A172 · ImageFlowNet forecasts disease progression in longitudinal medical images

Presenter: Chen Liu — Yale University

Authors: Chen Liu, Santanu Antu, Kishalay Das, Ke Xu, Xiangyu Zhang, Aryaman Mishra, Jason Shaye, Ethan Zhang, Liangbo L. Shen, Guillaume Huguet, Zilong Wang, Alexander Tong, Danilo Bzdok, Sanjay Aneja, Jay Stewart, Jay C. Wang, Lucian V. Del Priore, Smita Krishnaswamy

Session: Day 2 · Fri Oct 2, 2026 · 4:15–5:15 PM

Forecasting disease progression from longitudinal medical images, or the process of generating future images from past history, has become an active area of research. However, it remains unclear whether current models learn the dynamics of the disease or primarily reconstruct the appearance while underestimating disease-related changes. Because real progression is often slow and subtle, reconstruction-biased methods can appear accurate even when they do not reflect disease evolution. To isolate this limitation, we build a controlled benchmark from real retinal geographic atrophy images by computationally amplifying atrophy growth over time, strengthening the progressive signal while retaining native anatomy and texture. In this setting, we evaluate ImageFlowNet, which transports patient embeddings in a shared latent space with position-parameterized neural ordinary differential equations to predict progression rather than merely reproduce the input.

ImageFlowNet forecasts disease progression in longitudinal medical images

Chen Liu[1,*] Santanu Antu[1] Kishalay Das[1] Ke Xu[1] Xiangyu Zhang[1] Aryaman Mishra[1] Jason Shaye[1] Ethan Zhang[1] Liangbo L. Shen[2] Guillaume Huguet[3,4] Zilong Wang[3,5] Alexander Tong[3,4] Danilo Bzdok[3,5] Sanjay Aneja[1] Jay Stewart[2] Jay C. Wang[1,2,6] Lucian V. Del Priore[1] Smita Krishnaswamy[1]

[*]Presenting author [1]Yale University [2]University of California, San Francisco [3]Mila - Quebec AI Institute [4]Université de Montréal [5]McGill University [6]Northern California Retina Vitreous Associates

Correspondence: smita.krishnaswamy@yale.edu, lucian.delpriore@yale.edu

Introduction

Forecasting disease progression from longitudinal medical images, or the process of generating future images from past history, has become an active area of research. However, it remains unclear whether current models learn the dynamics of the disease or primarily reconstruct the appearance while underestimating disease-related changes. Because real progression is often slow and subtle, reconstruction-biased methods can appear accurate even when they do not reflect disease evolution. To isolate this limitation, we build a controlled benchmark from real retinal geographic atrophy images by computationally amplifying atrophy growth over time, strengthening the progressive signal while retaining native anatomy and texture. In this setting, we evaluate ImageFlowNet, which transports patient embeddings in a shared latent space with position-parameterized neural ordinary differential equations to predict progression rather than merely reproduce the input.

Methods

A UNet encodes an image x_i at time t_i as a latent state z_{t_i} (Fig. 1). We flow the latest state into its corresponding future state z_{t_j} by integrating along a learned vector field f_θ using a position-parameterized ODE formulation.

$$\frac{\mathrm{d}z_\tau}{\mathrm{d}\tau} = f_\theta(z_\tau, \tau) \qquad (1) \qquad\qquad \frac{\mathrm{d}z_\tau}{\mathrm{d}\tau} = f_\theta(z_\tau) \qquad (2)$$

Eq. (1) is the standard time-parameterized neural ODE, whereas Eq. (2) is our position-parameterized ODE. Future states follow by integrating over time, yielding $z_{t_j} = z_{t_i} + \int_{t_i}^{t_j} f_\theta(z_\tau)\,\mathrm{d}\tau$. With a position-parameterized field, trajectories from different patients populate one compact flow, states are organized by disease severity rather than calendar time, and arbitrary-horizon forecasts can be produced from a single prior visit by integrating over $t_j - t_i$. A Langevin SDE, $\mathrm{d}z_\tau = f_\theta(z_\tau)\,\mathrm{d}\tau + \sigma_\phi(\tau)\,\mathrm{d}W_\tau$, provides an optional stochastic variant. The decoded prediction $\widehat{x}_j$ is supervised by pixel-wise mean squared error loss with the true future image x_j.

Results

On the amplified-atrophy benchmark, predicted images from ImageFlowNet surpass recent methods in every metric (Fig. 2). Compared to the strongest baseline (I^2SB), ImageFlowNet's pixel fidelity improves from PSNR 20.26 dB to 23.23 dB and MAE from 0.071 to 0.051. Structural similarity (SSIM) of ImageFlowNet leads the strongest baseline by 0.027. The disease-region delineation improves even more sharply: Dice-Sørensen coefficient (DSC) increases from 0.826 to 0.941 and the mean absolute error within the region-of-interest (ROI-MAE) decreases from 0.148 to 0.088. These results indicate that ImageFlowNet recovers both the overall image appearance and the disease-region progression.

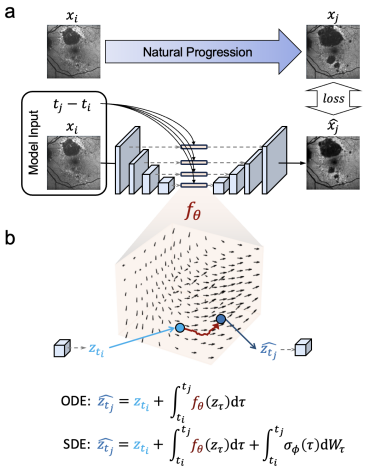

Figure 1: Overview of ImageFlowNet. (a) The model takes an earlier image x_i at time t_i and the time change $t_j - t_i$ to forecast the future image x_j at time t_j. (b) A vector field f_θ models the joint patient embedding space; trajectory inference is performed by integrating along this field. In practice, $t_j - t_i$ suffices for integration, while the absolute times t_i and t_j appear in the integral for notational clarity. An optional stochastic diffusion term yields the SDE variant.

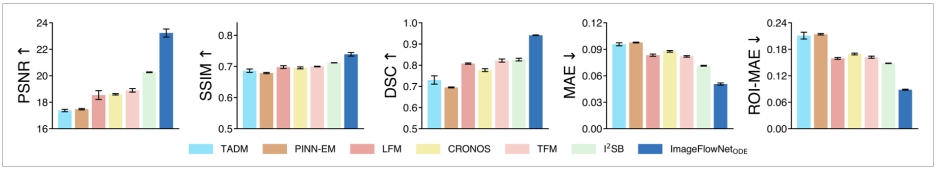

Figure 2: Amplified retinal atrophy benchmark. Future images predicted by ImageFlowNet not only demonstrate higher pixel fidelity (PSNR, MAE) and structural similarity (SSIM), but also delineate the disease region better (DSC, ROI-MAE), compared to methods including TADM, PINN-EM, LFM, CRONOS, TFM and I^2SB.