ICPR 2026

IMAGE PROCESSING LABORATORY

HypDeformNet

Edge-Deployable Deep Architecture with Jacobian-Stable Hyperbolic Deformation and Lipschitz Distillation for Immunotherapy Response Prediction

https://iplab.dmi.unict.it/ai4ilf/#/home

F. Rundo[1] · M. O. Spata[1] · G. L. Banna[2] · S. Battiato[1]

[1] *University of Catania, Dept. of Mathematics and Computer Science*

[2] *Portsmouth Hospitals University NHS Trust & University of Portsmouth*

CLINICAL MOTIVATION

Immunotherapy response is hard to predict from pre-treatment CT

30% of metastatic urothelial carcinoma (mUC) patients present with metastatic disease; 5-year survival is under 15%.

20–30% of patients achieve Complete/Partial Response or Stable Disease under PD-1/PD-L1 blockade — the rest progress.

Limited biomarkers and conventional radiomics fail to capture tumor–immune spatial organization, and cohorts are small and imbalanced.

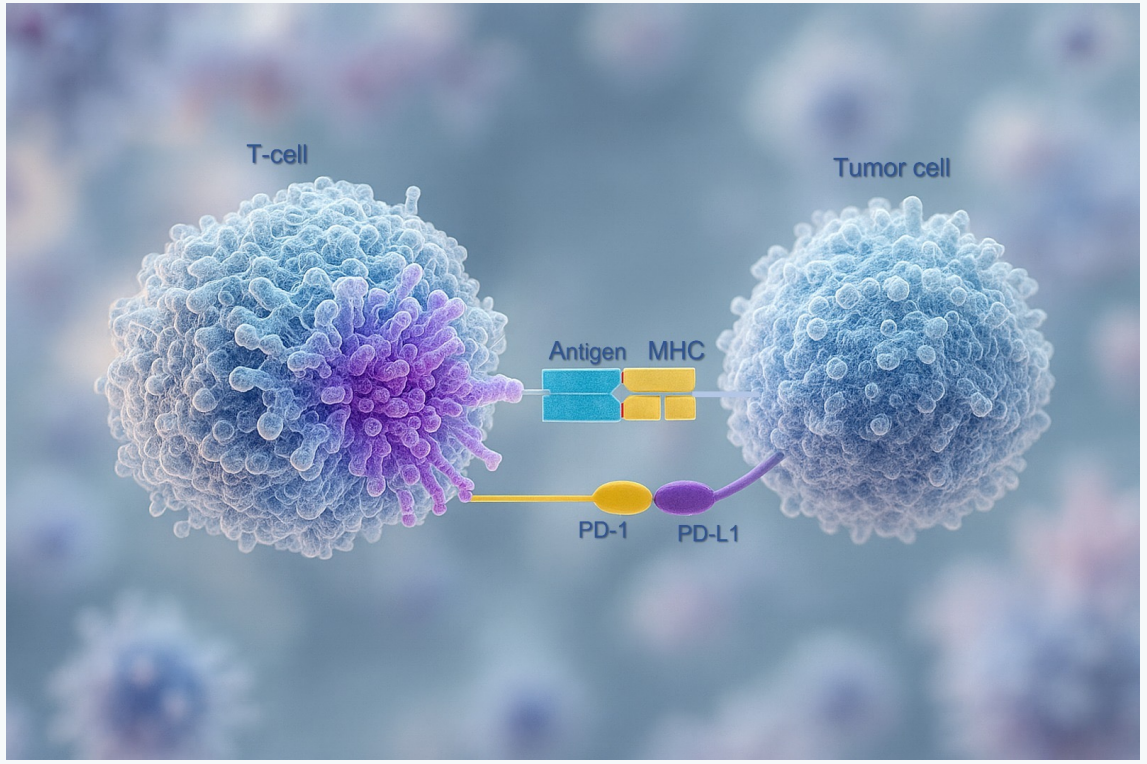


Fig. 1 — PD-1/PD-L1 checkpoint mechanism in tumor immune evasion

Goal: predict immunotherapy benefit directly from pre-treatment staging CT, enabling near real-time triage at acquisition time.

RELATED WORK

Prior approaches fall short of clinical-grade accuracy

Approach	Representative work	Reported performance
Radiomics (PET/MR)	Liebgott et al. — handcrafted features	94% accuracy
Weakly-supervised WSI	Ligero et al. — PD-L1 whole-slide images	AUC 0.88 (NSCLC), 0.80 (pan-cancer)
Multimodal fusion	Da-Ano et al. — PET/CT fusion	AUC 0.83
Spatio-temporal DL	Rundo et al. — densely connected non-local nets	93–94% accuracy
Multi-omics / genomics	Li et al., Yan et al., Wang	0.95 / 0.98 acc./AUC (exome features)
H&E whole-slide ICI	Rakaee et al.	AUC 0.75 (internal), 0.66 (external)

Common gaps: hand-crafted or Euclidean features that ignore the hierarchical, tree-like organization of tumor morphology; poor generalization on rare-cancer, small-cohort settings; and no path to on-device / real-time deployment.

