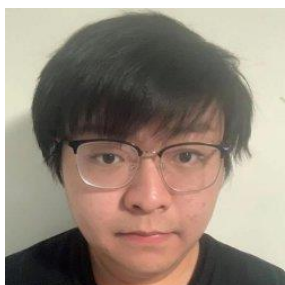


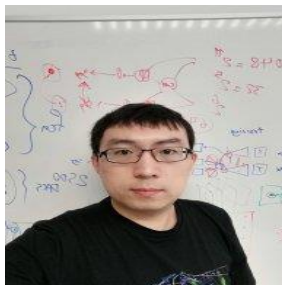
You Only Cut Once: Boosting Data Augmentation with a Single Cut



Junlin Han



Pengfei Fang



Weihao Li



Jie Hong



Ali Armin



Ian Reid



Lars Petersson



Hongdong Li



Australian
National
University



ICML | 2022

Thirty-ninth International Conference on
Machine Learning

Highlights

- A simple method for **performing** data augmentations
- ... benefits **a variety of** vision tasks, including classification, contrastive learning, and low-level vision, **for free**
- ... scales well to **almost all** augmentation operations
- ... applies to **multiple** neural network architectures

Data augmentation



Original



Horizontal flip



Color jitter



Random erasing



AutoAug



Mixup



Jitter + AutoAug (G1)



