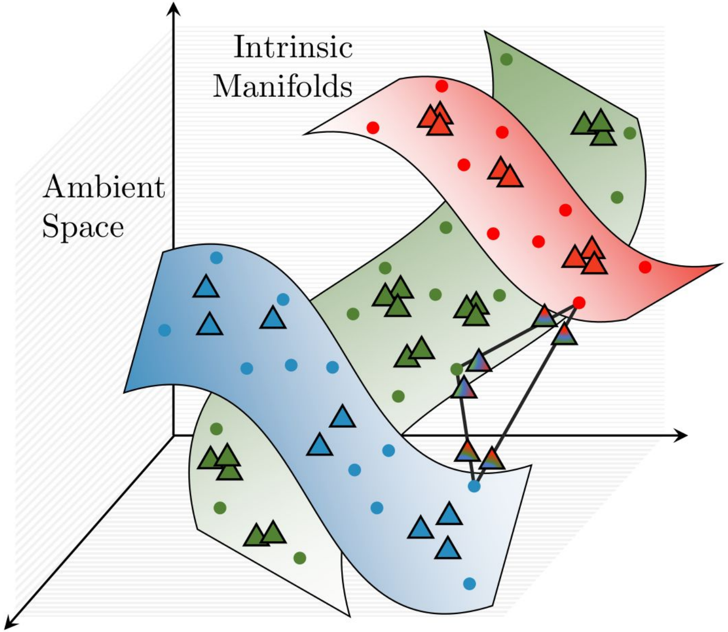
ζ-*mixup*: Richer, More Realistic Mixing of Multiple Images

Kumar Abhishek[†], Colin J. Brown[‡], Ghassan Hamarneh[†]



MIDL
Nashville 2023

SFU
SIMON FRASER UNIVERSITY
MIAL
Medical Image Analysis Lab

ζ-*mixup*: Richer, More Realistic Mixing of Multiple Images

Kumar Abhishek[†], Colin J. Brown[‡], Ghassan Hamarneh[†]

[†] Medical Image Analysis Lab, School of Computing Science, Simon Fraser University, Canada
[‡] Hinge Health, Canada

MIDL Nashville 2023

MIAL Medical Image Analysis Lab

SFU SIMON FRASER UNIVERSITY

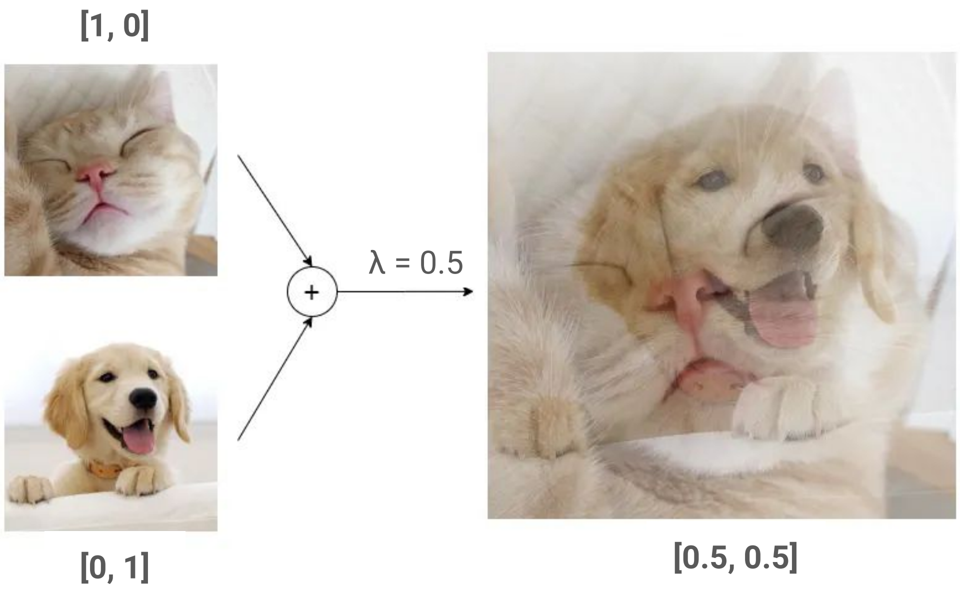
mixup Data Augmentation

Generate **convex combinations** of training samples and **linear interpolations** of labels.