PROPOSED METHOD

HypDeformNet: an end-to-end PD-L1 outcome prediction system

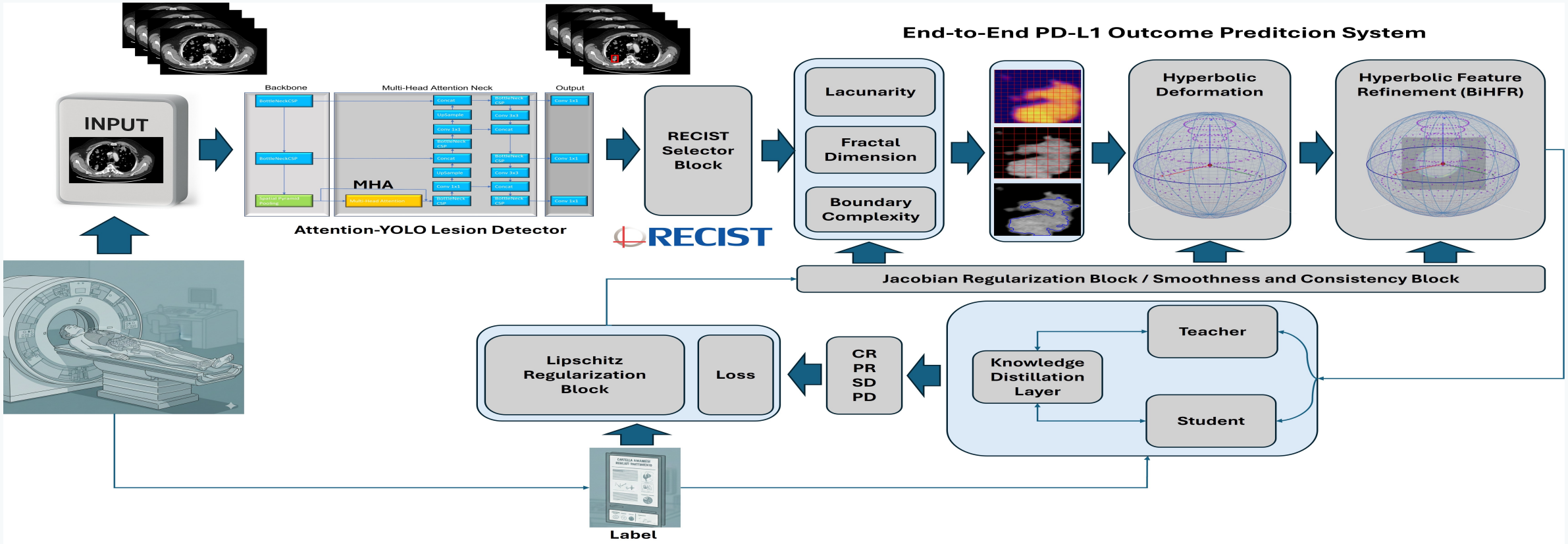


Fig. 2 — Overall HypDeformNet pipeline: CT input → Attention-YOLO lesion detection → morphology descriptors → hyperbolic deformation → BiHFR → Lipschitz-controlled teacher–student distillation → CR/PR/SD/PD outcome

Preprocess & detect | Describe morphology | Deform (hyperbolic) | Refine (BiHFR) | Distill (teacher→student)

STEP 1 · INPUT BLOCK

CT standardization and Attention-YOLO lesion detection

Geometric & intensity standardization

Each patient's CT volume V_n is resampled to a canonical grid via the voxel-to-world affine, using tri-linear interpolation:

$$\dot{S}(\tilde{x}) = \Sigma_u \; V(u) \cdot \Pi(1 - |\xi_d(\tilde{x}) - u_d|)$$

Intensities are windowed (μ_W, σ_W), clipped to (a,b), then normalized with quantiles Q_p1, Q_p2 and stabilizer ε into [0,1] before axial slicing.

Attention-YOLO Lesion Detector

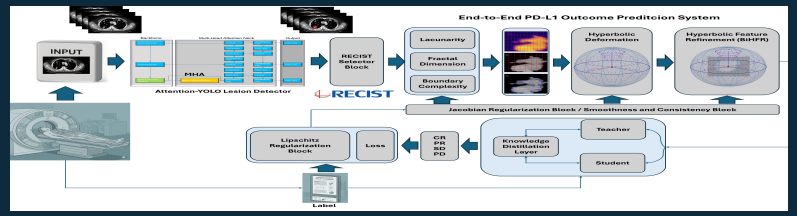
- YOLOv5s backbone extracts multi-scale features $F_{n,s,\ell}$ per axial slice.
- A Multi-Head Attention block is inserted before the feature-pyramid neck, so every spatial location attends to all others and suppresses spurious activations.
- Boxes are filtered with NMS (τ_nms = 0.5) and RECIST-1.1 admissibility, keeping the lesion of maximal physical diameter.
- **Detector trained on the institutional cohort + public DeepLesion dataset.**

92.7%

lesion-detection accuracy
(institutional cohort + DeepLesion)

