MICCAI 2026
29TH INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING & COMPUTER ASSISTED INTERVENTION
Sept. 27 - Oct. 1, 2026 • Strasbourg Convention Center, FRANCE
MICCAI

MLMI 2026 · WORKSHOP ON MACHINE LEARNING IN MEDICAL IMAGING · MICCAI SATELLITE EVENT

Sudha Gopalakrishnan Brain Centre · IIT Madras · jigneshkumar@htic.iitm.ac.in

Learn the Metric, Not the Deformation

Cross-modal histology–MRI registration with a 76k-parameter contrastive encoder: the similarity is learned once, the deformation is solved per pair.

Jigneshkumar Mistry Keerthi Ram Supriti Mulay Sivathanu Neelakantan Mohanasankar Sivaprakasam

MAIN FINDING

Predicting the deformation is what fails across brains. Learn only the similarity metric and solve each pair's warp on the spot.

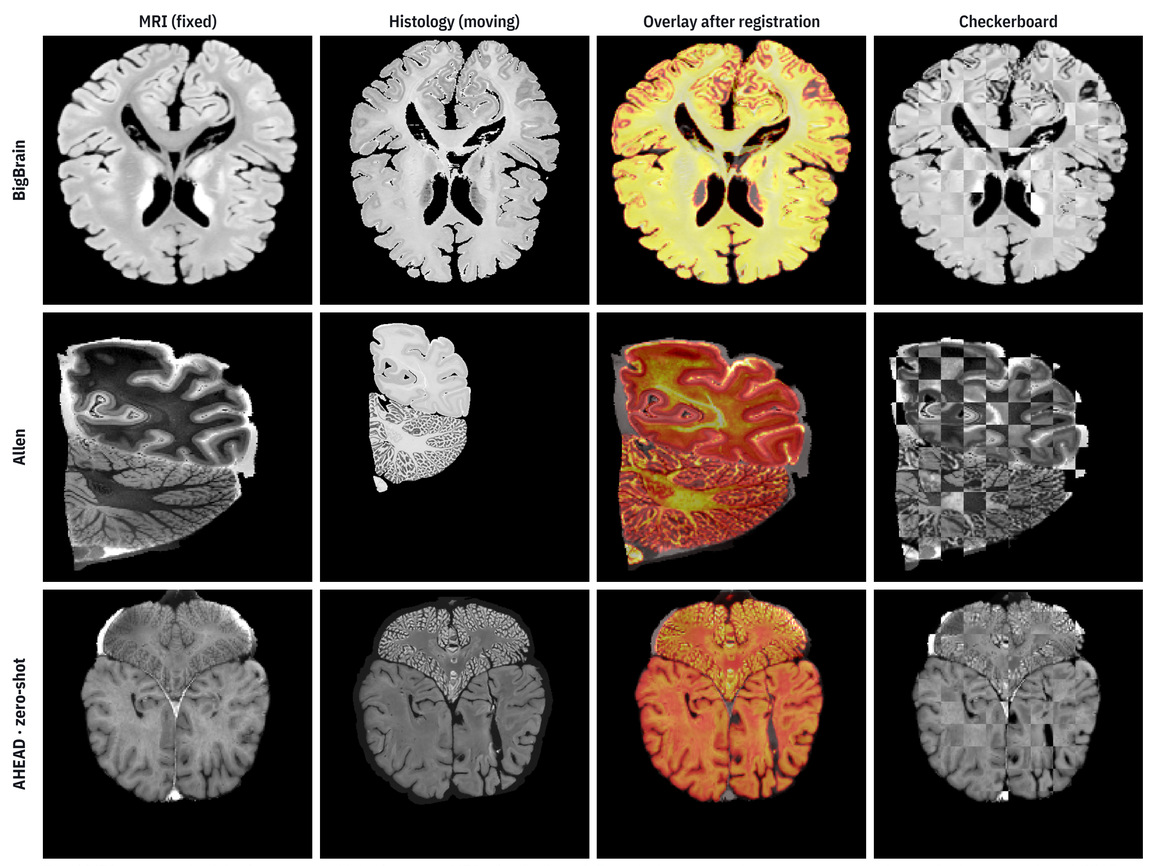
A 76k-parameter encoder reaches **1.47 mm** landmark error with **zero folding** on BigBrain, and transfers **zero-shot** to AHEAD brains it never saw — in 0.9 s per pair.