$$\hat{x} = \lambda x_1 + (1 - \lambda) x_2$$

$$\hat{y} = \lambda y_1 + (1 - \lambda) y_2$$

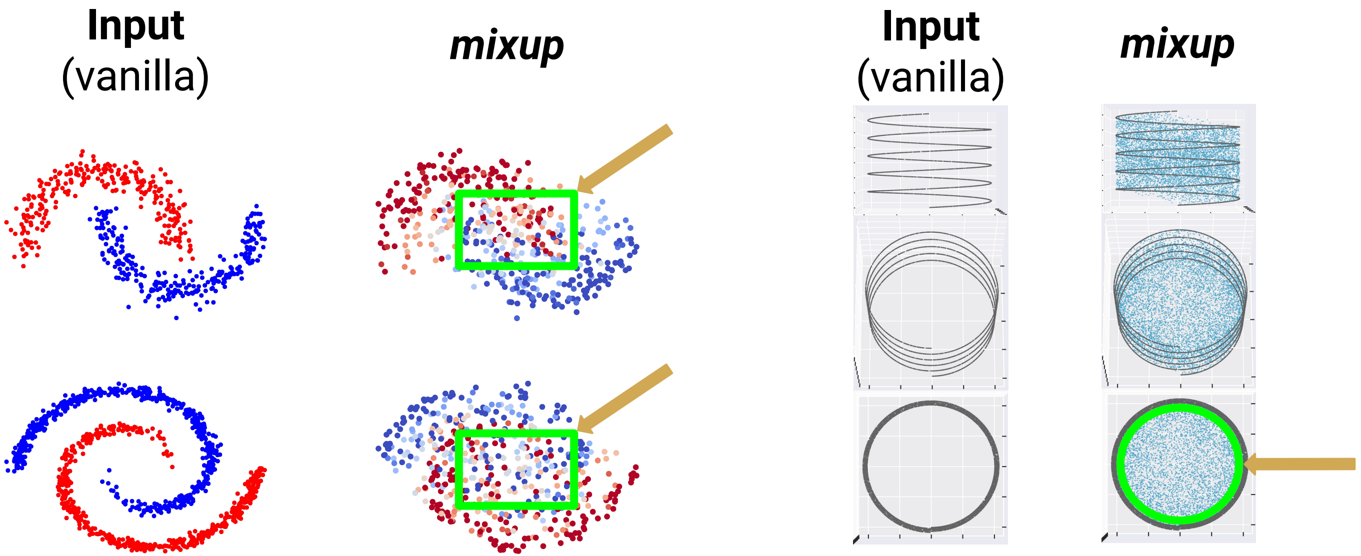
Assumption: a model should behave linearly between any two training samples, even if the distance between them is large.



Problems with *mixup*

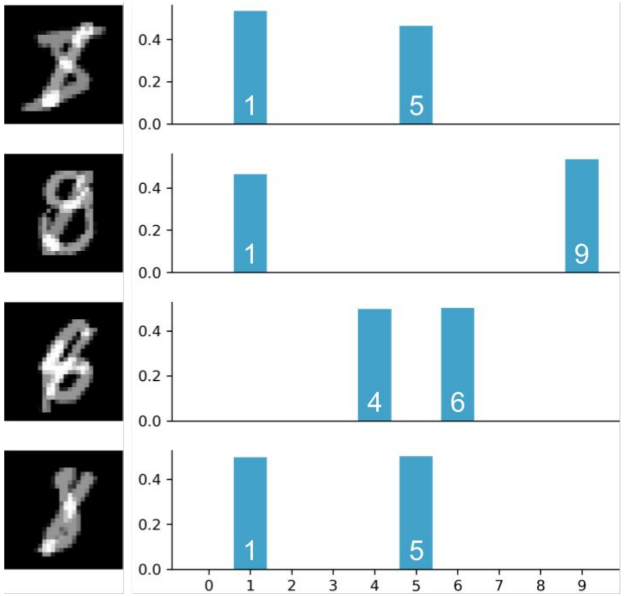
- Can sample data off the manifold, causing an inflated intrinsic dimensionality.

- Can generate samples with incorrect labels.



