SAGIVTECH IMVC 2026
15th Israel Machine Vision Conference (IMVC) 2026
Monday | April 27, 2026 | Pavilion 10, EXPO Tel Aviv

# Shifting the Retinal Foundation Models Paradigm from Slices to Volumes for Optical Coherence Tomography

*Raphael Judkiewicz, Technion - Israel Institute of Technology*

Optical Coherence Tomography (OCT) is essential in ophthalmology for cross-sectional imaging of the retina. Pretrained foundation models facilitate task-specific model development by enabling fine-tuning with limited labeled data. However, current foundation models rely on a single B-scan (usually the central slice), overlooking volumetric context. This research investigates video foundation models to capture full 3D retinal structure and improve diagnostic performance. V-JEPA, a state-of-the-art video foundation model, was benchmarked against retinal foundation models (RETFound, VisionFM) and a natural image foundation model (DINOv2). All were fine-tuned to detect Age-related Macular Degeneration or Glaucomatous Optic Neuropathy using five OCT datasets. V-JEPA consistently equaled or outperformed image-based models, achieving an average AUROC of 0.94 (0.80–0.99), versus 0.90 (0.76–0.98) for the best image model, a statistically significant improvement ($p < 0.001$). To our knowledge, this is the first application of transformer-based video models to volumetric OCT, highlighting their promise in 3D medical imaging.