

Monitoring Disease Progression with Stable Diffusion: A Visual Approach

Sourav Kumar^{1,2}, Shell Xu Hu³, Aaron S Coyner⁴, Wei-Chun Lin⁴, Susan Ostmo⁴, Deniz Erdogmus⁵, RV Paul Chan⁶, Michael F. Chiang^{7,8}, J.Peter Campbell⁴, Jayashree Kalpahty-cramer⁹, Timothy Hospedales^{3,10} and Praveer Singh^{9,11}

¹Radiology, MGH/Harvard Medical School, Charlestown MA ²Electrical And Computer Engineering, Brown University, Providence, RI ³Samsun AI Research, Cambridge, UK ⁴Ophthalmology, Oregon Health & Science University, Portland, OR ⁵Electrical and Computer Engineering, NorthEastern University, Boston, MA ⁶Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, IL ⁷National Eye Institute, National Institutes of Health, Bethesda, MD ⁸National Library of Medicine, National Institutes of Health, Bethesda, MD, ⁹Ophthalmology, University of Colorado School of Medicine, Aurora, CO ¹⁰School of Informatics, University of Edinburg, Cambridge, UK ¹¹Moorfields Eye Hospital NHS Foundation Trust, London, UK

Editors: Accepted for publication at MIDL 2024

Abstract

Monitoring disease progression is quintessential for effective healthcare by providing early diagnosis, timely interventions and improved patient outcomes. In the recent years, diffusion models have become extremely popular for synthetic medical image generation. In this work, we combine Stable Diffusion with image editing through textual guidance and cross-attention reweighing to predict disease progression for patients enrolled for a Retinopathy of Prematurity (ROP) disease screening program. Our technique provides effective visualization for monitoring disease severity over time on new subjects.

Keywords: Diffusion Models, modelling disease severity, medical image synthesis.

1. Introduction

Retinopathy of Prematurity (ROP) is the leading cause of childhood blindness in premature infants, which is caused by the growth of abnormal vasculature in retina. While only 5-10% of total cases progress to severe ROP, it is quite important to understand the physiology as well as disease progression for early diagnosis and effective and timely treatment of ROP.

Recent works (Chen et al., 2021; Thambawita et al., 2022; Coyner et al., 2022; Han et al., 2018) have shown the efficacy of generative models for synthetic medical image generation, either for data augmentation or for privacy-preserved training of AI models. However, to the best of our knowledge, none of the works showcase longitudinal progression in imaging for the same patient via synthetic generation. Stable diffusion models have recently garnered significant attention in medicine, given their effectiveness in generating high-quality synthetic medical imaging (Chambon et al., 2022; Kim and Ye, 2022).

In this study, we work towards developing a tool to visualize the progression of ROP disease even in normal subjects, to better understand the physiology of the disease as well as to distinguish progression of normal eyes from diseased eyes. We leverage the Stable Diffusion model (Rombach et al., 2022) to build a class-conditioned ROP image generation

tool, which further enables a free transformation of disease severity for a given ROP image by combining recent image editing techniques. With our contribution, we hope to pave the way for more accurate and efficient ways for diagnosing as well as early intervention and timely treatment of this condition.

2. Methods

We cast the problem of visualizing ROP progression as an image editing problem. Given an input image of any of the Normal/Pre-Plus/Plus class, the goal is to transform it into another class, while preserving the global structure & some local attributes so that the transformed image is still perceived as an image belonging to the same patient. We first fine-tune a Text-to-Image SD model on the iROP dataset (5500 images), thus allowing the model to generate realistic images of each class prompted by textual inputs. We next employ two image editing techniques: Null Text Inversion (Mokady et al., 2022) & Prompt-to-Prompt Image Editing (Hertz et al., 2022), essentially allowing image modification through attention swapping and reweighting, Fig 1 during the forward diffusion process. To evaluate SD-based image generation, we calculate Frechet Inception Distance (FID) score between real & generated images as well as Plus disease classifier accuracy when tested on class-conditioned generated images. To evaluate image editing, we calculate the correlation between ROP Vascular Severity Score (VSS) of a generated image & real image of increased severity for the same normal image.

