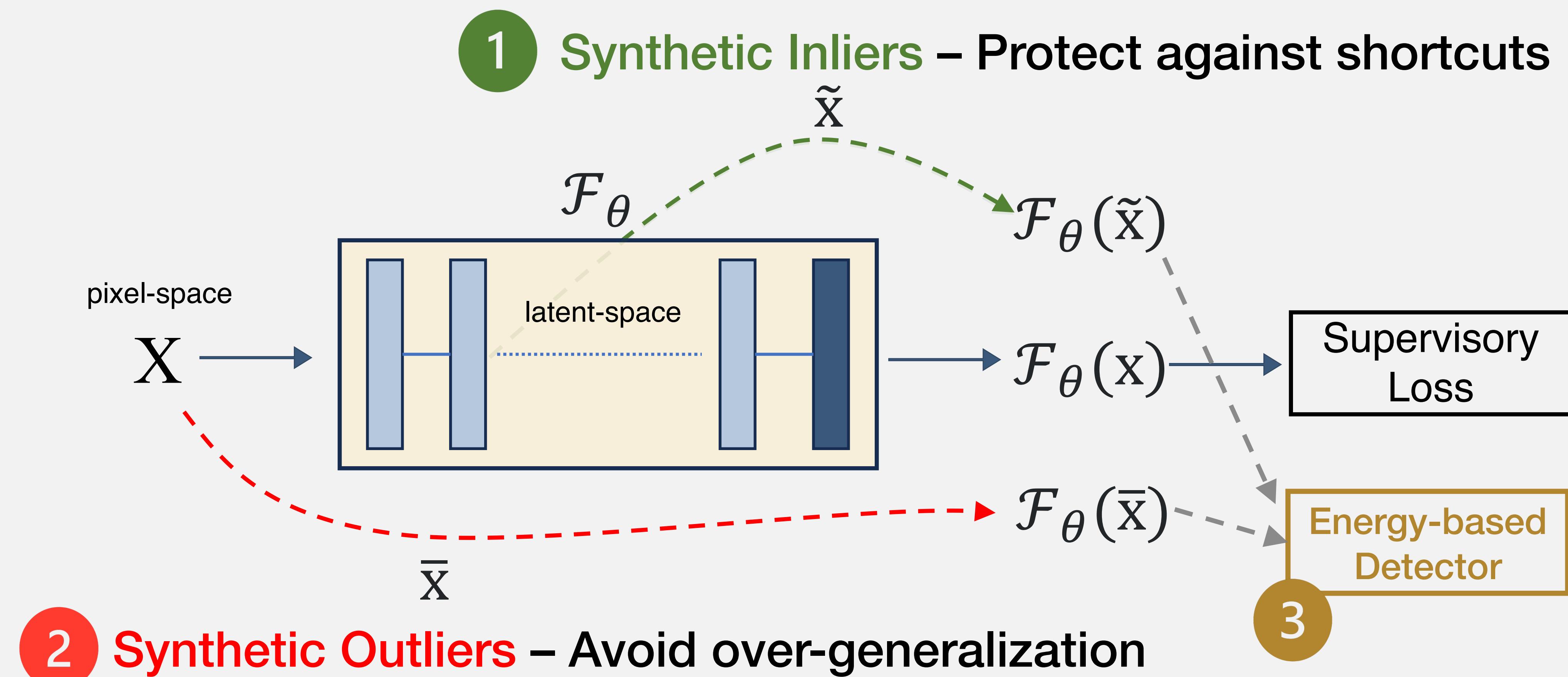




Vivek Narayanaswamy*, Yamen Mubarka*, Rushil Anirudh, Deepta Rajan, Andreas Spanias, Jayaraman J. Thiagarajan

