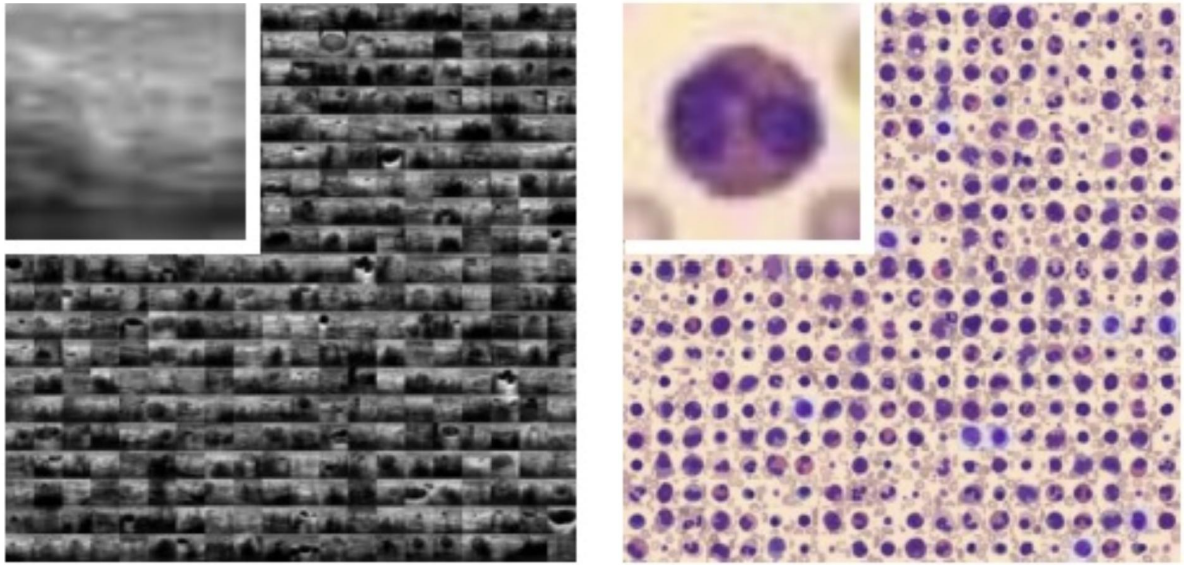
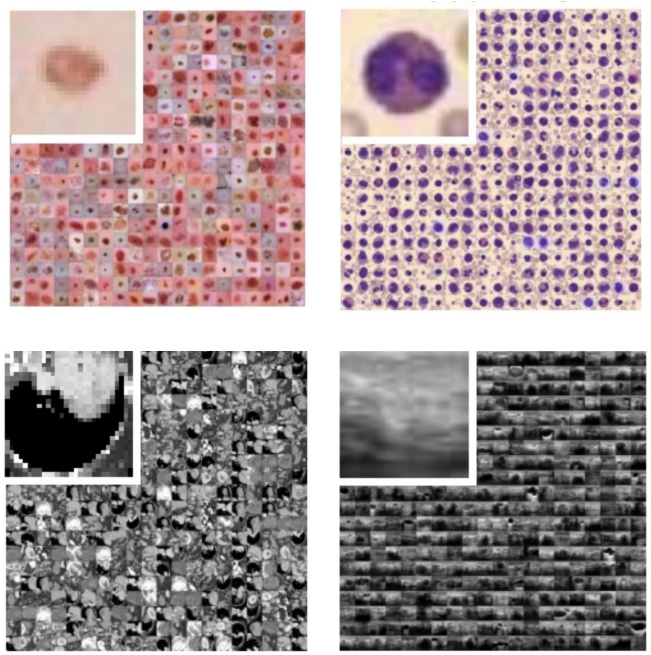
Exploring Minimally Sufficient Representation in Active Learning through Label-Irrelevant Patch Augmentation

Zhiyu Xue, Yinlong Dai, **Qi Lei**



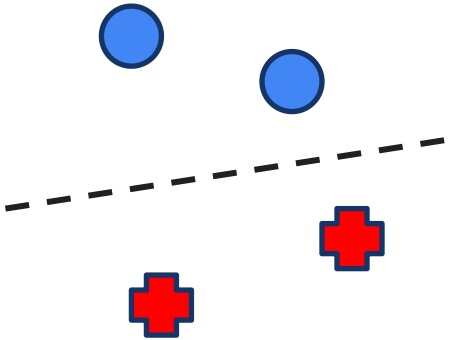
Annotating medical images **- costly!**
time-consuming!

Active learning to the rescue!



x

Feature representation

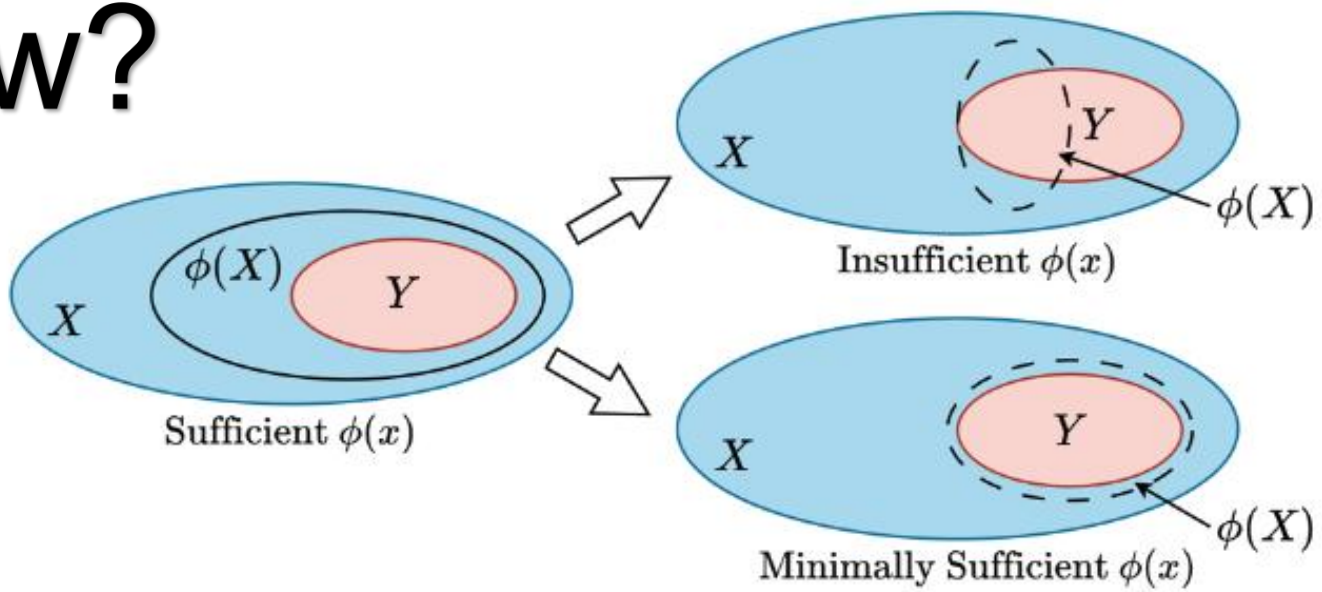


$\phi(x)$

Ideal ϕ – **linearly separable**

(computationally+sample) efficiently learned

How?