Mixing image



Vertical flip
Geometric
transformation



Gaussian blur
Photometric
transformation



Cutout
Information
dropping



RandAug
Search-based

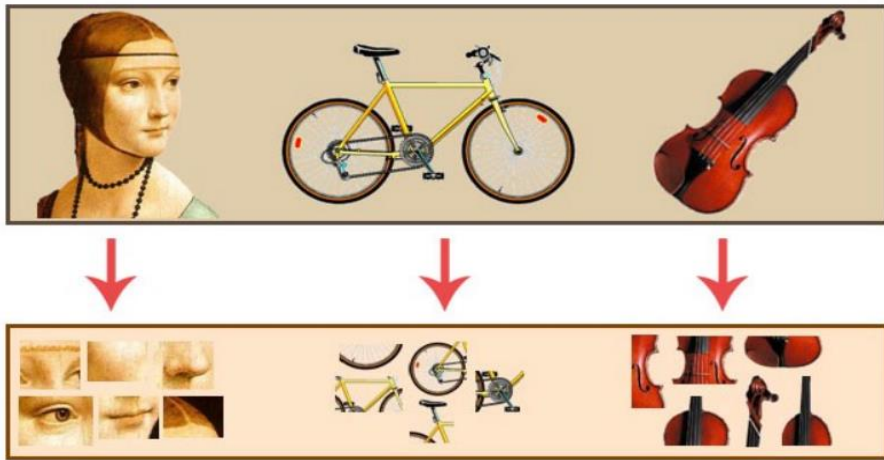


CutMix
Mix-based

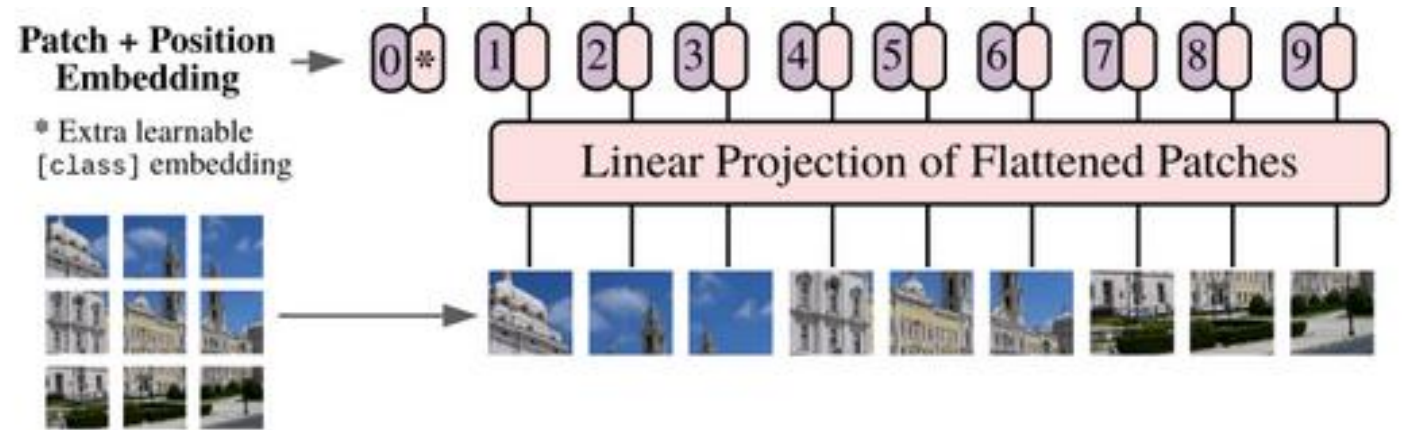


Mixup + Erasing (G2)
Group

Motivation and Intuition



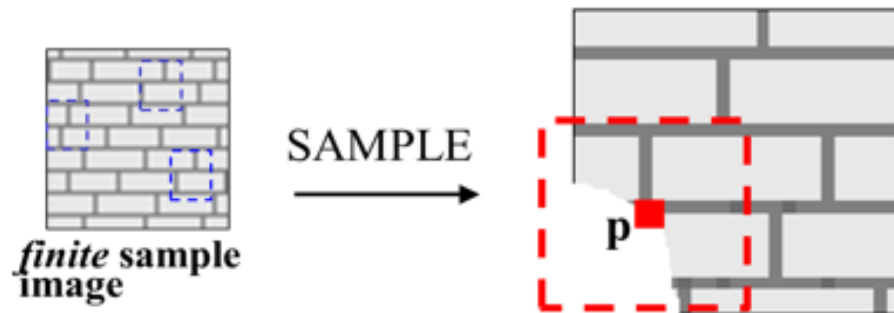
Bag of features



ViT

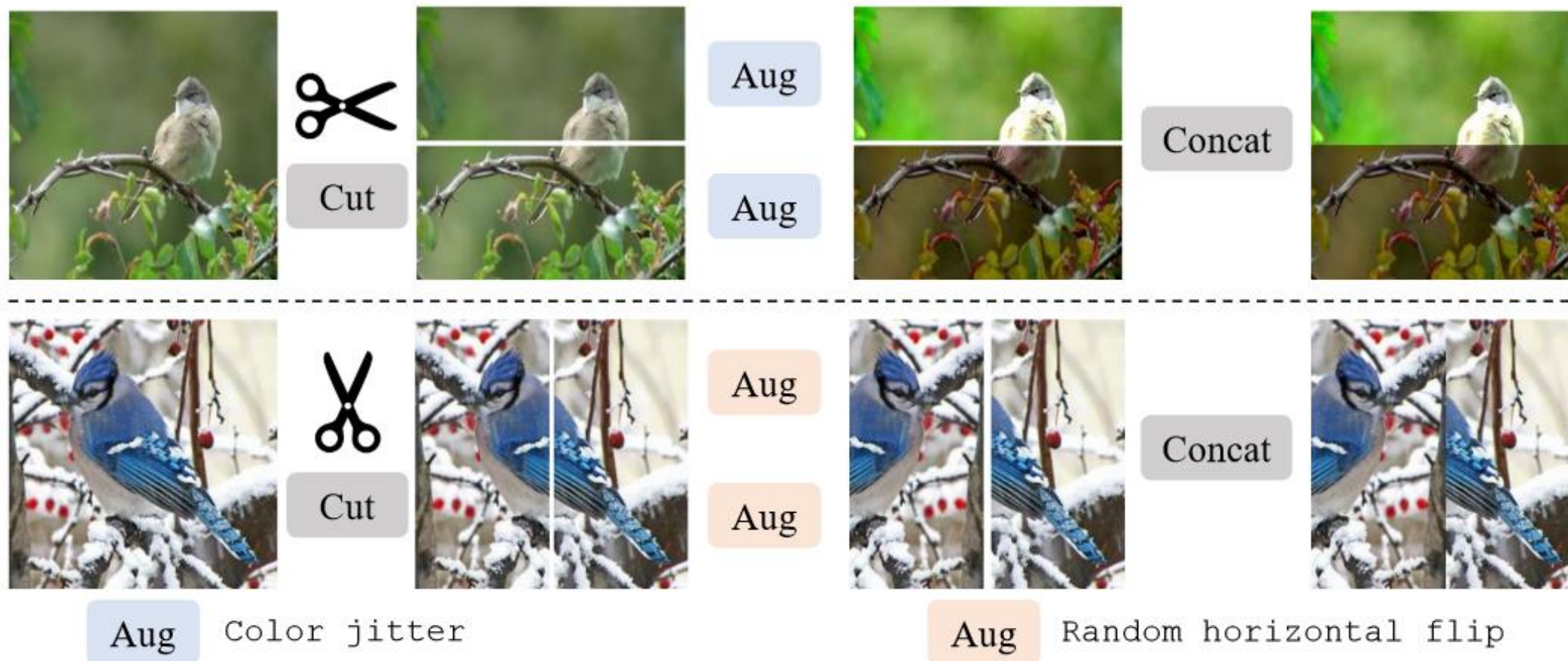


Human cognition



Texture synthesis

Our method: You Only Cut Once (YOCO)