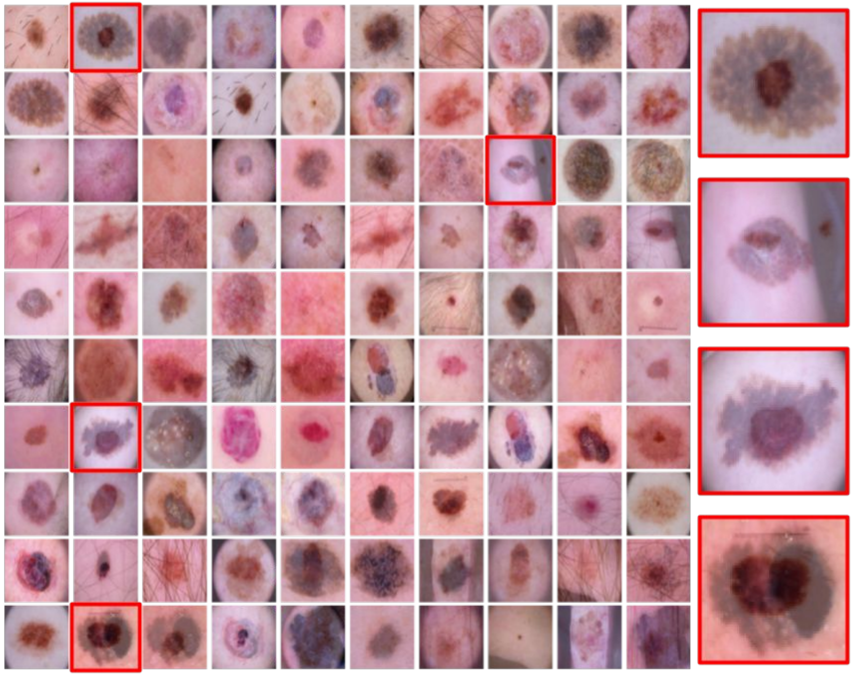
Problems with *mixup*

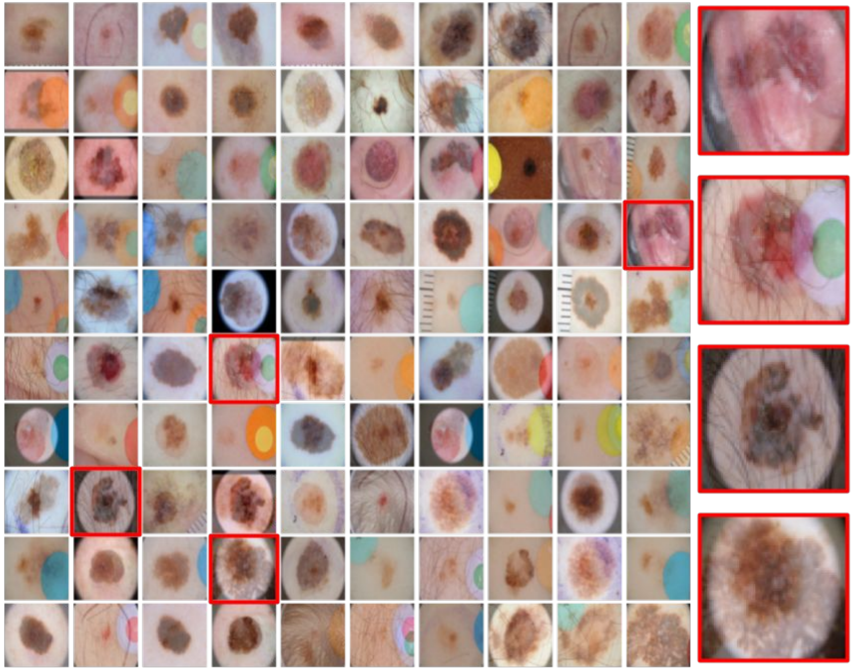
mixup outputs have ghosting artifacts and lower realism.



mixup outputs can contain incorrect soft labels.

Problems with *mixup*





Similar problems are observed with medical images.

Arguments

- We should only synthesize examples with **high confidence of realism**.
- A model should only behave **linearly nearby training samples**.

Goal: A better augmentation method

- **Realism:** synthesize samples close to the original samples
- **Diversity:** allow generating diverse samples by exploring the input space
- **Label richness:** generate samples with labels incorporating information from several classes while staying on the manifold of realistic samples
- **Valid probabilistic labels** for synthesized samples
- **Computationally efficient** to allow augmentation of training batches

Proposed Method

Synthesize new samples as **convex combinations of N samples** as

$$\hat{x} = \sum_{i=1}^{N} w_i x_i; \quad \hat{y} = \sum_{i=1}^{N} w_i y_i$$

where weights w_i should satisfy the desirable criteria.