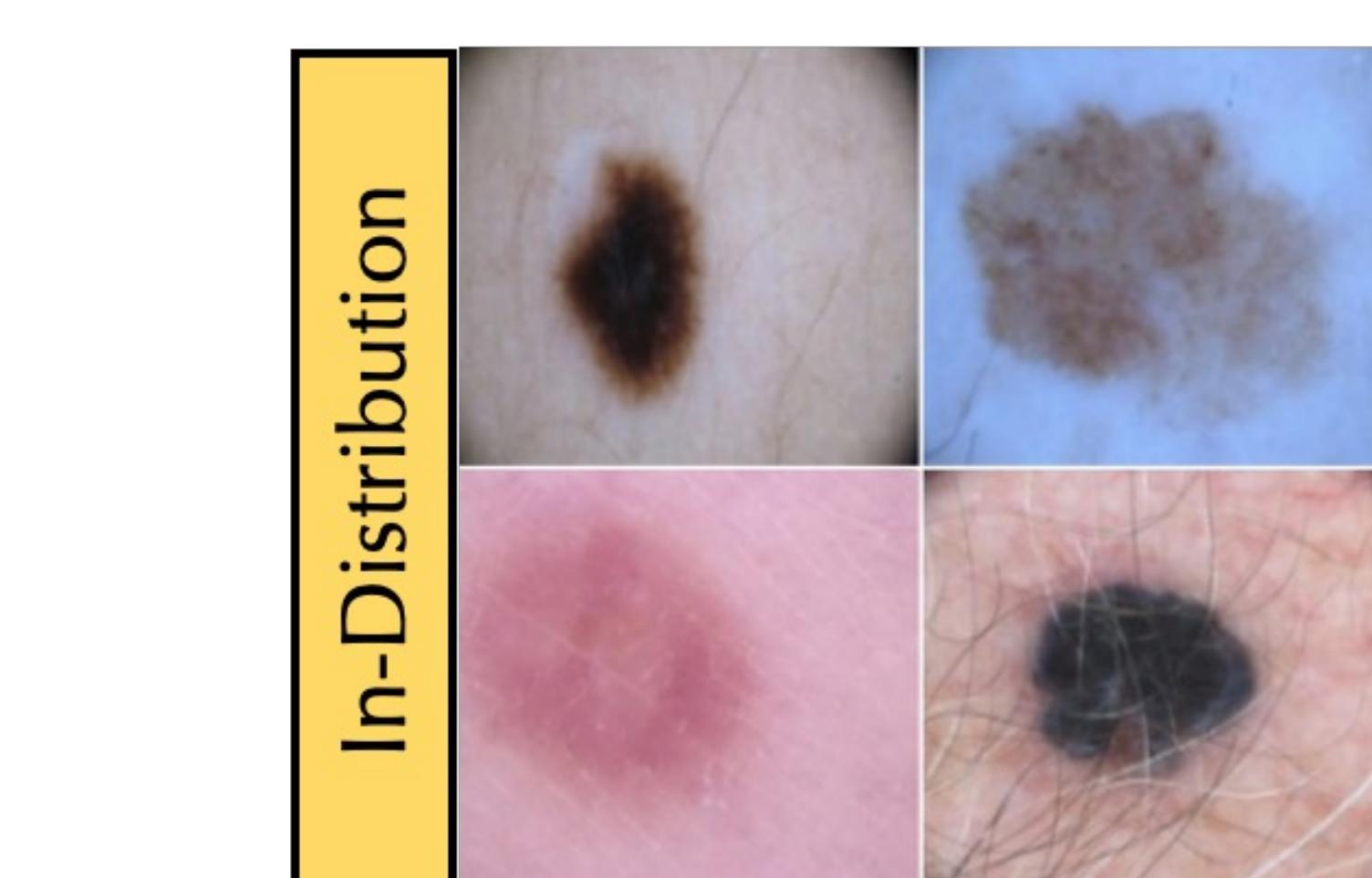
Synthetic Data Generation to Enable Open-Set Recognition without Hurting ID Accuracy



