SUMix: Mixup with Semantic and Uncertain Information

Huafeng Qin[1,*], Xin Jin[1,*], Hongyu Zhu[1,*], Hongchao Liao[1], Mounîm A. El-Yacoubi[2], and Xinbo Gao[3]

[1] Chongqing Technology and Business University, Chongqing, China
[2] Telecom SudParis, Institut Polytechnique de Paris, Paris, France
[3] Chongqing University of Posts and Telecommunications, Chongqing, China

• Introduction

- Mixup data augmentation approaches have been applied forvarious tasks of deep learning to improve the generalization ability of DNNs.
- Some approaches CutMix, SaliencyMix, FMix *etc*. randomly replace a patch in one image with patches from another to generate the mixed image, but those approaches will caused a problem "***Label MisMatch***". Shown in Fig 1.

Fig 1. The figure shows hand-crafted mixup methods with "*Label MisMatch*" problem.

- We proposed **SUMix**, a novel approach to learn the mixing ratio λ, as well as the uncertainty for the mixed samples duringthe training process. Extensive experiments on five image classification datasets verify that our proposed SUMix can remarkably improve performances of existing mixup augmentations (some shown in Fig 3) in a plug-and-play manner while achieving better robustness.

So our main contributions are as follows:

a) We propose a learnable metric to compute the mixed ratio by similarity between the mixed samples and the original samples.
b) We further consider the uncertainty and semantic information of the mixed samples and recalculate a reasonable feature vector, providing an additional regularization loss for model training.
c) Our SUMix helps mainstream Cutting-based mixup methods to improve classification tasks without spending too excessive extra time overhead.

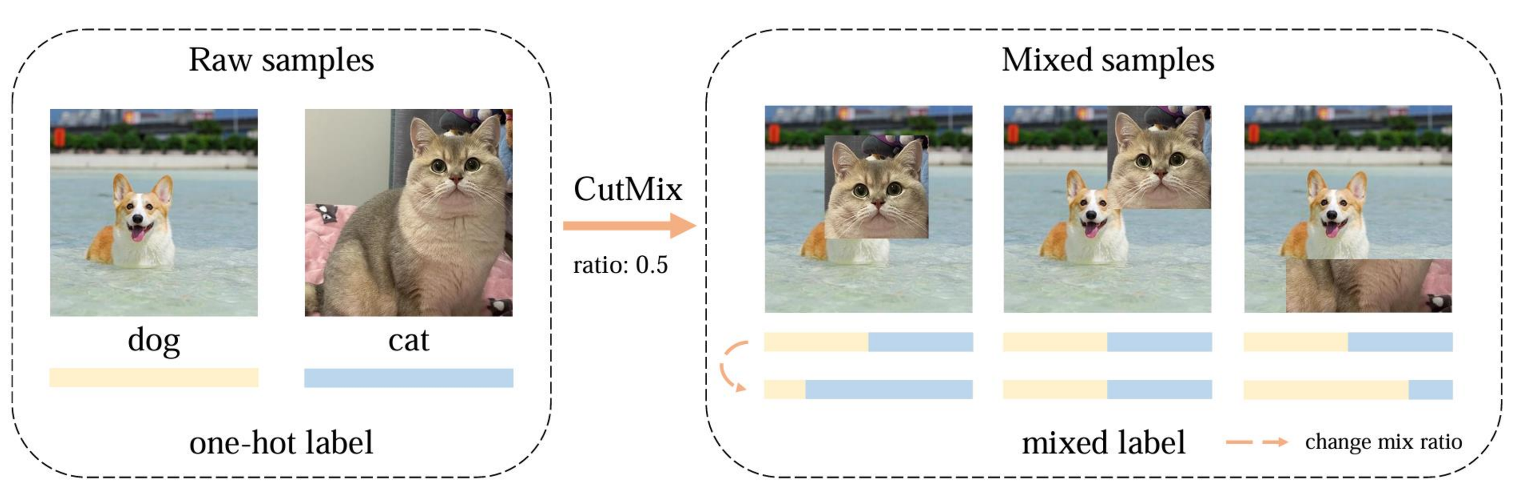

Fig 2. The figure shows different cases of raw samples that underwent the CutMix with a mixing ratio of 0.5 to obtain mixed samples, and right term shows the redefined mixing raitio $\tilde{\lambda}$.