$\tau\_nms = 0.5$

RECIST 1.1 admissibility filter
keeps the maximal-diameter target lesion



Quantifying lesion morphology to steer augmentation

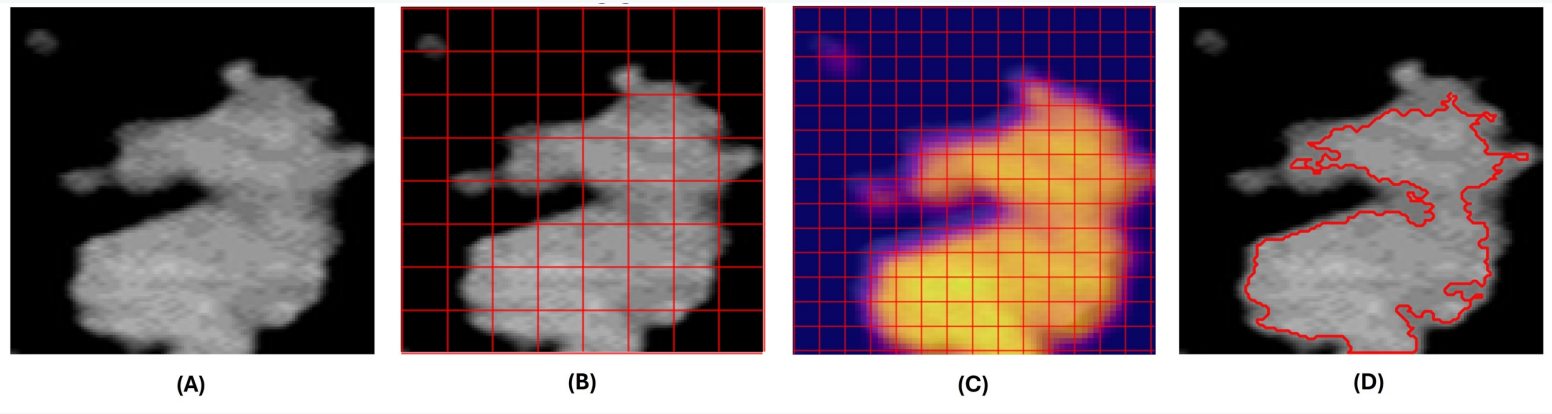


Fig. 3 — (A) RECIST lesion mask · (B) box-counting fractal grid · (C) lacunarity heat-map · (D) boundary-complexity contour

Fractal dimension D_n
$-\, d \log N_n(\varepsilon) / d \log \varepsilon$ from a differentiable box-counting surrogate: captures contour irregularity.

Lacunarity $\Lambda_n(\varepsilon)$
$\mathrm{Var}(\{m\_B\}) / (E[\{m\_B\}]+\varepsilon_0)^2$ over box masses: captures gap/texture heterogeneity.

Boundary complexity C_n
$P_n^2 / (4\pi A_n)$ from perimeter and area of the soft-mask level set: captures irregular rim shape.

Why it matters

The morphology descriptor $\mu_n \in \mathbb{R}^5$ (fractal dim., lacunarity, boundary complexity + intensity stats) conditions:

- Which hyperbolic deformation modes are sampled (Sec. 4.3)
- FiLM-style modulation inside BiHFR (Sec. 4.4)

Removing morphology conditioning drops patient-level accuracy from 90.9% to 85.0% (teacher).

Morphology-conditioned warping on the Poincaré disk

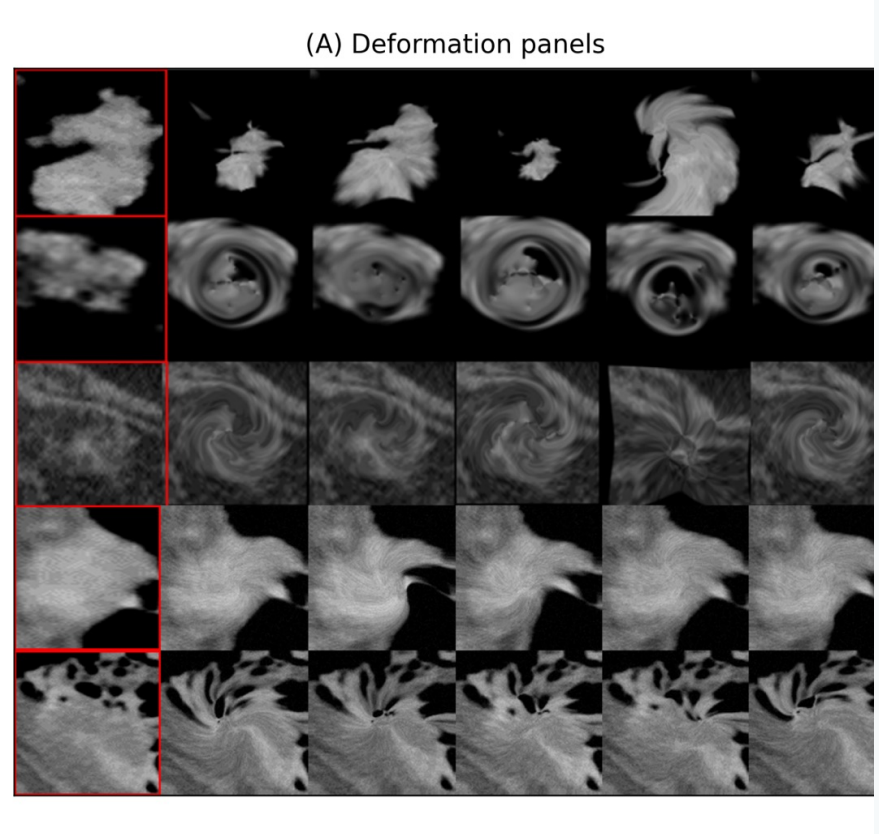


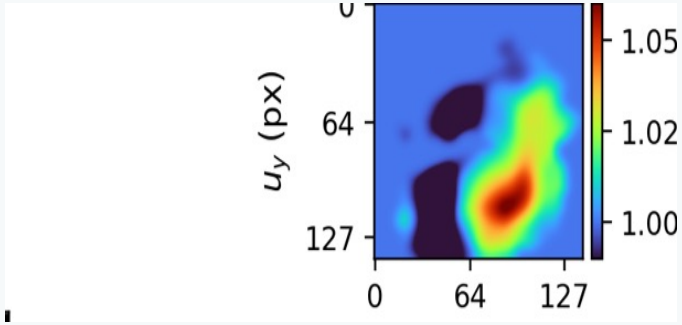
Fig. 4(A) — RECIST crop (red) with 5 morphology-conditioned deformations per lesion

