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

SFU SIMON FRASER UNIVERSITY

MIAL MEDICAL IMAGE ANALYSIS LAB

mixup Data Augmentation

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

$$\hat{x} = \lambda x_1 + (1 - \lambda) x_2 \qquad \hat{y} = \lambda y_1 + (1 - \lambda) y_2$$

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

Problems:

- Can sample data off the data manifold.
- Can generate samples with incorrect labels.

Proposed Data Augmentation: ζ-*mixup*

Arguments:

- Synthesized samples should have **high confidence of realism**.
- A model should only behave **linearly nearby training samples**.

Formulation

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

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

Sample weights from **terms of a *p*-series**, apply them to a **randomized ordering** s of training samples, and **normalize** the weights.

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

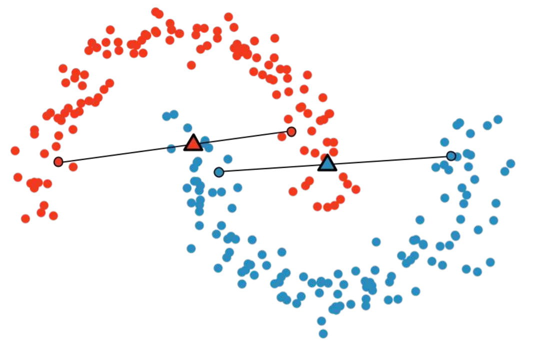
$C = \sum_{j=1}^{N} j^{-\gamma}$ is the N-truncated Riemann zeta function at γ, $\zeta(\gamma)$.

γ: hyperparameter to control **how far from the original samples** the synthetic samples are created.

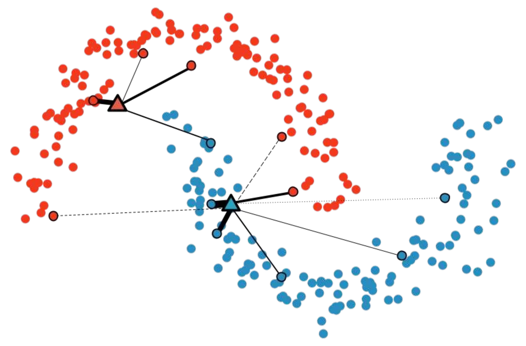
Key properties:

- Can synthesize ***N*! new samples** for a single value of γ.
- For $\gamma \geq 1.72865$, the weight assigned to one sample dominates all other weights.
- ***mixup* is a special case** of ζ-*mixup*.

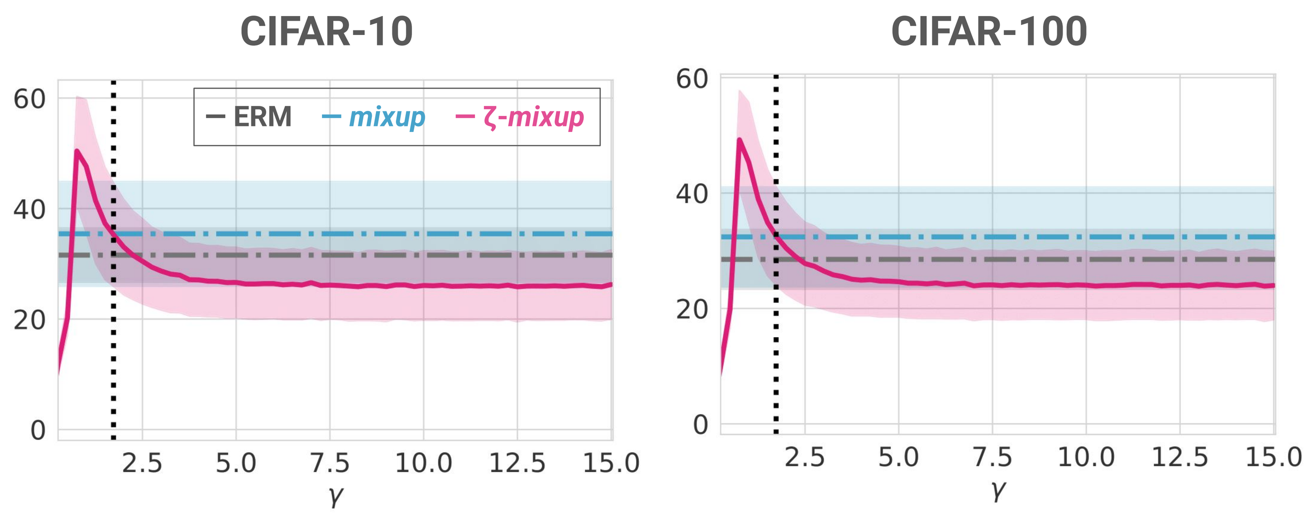
Results



mixup can only mix 2 samples.