Augmented images

H flip



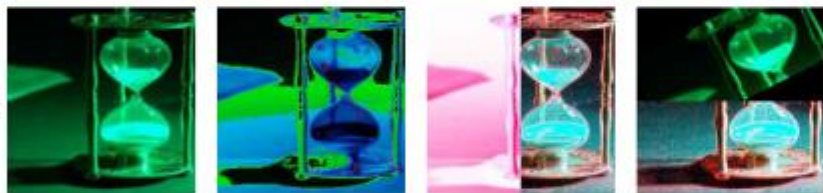
Color jitter



Erasing



AutoAug



Mixup



G1



V flip



Gaussian blur



Cutout



RandAug



CutMix



G2



Input

Image-level

YOCO

YOCO

Input

Image-level

YOCO

YOCO

Experiments

- Image Classification

CIFAR-10, CIFAR-100, and ImageNet

- Contrastive Learning

Transfer to classification, segmentation, and detection

- Low-level vision

Image deraining and image super-resolution.

Results: CIFAR-10 Classification

Models	Geometric trans		Photometric trans		Information dropping		Search-based		Mix-based		Group	
	H flip	V flip	Jitter	Blur	Erasing	Cutout	AutoAug	RandAug	Mixup	CutMix	G1	G2
PreResNet18	4.64	5.90	4.79	5.29	4.32	4.66	3.71	4.18	3.23	3.45	3.69	3.63
+ YOCO	5.05	5.52	4.70	5.00	4.28	4.36	3.38	4.08	3.11	3.38	3.56	3.25
Δ	+0.41	-0.38	-0.09	-0.29	-0.04	-0.30	-0.33	-0.10	-0.12	-0.07	-0.13	-0.38
Xception	4.62	5.20	4.52	5.31	4.07	4.15	3.56	4.01	3.70	3.66	4.04	3.41
+ YOCO	4.91	5.11	4.51	4.65	4.07	3.94	3.31	3.95	3.17	3.22	3.32	3.20
Δ	+0.29	-0.09	-0.01	-0.66	0.00	-0.21	-0.25	-0.06	-0.53	-0.44	-0.72	-0.21
DenseNet121	4.58	5.21	4.72	5.09	4.17	4.55	3.83	4.22	3.71	3.53	3.98	3.57
+ YOCO	5.09	5.18	4.70	4.76	4.13	4.25	3.67	3.91	3.28	3.50	3.50	3.36
Δ	+0.51	-0.03	-0.02	-0.33	-0.04	-0.30	-0.16	-0.31	-0.43	-0.03	-0.48	-0.21
ResNeXt50	4.69	5.12	4.74	5.66	4.18	4.93	3.65	4.09	3.69	3.52	3.79	3.66
+ YOCO	5.10	5.03	4.60	5.27	3.96	4.64	3.26	3.91	3.45	3.44	3.26	3.41
Δ	+0.41	-0.09	-0.14	-0.39	-0.22	-0.29	-0.39	-0.18	-0.24	-0.08	-0.53	-0.25
WRN-28-10	3.51	4.10	3.63	3.99	2.92	3.26	2.72	3.23	2.60	2.86	2.75	2.79
+ YOCO	3.53	4.09	3.44	3.97	2.89	2.96	2.62	3.05	2.46	2.82	2.65	2.35
Δ	+0.02	-0.01	-0.19	-0.02	-0.03	-0.30	-0.10	-0.18	-0.14	-0.04	-0.10	-0.44
ViT	4.69	5.51	4.74	5.49	4.32	4.44	3.76	4.13	3.67	3.49	3.82	3.27
+ YOCO	4.92	5.34	4.70	4.95	4.18	4.42	3.54	3.96	3.17	3.21	3.51	3.14
Δ	+0.23	-0.17	-0.04	-0.54	-0.14	-0.02	-0.22	-0.17	-0.50	-0.28	-0.31	-0.13
Swin	4.83	5.27	4.90	5.15	4.26	4.55	3.63	4.15	3.56	3.64	3.74	3.58
+ YOCO	4.85	5.25	4.40	5.12	4.23	4.03	3.54	4.13	3.16	3.22	3.44	3.46
Δ	+0.02	-0.02	-0.50	-0.03	-0.03	-0.52	-0.09	-0.02	-0.40	-0.42	-0.30	-0.12
Average Δ	+0.27	-0.11	-0.14	-0.32	-0.07	-0.28	-0.22	-0.15	-0.34	-0.19	-0.37	-0.25