WHAT IS NEW

- A 76k fully-convolutional PatchNCE encoder used as a **cross-modal similarity metric.**
- A velocity decoder used as an **initial-condition predictor,** not a deformation predictor.
- A per-pair diffeomorphic optimiser: sub-affine TRE, zero folding, zero-shot transfer — **all three at once.**

DATA & TRAINING

Allen ds003590	Nissl 250 µm ↔ 7T T2w 0.5 mm
BigBrain	donor histology 100 µm ↔ donor MRI
AHEAD 122017 / 152017	thionin ↔ quantitative R_1, R_2^* — never trained on
Landmarks	BigBrain AFIDs, 32 in ICBM2009sym; 28- and 30-fiducial protocols
Encoder objective	PatchNCE $\tau = 0.07$, 256 patches/img, other sites as hard negatives
Augmentation	paired-geometric (keeps correspondence) + independent-intensity (forces modality invariance)
Slice matching	DINOv2 layer-7 features + DTW — zero matching error on test pairs



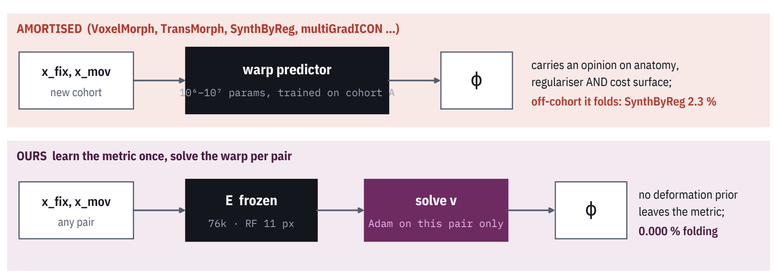
Fixed MRI · raw moving histology · warped histology over MRI · checkerboard. Boundaries stay continuous across every tile; det J > 0 at every pixel. AHEAD (thionin → quantitative R_1) was never seen in training.

1.47 mm	0.908	0.000 %	0.077	76 k	0.9 s
MEAN TRE · 28 AFIDS	**DICE**	**FOLDED PIXELS**	**MIND · AHEAD ZERO-SHOT**	**ENCODER PARAMETERS**	**PER PAIR · 122 MiB**
±0.016 mm over 50 seeds; affine alone 1.87, CoMIR 1.85	overlap on par with the best rows — without their folding	min det J 0.540 over 1,032,192 px, measured at inference	TIRL-FFD 0.150, SynthByReg 0.309 on the same 32 slices	28× smaller than CoMIR; both networks frozen at test time	one A100, no test-time training — only v is updated

12 methods, one protocol · 3 datasets · 4 MRI contrasts · 2 training brains, 2 never seen · Wilcoxon $p < 10^{-6}$ (Holm) · NMI 1.247 zero-shot · det J > 0 at every pixel · 192 × 192 working grid

THE PROBLEM

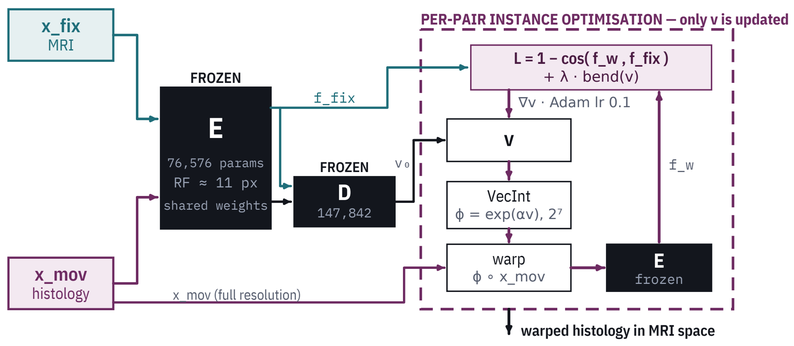
Why not predict the deformation?