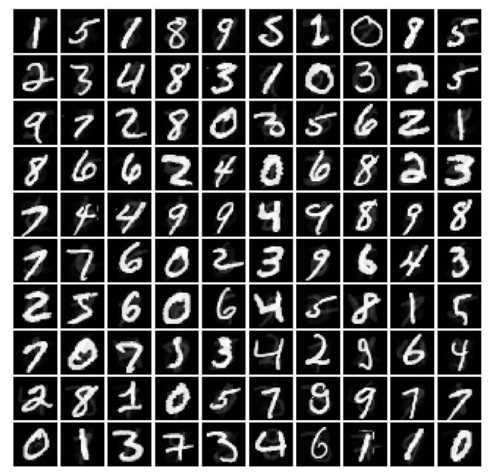


ζ-*mixup* can mix N samples (e.g., 4, 8) and **respects the data manifold.**

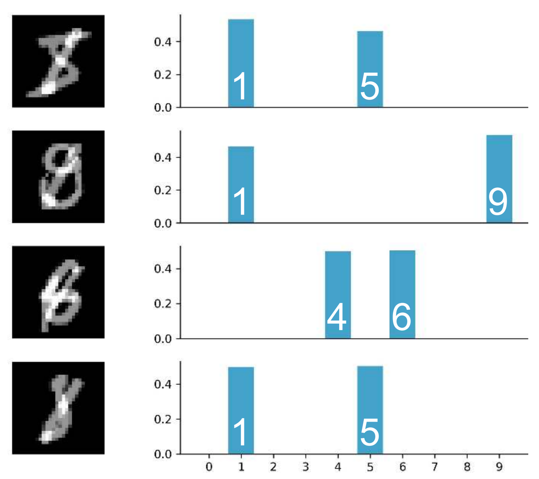


ζ-*mixup* better preserves the intrinsic dimensionality of datasets (estimated using 128 nearest neighbors).

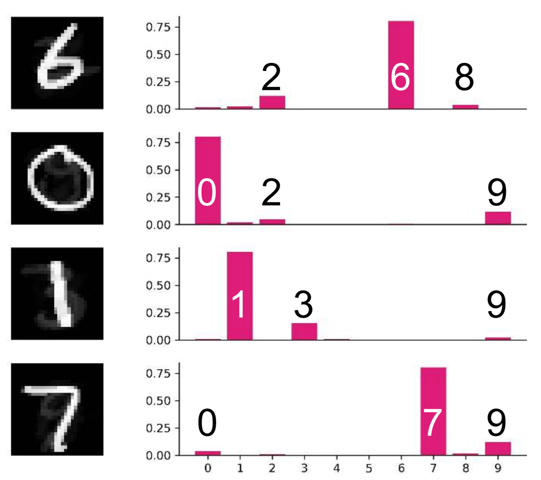
mixup outputs have ghosting artifacts and lower realism.



ζ-*mixup* outputs have a much higher realism.

