Retinovascular Biomarkers of Acute Ischemic Stroke: A Generative Modelling Framework for Synthetic Fundus Image Synthesis and Classification

Xin Wang[1], Sampath Rapuri[1], Zhiyuan Ding[1], Sanjukta Biswas[1], Kehui Ge[1], Xianzhe Tan[1], Siam Mohammed[2], Siamak Ardekani[1], Joseph L. Greenstein[1], Casey Overby Taylor[1], Kemar E. Green[3–5]

[1] Johns Hopkins University, Department of Biomedical Engineering, Baltimore, MD, USA [2] Johns Hopkins University, Department of Computer Science, Baltimore, MD, USA [3] NeuroAgent AI, Inc., USA [4] Johns Hopkins University, Data Science and Artificial Intelligence Institute, Baltimore, MD, USA [5] Johns Hopkins University, Institute for Data Intensive Engineering and Science, Baltimore, MD, USA

Introduction

- **Global Burden:** Stroke is a leading cause of death and disability; ischemic cases (87%) cost the U.S. ~$56B annually.[1]
- **The Resource Gap:** Life-saving early interventions are often limited by a lack of specialized diagnostic infrastructure.
- **Retinal Imaging:** Fundus photographs offer a non-invasive, low-resource alternative to CT/MRI via shared vascular pathways.[2]
- **Data Scarcity:** Clinical datasets for AI development are virtually non-existent and difficult to share due to privacy constraints.
- **Synthetic Solution:** We model physiological vascular changes during acute stroke to create a robust, patient-free synthetic dataset for AI classifier training.

Aims

- **AIM 1:** Arteriolar-to-Venular Ratio (AVR) as a Stroke Biomarker.
- **AIM 2:** Synthetic Dataset using Diffusion Modeling & Validation.
- **AIM 3:** Deep Learning Classifier Evaluation to classify stroke patterns.

Method & Analytics

- **We** develop an automated pipeline to extract the AVR from retinal images. A threshold of **AVR < 0.66** as a quantitative biomarker to identify high-risk vascular narrowing.
- **We trained** a generative **diffusion model** to synthesize realistic retinal vascular patterns, explicitly modeling retinal arcades and focusing on major vessels. The model jointly predicts vessel masks and AVR.
- **We** trained a **ResNet-34** classifier on the synthetic dataset to detect stroke-risk patterns.

AVR Computation

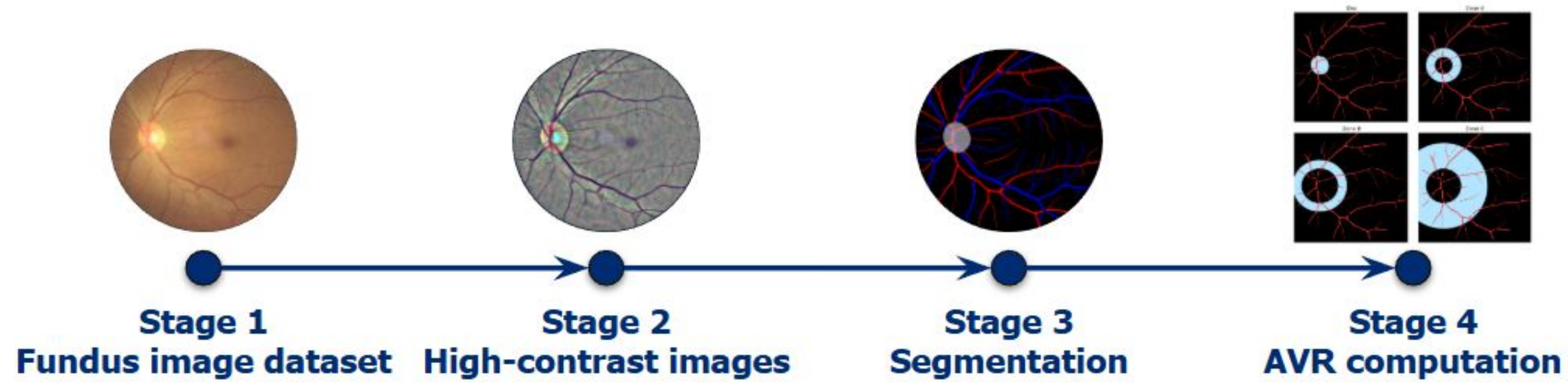

Figure 1: Pipeline for AVR Computation.