minimal: reduce irrelevant features to y

sufficient: keep relevant features to y

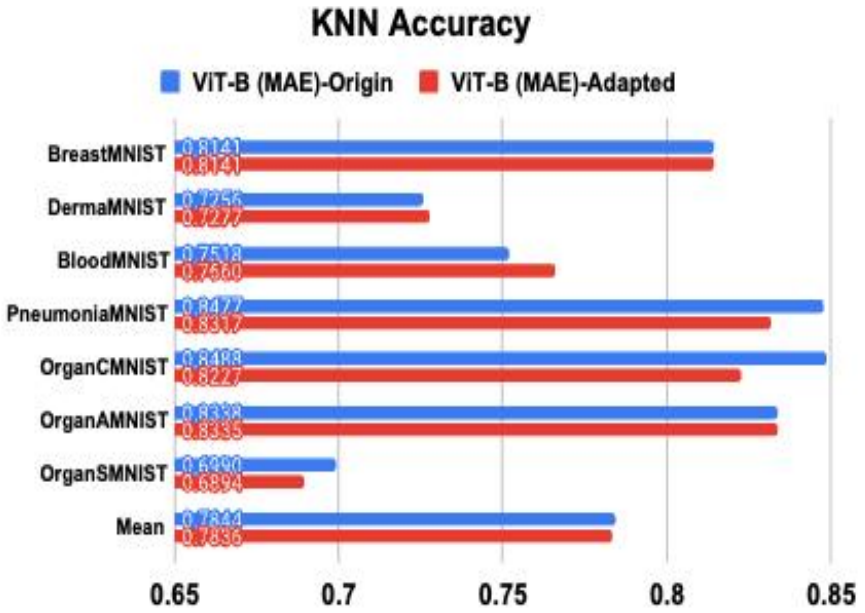
linearly separable

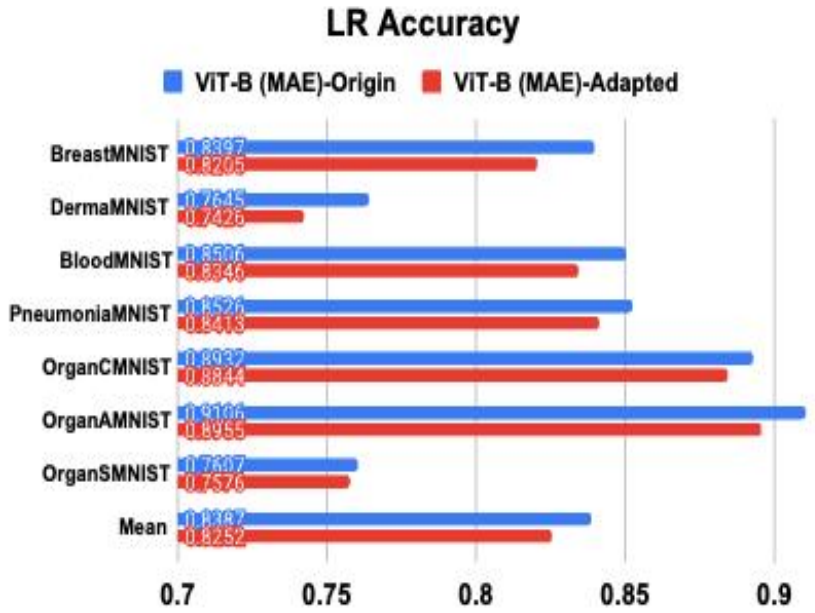
sample efficient: initially off-the-shelf model + finetuned by a few labeled samples

Standing point of our framework

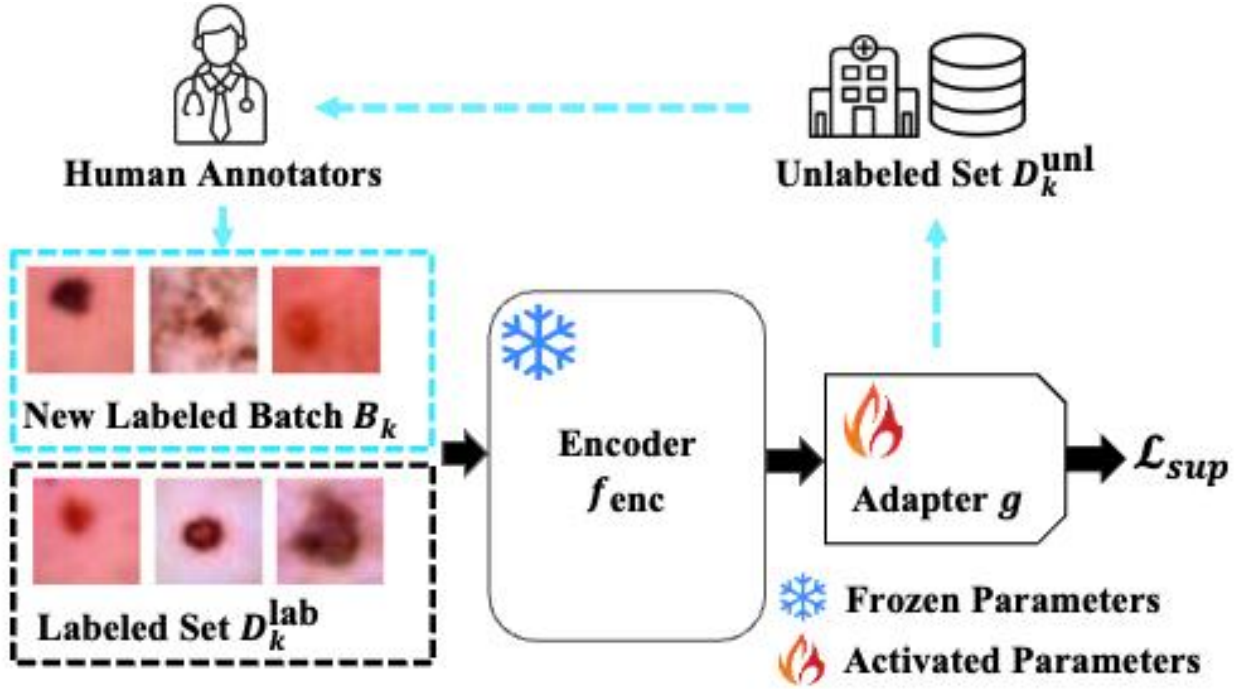
Off-the-shelf ViTs is good feature representation for starters

