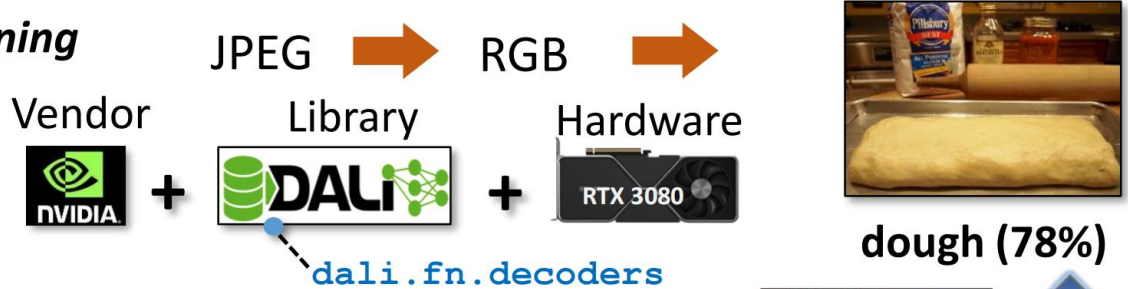


SYSNOISE: EXPLORING AND BENCHMARKING TRAINING-DEPLOYMENT SYSTEM INCONSISTENCY

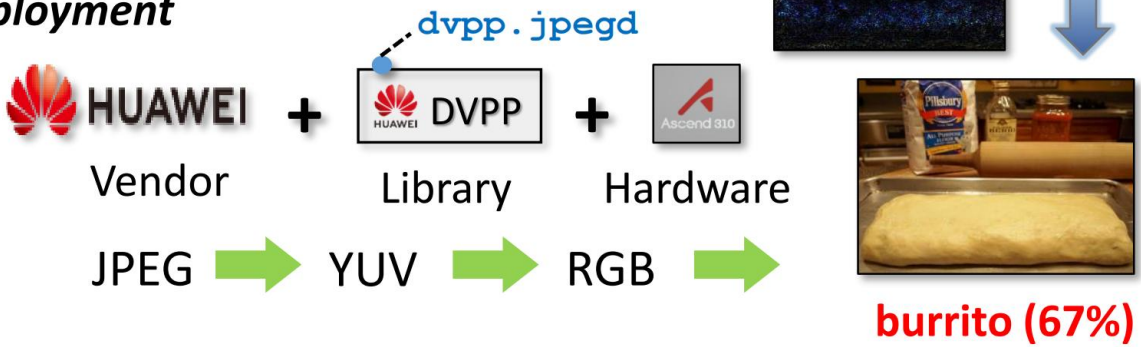
YanWang, YuhangLi, RuihaoGong, AishanLiu, YanfeiWang, JianHu, YongqiangYao, Yunchen
Zhang, Tianzi Xiao, Fengwei Yu, Xianglong Liu

What is SysNoise and how does it impact on practical applications?