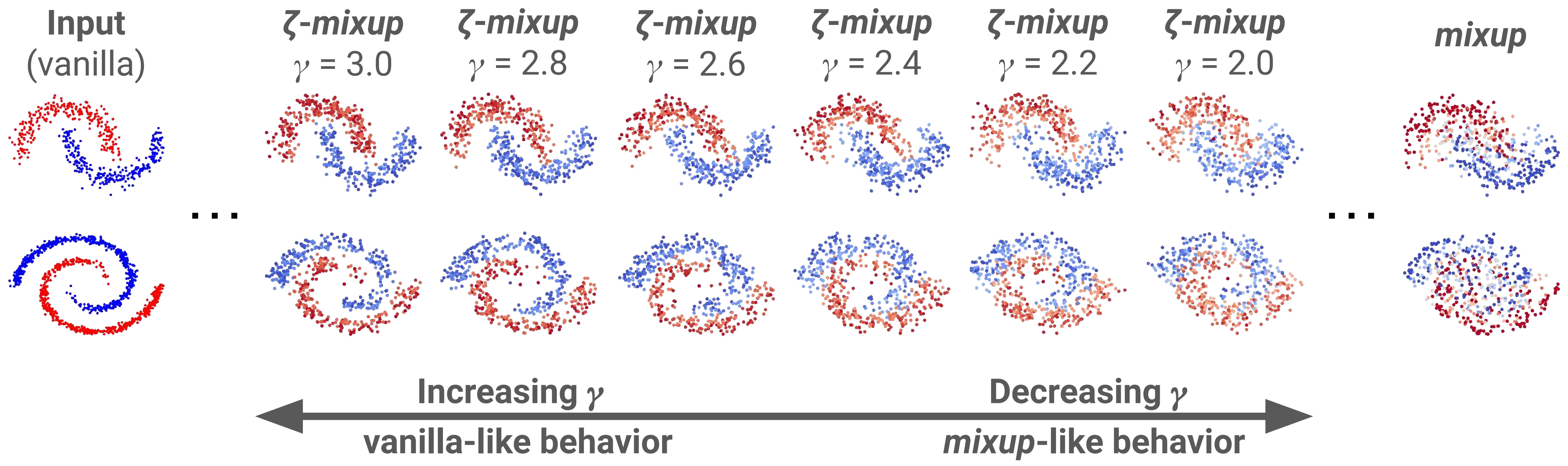


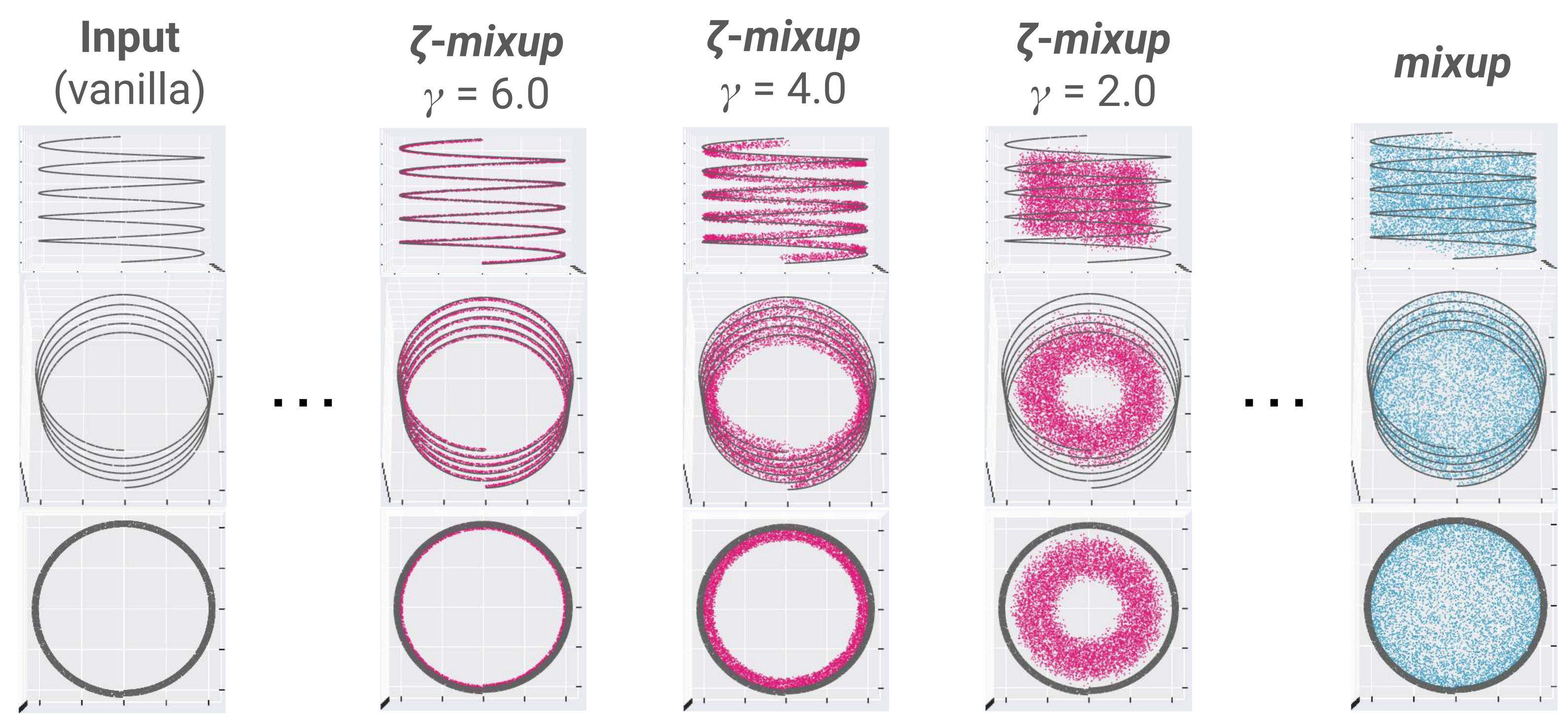
mixup outputs can contain incorrect soft labels.



ζ-*mixup* outputs contain correct and rich soft labels, incorporating information from multiple classes.

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





ζ-*mixup* yields realism and diversity in the synthesized samples.

Natural image classification (classification error rate)

Method	CIFAR-10 ResNet-18	CIFAR-100 ResNet-18	Method	CIFAR-10 ResNet-18	CIFAR-10 ResNet-50	CIFAR-100 ResNet-18	CIFAR-100 ResNet-50
ERM	5.48	23.33	CutMix	4.13	4.08	19.97	18.99
mixup	4.68	21.85	+ ζ-*mixup*	**3.84**	**3.61**	**19.54**	**18.86**
ζ-*mixup*	**4.42**	**21.35**					

Medical image classification (micro-averaged F1 score)

Method	ISIC 2016 ResNet-18	ISIC 2016 ResNet-50	ISIC 2017 ResNet-18	ISIC 2017 ResNet-50	ISIC 2018 ResNet-18	ISIC 2018 ResNet-50	DermoFit ResNet-18	DermoFit 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**

ζ-*mixup* improves classification performance on **natural** and **medical images (skin lesion**; measured by F1-micro) **datasets**, and can be **combined with other augmentation methods** (e.g., CutMix).

Acknowledgements

NSERC CRSNG · Digital Research Alliance of Canada · Alliance de recherche numérique du Canada · NVIDIA

Contact:
kabhishe@sfu.ca