---- Adapting to specific dataset hurts the performance!



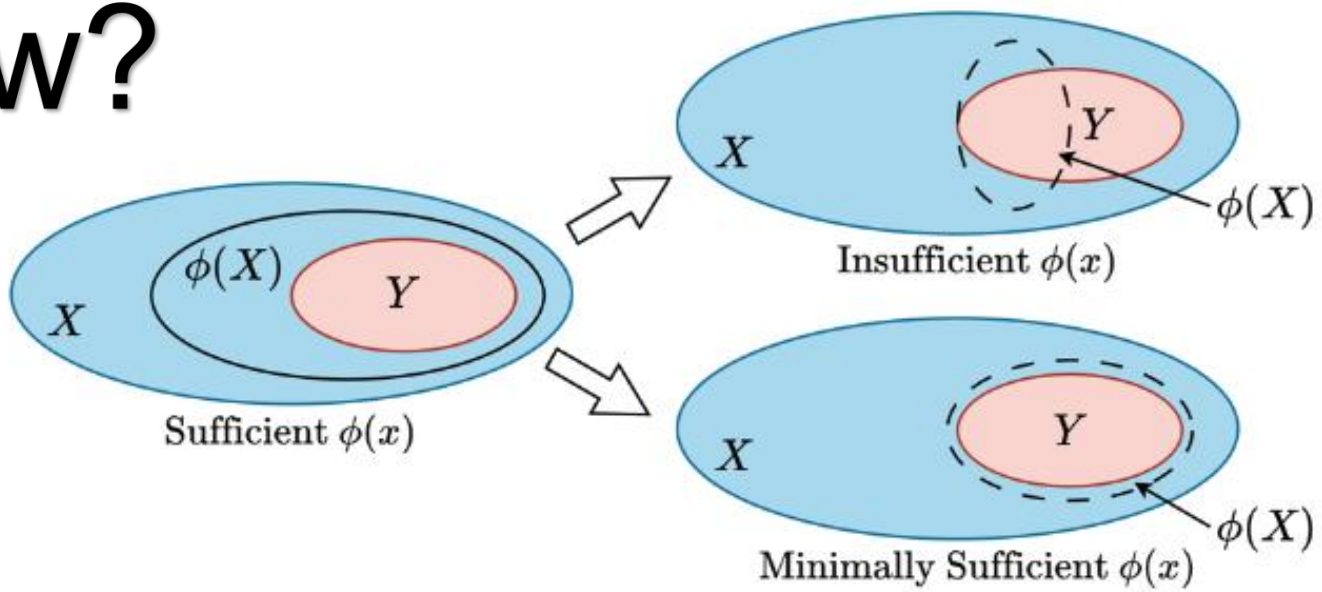


Parameter-efficient AL on Off-the-shelf ViT



Lightweight adapter: $\phi(x) = g(f_{enc}(x))$, update feature representation through g

How?



minimal: reduce irrelevant features to y

sufficient: keep relevant features to y

→ **linearly separable**

sample efficient: initially off-the-shelf model + finetuned by a few labeled samples

Label-Irrelevant Data Augmentation

minimal: reduce irrelevant features to y

➔ diverse and strong data augmentation

sufficient: keep relevant features to y

➔ label-irrelevant data augmentation

Step 1. **Label-Irrelevant Patches Localization***.

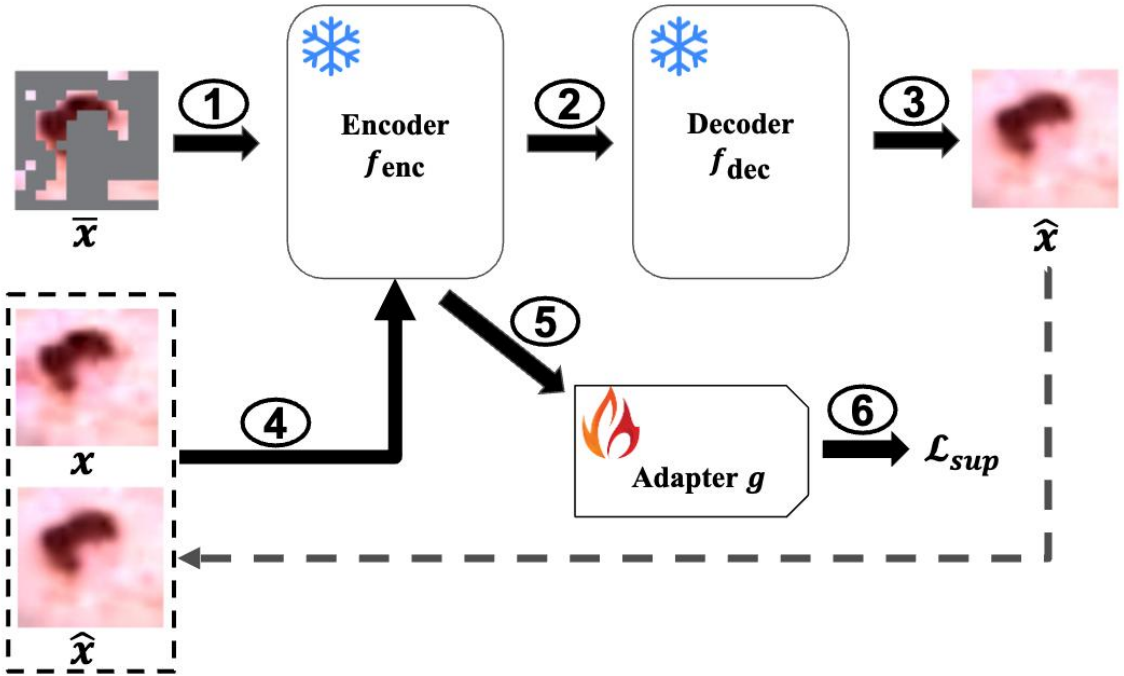
Step 2. **Label-Irrelevant Patches Augmentation:**

our proposals **--** SelfPatchAug and SubstitutivePatchAug

*LastAttentionMap[1], CosineAttentionMap[2], Saliency[3], DeepLIFT[4]