One such weighting scheme: sample weights from the **terms of a p-series**: $w_i = i^{-p}$

Given N samples and an $N \times N$ permutation matrix π, resulting in a randomized ordering of samples $s = \pi[1, 2, \ldots, N]^T$, the weights are

$$w_i = \frac{s_i^{-\gamma}}{C}, \quad i \in [1, N]$$

$C = \sum_{j=1}^{N} j^{-\gamma}$ is the N-truncated Riemann zeta function (hyperparameter γ), thus ***ζ-mixup***.

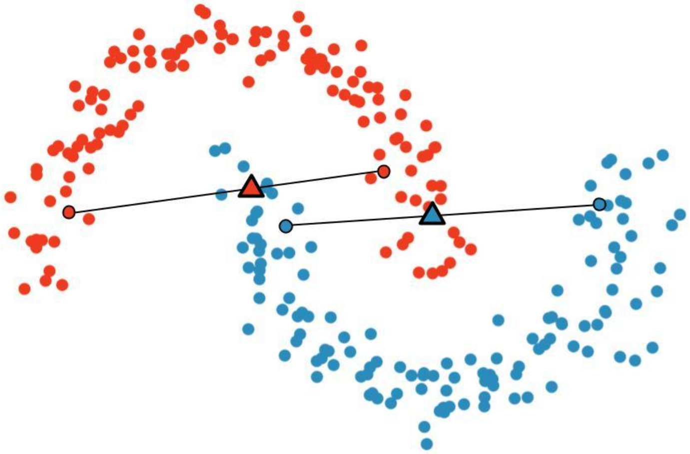
Properties of *ζ-mixup*

- Since there exist $N!$ possible $N \times N$ permutation matrix, we can generate $N!$ new samples for a single value of γ.
- For $\gamma \geq \gamma_{\text{min}} = 1.72865$, the weight assigned to one sample dominates all other weights, i.e.,

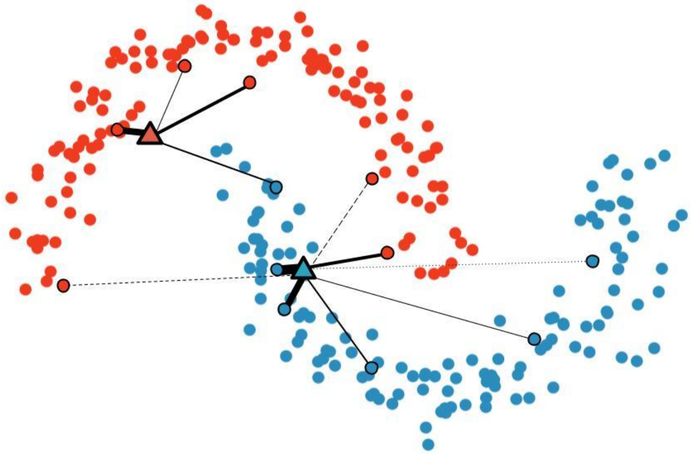
$$\forall \gamma \geq 1.72865, \;\; w_1 > \sum_{i=2}^{N} w_i$$

- For $N = 2$ and $\gamma = \log_2 \left(\frac{\lambda}{1-\lambda} \right)$, *ζ-mixup* simplifies to *mixup*.