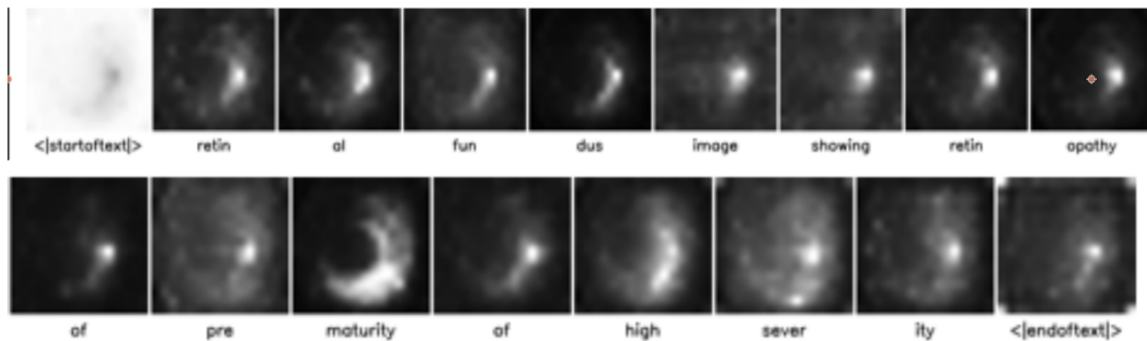


Figure 1: Prompt Attention Map showing image regions impacted by each word

3. Results

Our SD model achieves FID scores of 92, 114, and 137 for Normal, Pre-Plus, & Plus classes, respectively. We achieve 97% accuracy for generated plus class when tested with a Plus Classifier model. A correlation of 0.97 between ROP VSS of generated & real images indicates proportional attention weight adjustment enables the generation of required severity images. The predicted diseased samples of the real normal retinal images follow the same vascular structure, including dilation and tortuosity of blood vessels as they do in the natural progression of the disease, Fig 2, thus proving that this methodology can very closely predict the disease progression in any patient. Using this we can visualize the severity of

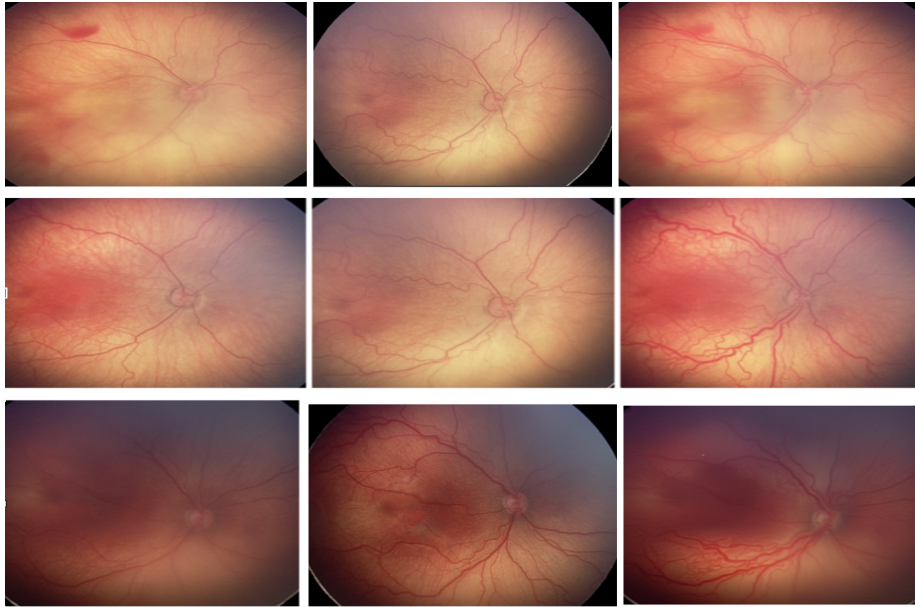


Figure 2: Actual (middle) vs Predicted disease progression (rightmost) of a normal retinal image(leftmost)

