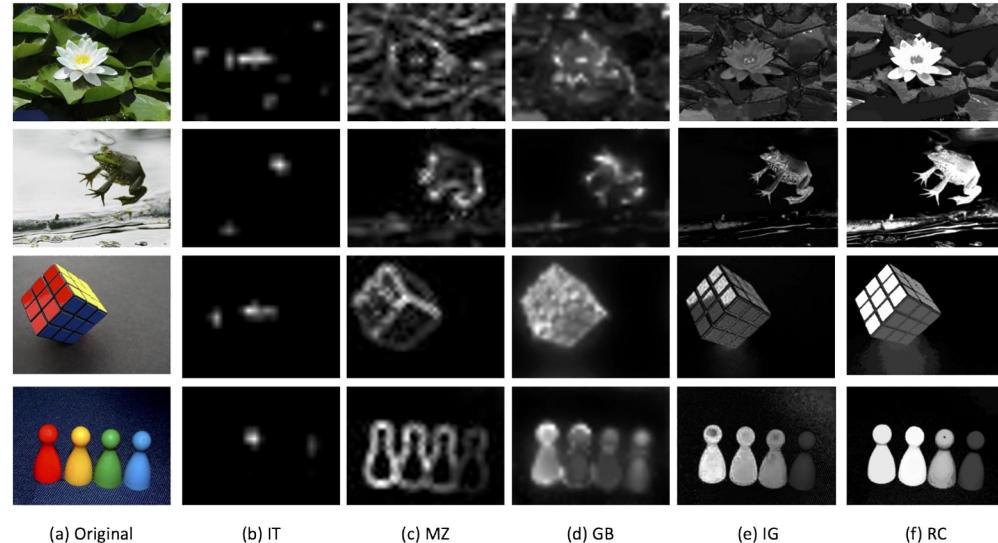


Figure 12. Visual comparison of saliency maps: (a) input image, (b) IT, (c) MZ, (d) GB, (e) IG, (f) region-contrast method. *Source: author*

image	IT		MZ		GB		IG		RC	
	MSE	SSIM	MSE	SSIM	MSE	SSIM	MSE	SSIM	MSE	SSIM
lotus	7779.45	0.05	5906.95	0.2	3394.31	0.36	4426.91	0.25	728.21	0.58
frog	378707.37	0	34303.75	0.02	35239.6	0.02	36923.46	0.05	39097.61	0.16
Rubik's cube	9927.65	0.01	8434.6	0.06	7849.82	0.15	8572.88	0.09	9575.38	0.16
pins	4285.96	0.04	3966.07	0.06	2242.94	0.25	2584.39	0.12	3184.76	0.25

Table 3. MSE and SSIM of the maps from the method proposed, IT, MZ, GB, and IG. *Source: author*

Functions	Algorithm Implementation
Edge detection with Canny operator	8.1 ms
Region contrast saliency object detection	0.0043 ms
Image merging	36ms
Color uniform quantization	340ms
RGB → Defined color space	3.6ms
Bicubic downsampling interpolation	0.34ms
Total Time	388.04 ms