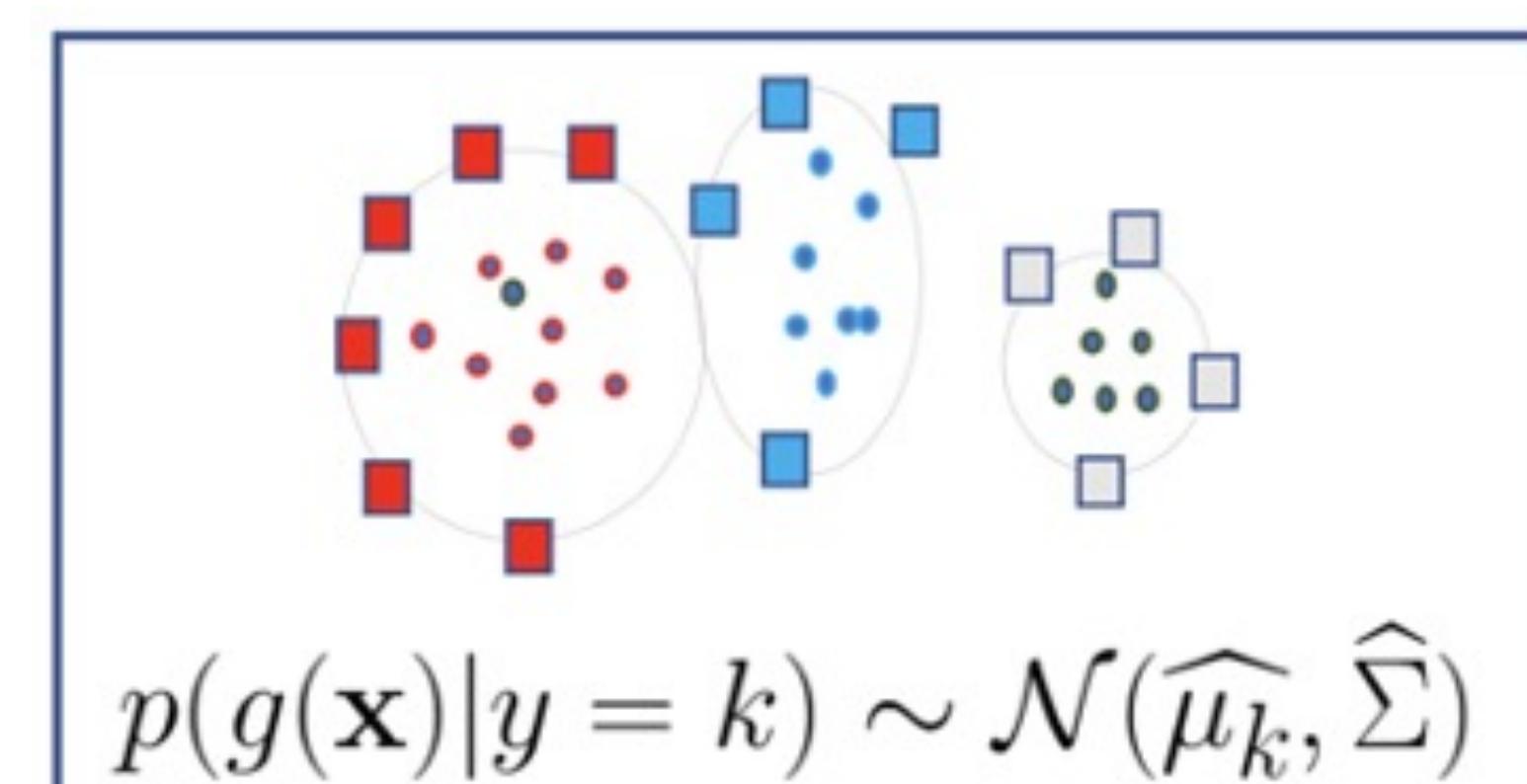
- 1 Synthesize inliers in the latent space**
Sample low-likelihood regions from class-specific feature distributions
- 2 Synthesize outliers in the pixel space**
Highly diverse set of outliers to ensure the OOD subspace does not overlap with the ID subspace
- 3 Train energy-based OOD detector**
Margin-based losses to calibrate the OOD detector to accept synthetic inliers and reject outliers

Open-Set Recognition

SoTA OOD calibration methods fail on medical open-set recognition!