ζ-*mixup* can **mix any number of samples**

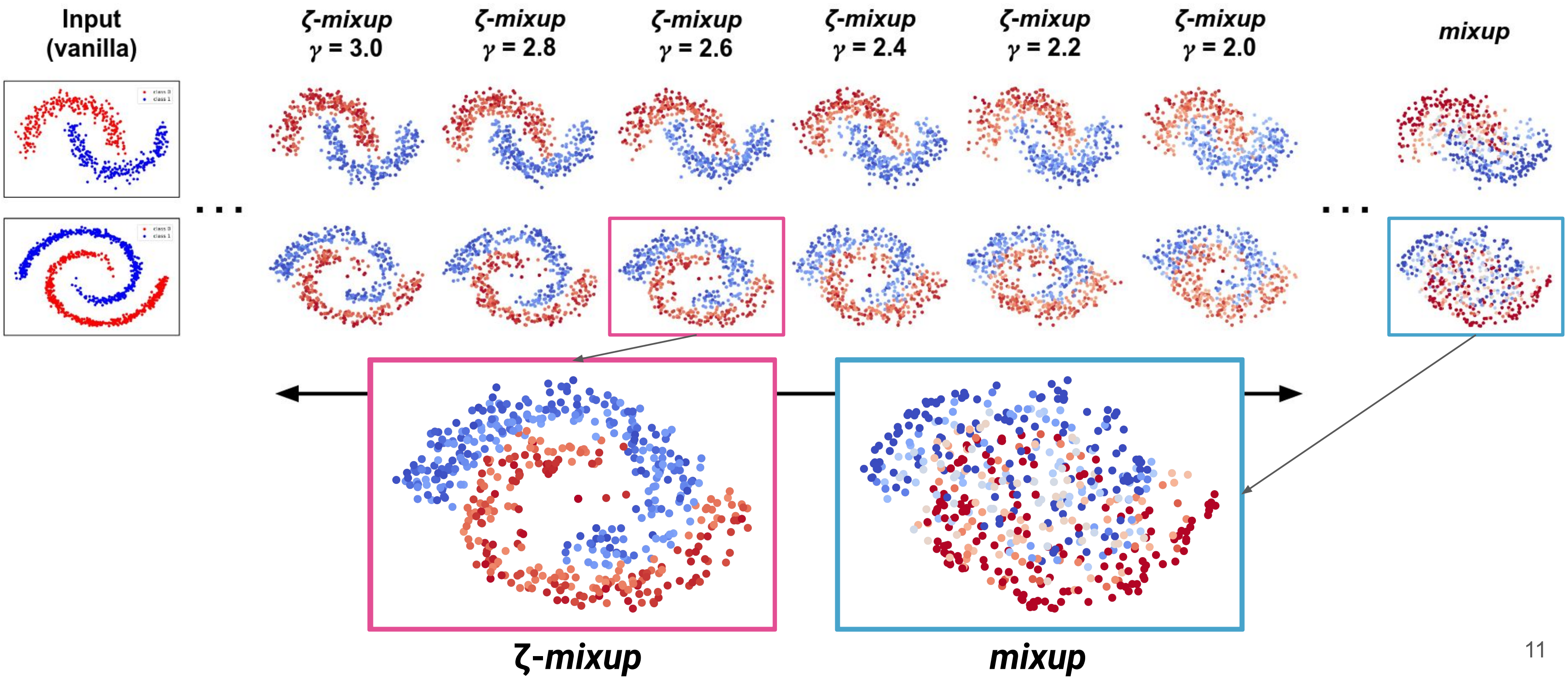


mixup

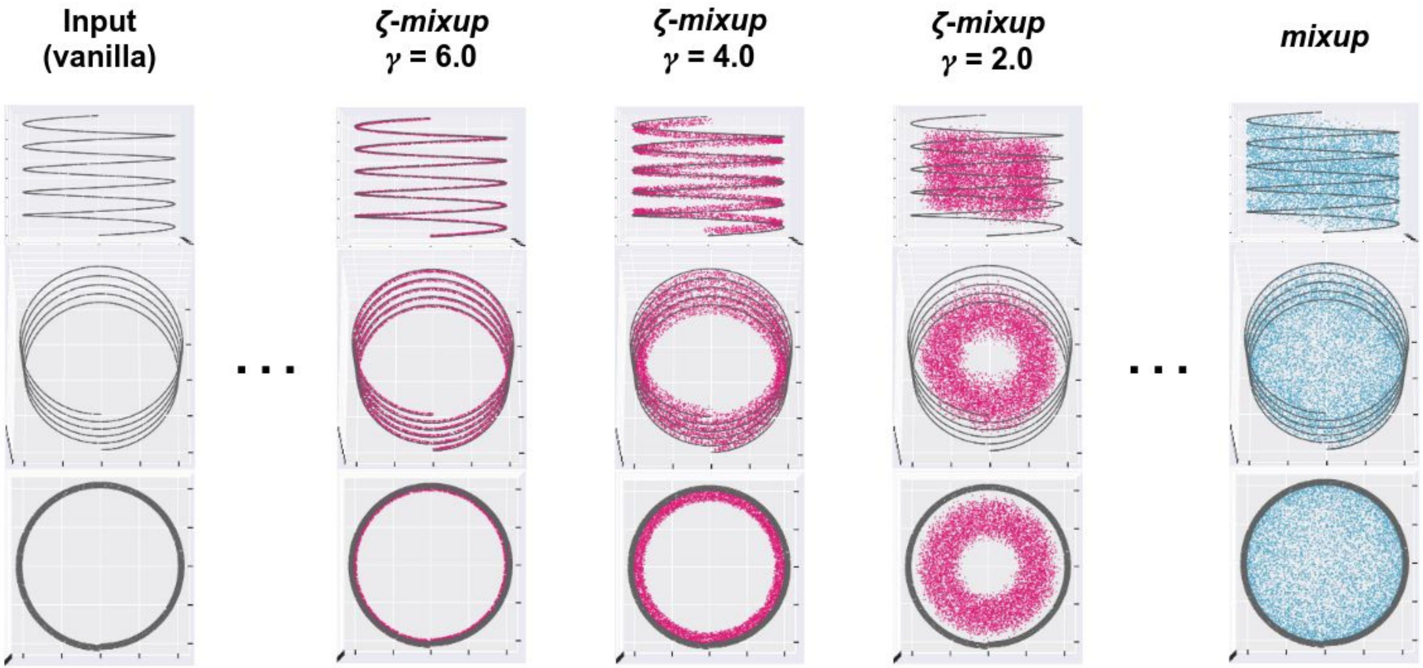


ζ-*mixup* can mix $\boldsymbol{N}$ samples (e.g., 4, 8).

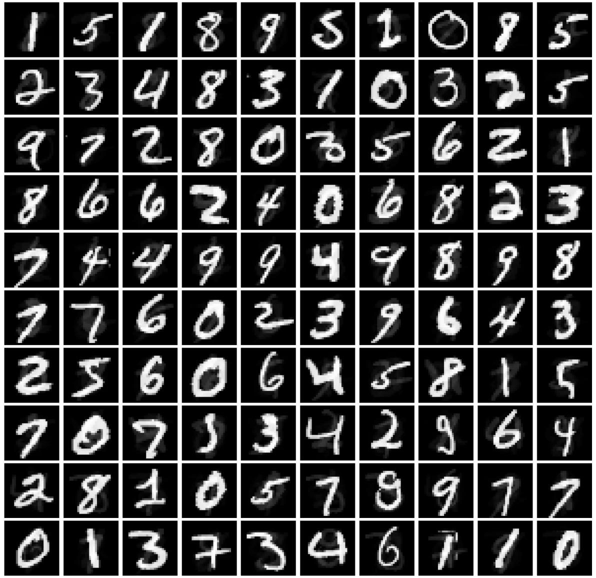
ζ-*mixup* yields **realism** and **diversity**

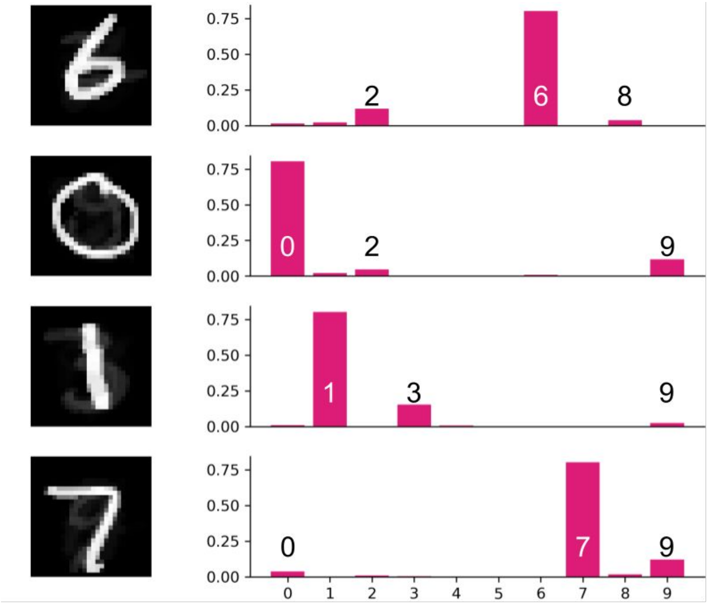


ζ-mixup yields **realism** and **diversity**



ζ-*mixup* outputs exhibit **label richness**, **realism**, and **label correctness**





ζ-***mixup*** outputs have a **higher degree of realism**, and contain **correct and rich soft labels**, incorporating **information from multiple classes**.