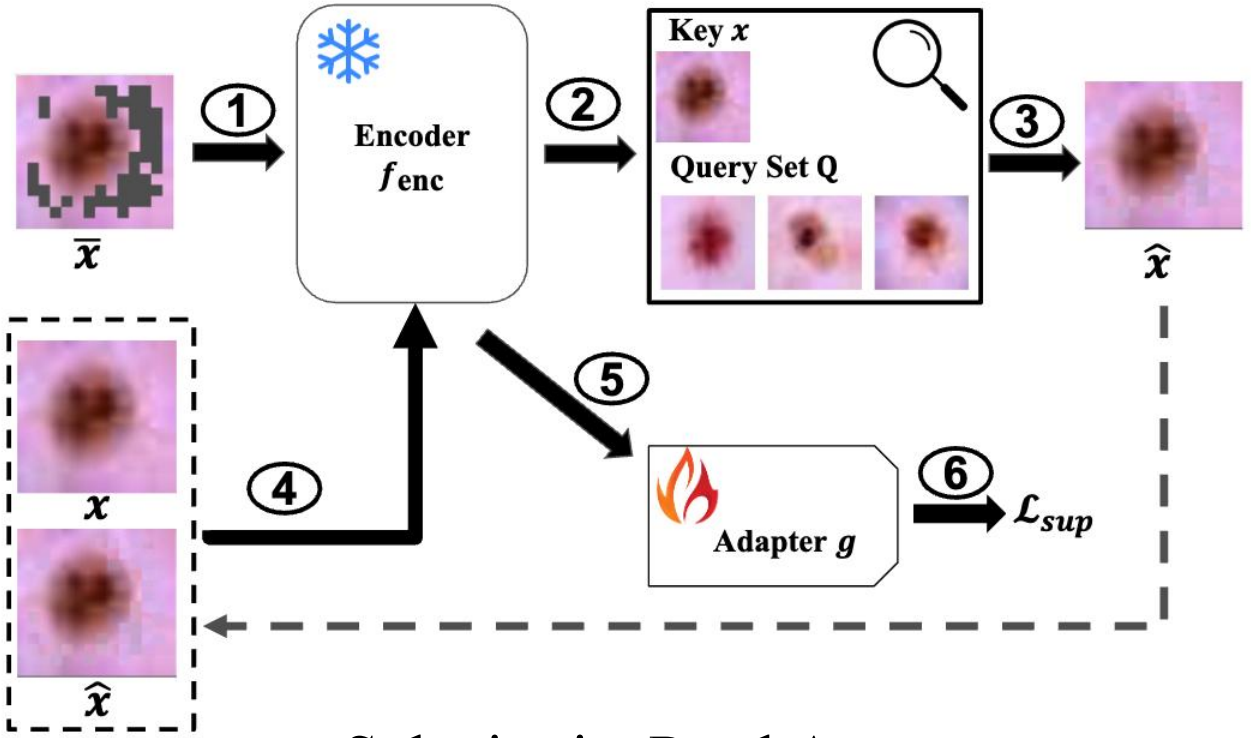
[1] Chefer et al. 2021. [2] Xu, et al. 2022. [3] Shrikumar et al. 2017. [4] Simonyan et al. 2014.

Label-Irrelevant Data Augmentation



SelfPatchAug

Label-Irrelevant Data Augmentation



SubstitutivePatchAug

Experimental Results

Datasets

	Data Modality	Number of Class	Training / Test
DermaMNIST	Dermatoscope	7	7007 / 2005
BloodMNIST	Blood Cell Microscope	8	11959 / 3421
PneumoniaMNIST	Chest X-Ray	2	4708 / 624
OrganAMNIST	Abdominal CT	11	34581 / 17778
OrganCMNIST	Abdominal CT	11	13000 / 8268
OrganSMNIST	Abdominal CT	11	13940 / 8829

