RetinaSynth: Diffusion-Based Synthetic Imaging

Recent advances in artificial intelligence (AI) have unlocked remarkable potential across a wide spectrum of application domains. For AI models to be successfully deployed, they usually need access to vast amounts of high-quality data. In healthcare, however, such data is very constrained due to the sensitive nature of patient information, and strict legal frameworks (e.g., HIPAA in the United States and GDPR in Europe) govern its use. This creates a major obstacle for building robust AI-driven medical solutions [1]. A promising approach to overcome this barrier is the generation of synthetic data that closely resembles real-world data distributions while protecting individual privacy. Synthetic datasets can therefore accelerate model development and validation while staying fully compliant with privacy regulations. Beyond augmentation, synthetic data unlocks other new opportunities: translating images between modalities, simulating imaging under different acquisition parameters, and producing realistic material for training and education [2,3]. In this paper, we investigate if generative models (VAEs, diffusion, Pix2Pix) can create realistic retinal images, while also assessing their clinical usefulness, fairness, and regulatory acceptability. We propose a three-stage pipeline, shown in Figure 1., that (i) extracts vascular structures via segmentation, (ii) generates synthetic vessel skeletons using either Variational Autoencoders (VAEs) or diffusion models, and (iii) translates skeletons into retinal images using Pix2Pix in both paired and unpaired settings. To the best of our knowledge, this work is first to apply diffusion models to vessel-segmented fundus images for synthetic retinal image generation.

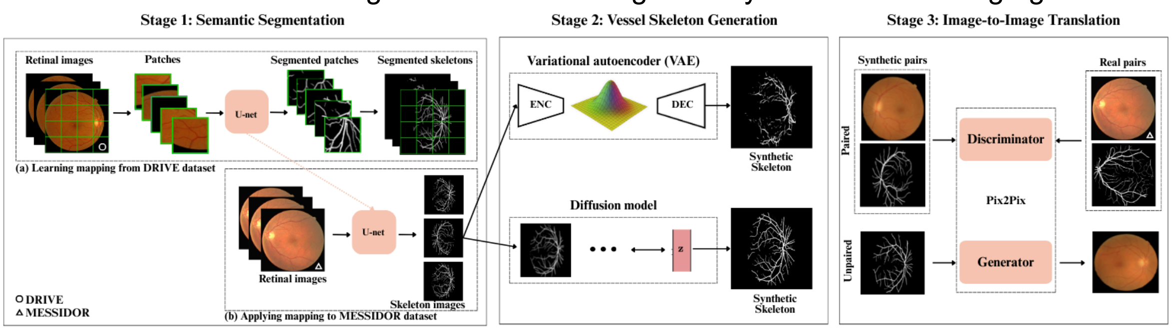

Figure 1. Illustration of stages in our pipeline for synthetic retinal images generation.

Through comparative evaluation, we demonstrated differences in fidelity, diversity, and realism across these paradigms. SSIM ≈ 0.85 indicates strong structural similarity between synthetic and real images; PSNR ≈ 23 dB reflects visible but moderate photometric differences; MAE on [0, 255] corresponds to roughly 3-4% average per-pixel deviation. Across our experiments, diffusion models delivered the most photorealistic retinal images-sharp textures and clean vessel boundaries-at the cost of slower sampling and a need for light postprocessing. VAEs are fast and stable but characteristically over-smooth, serving mainly as a lightweight baseline or pretraining aid.

Synthetic retinal images hold significant potential for augmenting medical datasets, enabling educational use cases, and supporting anonymization without compromising patient privacy. Our study provides an early but systematic view of how generative models can contribute to these goals and identifies promising directions for further exploration. Our work opens new opportunities for the medical AI ecosystem. Researchers in medical imaging can use it to test and refine generative models, such as VAEs, diffusion models, and Pix2Pix, for expanding scarce datasets. Clinicians and educators gain practical tools to generate lifelike, pathology-representative images that make training and teaching more effective. Beyond that, the wider medical AI community can benefit from our insights on evaluation metrics, privacy safeguards, and promising future directions for bringing synthetic images into everyday diagnostic and educational practice.

[1] Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. Medical image analysis. 2019 Dec 1;58:101552.

[2] Costa P, Galdran A, Meyer MI, Niemeijer M, Abràmoff M, Mendonça AM, Campilho A. End-to-end adversarial retinal image synthesis. IEEE transactions on medical imaging. 2017 Oct 2;37(3):781-91.

[3] Goceri E. Medical image data augmentation: techniques, comparisons and interpretations. Artificial intelligence review. 2023 Nov;56(11):12561-605.