Visualization of Synthetic Vessel Mask

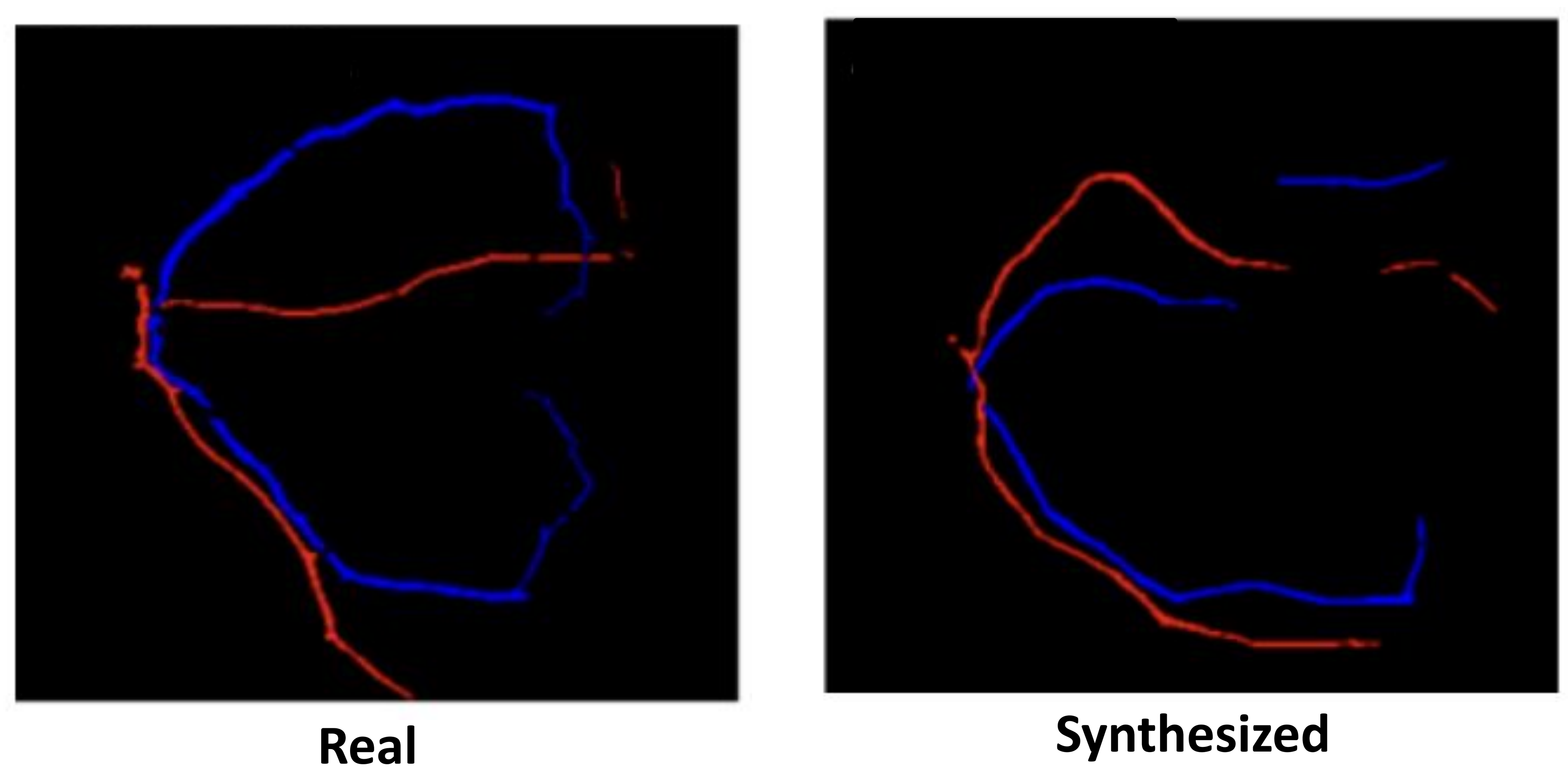

Figure 2: An example of a real and synthesized retinal vascular arcades.