Scaling-and-squaring only approximates a diffeomorphism while the velocity stays small relative to the grid. A predictor pushed out of distribution leaves that regime — **architectural intent is not a guarantee.**

OUR RECIPE

Frozen metric, per-pair solve



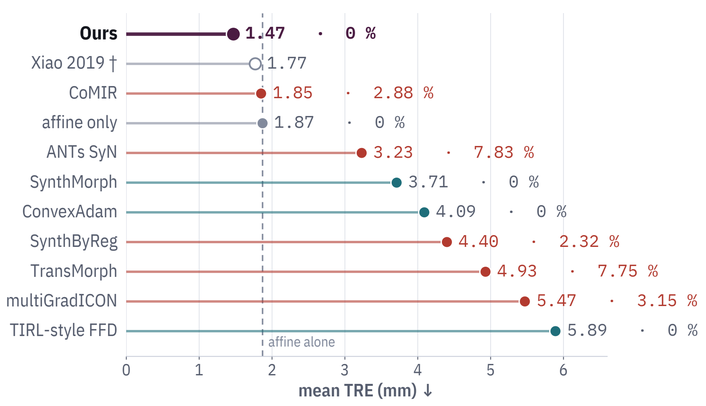
$$\hat{v} = \arg\min_{v} \; 1 - \cos\big(E(x_fix),\, E(\phi_v \circ x_mov) \big) + \lambda \sum \|\Delta v\|^2$$

— learned cross-modal metric — bending energy, λ = 10

E (3 × 3×3 conv, 64 ch, PatchNCE $\tau = 0.07$) scores local texture, not global shape — that is what lets one metric cross brains and contrasts. **D** only supplies v_0: it lifts the optimiser out of the rotation local minimum (recovery 75 % → ≥ 91 % up to 40°). Trained jointly on Allen + BigBrain, 10k steps, then frozen forever.

BIGBRAIN → MRI · AFID PROTOCOL

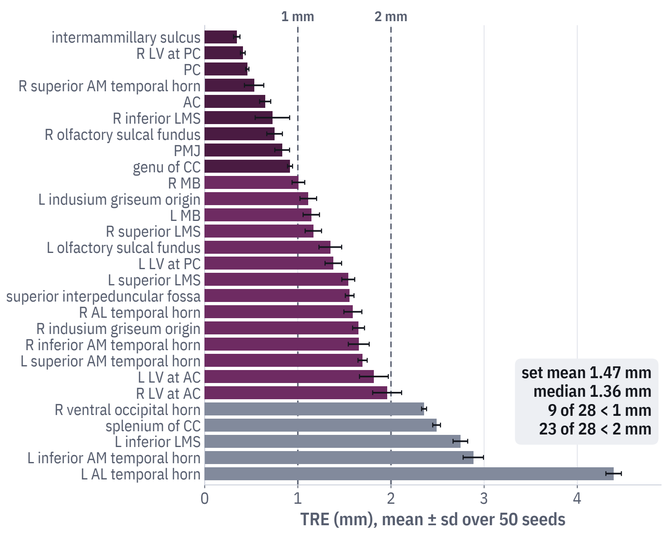
Lowest landmark error of 12 methods



Red tears tissue, teal is fold-free; the second number is the folded fraction. Same pair, grid and AFID slices for every row. † literature value on the 30-fiducial set (ours: 1.70 mm there).

PER-LANDMARK ERROR

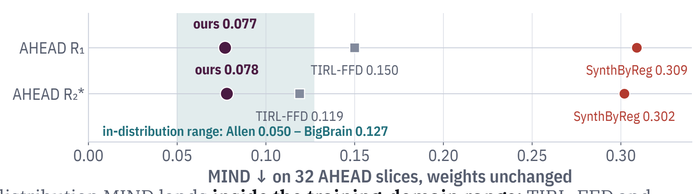
All 28 fiducials, 50 random seeds



Deep midline landmarks (AC, PC, intermammillary sulcus) land well under 1 mm; the tail is cortical-margin and temporal-horn points, worst the **L anterior temporal horn at 4.4 mm.** Error bars are the seed spread.