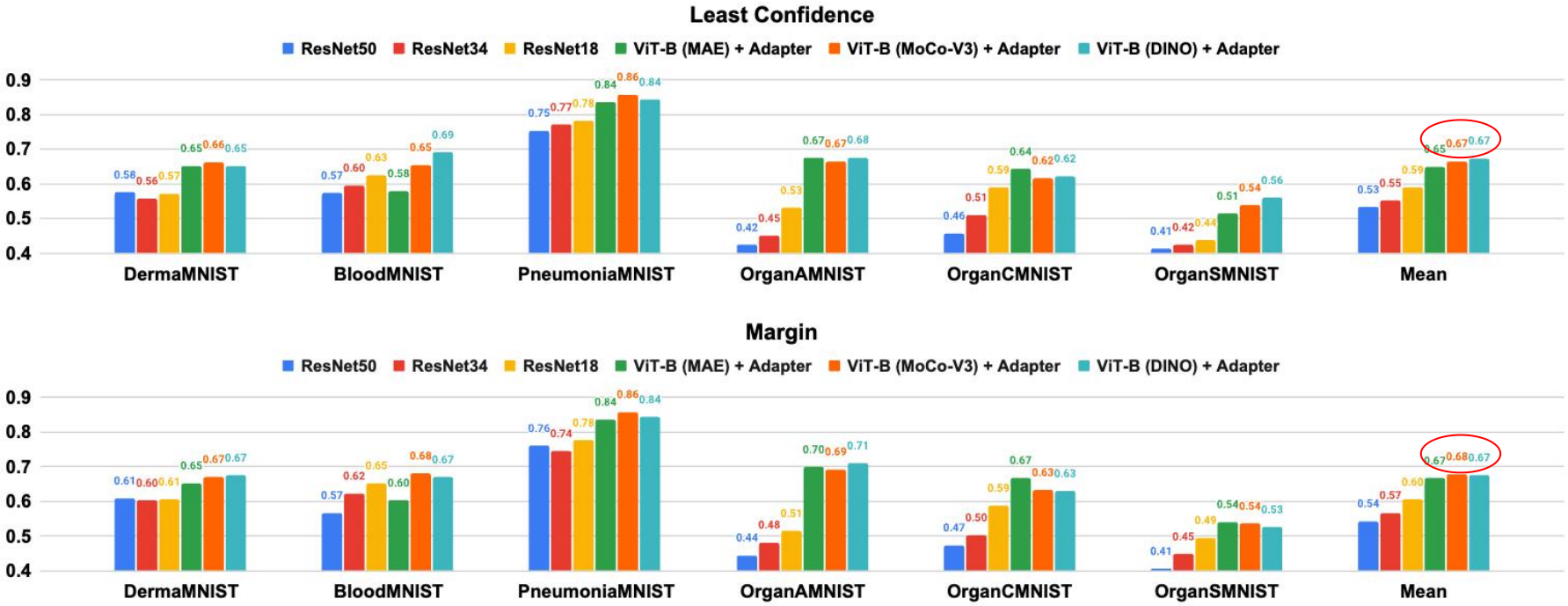
AL Settings

Few-shot AL ($N_0^{lab} = 10, K = 50, b = 5$)

Many-shot AL ($N_0^{lab} = 1000, K = 5, b = 500$)

