You Only Cut Once
Boosting Data Augmentation with a Single Cut

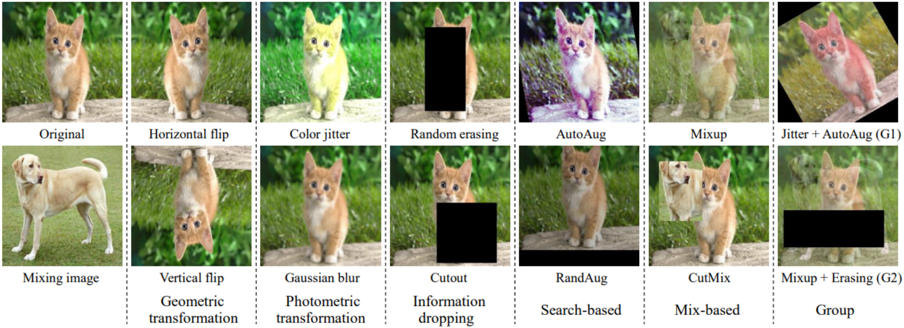
Yi-Cheng Hung[1], Wei-Chih Wong[2],
Tsung-Tse Feng[3], Ming-Chieh Kuo[3]
Group 8

Department of [1] IMM, [2] AIG, [3] IAM
National Yang Ming Chiao Tung University

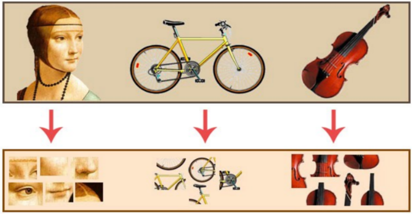
December 12, 2022

Outline

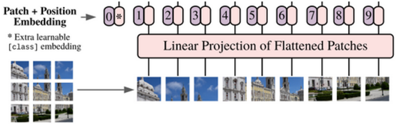
Data Augmentation



Motivation

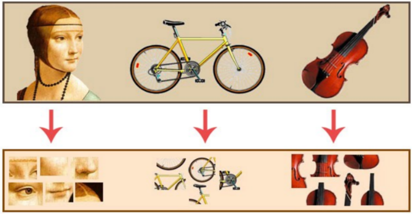


Bag of features

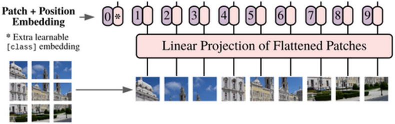


ViT

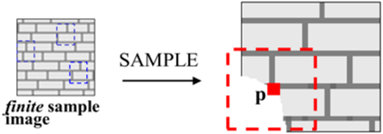
Motivation



Bag of features



ViT



Texture synthesis

Human cognition

Method (You Only Cut Once)

The augmentation $a(\cdot) : \mathrm{R}^{C\times H\times W} \longrightarrow \mathrm{R}^{C\times H\times W}$,
is applied separately with in each piece.

$$[X_1, X_2] = cut_H(X), \text{ if } 0 < p \leq 0.5, \text{ s.t. } X_i \in \mathrm{R}^{C\times \frac{H}{2}\times W}$$

$$[X_1, X_2] = cut_W(X), \text{ if } 0 < p \leq 0.5, \text{ s.t. } X_i \in \mathrm{R}^{C\times H\times \frac{W}{2}}$$

$$X' = concate[a_1(X_1), a_2(X_2)]$$

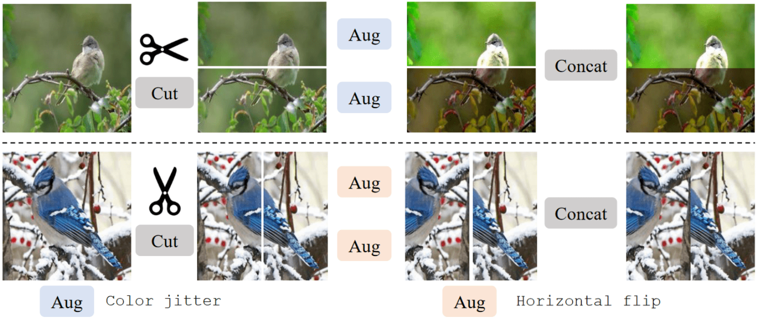
Method (You Only Cut Once)

The augmentation $a(\cdot) : \mathrm{R}^{C\times H\times W} \longrightarrow \mathrm{R}^{C\times H\times W}$,
is applied separately with in each piece.