• SUMix

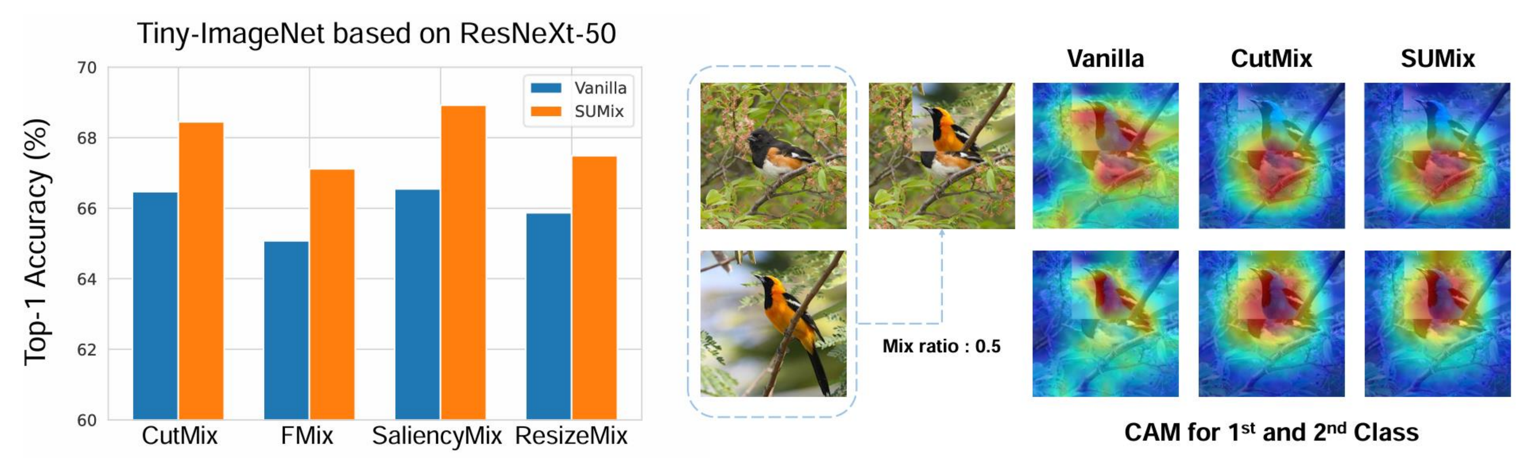

Fig 3. *Left:* Classification results of Vanilla mixups and with SUMix; *Right:* Comparison of Vanilla method,CutMix, and SUMix for CAM visualization.

1. Uncertainty Classifier

SUMix combines two losses as the final loss function: vanilla mixup cross entropy ($\mathcal{L}_{MCE}$) and a regularized loss:

$$\mathcal{L}_{su} = \frac{1}{N}\left(\sum_{1=0}^{N} \mathcal{L}_{MCE}\big(f_\theta(\tilde{X}_i), Y_i, \tilde{\lambda}_i\big)\right) + \xi * \mathcal{L}_{MCE}\big(SU(f_\theta(\tilde{X}_i), U_\omega), \tilde{Y}\big).$$

2. Mix ratio Learning

We got a triple feature pairs $(\tilde{z}, z_a, z_b)$ from raw samples and mixed sample by the encoder. and normalized to modify their similarity as the fixed mixing ratio $\tilde{\lambda}$:

$$\tilde{\lambda}_a = \frac{\lambda * e^{-\|\sigma(\tilde{z}-z_a)\|_2}}{\lambda * e^{-\|\sigma(\tilde{z}-z_a)\|_2} + (1-\lambda) * e^{-\|\sigma(\tilde{z}-z_b)\|_2}},$$

where $\sigma(\cdot)$ denotes softmax function, e denotes exp and $\|\cdot\|_2$ denotes l_2 norm.