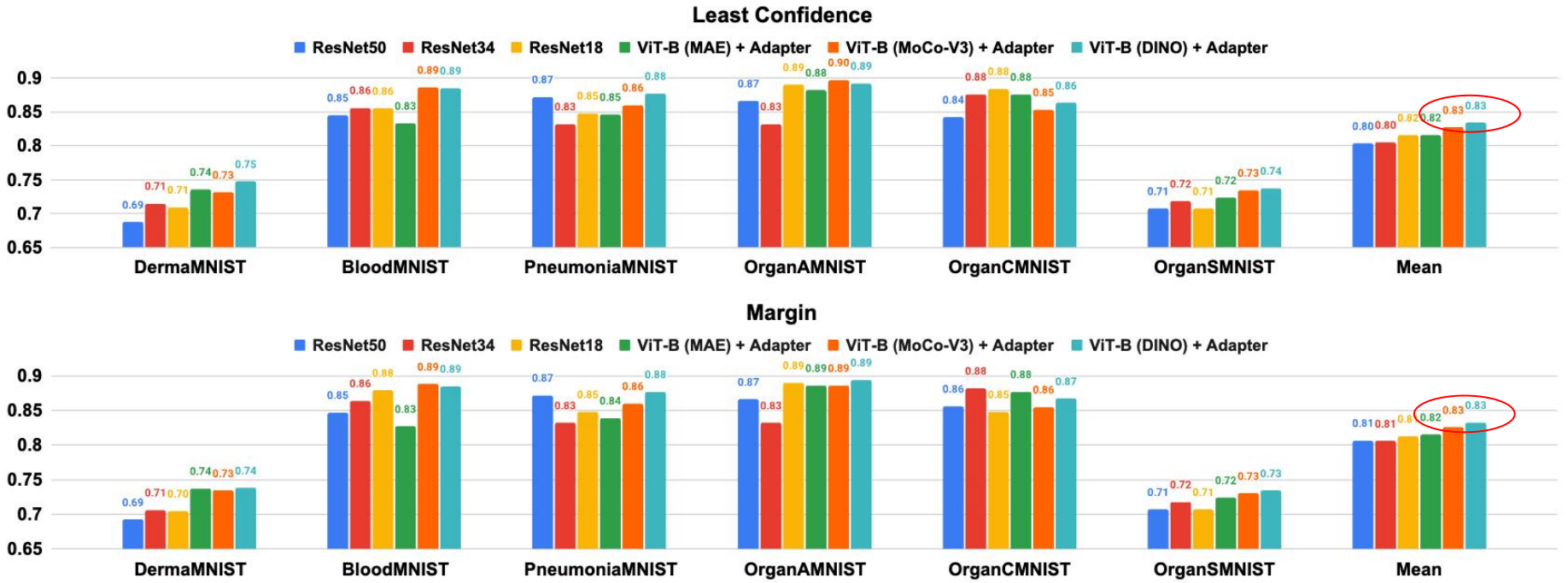
Initial #samples AL rounds Batch size

Experimental Results



AUBC Results Few-shot AL for Different Architectures

Experimental Results



AUBC Results Many-shot AL for Different Architectures

Experimental Results

	DermaMNIST	BloodMNIST	PneumoniaMNIST	OrganAMNIST	OrganCMNIST	OrganSMNIST	**Mean**
			Least Confidence				
ViT-B (MAE)	0.6513	0.5790	0.8350	0.6740	0.6420	0.5140	0.6492
+ RandAug	0.6400	0.5800	**0.8410**	0.6840	**0.6760**	0.5290	0.6583
+ AutoAug	0.6410	0.6090	0.8160	**0.7050**	0.6410	0.5120	0.6540
+ NormalAug	0.6330	0.6220	0.8400	0.5440	0.5410	0.5110	0.6152
+ SubstitutivePatchAug	0.6700	**0.6340**	0.8280	0.6850	0.6710	**0.5360**	**0.6707**
+ SelfPatchAug	**0.6920**	0.6270	0.8330	0.6840	0.6390	0.5170	0.6653
ViT-B (MoCo-V3)	0.6620	0.6540	0.8570	0.6650	0.6160	0.5380	0.6653
+ RandAug	0.6420	**0.7020**	**0.8610**	0.6790	0.6380	0.5430	0.6775
+ AutoAug	0.6690	0.6960	0.8490	0.6870	0.6140	0.5510	0.6777
+ NormalAug	**0.6720**	0.6750	0.8220	0.5560	0.5740	0.5460	0.6408
+ SubstitutivePatchAug	0.6640	0.6890	0.8270	**0.6930**	**0.6480**	**0.5830**	**0.6840**
ViT-B (DINO)	0.6510	0.6900	0.8420	0.6760	0.6220	0.5590	0.6733
+ RandAug	0.6590	0.6920	0.8670	0.6980	0.6490	0.5340	0.6832
+ AutoAug	0.6440	0.6920	**0.8700**	0.6900	0.6490	0.5510	0.6827
+ NormalAug	0.6740	**0.6980**	0.8250	0.5670	0.5540	0.5260	0.6407
+ SubstitutivePatchAug	**0.6760**	0.6940	0.8600	**0.7020**	**0.6750**	**0.5770**	**0.6973**
			Margin				
ViT-B (MAE)	0.6517	0.6020	0.8350	0.6980	0.6670	0.5400	0.6656
+ RandAug	0.6737	0.6240	**0.8410**	0.7200	0.6980	0.5500	0.6845
+ AutoAug	0.6463	0.6170	0.8160	0.7190	0.6690	0.5400	0.6679
+ NormalAug	0.6427	0.6230	0.8400	0.6427	0.5940	0.5270	0.6449
+ SubstitutivePatchAug	0.6640	0.6410	0.8280	**0.7240**	**0.7270**	**0.5850**	**0.6948**
+ SelfPatchAug	**0.6880**	**0.6470**	0.8330	0.7220	0.7000	0.5650	0.6925
ViT-B (MoCo-V3)	0.6700	0.6800	**0.8570**	0.6920	0.6320	0.5350	0.6777
+ RandAug	0.6610	0.7100	0.8560	0.7280	0.6670	0.5700	0.6987
+ AutoAug	0.6570	0.7150	0.8490	0.7160	0.6610	0.5590	0.6928
+ NormalAug	0.6660	0.6970	0.8320	0.5500	0.5910	0.5500	0.6477
+ SubstitutivePatchAug	**0.6940**	**0.7230**	0.8270	**0.7500**	**0.6880**	**0.5880**	**0.7117**
ViT-B (DINO)	0.6740	0.6690	0.8420	0.7090	0.6290	0.5250	0.6747
+ RandAug	0.6610	0.7110	0.8470	0.7220	0.6670	0.5450	0.6922
+ AutoAug	0.6760	0.7070	**0.8690**	0.7280	0.6680	0.5640	0.7020
+ NormalAug	0.6770	0.7090	0.8250	0.5740	0.5780	0.5560	0.6532
+ SubstitutivePatchAug	**0.6810**	**0.7200**	0.8560	**0.7610**	**0.7040**	**0.6050**	**0.7212**

AUBC Results Few-shot AL for Different DA methods

Experimental Results

	DermaMNIST	PneumoniaMNIST	OrganAMNIST	OrganCMNIST	OrganSMNIST	**Mean**
Least Confidence						
ViT-B (MAE)	0.7353	0.8460	0.8820	0.8757	0.7230	0.8124
+ RandAug	**0.7450**	**0.8810**	0.9020	**0.8903**	0.7350	**0.8307**
+ AutoAug	0.7367	0.8600	**0.8960**	0.8860	0.7400	0.8237
+ NormalAug	0.7397	0.8440	0.7730	0.7887	0.7190	0.7729
+ SubstitutivePatchAug	0.7447	0.8530	**0.9050**	**0.8940**	**0.7530**	0.8299
+ SelfPatchAug	**0.7420**	0.8480	0.8890	0.8910	0.7400	0.8220
ViT-B (MoCo-V3)	0.7310	0.8600	0.8970	0.8530	0.7340	0.8150
+ RandAug	0.7380	**0.8830**	0.8930	0.8680	0.7530	0.8270
+ AutoAug	0.7430	0.8760	0.8970	0.8660	0.7500	0.8264
+ NormalAug	**0.7410**	0.8710	0.7530	0.7660	0.7370	0.7736
+ SubstitutivePatchAug	0.7430	0.8670	**0.9070**	**0.8800**	**0.7610**	**0.8316**
ViT-B (DINO)	0.7470	0.8770	0.8920	0.8630	0.7370	0.8232
+ RandAug	0.7530	**0.8820**	0.9040	0.8780	0.7560	0.8346
+ AutoAug	0.7510	0.8810	0.9060	0.8830	0.7510	0.8344
+ NormalAug	0.7500	0.8620	0.7530	0.7730	0.7270	0.7730
+ SubstitutivePatchAug	**0.7620**	0.8680	**0.9160**	**0.8870**	**0.7590**	**0.8384**
Margin						
ViT-B (MAE)	0.7373	0.8390	0.8860	0.8770	0.7240	0.8127
+ RandAug	0.7433	**0.8770**	0.9020	0.8917	0.7450	**0.8318**
+ AutoAug	0.7417	0.8670	0.8960	0.8910	0.7410	0.8273
+ NormalAug	0.7427	0.8420	0.7670	0.7893	0.7190	0.7720
+ SubstitutivePatchAug	0.7433	0.8620	**0.9080**	**0.8960**	**0.7490**	0.8317
+ SelfPatchAug	**0.7440**	0.8510	0.8950	0.8850	0.7340	0.8218
ViT-B (MoCo-V3)	0.7340	0.8600	0.8860	0.8550	0.7300	0.8130
+ RandAug	**0.7410**	0.8750	0.8930	0.8640	0.7520	0.8250
+ AutoAug	0.7200	**0.8770**	0.8950	0.8710	0.7500	0.8226
+ NormalAug	0.7350	0.8730	0.7500	0.7620	0.7440	0.7728
+ SubstitutivePatchAug	**0.7410**	0.8670	**0.8960**	**0.8850**	**0.7580**	**0.8294**
ViT-B (DINO)	0.7380	0.8770	0.8940	0.8670	0.7340	0.8220
+ RandAug	0.7490	**0.8810**	0.9090	0.8790	0.7520	0.8340
+ AutoAug	0.7510	0.8800	0.9060	0.8790	0.7510	0.8334
+ NormalAug	0.7500	0.8650	0.7550	0.7690	0.7300	0.7738
+ SubstitutivePatchAug	**0.7550**	0.8700	**0.9160**	**0.8880**	**0.7590**	**0.8376**