Given lesion descriptor μ_n, a latent code ζ drives a near-diffeomorphic warp ϕ in the Poincaré ball $\mathbb{B}^2\_c$ via M=16 control points and a stationary velocity field:

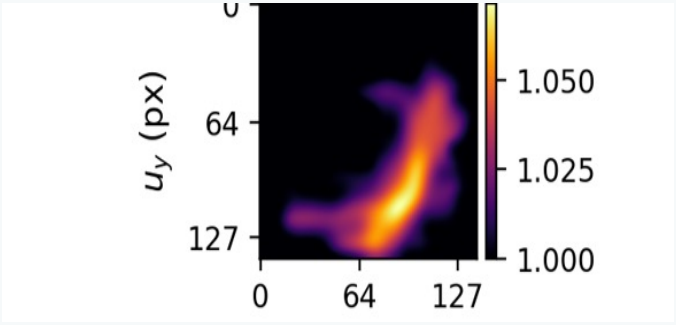
$$d\_c(p,q) = (1/\sqrt{c}) \cdot \operatorname{arcosh}(1 + 2c\|p-q\|^2 / ((1-c\|p\|^2)(1-c\|q\|^2)))$$

Jacobian regularization discourages foldings and volume blow-up:

$$L\_jac(\varphi) = \int (\log(J\_\varphi+\varepsilon))^2 du + \int [\max(0,-J\_\varphi)]^2 du$$



Jacobian determinant det(∇φ) — folding check



Local stretch σ_max(∇φ) — bounded anisotropic distortion

STEP 3 · FEATURE REFINEMENT

Biological-Inspired Hyperbolic Feature Refinement (BiHFR)

Each warped view is encoded, then every pixel is mapped to a center-to-periphery position on the Poincaré disk:

$$z(u) = \tanh(\sqrt{c}\cdot r(u)) \cdot \bar{u}/r(u) \in \mathbb{B}^2_c$$

Local self-attention is then restricted to hyperbolically-nearby pixels, mimicking biological center-surround receptive fields:

$$N(u) = \{u' : \chi(z(u), z(u')) < \rho\}$$

The morphology descriptor μ_n enters via FiLM modulation (γ_n, β_n), and a lightweight decoder predicts a bounded residual added back to the view.

Regularized for stability

- Feature-drift penalty (stay close to original features)
- Roughness penalty on the refined image gradients
- Cross-view consistency across the K deformation views

Ablation impact

- Without BiHFR refinement: 90.9% → 89.5% (teacher)
- Without its regularizer: 90.9% → 90.2% (teacher)
- *Moderate but consistent contribution to accuracy and robustness.*

STEP 4 · DISTILLATION

Online hyperbolic knowledge distillation with Lipschitz control

Teacher
ResNet-101

mutual distillation

Student
ResNet-18

Both backbones embed features into the Poincaré ball $\mathbb{B}^d_c$ and classify in tangent space:

$$\exp_0(v) = \tanh(\sqrt{c}\|v\|)\cdot v/(\sqrt{c}\|v\|) \qquad \ell = W\cdot \log_0(z) + b$$

Distillation loss (averaged over K views):

$$L_{KD} = (1-\alpha)[L_{CE}^T + L_{CE}^S] + \alpha T^2 \cdot SKL(p_T, p_S) + \lambda_h \|\log_0 z_T - \log_0 z_S\|^2 + \lambda_L \cdot L_{Lip}$$

Layer-wise Lipschitz penalty bounds sensitivity to view perturbations:

$$L_{Lip}(\theta) = \Sigma_\ell \max\{0, \sigma_{max}(W\ell) - \tau\ell\}^2$$

97.5%
of teacher accuracy
retained by the student

3.8×
fewer parameters
in the deployed student

K=500
morphology-guided views
averaged at inference

EXPERIMENTAL SETUP

Dataset, cohort composition, and validation protocol

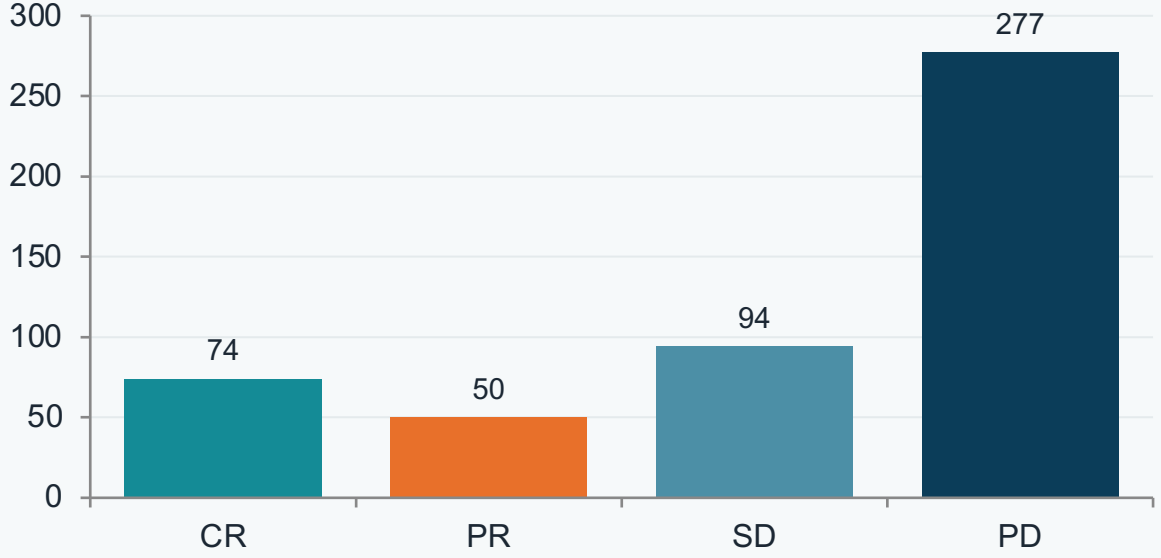
44
mUC patients
(IRB Catania 1)

495
RECIST target lesions

3-fold
patient-stratified
cross-validation

4
outcome classes
CR / PR / SD / PD

Lesion label distribution (imbalanced)

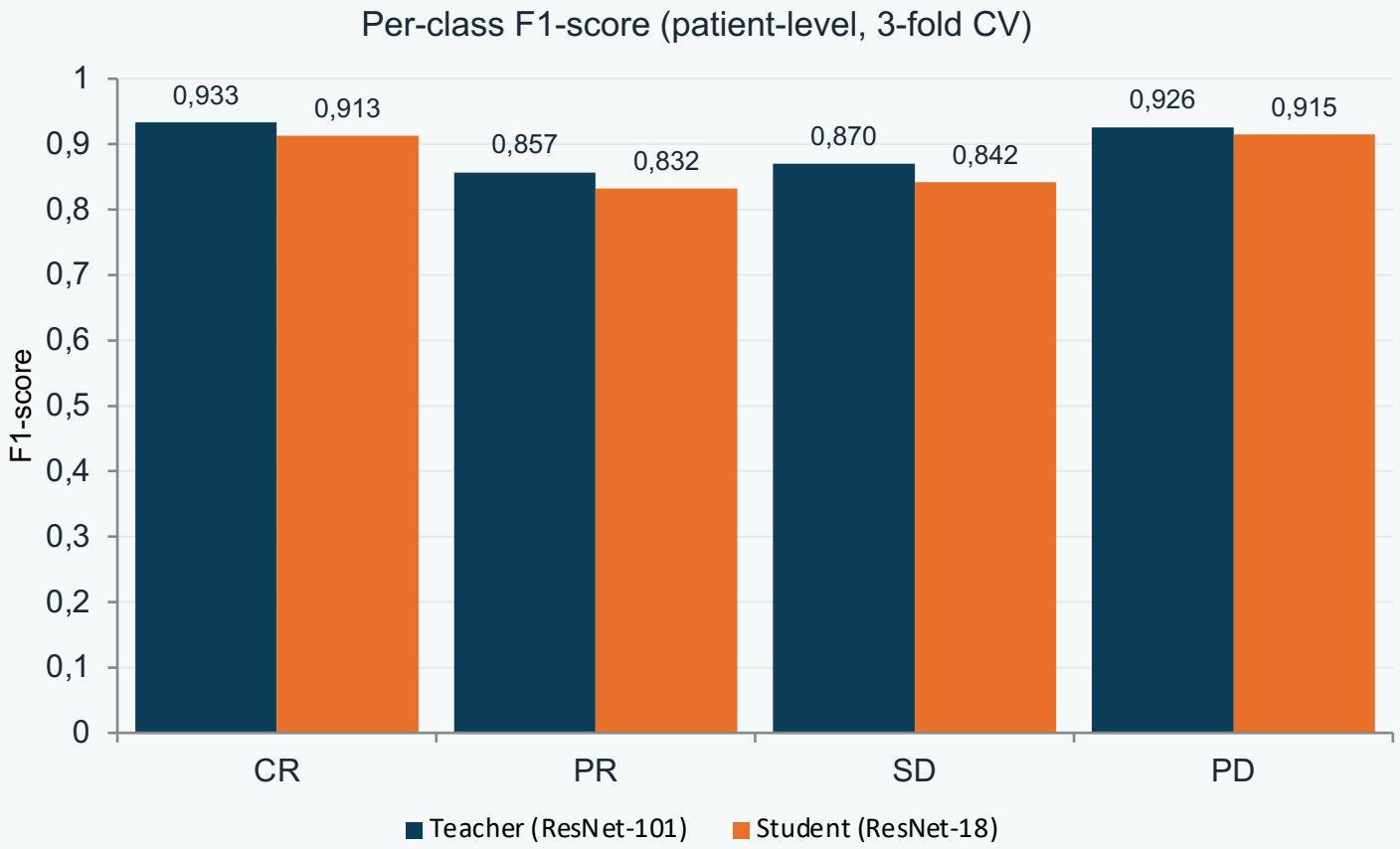


Protocol highlights

- Strict patient-level partitioning — no lesion leakage across folds.
- K = 500 morphology-guided hyperbolic deformations per training lesion (curvature c = 0.1).
- SGD, momentum 0.9; LR 1e-2 (teacher) / 5e-3 (student); cosine annealing, 100 epochs, batch 16.
- Distillation temperature T = 4.0, α = 0.5; Lipschitz bound τℓ = 1.5.
- Primary metric: patient-level accuracy (majority vote); lesion-level = upper bound.

