Intelligent Ocular Image Generation Milestone 1 Report

Principal Investigator: William Thistleton

Project Manager: Juana Attieh

Photrek, LLC

admin@photrek.io

11 October, 2024

Executive Summary

An international team of scientists, developers, and management was assembled and collaborative workflows successfully established. The team is achieving its goals with the New England Research Cloud (NERC) supplying *Infrastructure as a Service* (IaaS). Retinal Image databases have been aggregated and harmonized, primarily from Kaggle, allowing for training and testing with several dozens of thousands of images. Preliminary runs have been accomplished on a small test bed of 1000 images to ensure smooth functioning of code via unit, integration, and system tests. As expected, small sample results show very little detail. Runs with 35K+ images have been performed, with good results from the GAN approach and challenges remaining on the VAE approach.

Milestone Description and Deliverables

Milestone Description:

Image Database Exploration and Needs Assessment

Objectives: Assess image libraries and incorporate several relevant libraries into our coding environment. Preliminary image generation with VAE and Photrek cVAE. Performance evaluation on synthetic image generation.

Deliverable Description:

- Technical Report on design, implementation, and initial results

Milestone Accomplishments

☑ Technical Report on design, implementation, and initial results.

Photrek has successfully built a development team across four continents (Africa, Asia, North America, and South America), supporting talent development at four separate universities. We use a combination of Google Tools for meetings, communications, document preparation, and collaboration, as well as writing and running code on the New England

Research Cloud (NERC). The NERC provides Software as a Service (SaaS), Platform as a Service (PaaS), and particularly important for us, Infrastructure as a Service (IaaS).

We are pursuing two rival approaches for synthetic image generation, Variational Auto-Encoders (VAEs) - see, for example [12] and Generative Adversarial Networks (GANs) [7].

Preprocessing the ocular image data from various sources was performed to ensure consistency in image size, color space, and other relevant properties.

Code in both the VAE and GAN environments has been developed and tested on small training sets of approximately 1000 images to perform unit, integration, and system tests. A larger curated data set of approximately 35,000 curated images has been used to conduct User Acceptance Testing in a preliminary way, with acceptable results for the GAN (to be discussed in Milestone 2) and showing some challenges to be overcome with the VAE.

The results using the VAE can be seen in Figure 1. VAEs are powerful but can struggle with complex, high-resolution images like fundus images due to their tendency to generate blurry reconstructions. While VAEs can generate good samples of the celebA dataset [15], moving to a GAN, VAE-GAN, or a more advanced encoder-decoder architectures is necessary to create high-quality reconstructions and samples for the complex data set of ocular fundus images.

In particular, VAEs learn to compress data into a lower-dimensional latent space by focusing on the most critical features of the images. Features that vary between images carry more information for the model because they help distinguish one image from another.

The reconstruction loss incentivizes the model to spend more capacity on encoding the features harder to predict and require more precise encoding in the latent space.

GANs, however, where our reconstructions have been more successful, have a discriminator, forcing the generator to reconstruct all parts of the image with high fidelity, including the constant black background. The discriminator would notice if the black region was not correctly reconstructed, and it would penalize the generator accordingly.

We will continue to adapt the VAE architecture but do not anticipate substantially improved results. We anticipate, given the current successes, that the GAN approach will be the “workhorse” of this project.

(a) (b)

Figure 1. VAE results trained on the 35K ocular fundus image dataset after 1000 epochs. By comparison, in (b) we show the VAE results of the celebA dataset after only 76 epochs. Fundus images are problematic for the VAE approach.