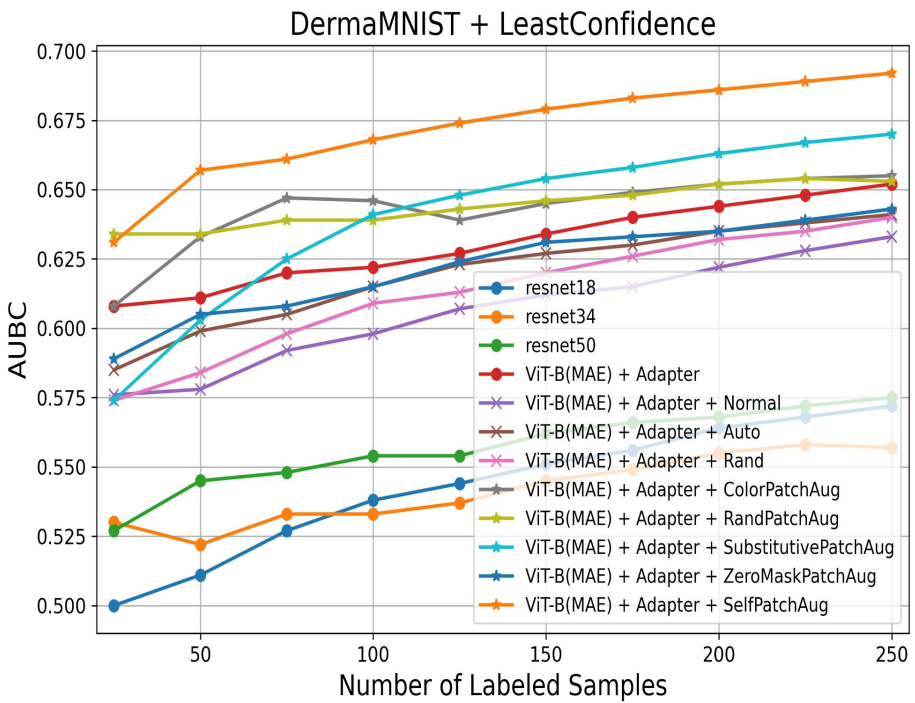
AUBC Results Many-shot AL for Different DA methods

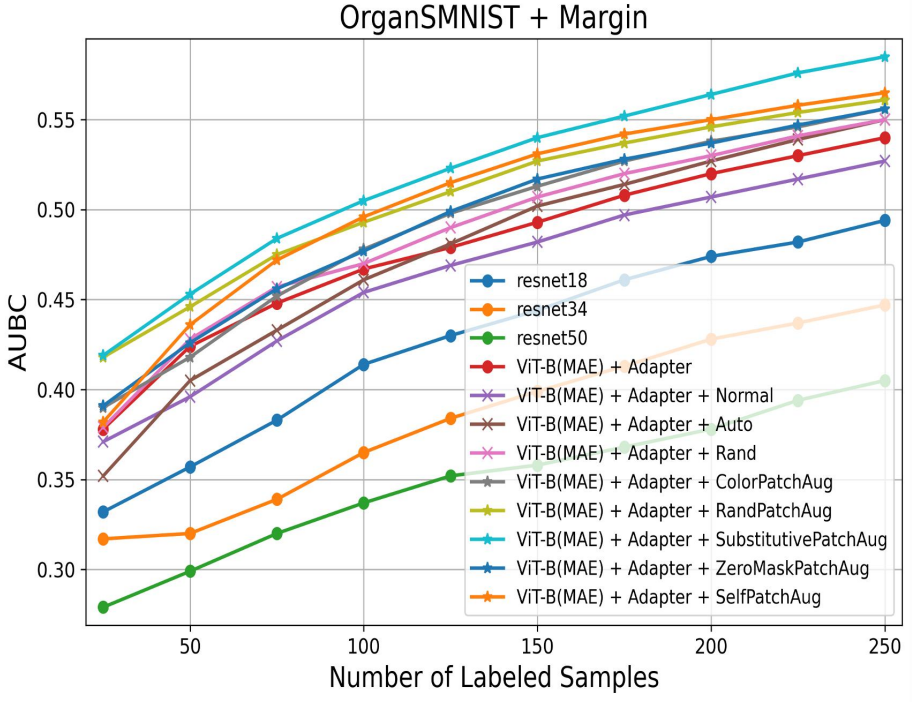
Experimental Results

	DermaMNIST	BloodMNIST	PneumoniaMNIST	OrganAMNIST	OrganCMNIST	OrganSMNIST	**Mean**
			Coreset				
ViT-B (MoCo-V3)	0.6170	0.6660	0.7950	0.6630	0.6100	0.5260	0.6462
+ RandAug	0.6520	0.6570	0.8170	0.6920	0.6130	0.5410	0.6620
+ AutoAug	0.6290	0.6900	**0.8390**	0.6990	0.6170	0.5390	0.6688
+ NormalAug	0.6280	0.6340	0.7720	0.5560	0.5800	0.5470	0.6195
+ SubstitutivePatchAug	**0.6610**	**0.7190**	0.8260	**0.7170**	**0.6860**	**0.5480**	**0.6928**
ViT-B (DINO)	0.6310	0.6470	0.8380	0.6880	0.6060	0.5560	0.6610
+ RandAug	0.6640	0.6440	0.8210	0.6790	0.6220	0.4700	0.6500
+ AutoAug	0.6430	0.6500	0.8340	**0.7200**	**0.6490**	0.5410	0.6728
+ NormalAug	0.6320	0.6720	0.8310	0.5570	0.5640	0.5170	0.6288
+ SubstitutivePatchAug	**0.7120**	**0.6940**	**0.8350**	0.7140	0.6360	**0.5940**	**0.6975**
			Entropy				
ViT-B (MoCo-V3)	0.6730	0.6460	**0.8570**	0.6290	0.5970	0.5120	0.6523
+ RandAug	0.6560	0.6650	0.8320	**0.6810**	0.6130	0.5230	0.6617
+ AutoAug	0.6460	0.6760	0.8420	0.6700	0.6140	0.5160	0.6607
+ NormalAug	0.6550	0.6980	0.8240	0.5580	0.5540	0.5390	0.6380
+ SubstitutivePatchAug	**0.6860**	**0.6920**	0.8240	0.6410	**0.6450**	**0.5450**	**0.6722**
ViT-B (DINO)	0.6670	0.6690	0.8510	0.6690	0.6170	0.5440	0.6695
+ RandAug	0.6440	**0.6960**	0.8610	**0.6800**	0.6370	0.5240	0.6737
+ AutoAug	0.6380	0.6590	**0.8680**	0.6720	0.6170	0.5270	0.6635
+ NormalAug	0.6510	0.6790	0.8250	0.5430	0.5370	0.5450	0.6300
+ SubstitutivePatchAug	**0.6830**	0.6920	0.8560	0.6660	**0.6620**	**0.5550**	**0.6857**