Metric: Top-1 error rate

76/84 results are improved

Results: CIFAR-100 Classification

Models	Geometric trans		Photometric trans		Information dropping		Search-based		Mix-based		Group	
	H flip	V flip	Jitter	Blur	Erasing	Cutout	AutoAug	RandAug	Mixup	CutMix	G1	G2
PrcResNet18 + YOCO Δ	23.36 24.66 +1.30	25.16 24.73 -0.43	23.95 23.62 -0.33	25.96 25.11 -0.85	24.03 23.94 -0.09	25.01 23.41 -1.60	21.94 20.55 -1.39	22.70 22.70 0.00	21.32 19.62 -1.70	20.05 19.30 -0.75	22.97 21.06 -1.91	21.23 20.92 -0.31
Xception + YOCO Δ	23.34 24.46 +1.12	25.16 24.65 -0.51	23.39 23.42 +0.03	26.17 25.51 -0.66	25.01 23.74 -1.27	25.44 23.56 -1.88	22.50 20.29 -2.21	22.87 22.45 -0.42	21.41 19.96 -1.45	20.17 19.25 -0.92	23.02 20.75 -2.27	21.79 20.68 -1.11
DenseNet121 + YOCO Δ	23.14 24.21 +1.07	23.71 24.62 +0.91	24.25 23.14 -1.11	26.21 24.68 -1.53	23.97 23.34 -0.63	24.71 23.83 -0.88	22.43 20.63 -1.80	22.66 23.11 +0.45	21.17 19.87 -1.30	19.50 19.26 -0.24	22.80 21.13 -1.67	21.23 20.95 -0.28
ResNeXt50 + YOCO Δ	24.60 25.11 +0.51	24.61 24.22 -0.39	23.58 23.76 +0.18	25.86 25.66 -0.20	24.50 23.51 -0.99	24.84 23.77 -1.07	22.32 21.01 -1.31	22.77 22.47 -0.30	20.78 19.74 -1.04	19.72 19.68 -0.04	22.69 20.58 -2.11	21.69 20.55 -1.14
WRN-28-10 + YOCO Δ	19.90 20.44 +0.54	21.70 21.34 -0.36	19.54 19.59 +0.05	21.95 21.17 -0.78	19.79 19.65 -0.14	19.73 19.68 -0.05	18.07 17.17 -0.90	19.25 19.29 +0.04	18.93 17.64 -1.29	17.03 17.29 +0.26	18.39 18.89 +0.50	17.74 17.72 -0.02
ViT + YOCO Δ	23.67 24.52 +0.85	24.01 24.74 +0.73	24.53 23.49 -1.04	26.06 25.04 -1.02	24.09 23.20 -0.89	25.10 23.38 -1.72	22.14 20.61 -1.53	22.79 22.34 -0.45	21.15 19.26 -1.89	20.27 19.17 -1.10	22.75 20.85 -1.90	21.64 21.20 -0.44
Swin + YOCO Δ	23.57 24.66 +1.09	24.51 24.37 -0.14	24.04 23.96 -0.08	25.49 25.39 -0.10	24.81 23.22 -1.59	25.11 23.86 -1.25	22.30 20.70 -1.60	23.14 22.62 -0.52	21.42 19.43 -1.99	19.78 19.74 -0.04	22.92 20.75 -2.17	21.26 20.88 -0.38
Average Δ	+0.93	-0.03	-0.33	-0.73	-0.80	-1.21	-1.53	-0.17	-1.52	-0.40	-1.79	-0.53

