SIIM CAIMI26
Conference on Artificial Intelligence in Medical Imaging
OCT 26-27 | University of Pennsylvania, Philadelphia, PA

Anatomy-Aware ROI Input Outperforms In-Domain Self-Supervised Pretraining for MRI-Based Adenomyosis Classification

Delaram Ghadimi, MD, Mayo Clinic; Amirali Khosravi, MD; Tubo Shi, MB, ChB; Ivana Isgum, PhD; Bradley J. Erickson, MD, PhD, CIIP, FSIIM; Wendaline VanBuren, MD

Introduction/Background

Adenomyosis remains challenging to diagnose on pelvic MRI, and limited labeled datasets constrain deep learning performance. In-domain self-supervised learning (SSL) is increasingly used to improve representation learning in low-label settings, but its benefit relative to anatomy-aware input design has not been established. We compared these two strategies for MRI-based adenomyosis classification.

Methods/Intervention

We retrospectively analyzed pelvic MRI from 420 patients (132 adenomyosis, 288 controls) with automated uterus segmentations and registered sagittal T2, T1 fat-suppressed, and T1 contrast-enhanced images. Five classifier configurations were evaluated: (1) whole-image DINOv3 vision transformer, (2) the same model after continued in-domain SSL pretraining (on 24000 in-house sagittal three-sequence gynecologic MRI series), (3) the same backbone confined to the segmented uterine region (masking, cropping, intensity normalization to reduce organ-size shortcuts), (4) a three-sequence extension of this masked-ROI design, each masked with its own segmentation, and (5) a 3D MRI foundation model (trained on 129k out-of-domain MRI series) with and without continued in-domain pretraining. Performance was assessed via bootstrap resampling (20 iterations) and cross-validation.

Results/Outcome

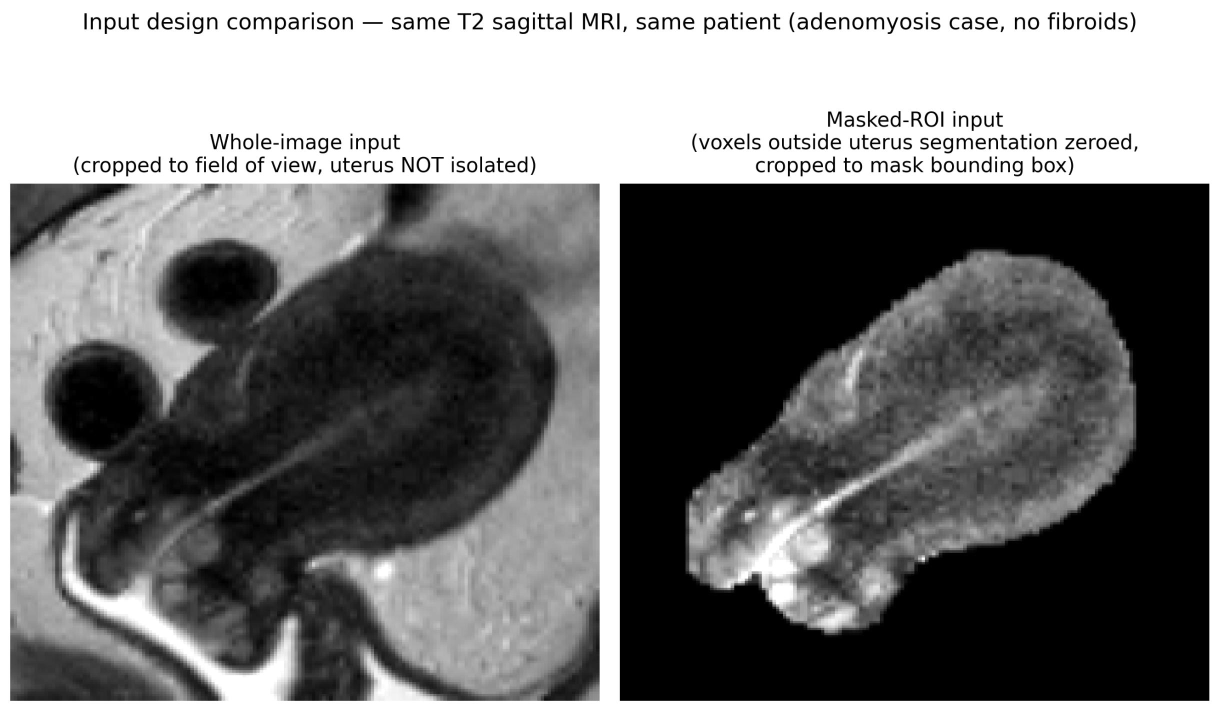
Continued in-domain SSL pretraining did not improve performance over released foundation-model weights in either family. Whole-image DINOv3 classifiers achieved comparable AUROC (~0.70) with and without continued pretraining, while the 3D foundation model showed similarly unchanged performance (AUROC 0.67–0.69). In contrast, the anatomy-aware ROI approach achieved the highest performance (AUROC 0.756±0.061, balanced accuracy 0.72, F1 0.70, MCC 0.42) on bootstrap validation. A three-sequence extension reproduced this performance (AUROC 0.760±0.042) with lower bootstrap-to-bootstrap variance, indicating the gain reflects anatomy-aware masking rather than sequence count. Independent held-out testing is ongoing.