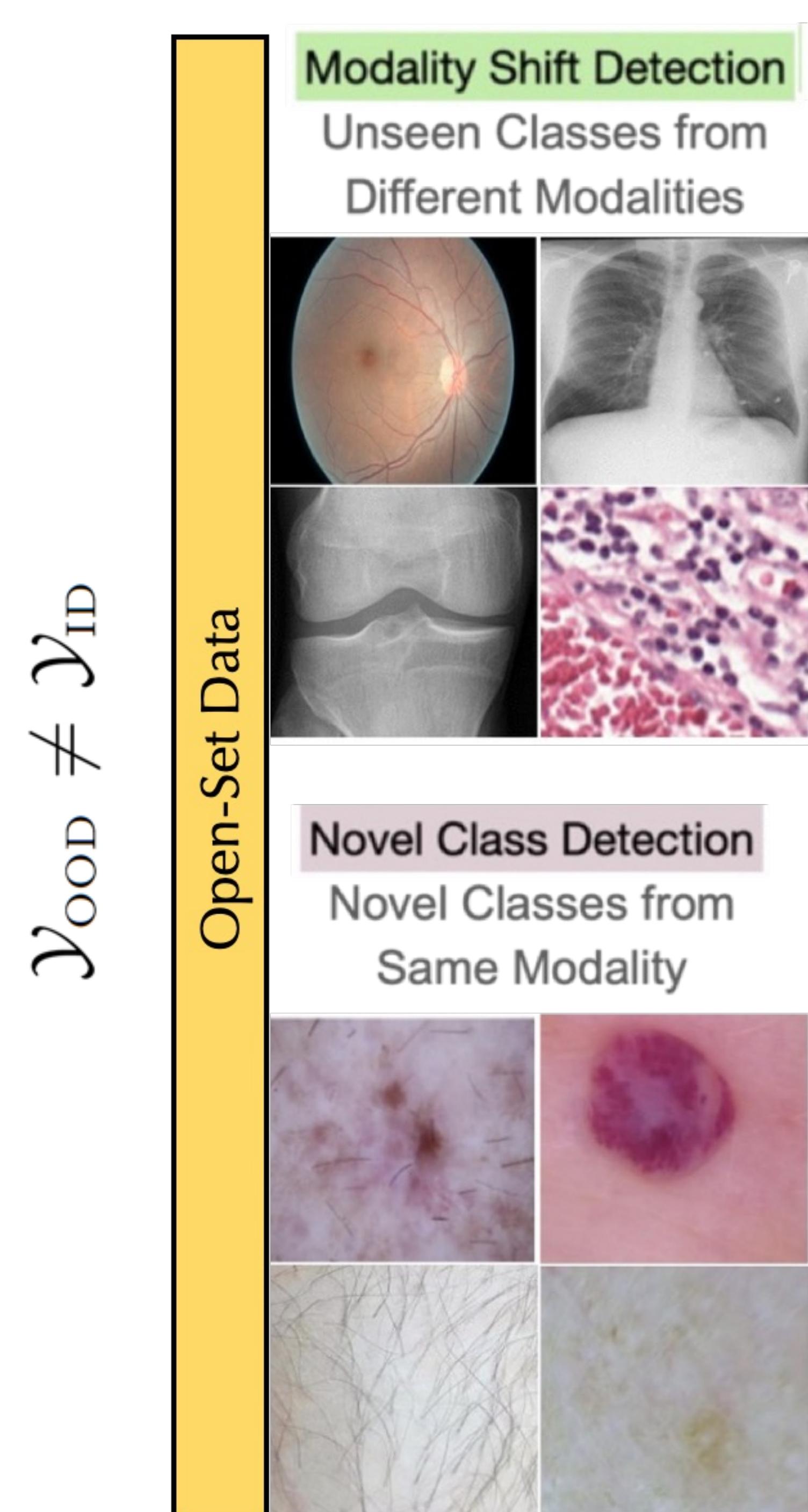
