How Useful Are Vision Foundation Model Features for Out-of-the-Box Disease Progression Prediction?

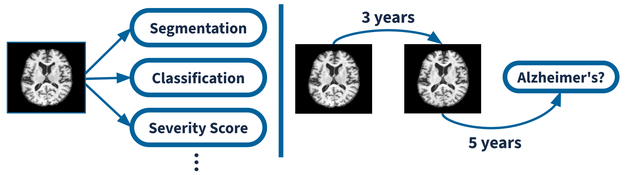
Basar Demir[1], Soumitri Chattopadhyay[1], Thomas Hastings Greer[2], Boqi Chen[2], Marc Niethammer[1]

[1]University of California San Diego, [2]University of North Carolina at Chapel Hill

CVPR JUNE 3-7, 2026 DENVER COLORADO

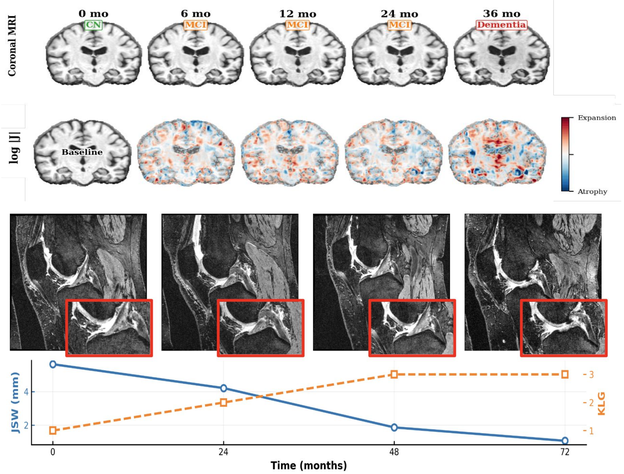
Motivation

- Beyond detecting disease: can we predict its future course?
- Disease progression often appears as subtle anatomical change over time.
- We test whether frozen medical vision foundation models already capture these signals.
- This enables progression prediction without task-specific fine-tuning.



Datasets

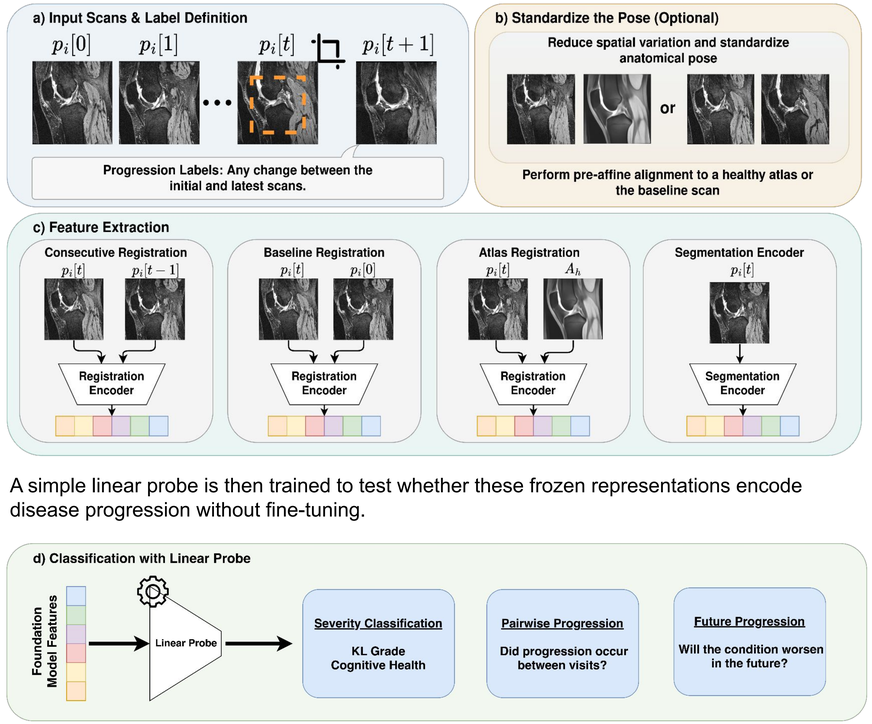
We study knee OA and Alzheimer's disease, two settings where disease progression can sometimes be seen by eye directly in longitudinal scans.