Table 4: Speed comparison for images downsampled to 55 x 37. *Source: author*

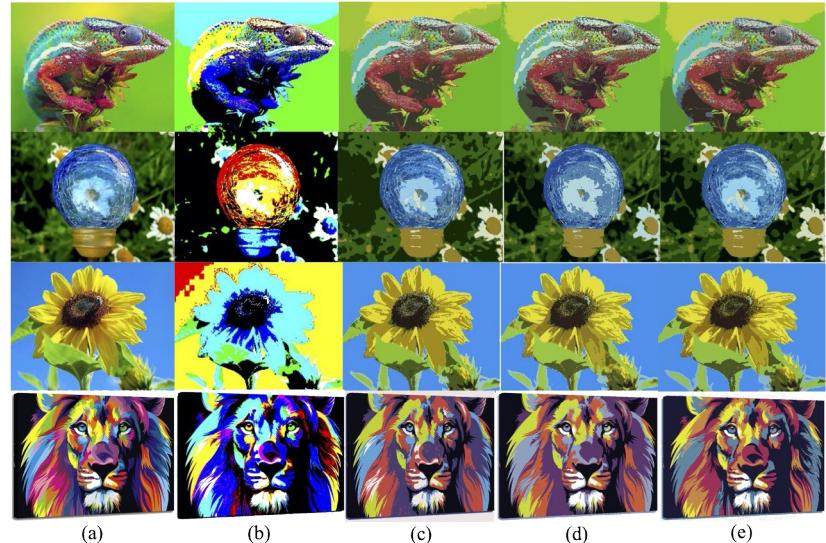


Figure 13. Color clustering comparison: (a) original image, (b) manually defining colors, (c) RGB, (d) CIELAB, (e) YCbCr. *Source: author*

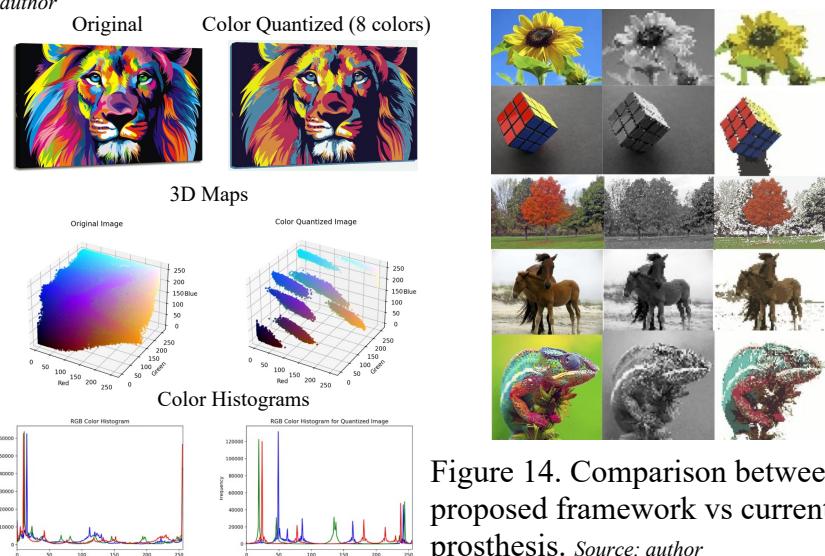


Figure 14. Comparison between proposed framework vs current prosthesis. *Source: author*

Results (Contin.)