RESULTS

Patient-level performance: teacher vs. distilled student



Teacher — patient-level

90.9%

95% CI [84.3%, 97.5%] · macro-F1 0.897

Student — patient-level

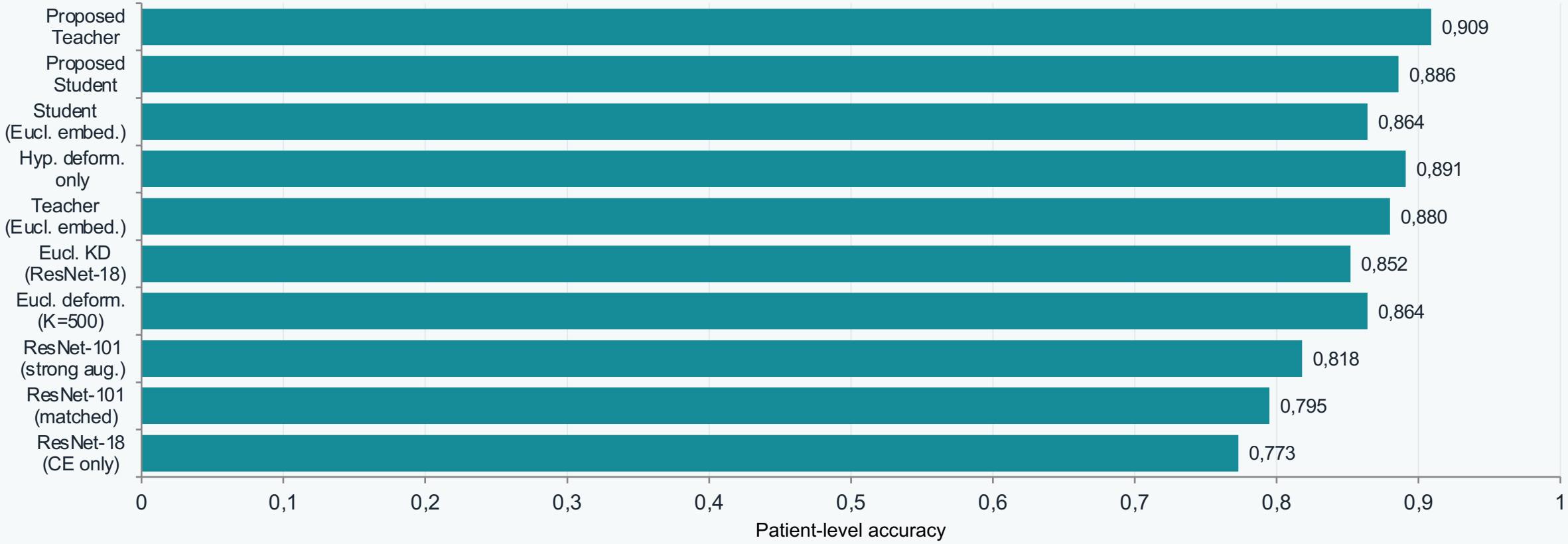
88.6%

95% CI [82.1%, 95.1%] · macro-F1 0.877

No cross-extreme (CR ↔ PD) misclassifications — model preserves RECIST response ordering.