Generative quality: FID 57.8, KID 0.0648 (validated on **2000** synthetic samples)

Validation of Conditioning Accuracy

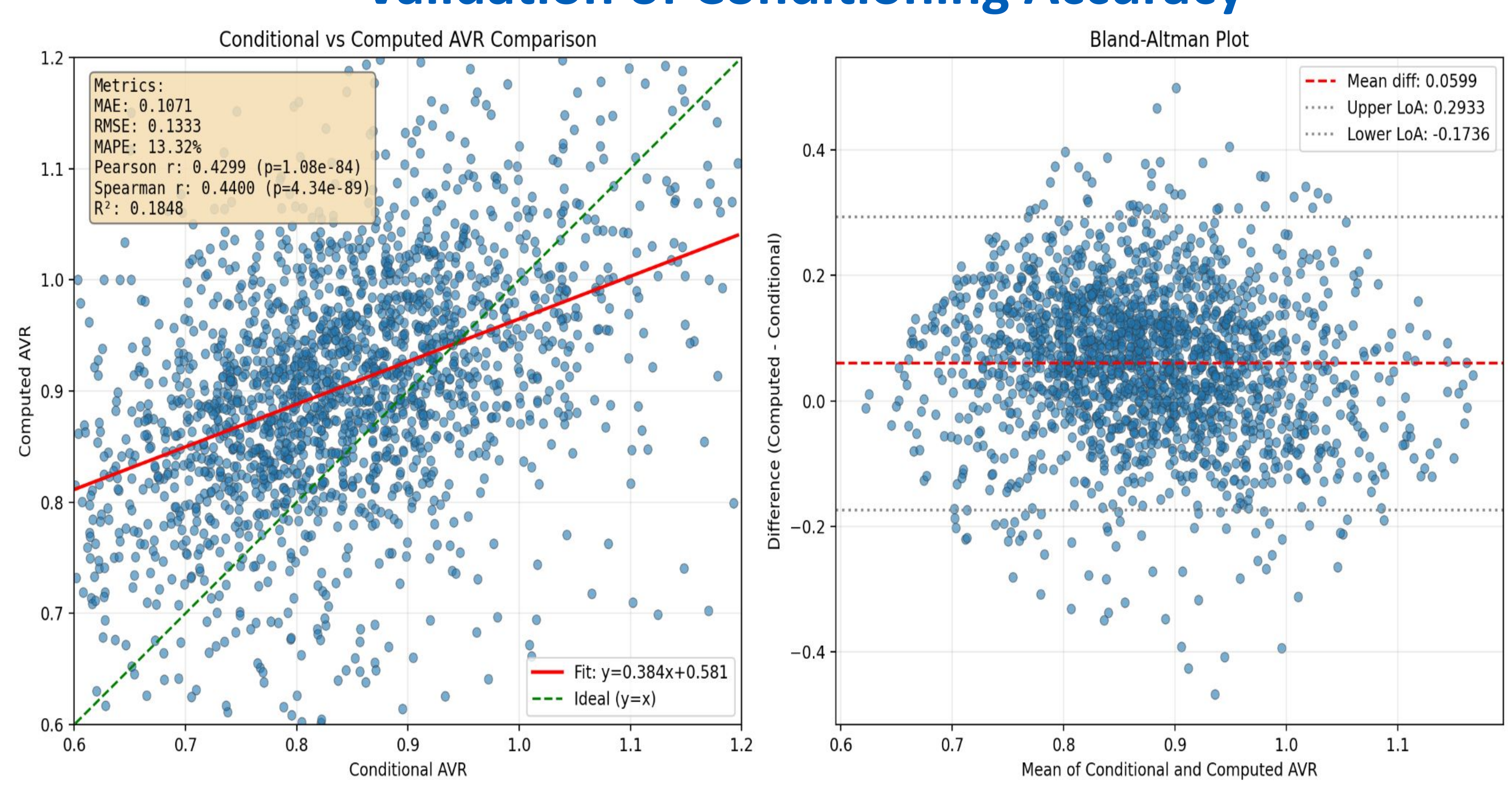

Figure 3: Agreement between input conditional variable and computed AVR from the synthesized retinal masks.