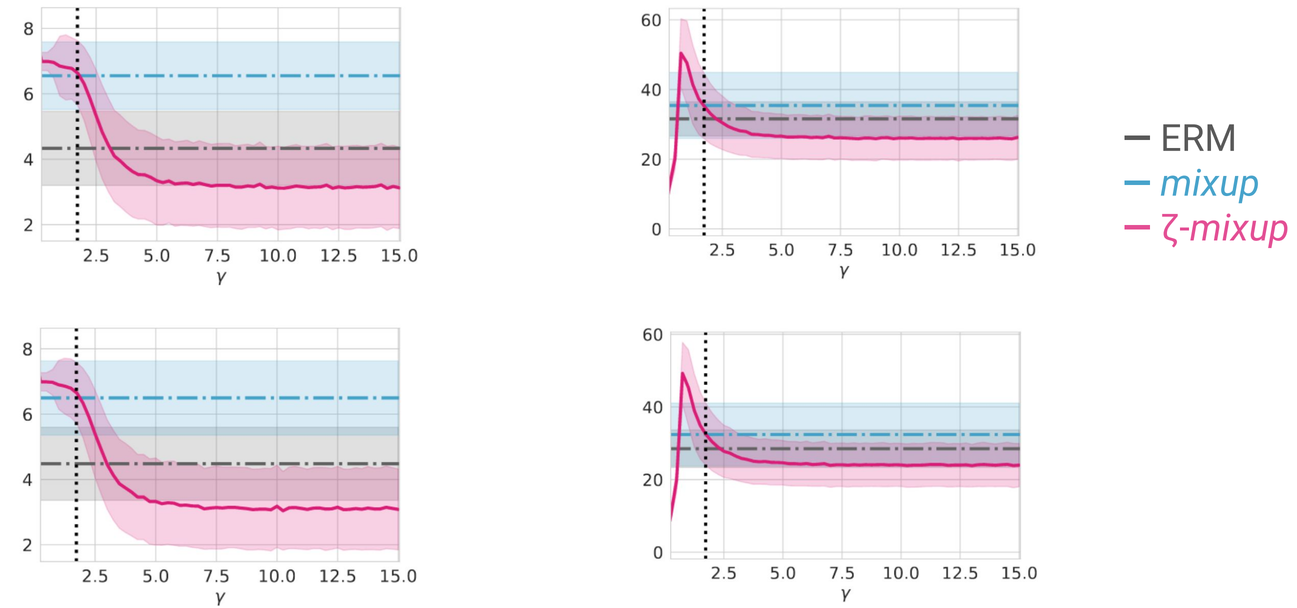
ζ-*mixup* better preserves the intrinsic dimensionality

Intrinsic dimensionality estimated using

8 nearest neighbors | 128 nearest neighbors

CIFAR-10

CIFAR-100



ζ-*mixup* achieves **better classification performance**

Natural image datasets (classification error rate)

Method	CIFAR-10 ResNet-18	CIFAR-100 ResNet-18
ERM	5.48	23.33
mixup	4.68	21.85
ζ-*mixup*	**4.42**	**21.35**

Method	CIFAR-10		CIFAR-100	
	ResNet-18	ResNet-50	ResNet-18	ResNet-50
CutMix	4.13	4.08	19.97	18.99
+ ζ-*mixup*	**3.84**	**3.61**	**19.54**	**18.86**

Skin lesion image datasets (micro-averaged F1 score)

Method	ISIC 2016		ISIC 2017		ISIC 2018		DermoFit	
	ResNet-18	ResNet-50	ResNet-18	ResNet-50	ResNet-18	ResNet-50	ResNet-18	ResNet-50
ERM	0.7836	0.8127	0.7383	0.6867	**0.8756**	0.8653	0.8269	0.8500
mixup	0.7968	0.8179	0.7333	0.7433	0.8394	0.8601	0.8577	0.8500
ζ-*mixup*	**0.8654**	**0.8602**	**0.7633**	**0.7733**	**0.8756**	**0.9016**	**0.8731**	**0.8962**

Thank you.

Questions?

kabhishe@sfu.ca