RESULTS

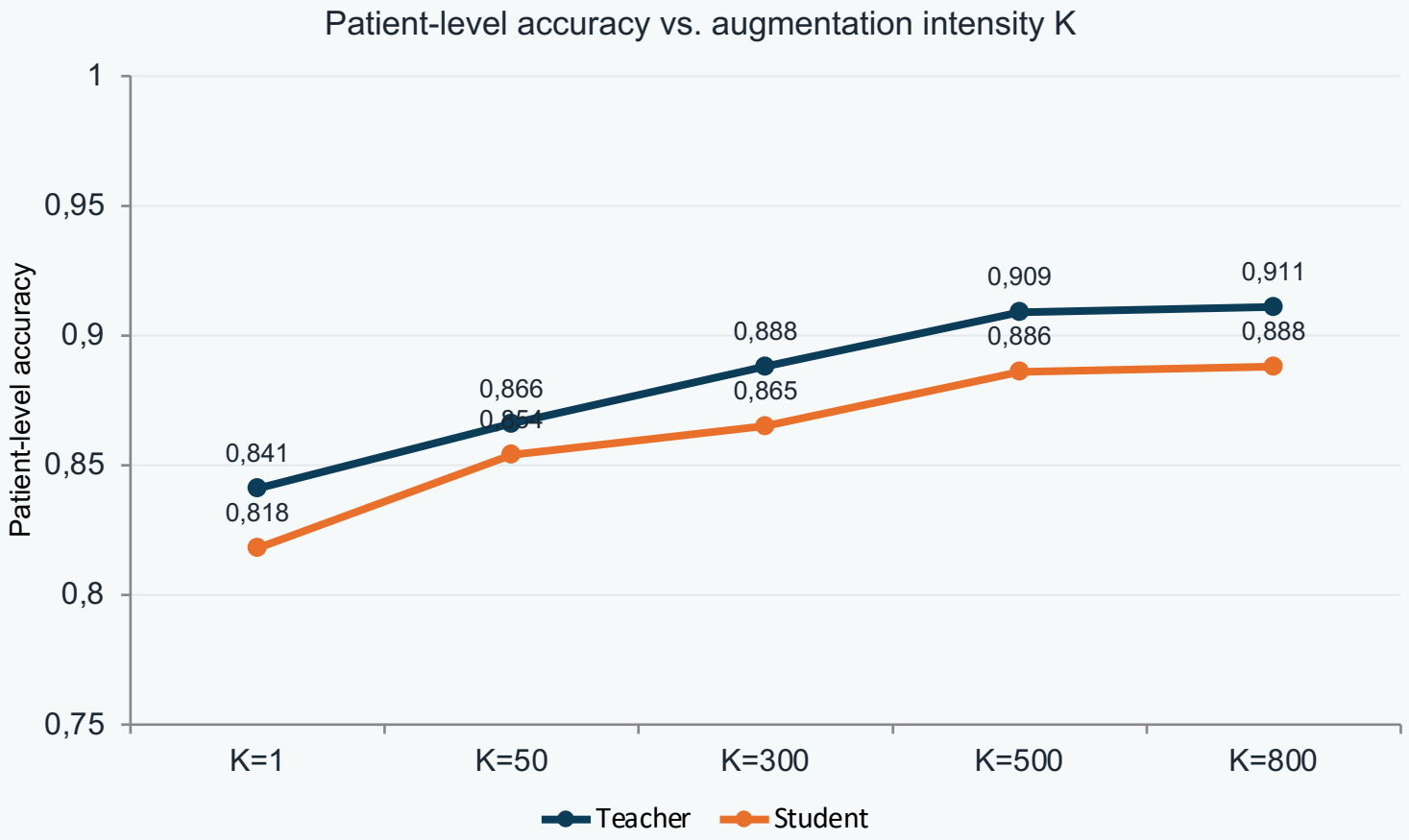
HypDeformNet outperforms protocol-matched baselines



Full pipeline: +4.5% over best protocol-matched baseline (p=0.012, teacher); hyperbolic embedding alone contributes +2.9% over Euclidean embedding. Best literature baseline overall: SwinTransformer-Tiny at 84.3% — still 6.6 pts below the proposed teacher.

ABLATION STUDY

Augmentation intensity K: more morphology-guided views help — up to a point

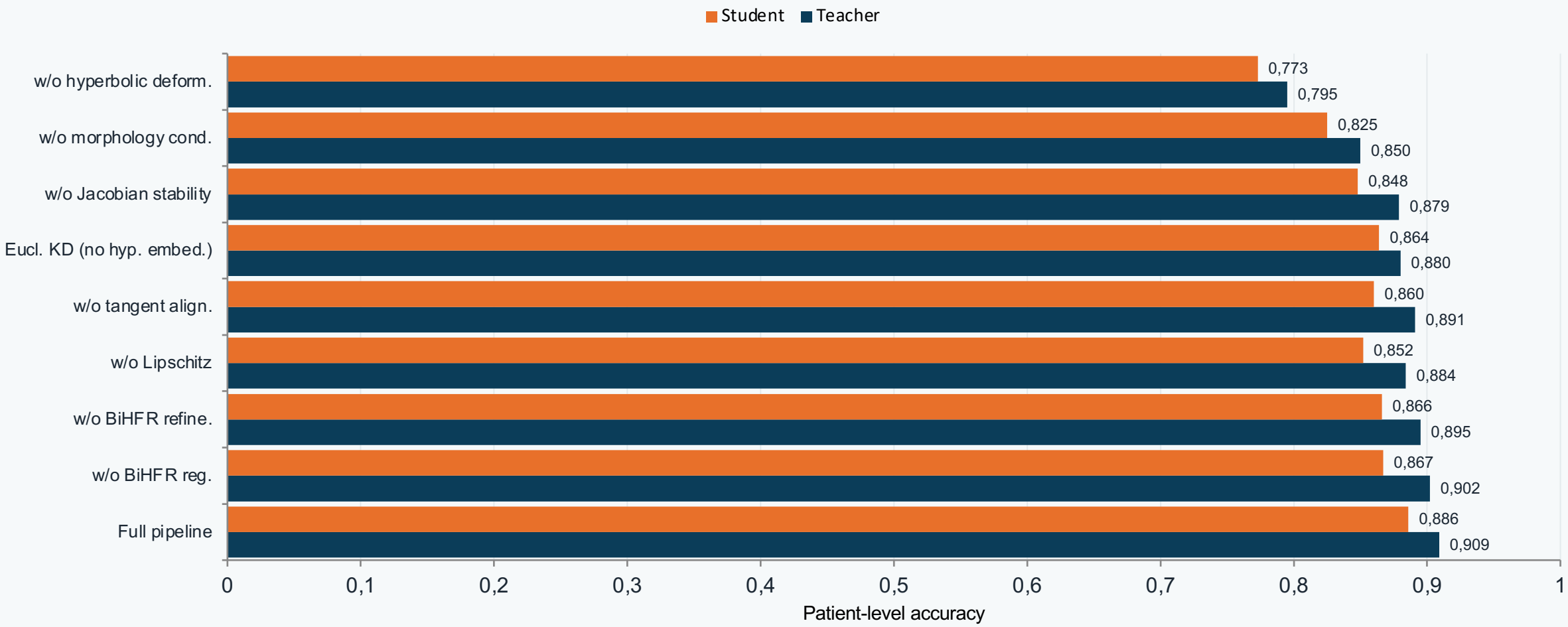


Key takeaway

- Accuracy rises monotonically from K=1 (no augmentation) to K=500.
- **K=1 → K=500: +6.8 pts (teacher), +6.8 pts (student)**
- K=500 → K=800 gives only marginal gains (+0.2 pt teacher, +0.2 pt student) at 60% higher augmentation cost.
- *K=500 is selected as the optimal performance/cost trade-off used throughout the paper.*