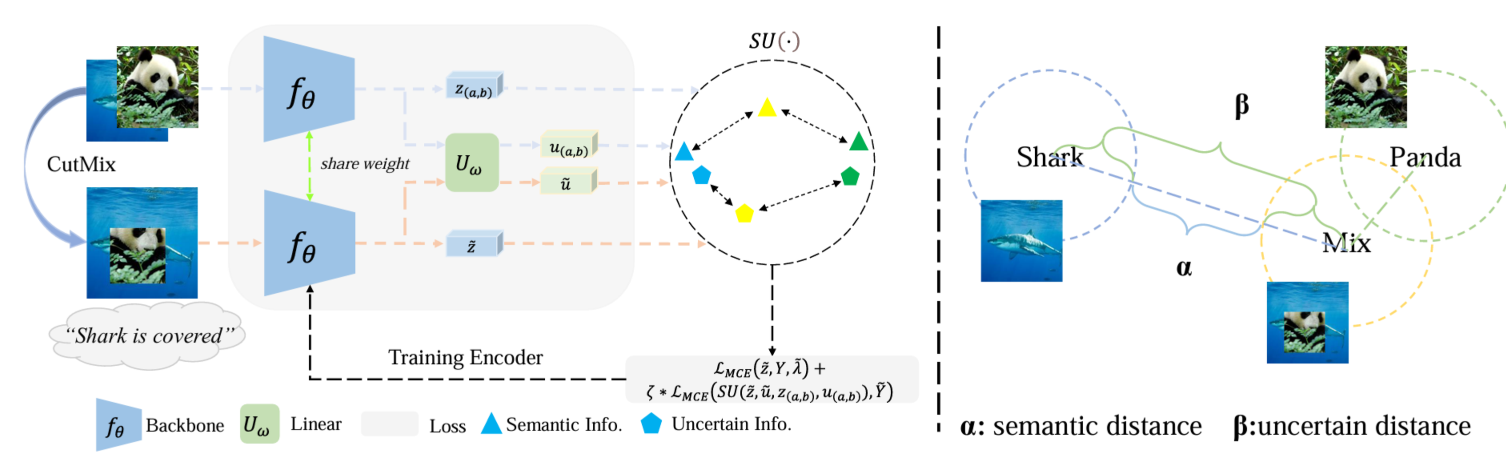

Fig 4. The pipeline of SUMix.

3. Uncertainty Estimation

SUMix uses a MLP to caputer sample uncertain information, combine the semantic information to reformulated adaptive feature vectory:

$$Z_{su} = e^{-(\beta + \|\sigma(\tilde{z}-Z)\|_2)}, \beta = \tilde{u} + u,$$

where $u = \|\sigma(MLP(z))\|_2$. When there is significant uncertainty or semantic information has a large difference, Z_{su} receives a small gradient.

• Experiments

• Image Classification

Tab 1. Top-1 accuracy(%)↑ of mixup methods on CIFAR-100, Tiny-ImageNet and ImageNet-1K. * denotes mixup methods with SUMix.

	CIFAR100			Tiny-ImageNet		ImageNet-1K
Method	ResNet18	ResNeXt50	W-ResNet28-8	ResNet18	ResNeXt50	ResNet18
CutMix	78.17	78.32	84.45	65.53	66.47	68.95
FMix	79.69	79.02	84.21	63.47	65.08	69.96
SaliencyMix	79.12	78.77	84.35	64.60	66.55	69.16
ResizeMix	80.01	80.35	84.87	63.74	65.87	69.50
CutMix*	**79.78**	**79.91**	**84.56**	**65.71**	**68.74**	**69.71**
FMix*	**80.20**	**80.79**	**84.32**	**63.69**	**67.12**	**70.48**
SaliencyMix*	**79.91**	**79.32**	**84.58**	**65.68**	**68.92**	**69.52**
ResizeMix*	**80.38**	**80.72**	**84.91**	**65.30**	**67.49**	**69.76**
Avg. Gain	**+0.82**	**+1.07**	**+0.12**	**+0.81**	**+2.07**	**+0.47**

Tab 2. Top-1 accuracy(%)↑ of mixup methods on CUB2-00 and FGVC-Aircrafts.

	CUB200		FGVC-Aircrafts	
Method	ResNet18	ResNeXt50	ResNet18	ResNeXt50
CutMix†	77.70	83.67	78.84	84.55
FMix	77.28	84.06	79.36	84.10
SaliencyMix†	75.77	82.83	79.78	84.31
ResizeMix	78.50	84.16	78.10	84.08
CutMix*	**78.20**	**83.71**	**79.72**	**85.84**
FMix*	**79.24**	**84.33**	**79.48**	**84.64**
SaliencyMix*	**76.98**	**82.84**	**79.90**	**84.49**
ResizeMix*	**78.56**	**84.23**	**80.29**	**85.12**
Avg. Gain	**+0.93**	**+0.10**	**+0.82**	**+0.67**

Tab 3. Top-1 accuracy(%)↑ of mixup methods on CIFAR100 based on ViTs.

	CIFAR100			
Method	DeiT-Small		Swin-Tiny	
CutMix	74.12		80.64	
FMix	70.41		80.72	
SaliencyMix	69.78		80.40	
ResizeMix	68.45		80.16	
CutMix*	**75.26**	+1.14	**80.83**	+0.19
FMix*	**70.69**	+0.28	**80.73**	+0.01
SaliencyMix*	**70.31**	+0.53	**80.71**	+0.29
ResizeMix*	**68.78**	+0.33	**80.59**	+0.43

Tab 4. Top-1 acc(%)↑ and FGSM error(%)↓ of ResNet18 without and with SUMix.