Metric: Top-1 error rate

68/84 results are improved

Results: ImageNet Classification

Augs	Generalization	Partial	Calibration	Adversarial attacks		Corruptions		Distribution shift
	Clean	Clean	Clean, RMS↓	FGSM attack	PGD attack	Random replace	Gaussian noise	ImageNet-A
H flip	77.10/76.59	55.21/56.14	6.42/8.81	17.28/21.01	9.37/14.98	58.90/58.77	73.58/72.92	3.39/4.29
+ YOCO	77.28/77.01	56.27/56.44	7.87/8.19	20.59/21.88	14.21/15.21	59.11/58.91	73.63/72.96	3.80/4.32
Δ	+0.18/+0.42	+1.06/+0.30	+1.45/-0.62	+3.31/+0.87	+4.84/+0.23	+0.21/+0.14	+0.05/+0.04	+0.41/+0.03
Jitter	77.15/76.87	56.04/56.34	8.06/8.19	18.31/18.80	11.07/11.89	59.11/58.82	73.74/73.38	3.98/4.13
+ YOCO	77.35/77.12	56.41/56.60	7.50/8.07	16.32/17.77	5.96/6.90	59.40/59.32	74.46/74.14	4.29/4.56
Δ	+0.20/+0.25	+0.37/+0.26	-0.56/-0.12	-1.99/-1.03	-5.11/-4.99	+0.29/+0.50	+0.72/+0.76	+0.31/+0.43
Erasing	77.40/77.09	56.75/56.82	7.67/7.88	22.01/22.67	12.74/13.31	58.75/58.44	73.81/73.48	4.11/4.32
+ YOCO	77.29/77.20	56.66/56.76	7.57/8.00	22.34/22.82	12.93/14.00	58.80/58.85	73.95/73.59	4.12/4.28
Δ	-0.11/+0.11	-0.09/-0.06	-0.10/+0.12	+0.33/+0.15	+0.19/+0.69	+0.05/+0.41	+0.14/+0.11	+0.01/-0.04
AutoAug	77.55/76.93	55.61/55.05	6.17/6.73	11.51/11.62	2.05/2.18	58.36/58.12	74.45/73.85	4.54/4.69
+ YOCO	77.88/77.65	56.48/56.52	6.04/6.25	12.09/11.73	1.11/0.86	58.90/58.29	76.03/75.70	5.01/4.80
Δ	+0.33/+0.72	+0.87/+1.47	-0.13/-0.48	+0.58/+0.11	-0.94/-1.32	+0.54/+0.17	+1.58/+1.85	+0.47/+0.11
Mixup	77.72/77.67	55.27/55.40	8.23/9.03	35.92/35.41	7.90/7.44	61.46/61.07	75.47/75.12	8.37/8.59
+ YOCO	77.81/77.74	56.00/55.97	4.04/4.42	39.63/39.67	13.02/13.63	61.14/60.85	75.10/74.88	8.27/8.55
Δ	+0.09/+0.07	+0.73/+0.57	-4.19/-4.61	+3.71/+4.26	+5.12/+6.19	-0.32/-0.22	-0.37/-0.24	-0.10/-0.04
G1	77.61/77.19	55.57/55.24	6.29/6.19	13.11/13.22	2.34/2.87	58.94/58.99	75.22/74.94	5.01/5.22
+ YOCO	77.76/77.64	56.00/55.97	6.17/6.02	13.62/13.61	1.82/2.61	59.15/59.03	75.73/75.59	5.41/5.47
Δ	+0.15/+0.45	+0.43/+0.73	-0.12/-0.17	+0.51/+0.39	-0.52/-0.26	+0.21/+0.04	+0.51/+0.65	+0.40/+0.25

Metric: Top-1 accuracy
best/last results

YOCO **outperforms** image-level augmentation

Results: Contrastive Learning

Method	ImageNet classification			VOC detection			COCO detection			COCO instance seg		
	linear protocol	1% label	10% label	AP ₅₀	AP	AP ₇₅	AP ₅₀	AP	AP ₇₅	AP ₅₀ ^{mask}	AP ^{mask}	AP ₇₅ ^{mask}
MoCo v2	67.5	34.5	61.1	81.9	56.8	62.9	57.3	38.0	40.8	54.1	33.3	35.5
+ YOCO	67.6	34.9	61.5	82.4	56.6	63.6	57.4	38.1	41.6	54.1	33.4	35.5
SimSiam	68.1	17.1	57.3	80.2	54.5	60.0	51.4	33.5	36.1	48.7	29.9	32.0
+YOCO	68.3	17.3	57.7	80.2	54.5	60.1	53.1	34.6	37.3	50.1	30.8	32.9

YOCO results in more powerful representation

Results: Low-level Vision

Methods	Test100	Rain100H	Rain100L	Test2800	Test1200	Average
MPRNet	30.27	30.41	36.40	33.64	32.91	32.73
+ YOCO	30.33	30.53	37.13	33.64	32.84	32.89

Deraining

Aug	Synthetic				Realistic RealSR
	DIV2K	Set14	Urban100	Manga109	
No aug	28.83	28.49	25.82	30.11	28.78
+ Image-level	28.83	28.48	25.82	30.08	28.99
+ YOCO	28.84	28.50	25.83	30.11	28.99

Super-resolution

Metric: PSNR

YOCO **generalizes** well to low-level vision tasks

Conclusion

- YOOCO boosts multiple augmentations for free
- How to perform data augmentations can be further explored
- Code available: <https://github.com/JunlinHan/YOOCO>