AHEAD · THIONIN → QUANTITATIVE MRI

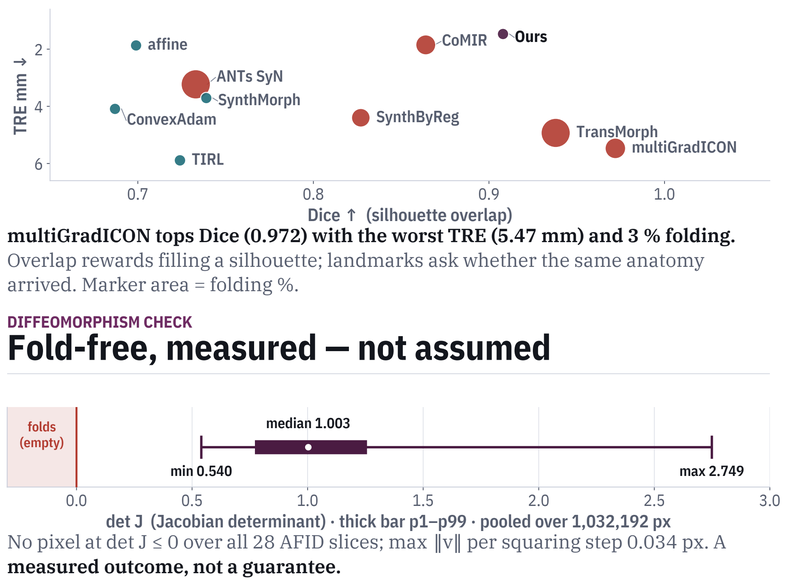
Zero-shot to brains never seen



Out-of-distribution MIND lands **inside the training-domain range;** TIRL-FFD and SynthByReg are 2–4× worse on the same slices (Wilcoxon $p < 10^{-6}$). NMI 1.247 / 1.240.

DICE VS LANDMARK ERROR · 10 METHODS

High Dice, wrong anatomy



multiGradICON tops Dice (0.972) with the worst TRE (5.47 mm) and 3 % folding. Overlap rewards filling a silhouette; landmarks ask whether the same anatomy arrived. Marker area = folding %.

DIFFEOMORPHISM CHECK

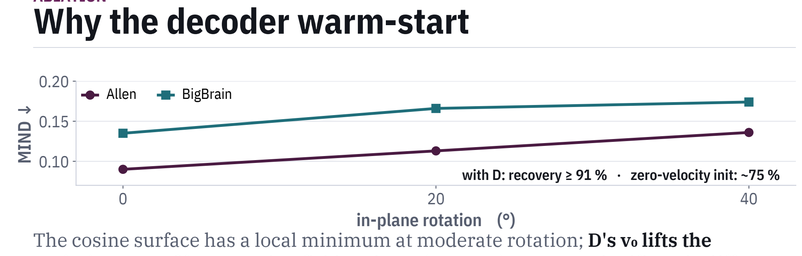
Fold-free, measured — not assumed



No pixel at det J ≤ 0 over all 28 AFID slices; max ‖v‖ per squaring step 0.034 px. A **measured outcome, not a guarantee.**

ABLATION

Why the decoder warm-start



The cosine surface has a local minimum at moderate rotation; **D's v_0 lifts the optimiser out of it.** Receptive-field variants (5–33 px) were not significantly different and are omitted.

TAKE HOME

1 Learn the metric, solve the warp.

A frozen 76k-parameter contrastive encoder gives the cost surface; Adam solves v for each pair in 0.9 s. Nothing about the deformation is amortised.

2 Accurate and fold-free at once.

1.47 mm mean TRE on 28 AFIDs — below the affine (1.87) and every deformable baseline — with 0.000 % folding measured at inference.

3 It transfers without retraining.

Zero-shot AHEAD MIND 0.077 lands between the two training sets; foundation models with 10^7 parameters do worse on this protocol.

HONEST LIMITS

- 2D per-slice: deformations are not coupled across sections; a volumetric extension with inter-slice regularisation is next.
- Presupposes an accurate global affine; the two posterior-fossa fiducials the affine misses stay out of reach.
- AHEAD evidence is image-similarity only; landmark tests rest on one BigBrain donor (within-subject).

SELECTED REFERENCES

Xiao 2019 — BigBrain–MNI AFIDs
Amunts 2013 — BigBrain
Casamitjana 2021 — Allen / SynthByReg
Pielawski 2020 — CoMIR
Tian 2024 — multiGradICON
Hoffmann 2022 — SynthMorph
Heinrich 2012 — MIND
Park 2020 — PatchNCE
Balakrishnan 2019 — VoxelMorph / VecInt
Rohlfing 2012 — overlap ≠ accuracy