	Clean Acc(%)↑		Corruption Acc(%)↑		FGSM Error(%)↓	
Method	MCE	**SUMix**	MCE	**SUMix**	MCE	**SUMix**
CutMix	78.17	**79.78**	43.06	**44.31**	91.15	90.41
FMix	79.69	**80.20**	48.79	**49.14**	89.16	89.08
SaliencyMix	79.12	**79.91**	43.73	**44.36**	89.64	91.49
ResizeMix	80.01	**80.38**	46.12	**46.28**	90.04	91.05

Tab 5. Top-1 accuracy(%)↑ of saliency-based mixup methods on CIFAR100.

	CIFAR100			
Method	ResNet18		ResNeXt50	
PuzzleMix	81.13		81.69	
AutoMix	82.04		82.84	
PuzzleMix*	**81.43**	+0.30	**82.60**	+0.91
AutoMix*	**82.30**	+0.26	**83.82**	+0.98
AdAutoMix	82.32		83.81	

• Occlusion Robustness

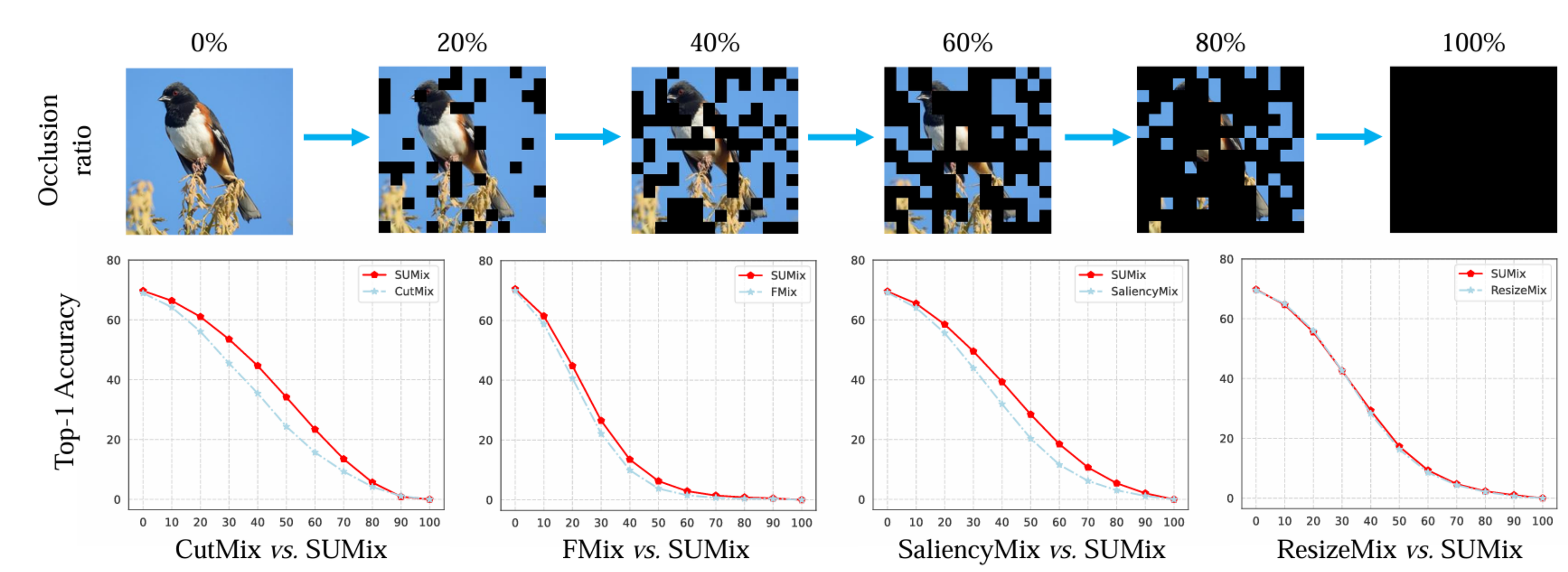

Fig 4. The top of the figure shows a visualization of the sample at 0% to 100% occlusion ratio. The lower four subfigures show the classification accuracy comparison of CutMix, FMix, SaliencyMix, ResizeMix with SUMix on ImageNet-1K using ResNet18 for 100 epochs of training.

Notes: If you'd like to know more about mixup methods, see our new work "A Survey on Mixup Augmentations and Beyond"
arXiv link: https://arxiv.org/abs/2409.05202
Github link: https://github.com/Westlake-AI/Awesome-Mixup

Survey | Project | Paper | Code | HomePage | WeChat