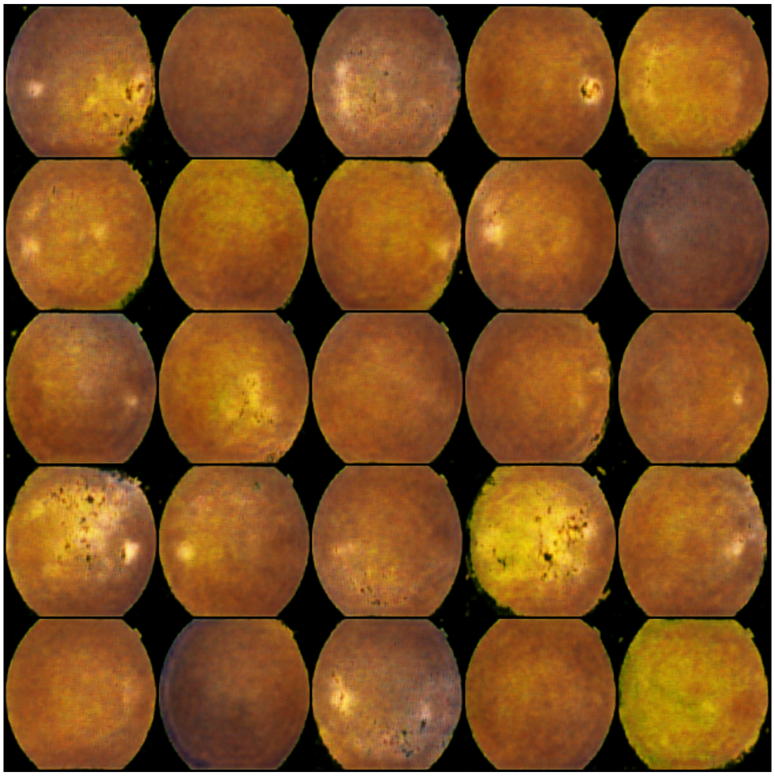

Figure 2. Preliminary GAN results using the small 1K ocular fundus image dataset and after training for only 49 epochs. Innovative approach applies architecture previously trained on CIFAR10 and refines coefficients by training with fundus images. Approach discussed and results presented in Milestone 2 Report.

Bibliography

[1] V. Gulshan et al., “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” JAMA, vol. 316, no. 22, p. 2402, Dec. 2016, doi: 10.1001/jama.2016.17216.

[2] S. K. Sattigeri, “EYE DISEASE IDENTIFICATION USING DEEP LEARNING,” vol. 09, no. 07, 2022.

[3] D. S. W. Ting et al., “Artificial intelligence and deep learning in ophthalmology,” Br. J. Ophthalmol., vol. 103, no. 2, pp. 167–175, Feb. 2019, doi: 10.1136/bjophthalmol-2018-313173.

[4] S. Kumar, T. Arif, A. S. Alotaibi, M. B. Malik, and J. Manhas, “Advances Towards Automatic Detection and Classification of Parasites Microscopic Images Using Deep

Convolutional Neural Network: Methods, Models and Research Directions," Arch. Comput. Methods Eng., vol. 30, no. 3, pp. 2013–2039, Apr. 2023, doi: 10.1007/s11831-022-09858-w.

[5] A. A. Marouf, M. M. Mottalib, R. Alhajj, J. Rokne, and O. Jafarullah, "An Efficient Approach to Predict Eye Diseases from Symptoms Using Machine Learning and Ranker-Based Feature Selection Methods," Bioengineering, vol. 10, no. 1, p. 25, Dec. 2022, doi: 10.3390/bioengineering10010025.

[6] T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer, and D. Won, "Eye Disease Classification Using Deep Learning Techniques." arXiv, Jul. 19, 2023. Accessed: Mar. 27, 2024. [Online]. Available: http://arxiv.org/abs/2307.10501

[7] M. Fan, X. Peng, and X. Gong, "Ocular Disease Recognition and Classification using TripleGAN," in Proceedings of the 2023 8th International Conference on Biomedical Signal and Image Processing, Chengdu China: ACM, Jul. 2023, pp. 7–11. doi: 10.1145/3613307.3613309.

[8] "Jin et al. - 2022 - FIVES A Fundus Image Dataset for Artificial Intel.pdf."

[9] Z. Yu, Q. Xiang, J. Meng, C. Kou, Q. Ren, and Y. Lu, "Retinal image synthesis from multiple-landmarks input with generative adversarial networks," Biomed. Eng. OnLine, vol. 18, no. 1, p. 62, Dec. 2019, doi: 10.1186/s12938-019-0682-x.

[10] "Shang et al. - 2024 - SynFundus-1M A High-quality Million-scale Synthet.pdf."

[11] R. Nuzzi, G. Boscia, P. Marolo, and F. Ricardi, "The Impact of Artificial Intelligence and Deep Learning in Eye Diseases: A Review," Front. Med., vol. 8, p. 710329, Aug. 2021, doi: 10.3389/fmed.2021.710329.

[12] D. P. Kingma and M. Welling, "An Introduction to Variational Autoencoders," Found. Trends® Mach. Learn., vol. 12, no. 4, pp. 307–392, 2019, doi: 10.1561/2200000056.

[15] Advanack Klys, Jake Snell, Richard Zemel, in Advances Neural Information Processing Systems 31 (NeurIPS 2018) Edited by: S. Bengio and H. Wallach and H. Larochelle and K. Grauman and N. Cesa-Bianchi and R. Garnett

Budget & Schedule

August 15, 2024 - October 15, 2024: Design, Implementation, Preliminary Results

Budget: $22,064 USD

Computer Supply Budget (overhead): 1000 USD

Total = 6000 USD

Milestone	Description	Budget	Status
1	Image Database Exploration and Needs Assessment	$22,064 USD	Submitted - Oct 21
2	Exploration of alternative generation processes	$17,648 USD	Submitted - Nov 22
3	Initial Hosting and Community Feedback	$6,468 USD	Planned - Nov
4	Final Report with Next Steps	$3,820 USD	Planned - Dec

Future Plans & Change Notifications

References