ABLATION STUDY

Every component contributes — hyperbolic deformation matters most



DEPLOYMENT

From cloud-scale teacher to on-scanner, real-time inference

11.7 MB	0.29 GMACs	0.6–1.2 s	0.2–0.7 J
INT8 model footprint (SRAM-feasible)	compute cost per inference	latency at 480 MHz on STM32H7-class MCU	energy per inference

The distilled ResNet-18 student — quantized to INT8 — is deployed directly on an STM32H7-class microcontroller, taking a 128×128 RECIST crop and outputting a 4-class prediction (CR/PR/SD/PD) with **activation buffers small enough for on-chip SRAM.** This makes near real-time, point-of-care immunotherapy response assessment feasible directly at the CT scanner — no cloud round-trip required.

Prior art

- Rundo F, Banna GL, Trenta F, Battiato S. Advanced deep network with attention and genetic-driven reinforcement learning layer for an efficient cancer treatment outcome prediction. In: Proc. IEEE Int. Conf. Image Process. (ICIP). 2021. p. 294–298. doi: 10.1109/ICIP42928.2021.9506620.

- F. Rundo, M.O. Spata, A. Calvagna, A. Caruso, E. Tramontana, S. Battiato, "HYPERNEST-TTA: HYPERBOLIC NESTED LEARNING WITH TEST-TIME ADAPTATION FOR DIABETIC RETINOPATHY ASSESSMENT", Accepted paper to the IEEE International Conference on Image Processing 2026 (ICIP, 2026) , 13-17 September 2026, Tampere, Finland

- F, Rundo, M. O. Spata, C. Pino, S. Battiato, "Dopamine-Inspired Neuro-Modulated Radiomic Pipeline for Computed-Tomography-Driven Prediction of Immunotherapy Response in Metastatic Cancer Treatment' , IEEE CVF/Proceedings of the "PHAROS - Adaptation, Fairness, Explainability in AI Medical Imaging (PHAROS-AFE-AIMI) and Competition" Workshop in conjunction with the International Conference on Computer Vision (ICCV), 2025 in Hawaii, USA, 19 – 23 October, 2025

- A. Rondinella, F. Guarnera, F. Rundo, S. Battiato "DeepSeg-MS: A 3D network based on hierarchical multi-scale learning for MRI multiple sclerosis lesion segmentation" , IEEE CVF/Proceedings of the "PHAROS - Adaptation, Fairness, Explainability in AI Medical Imaging (PHAROS-AFE-AIMI) and Competition" Workshop in conjunction with the International Conference on Computer Vision (ICCV), 2025 in Hawaii, USA, 19 – 23 October, 2025

CONCLUSION & FUTURE WORK

HypDeformNet: geometry-aware, edge-ready immunotherapy prediction

- Morphology-conditioned hyperbolic deformation + Jacobian stability turns a 44-patient cohort into a robust training signal.
- Biological-inspired hyperbolic feature refinement (BiHFR) amplifies clinically relevant rim/core and boundary cues.
- Lipschitz-controlled online distillation compresses a 90.9%-accurate teacher into an 88.6%-accurate, 3.8× smaller student.
- The student is MCU-feasible (11.7 MB, INT8), enabling real-time on-scanner deployment.

Limitations & future work: single-center, 44-patient cohort — next steps are multi-center validation and robustness testing across clinical settings.

Headline numbers

90.9%
teacher accuracy

88.6%
student accuracy

97.5%
teacher perf. retained

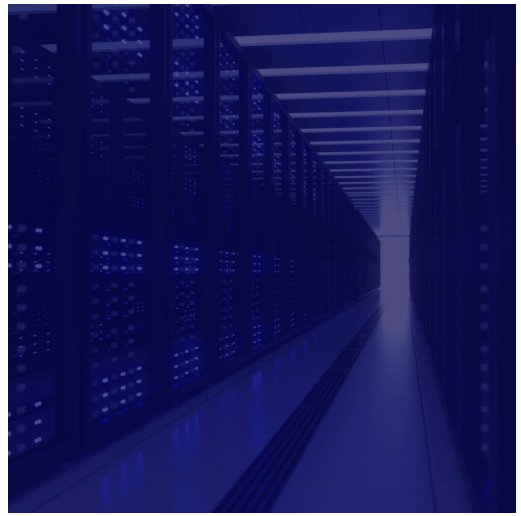
3.8×
parameter reduction

p=0.012
vs. best baseline

Thank You

Questions & Discussion

https://iplab.dmi.unict.it/ai4ilf/#/home



Sebastiano Battiato

sebastiano.battiato@unict.it

Dipartimento di Matematica e Informatica, University of Catania