AVR correlation: **Pearson r = 0.4399, Spearman ρ = 0.4400**

Classification Performance

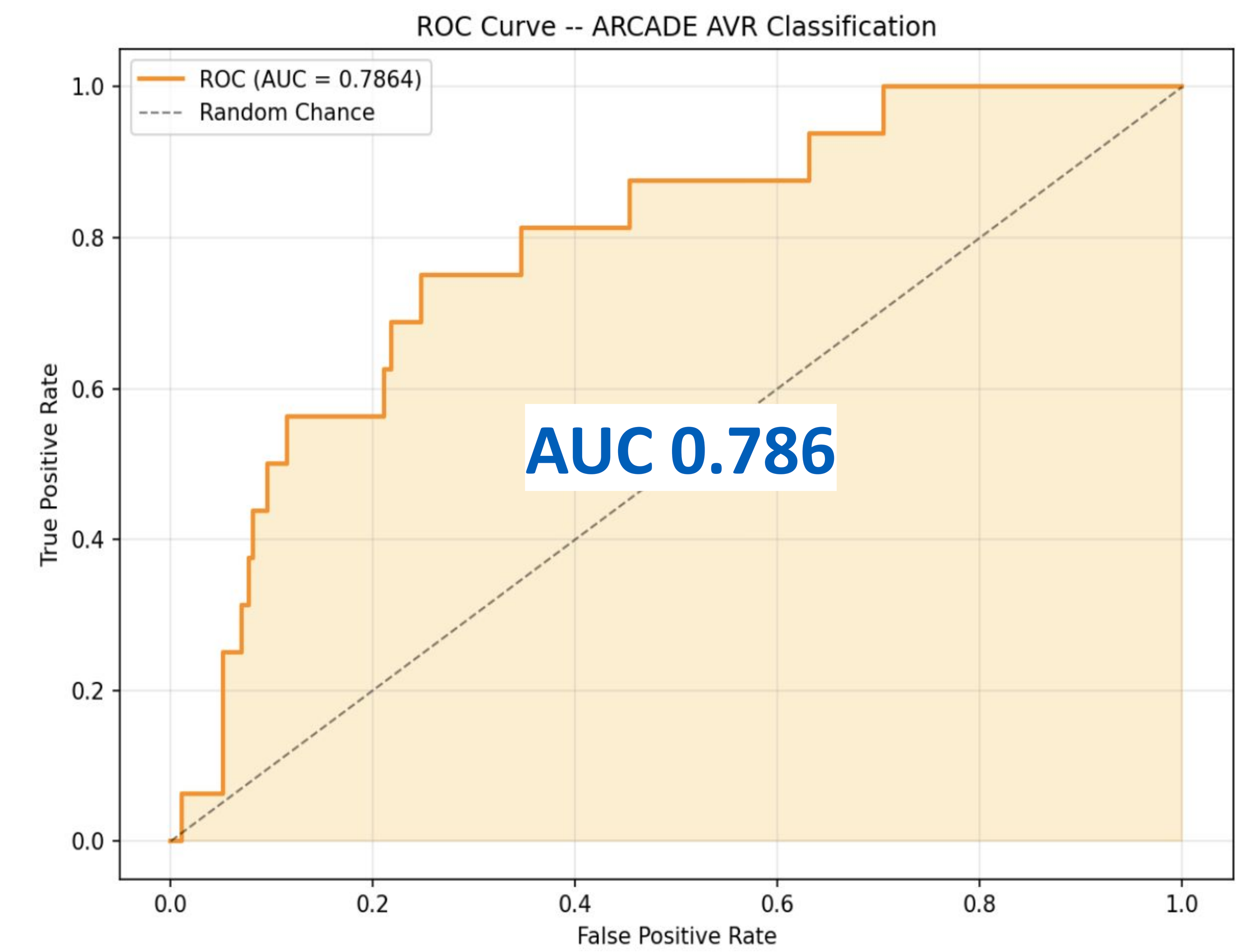

Figure 4: ROC curve for arcade-based AVR classification.

Conclusion

- We modeled physiological vascular changes to generate a robust, **patient-free** retinal dataset.
- This approach bypasses clinical data scarcity and privacy barriers, providing a **scalable foundation** for training stroke-specific AI classifiers.
- Future work will focus on validating the classifier on real clinical stroke fundus datasets and extending this biomarker-conditioned generative framework to other retinal image–based disease screening tasks.

References

1. National Center for Health Statistics. Multiple Cause of Death 2018–2023 on CDC WONDER Database. Accessed February 1, 2025.
2. Zhao Y, et al. Retinal microvascular changes in subtypes of ischemic stroke[J]. Frontiers in Neurology, 2021.