AUBC Results Few-shot AL for Different DA methods (More AL Methods)

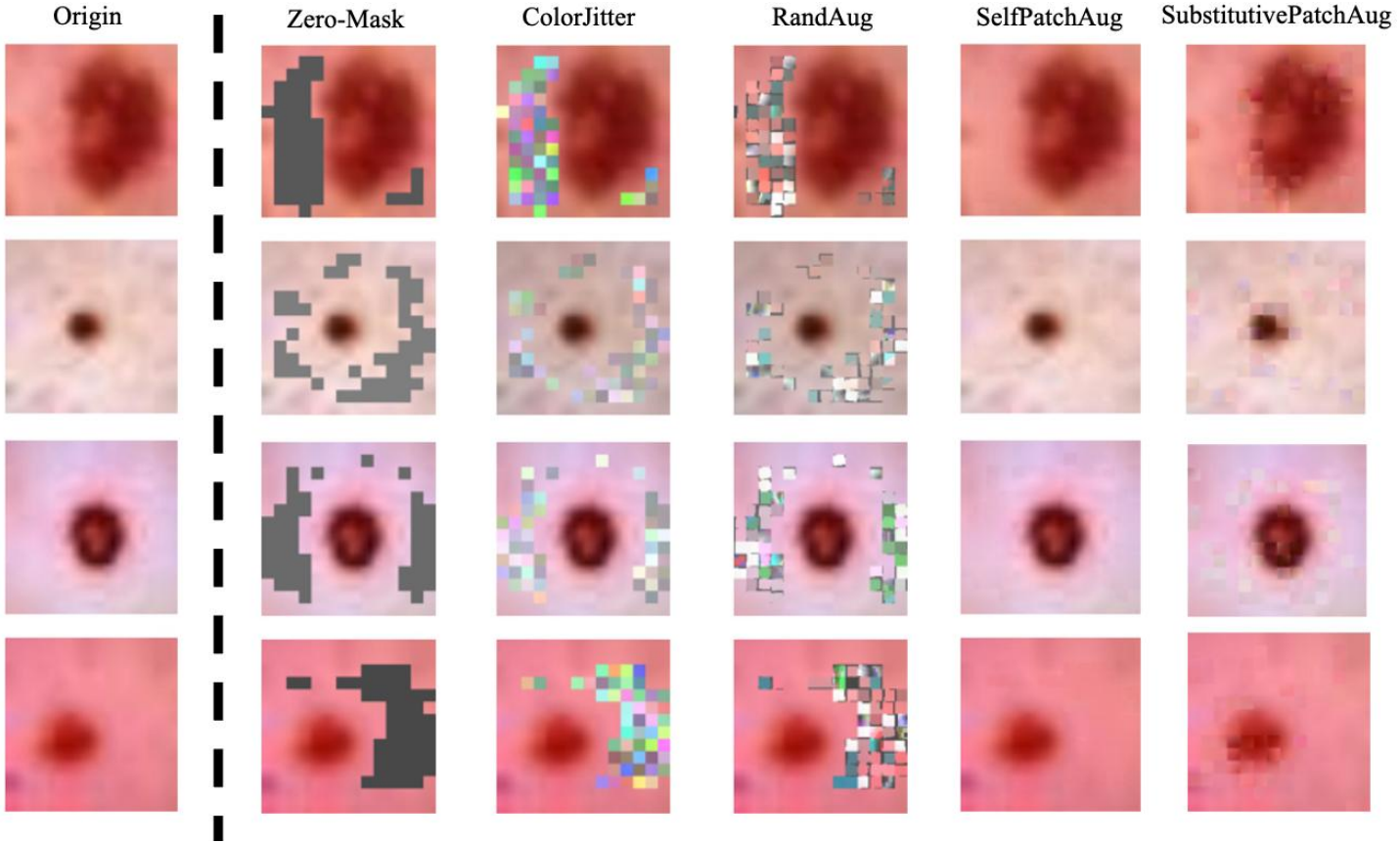
Experimental Results





Flexibility: We can also plug other DA methods for the patch augmentation!

Experimental Results



Visualization: Augmented Samples for Different Patch Augmentation Methods

Experimental Results

ViT-B (MAE)	Least Confidence				Margin			
	LastAtt	CosineAtt	Saliency	DeepLIFT	LastAtt	CosineAtt	Saliency	DeepLIFT
$r = 25\%$	0.675	**0.678**	0.648	0.668	0.640	0.653	0.646	0.613
$r = 50\%$	0.655	0.665	0.654	0.657	0.687	0.671	0.663	0.658
$r = 75\%$	0.643	0.670	0.675	0.638	0.672	0.664	0.662	**0.681**

Ablation Study: Different Localization Methods

ViT-B (MAE) + Adapter	Least Confidence		Margin		LeastConfidence MC		MeanSTD	
	With LaLS	Without LaLS	With LaLS	Without LaLS	With LaLS	Without LaLS	With LaLS	Without LaLS
BloodMNIST	**0.6270**	0.6040	**0.6470**	0.5970	**0.6200**	0.5970	0.5650	**0.5950**
DermaMNIST	**0.6920**	0.6570	**0.6880**	0.6540	**0.6570**	0.6540	0.6260	**0.6370**
PneumoniaMNIST	**0.8330**	0.8230	**0.8330**	0.8230	**0.8530**	0.8230	0.8070	**0.8320**
OrganAMNIST	0.6840	**0.6930**	0.7220	**0.7260**	0.6750	0.7100	0.5680	**0.6030**
OrganCMNIST	0.6390	**0.6530**	**0.7000**	0.6810	0.6470	**0.6740**	**0.6050**	0.6040
OrganSMNIST	0.5170	**0.5410**	**0.5650**	0.5620	0.5230	**0.5460**	**0.4620**	0.4500
Mean	**0.6653**	0.6618	**0.6925**	0.6738	0.6625	**0.6673**	0.6055	**0.6202**

Ablation Study: Instance-adaptive Label Smoothing

Conclusions

1. We design an **efficient AL framework** based on off-the-shelf ViTs to gain nearly minimally sufficient representations.

2. We design a **label-irrelevant patch augmentation** scheme that preserves semantic information better than prior data augmentation methods during AL.

3. Few-shot AL: improve performance by 5%-7%. Our proposed data augmentation scheme: improve performance by 1%-4%.

Thank you for listening!

Q&A