3. Autoencoder-OT: Image Augmentation for Patient Training

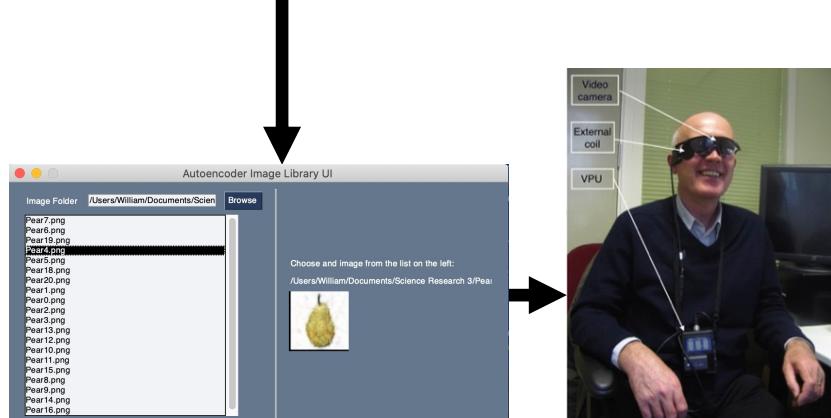
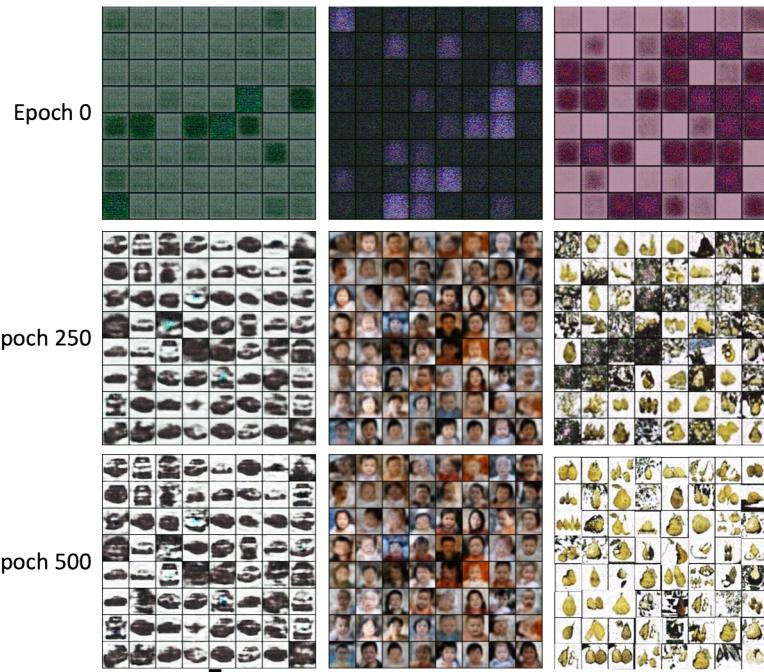


Figure 16. Image library and UI. Source: bottom right image (Trinh et. al, 2020)

Prototype Testing System

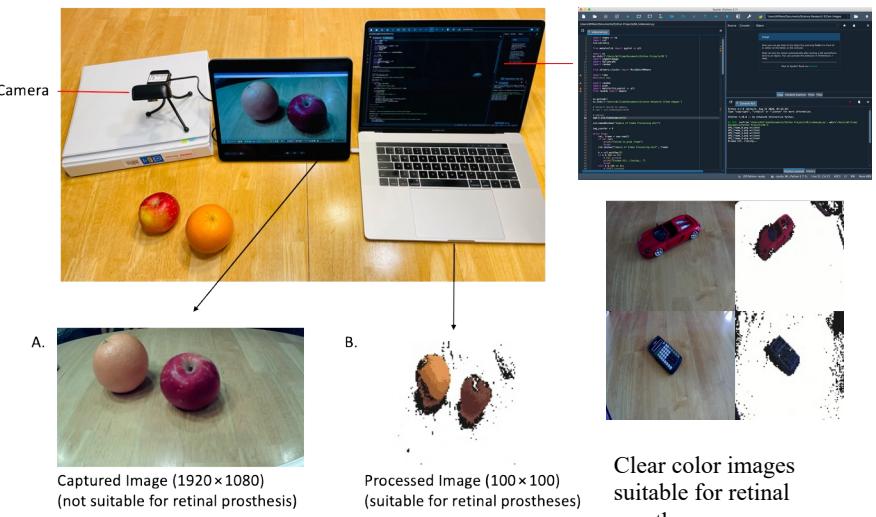


Figure 15. Illustration of the simulated VPU platform. Source: author

A prosthetic vision simulation platform captured images through a webcam mounted on a tripod and transmitted the result to a computer.

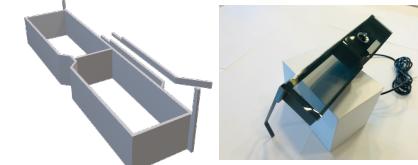


Figure 17. CAD prototype model. Source: author

Conversion of Images to Biphasic Pulse Trains

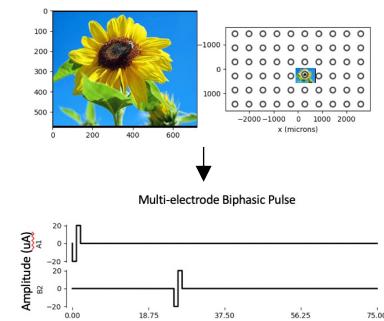
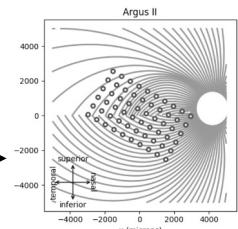