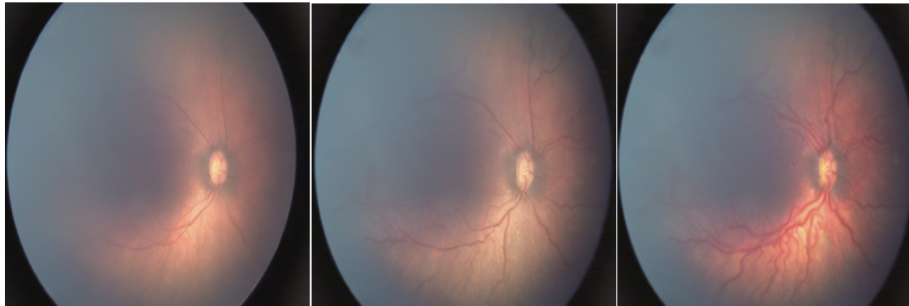


Figure 3: Predicting disease progression on a new sample with no progression yet. From left (Normal) to Pre-Plus(moderate severity) to Plus(Extreme severity of ROP- (right).

the disease at a future date on new data samples which have no disease development yet or are in the moderate disease severity. Fig 3.

4. Conclusion

Our method allows for visualization of the severity of the disease at a future date on new data samples, which may have no disease development yet or are at a moderate severity stage of the disease. This can provide clinicians with valuable insights into the potential progression of the disease and aid in early diagnosis and treatment, ultimately improving outcomes for premature infants.

References

- Pierre Chambon, Christian Bluethgen, Jean-Benoit Delbrouck, Rogier Van der Sluijs, Małgorzata Połacin, Juan Manuel Zambrano Chaves, Tanishq Mathew Abraham, Shivanshu Purohit, Curtis P. Langlotz, and Akshay Chaudhari. Roentgen: Vision-language foundation model for chest x-ray generation, 2022.
- Jimmy S. Chen, Aaron S. Coyner, R.V. Paul Chan, M. Elizabeth Hartnett, Darius M. Moshfeghi, Leah A. Owen, Jayashree Kalpathy-Cramer, Michael F. Chiang, and J. Peter Campbell. Deepfakes in ophthalmology: Applications and realism of synthetic retinal images from generative adversarial networks. *Ophthalmology Science*, 1(4):100079, 2021. ISSN 2666-9145. doi: <https://doi.org/10.1016/j.xops.2021.100079>. URL <https://www.sciencedirect.com/science/article/pii/S2666914521000816>.
- Aaron S Coyner, Jimmy S Chen, Ken Chang, Praveer Singh, Susan Ostmo, R V Paul Chan, Michael F Chiang, Jayashree Kalpathy-Cramer, J Peter Campbell, and Imaging and Informatics in Retinopathy of Prematurity Consortium. Synthetic medical images for robust, privacy-preserving training of artificial intelligence: Application to retinopathy of prematurity diagnosis. *Ophthalmol. Sci.*, 2(2):100126, June 2022.
- Changhee Han, Hideaki Hayashi, Leonardo Rundo, Ryosuke Araki, Wataru Shimoda, Shinichi Muramatsu, Yujiro Furukawa, Giancarlo Mauri, and Hideki Nakayama. Gan-based synthetic brain mr image generation. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 734–738, 2018. doi: 10.1109/ISBI.2018.8363678.
- Amir Hertz, Ron Mokady, Jay Tenenbaum, Kfir Aberman, Yael Pritch, and Daniel Cohen-Or. Prompt-to-prompt image editing with cross attention control, 2022.
- Boah Kim and Jong Chul Ye. Diffusion deformable model for 4d temporal medical image generation, 2022.
- Ron Mokady, Amir Hertz, Kfir Aberman, Yael Pritch, and Daniel Cohen-Or. Null-text inversion for editing real images using guided diffusion models, 2022.
- Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models, 2022.
- Vajira Thambawita, Pegah Salehi, Sajad Amouei Sheshkal, Steven A. Hicks, Hugo L. Hammer, Sravanthi Parasa, Thomas de Lange, Pål Halvorsen, and Michael A. Riegler. Singanseg: Synthetic training data generation for medical image segmentation. *PLOS ONE*, 17(5):1–24, 05 2022. doi: 10.1371/journal.pone.0267976. URL <https://doi.org/10.1371/journal.pone.0267976>.