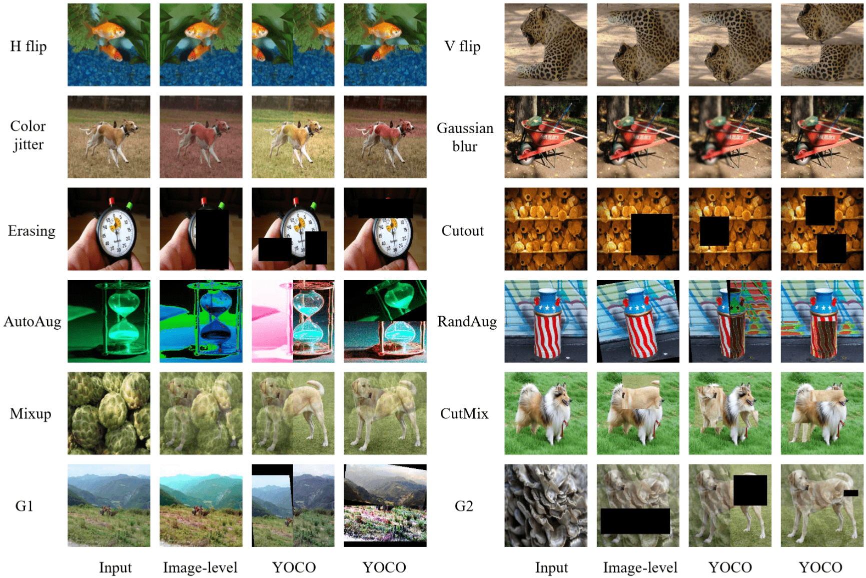
$$[X_1, X_2] = cut_H(X), \text{ if } 0 < p \leq 0.5, \text{ s.t. } X_i \in \mathrm{R}^{C\times \frac{H}{2}\times W}$$

$$[X_1, X_2] = cut_W(X), \text{ if } 0 < p \leq 0.5, \text{ s.t. } X_i \in \mathrm{R}^{C\times H\times \frac{W}{2}}$$

$$X' = concate[a_1(X_1), a_2(X_2)]$$



Visualization



Comparison (Classification CIFAR-10)

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
△	+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
△	+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
△	+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
△	+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
△	+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
△	+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
△	+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 △	+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 are improved.

Comparison (Classification CIFAR-100)

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

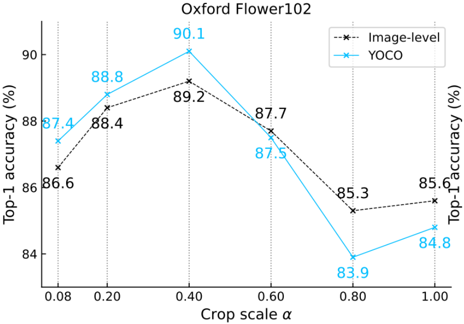
68/84 are improved.

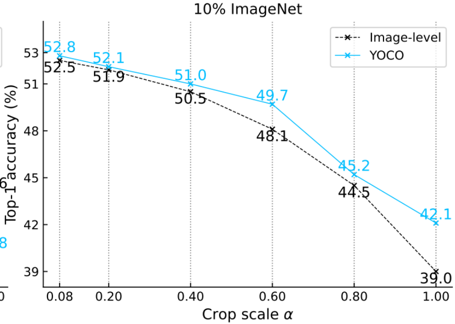
Number of cuts

Number H,W	Augmentations H flip	Jitter	Erasing	AutoAug	Mixup
1,1	70.8	76.1	76.0	77.8	79.4
1,2	69.1	75.9	76.0	77.2	79.2
1,3	68.0	75.5	75.9	76.7	79.2
2,1	69.7	76.5	75.9	76.9	80.1
2,2	68.7	75.9	75.6	76.5	79.6
2,3	67.9	75.7	75.6	76.2	79.5
3,1	69.9	75.9	75.8	77.0	80.2
3,2	67.6	75.8	75.5	76.3	79.8
3,3	67.4	75.2	75.5	76.0	79.6
YOCO	75.6	76.6	76.4	79.4	80.3

The number of cut in picture.

When to employ YOCO?





Low-Level Vision (Image Deraining)

(a) Input (b) Groundtruth (c) Derained results (d) Rain layer

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

Metric : PSNR

Low-Level Vision (Image super-resolution)

Aug	Synthetic DIV2K	Set14	Urban100	Manga109	Realistic RealSR
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

Model: CARN (Ahn et al., 2018) model
Metric : PSNR

Conclusion

Advantages

- A simple method for performing data augmentations
- Benefits in a variety of vision tasks:
  1. classification
  2. image deraining
  3. image super-resolution

Disadvantages

- Applying YOCO can be harmful to certain classes in the supervised classification task, as the effect of data augmentaion is class dependent.
- Designing a class-specific strategy of applying YOCO is desirable.

Reference

1. Junlin Han, at el, You Only Cut Once: Boosting Data Augmentation with a Single Cut , *ICML, 2022*

Student Information

Name	ID	Dept.
Yi-Cheng Hung	310653005	IMM
Wei-Chih Wong	311505013	AIG
Tsung-Tse Feng	310652025	IAM
Ming-Chieh Kuo	310652002	IAM

Thanks for listening!