Figure 18. Conversion of image to electrical stimuli. Source: author



Discussion/Conclusion

Virtual Magnifier

- ❖ The results show that the magnifier is accurate, efficient, and scalable with minimal angle and area distortions.
- ❖ “Digital knob” interface will allow users to easily select important ROIs while preserving detailed features in a scene

Image Optimization

- ❖ The region-contrast maps achieved the highest SSIM and MSE when compared with other salient map computation techniques.
- ❖ The developed color quantization algorithm performed pixel color depth mapping through a MiniBatchKmeans clustering algorithm in 3 color spaces.

Autoencoder-OT

- ❖ This research initiates the idea to create a training program for patients.
- ❖ A library GUI was developed to make the generated image library easily accessible.
- ❖ The library can serve as a patient training tool for first-time prosthesis users, a valuable solution to combat the steep learning curve.

Prototype System

- ❖ Demonstrated effectiveness of methodological framework; Average run-time ~ 400 ms

Future Work

- ❖ Integrating infrared thermography as an additional cue may further improve the saliency maps
- ❖ Adding the application of real-time video processing
- ❖ Enhancing the performance of the AE-OT model by increasing the dimension of the latent space and number of epochs.
- ❖ Constructing a leg prosthesis with “smart vision” to guide retinal prosthesis users
- ❖ Integrating frameworks on patients and provide doctors with low resolution training library

- A portion of this work (Enhancing the Bionic Eye: A Real-Time Image Optimization Framework to Encode the Maximum Amount of Spatial and Color Information into Retinal Prostheses) was recently published to IEEE Xplore as sole author
- Presented in peer-reviewed conferences

References

1. Brindley, G. S., & Lewin, W. S. (1968). The sensations produced by electrical stimulation of the visual cortex. *The Journal of physiology*, 196(2), 479–493. <https://doi.org/10.1113/jphysiol.1968.sp008519>
2. Felzenszwalb, P.F., Huttenlocher, D.P. Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision* 59, 167–181 (2004). <https://doi.org/10.1023/B:VISI.0000022288.19776.77>
3. Horsager, A., Greenwald, S.H., Weiland, J.D., Humayun, M.S., Greenberg, R.J., McMahon, M.J., Boynton, G.M., Fine, I., 2009. Predicting visual sensitivity in retinal prosthesis patients. *Invest. Ophthalmol. Vis. Sci.* 50, 1483e1491. <http://dx.doi.org/10.1167/iovs.08-2595>.
4. Humayan, M.S., Weiland, J.D., Fuji, G.Y., Greenberg, R., Williamson, R., Little, J., Mech, B., Cimmarusti, V., Van Boemel, G., de Juan, E., 2004. Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vis. Res.* 43, 2573-2581
5. L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254– 1259, 1998.
6. M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr and S. Hu, "Global Contrast Based Salient Region Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569-582, 1 March 2015, doi: 10.1109/TPAMI.2014.2345401.
7. Ps, Ramaiah & Swamynadhan, G & Nukapeyi, Sharmilli. (2011). Image Compression and Resizing for Retinal Implant in Bionic Eye. *International Journal of Computer Science & Engineering Survey (IJCSES)*. 2. 10.5121/ijcses.2011.2204.
8. van Rheede, J. J., Kennard, C., & Hicks, S. L. (2010). Simulating prosthetic vision: Optimizing the information content of a limited visual display. *Journal of vision*, 10(14), 10.1167/10.14.32 32. <https://doi.org/10.1167/10.14.32>
9. Yue, L.; Weiland, J.D.; Roska, B.; Humayun, M.S. Retinal stimulation strategies to restore vision: Fundamentals and systems. *Prog. Retin. Eye Res.* 2016, 53, 21–47.