Conclusion

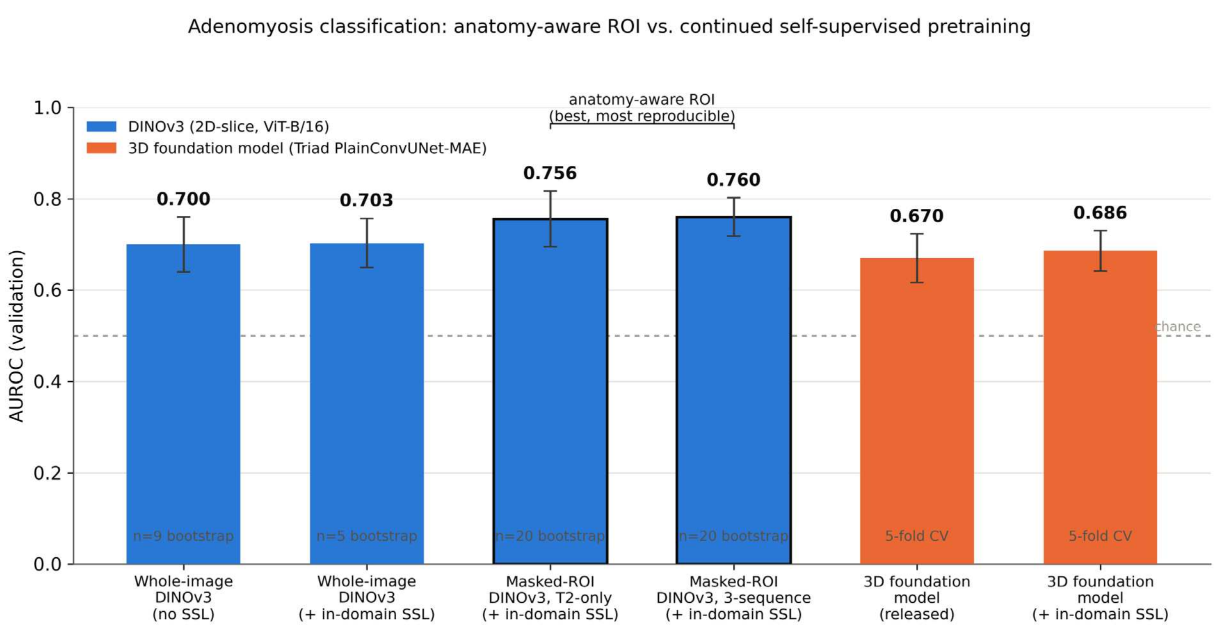
Across two foundation models, continued in-domain SSL pretraining provided no measurable benefit for adenomyosis classification, while anatomically constrained input design consistently produced the largest, most reproducible gains, whether applied to one sequence or extended across multiple sequences with independent masks. Anatomy-aware preprocessing may outperform additional SSL pretraining in small labeled MRI cohorts.