Methodology

We use frozen medical vision foundation models as feature extractors for longitudinal MRIs.

Registration features are extracted using different scan-pair strategies, while segmentation features are extracted from individual scans.

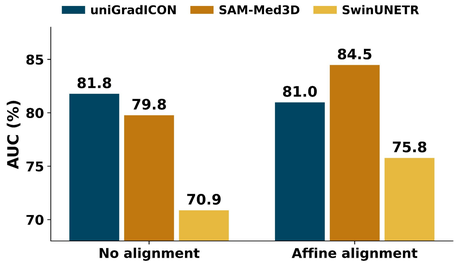


A simple linear probe is then trained to test whether these frozen representations encode disease progression without fine-tuning.

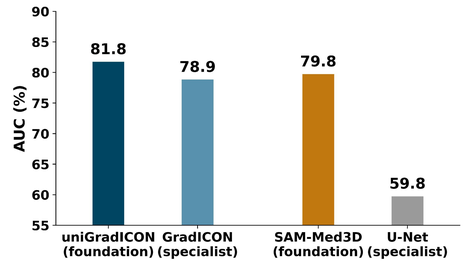
Key Takeaways

- Pick registration features (uniGradICON) for robustness on messy / unaligned data
- Pick segmentation features (SAM-Med3D) when it is possible to standardize the data.
- Future work should develop modular adaptation layers on pretrained models to better capture high-dimensional clinical features.

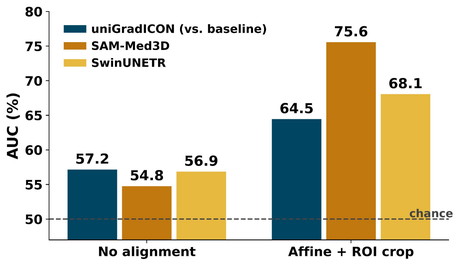
Results and Findings



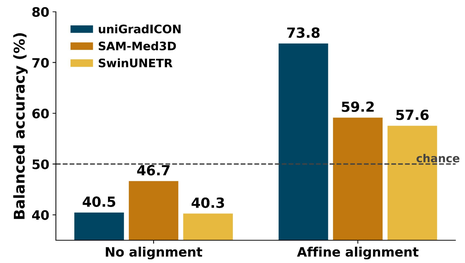
Registration is robust to spatial misalignment, segmentation requires affine pre-alignment to compete.



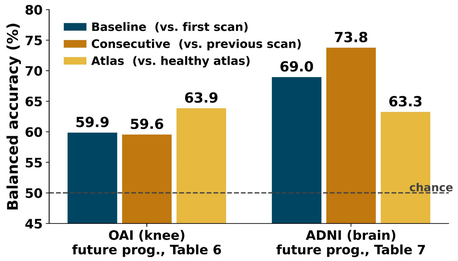
Foundational models beat specialist. Diverse multi-organ pretraining beats narrow specialist training.



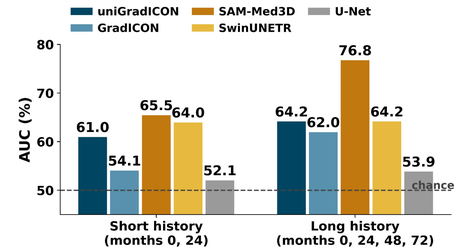
Affine alignment + joint cropping massively boosts segmentation features. Registration stays robust.



Affine pre-alignment is a necessity for longitudinal brain MRI. Without this, models collapse. When pre-aligned, it heavily boosts all model features.



There is no general registration pair that works best. For OAI, atlas-registered features give best results. For ADNI, consecutive pairs of scans work best.



Temporal context length matters. All models show improvement when given the full longitudinal trajectory.

Acknowledgements: The presented research was, in part, funded by the National Institutes of Health (NIH) under other transactions 1OT2OD038045- 01 and NIAMS 1R01AR082684. Data used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. ADNI data collection and sharing are funded by the National Institute on Aging, NIH Grant U19AG024904.

★ ORAL PRESENTATION ★

Email: bdemir@ucsd.edu
Code & resources: github.com/ucsdbiag