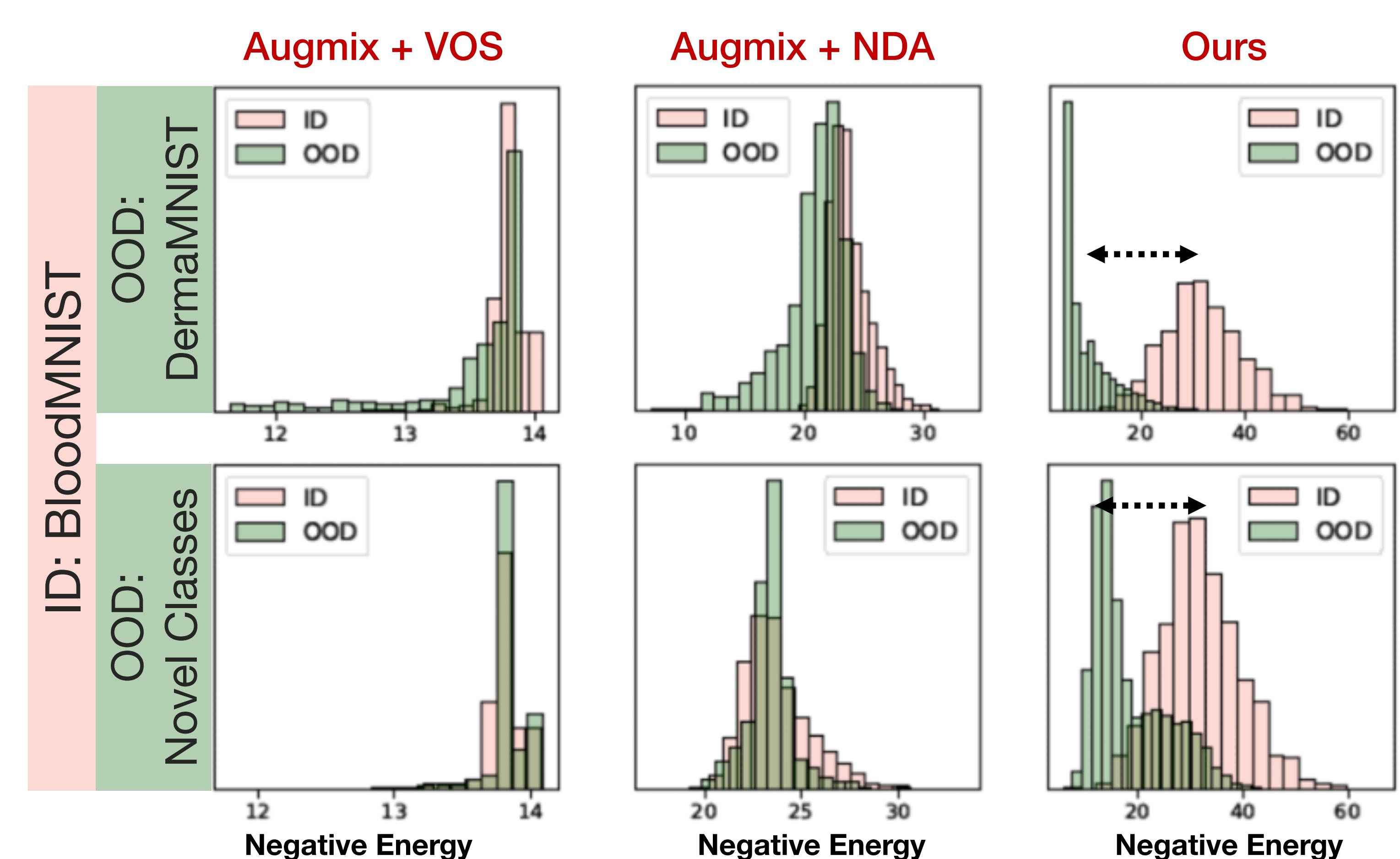