Statement of Impact

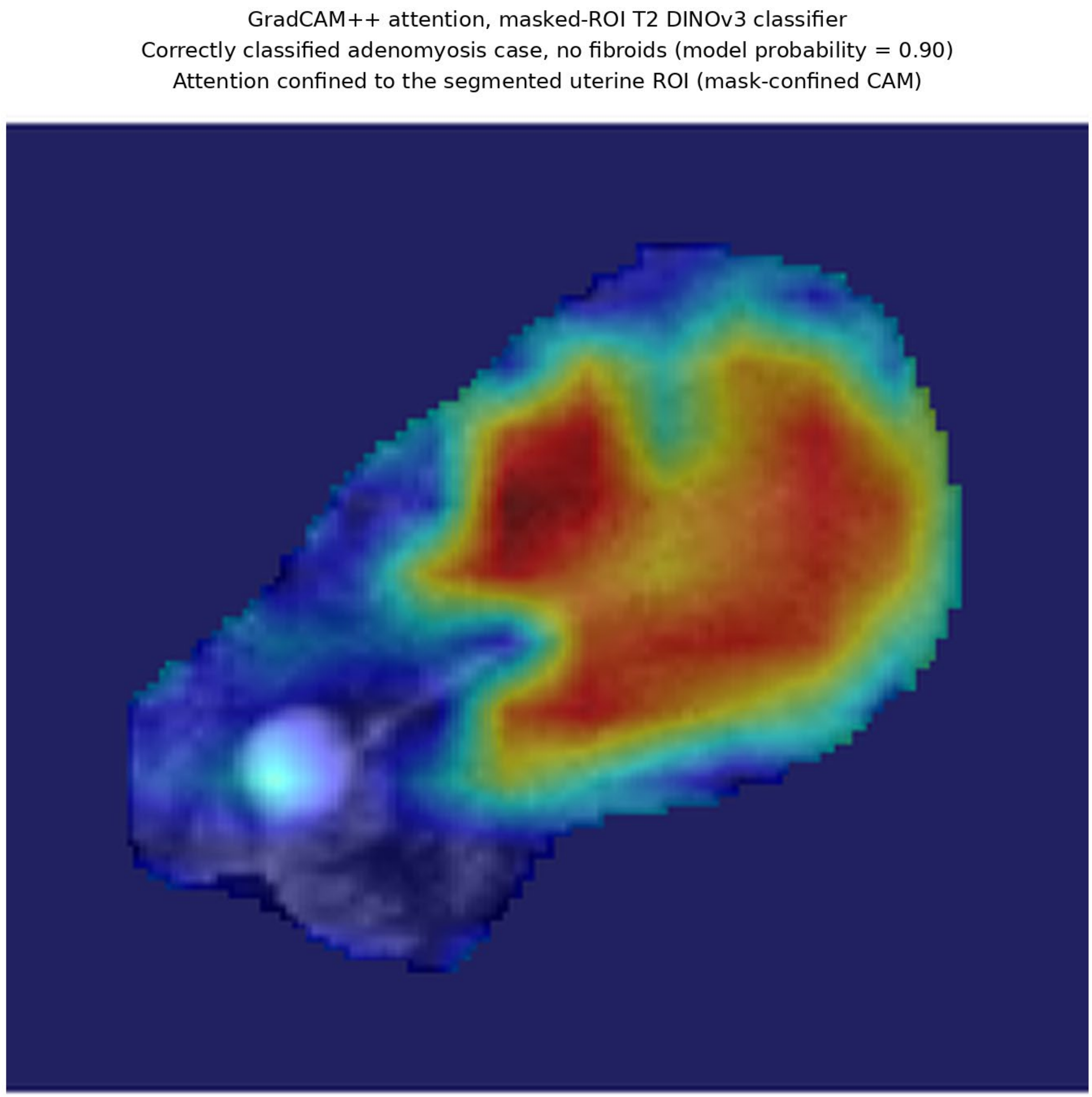
For AI developers in data-limited medical imaging domains, these results suggest prioritizing anatomy-aware input engineering over costly in-domain SSL pretraining as the more reliable path to performance gains. Clinically, an accurate, interpretable, ROI-confined adenomyosis classifier could support more consistent detection on routine pelvic MRI.



Input design comparison for the same patient (adenomyosis case, no fibroids), shown as a central sagittal T2 slice. Left: whole-image input, cropped to the field of view — the uterus is visible but not isolated from surrounding pelvic anatomy. Right: masked-ROI input — voxels outside the automated uterine segmentation are zeroed and the image is cropped to the mask bounding box, so the classifier sees only the segmented organ.



Validation AUROC across all six evaluated configurations, grouped by backbone architecture (blue: DINOv3 2D-slice ViT-B/16; orange: 3D foundation model). Error bars show ±1 SD across bootstrap iterations (DINOv3 models) or cross-validation folds (3D foundation models); n per bar is annotated. The two anatomy-aware masked-ROI configurations (outlined) achieved the highest and most reproducible performance, while continued in-domain SSL pretraining produced no consistent gain in either architecture family.



GradCAM++ attention map for the masked-ROI T2 DINOv3 classifier on a correctly classified adenomyosis case (no fibroids; model probability = 0.90). Attention is confined to the segmented uterine region, concentrating within the myometrium rather than at the image periphery, consistent with the model relying on anatomically relevant signal rather than a size- or border-based shortcut.

Keywords

Foundation Models; Vision Transformer Model; Self-Supervised Learning; Adenomyosis; Anatomy-aware AI; MRI