Inlier Synthesis

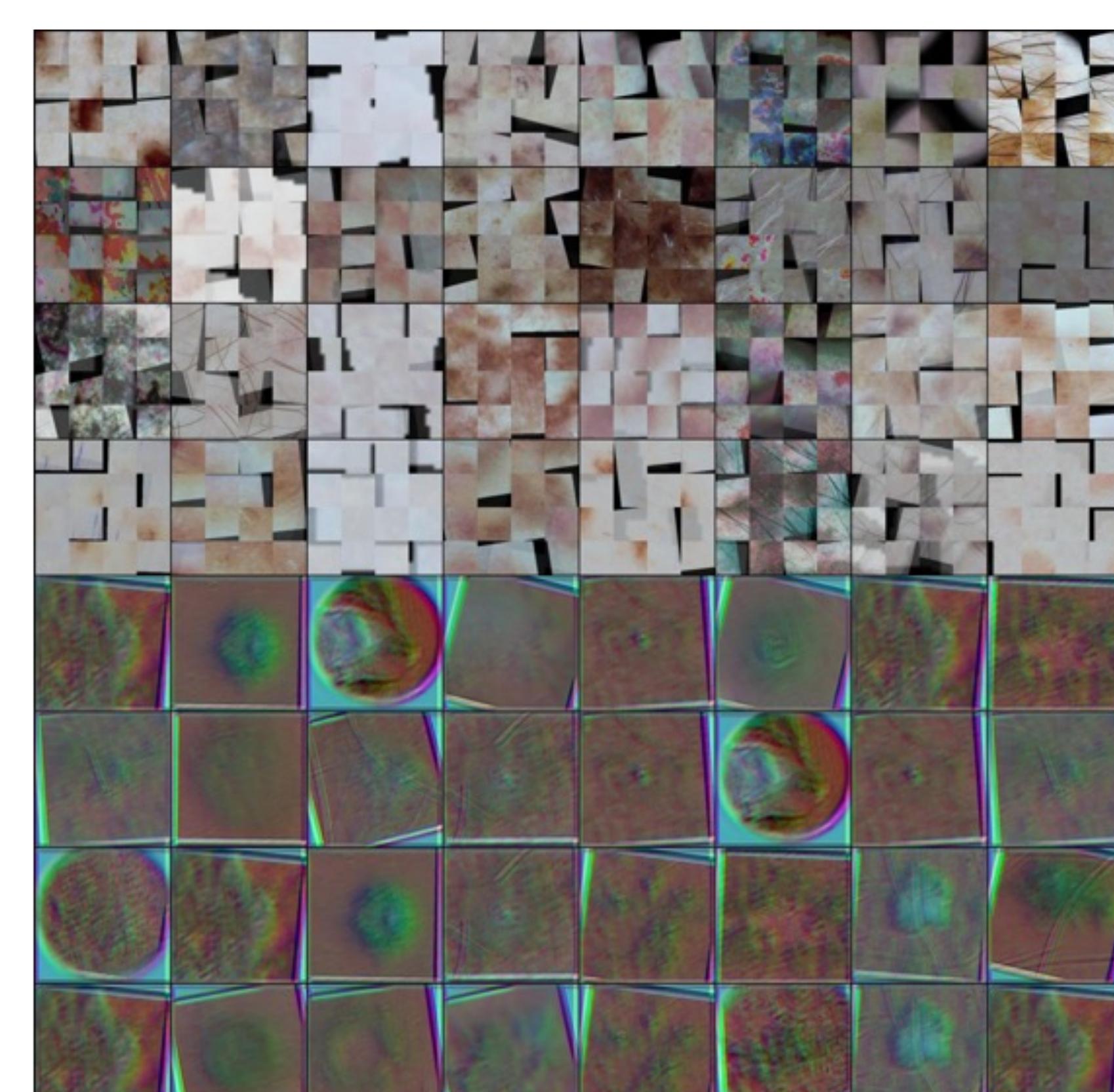


Push the tail samples closer to the class-specific prototypes

Results



Outlier Synthesis



High-severity compositional image manipulations
(e.g., Augmix, RandConv)

OOD Rejection AUROC (%)	Modality Shift		Novel Class	
	OCTMNIST TRAIN	TissueMNIST TEST	PathMNIST TRAIN	DermaMNIST TEST
G-ODIN	ID	OOD	ID	OOD
Augmix + VOS	49.7	75.4	53.9	38.2
Augmix + NDA	62.0	39.8	43.5	53.5
Ours	97.5	98.1	89.1	

Across a large suite of benchmarks, we achieve 15%-25% AUROC improvement over SoTA methods.