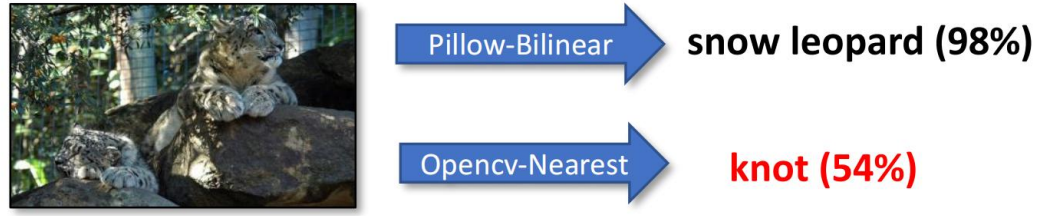
● **Training**



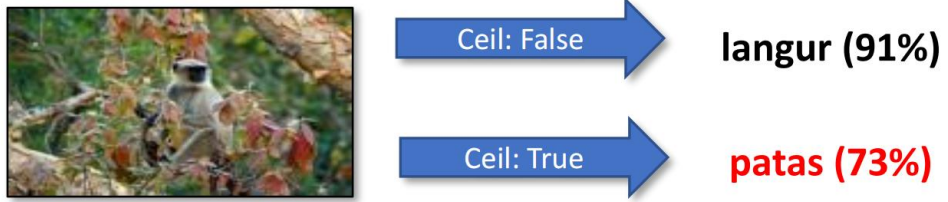
● **Deployment**



(a) Illustration of SysNoise



(b) Different Resize

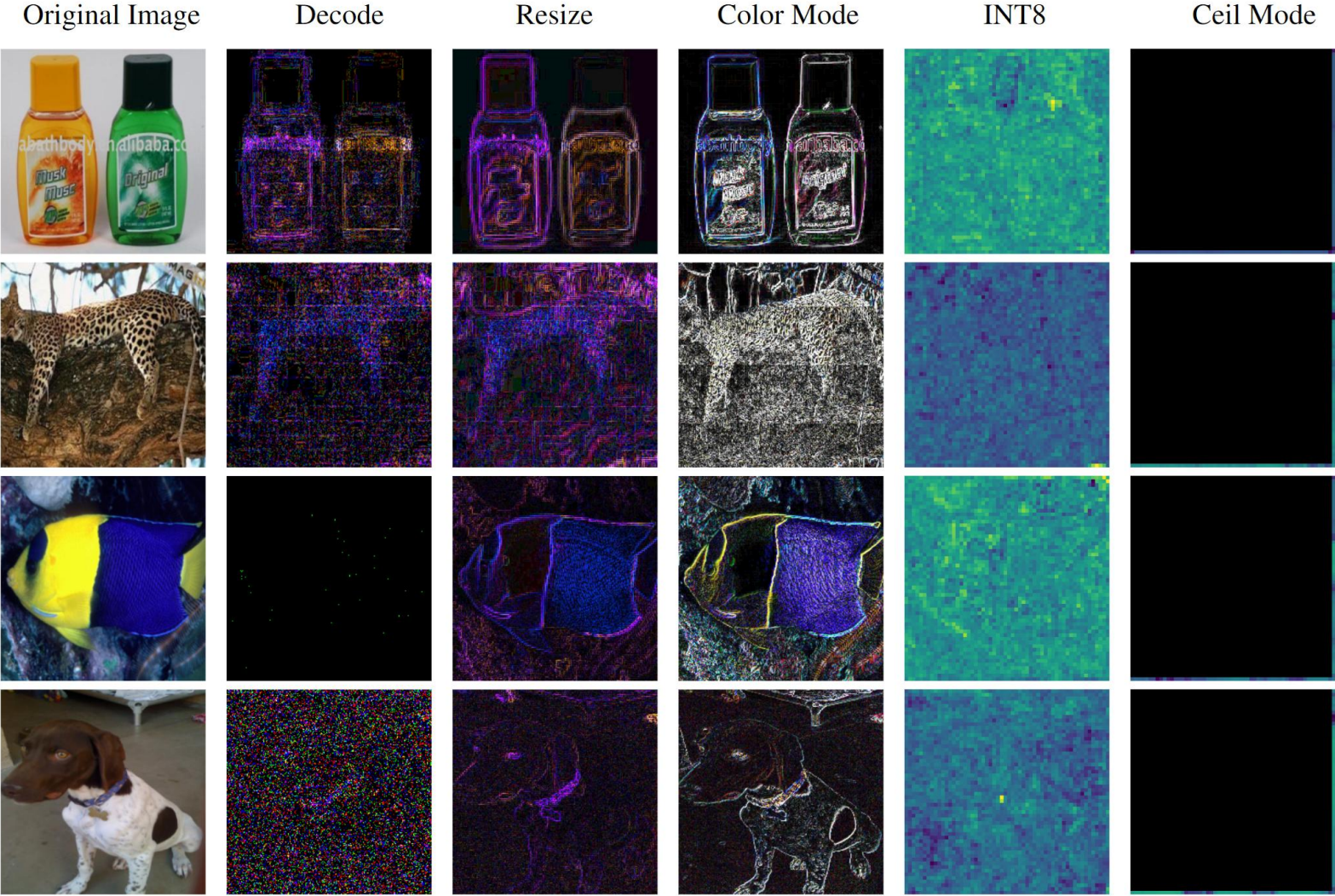


(c) Different Ceil Mode



(d) Different Detection Post-processing

What is SysNoise and how does it impact on practical applications?



Visualization of SysNoise. To make the noise more perceptible, we scale it to [0, 255].

What is SysNoise and how does it impact on practical applications?

- Classify the system noise that may be encountered during the ML Pipeline process
- Evaluate the impact on image classification, object detection, instance segmentation, NLP tasks
- In addition to studying individual noise, we study the impact of noise accumulation on multiple SysNoise on the model
- Explore whether the existing model robustness enhancement methods can alleviate system noise

Stage	Pre-processing			Model inference			Post-processing
	Decoder	Resize	Color Space	Ceil Mode	Upsample	Data Prec.	Detection Proposal
Type							
Task	Cls/Det/Seg	Cls/Det/Seg	Cls/Det/Seg	Cls/Det/Seg	Det/Seg	Cls/Det/Seg/NLP	Det
Input Dependence	✗	✗	✓	✗	✗	✓	✗
Noise Effect Level	High	Very High	Middle	High	Very High	High	Middle
Number of Categories	4	11	2	2	2	3	2
Occurrence Frequency	Very High	Very High	High	High	Middle	High	Middle

Effect of a single SysNoise on model performance: Image Classification Task

Lightweight models suffer hugely

Architecture	Trained	Decode	Resize	Color Mode	Precision (FP16/INT8)		Ceil Mode	Combined
	ACC	Δ ACC	Δ ACC	Δ ACC	Δ ACC	Δ ACC	Δ ACC	Δ ACC
MCUNet-293KB	63.40	0.41 (0.42)	4.02 (9.31)	0.20	0.01	0.04	-	9.97
ResNet18x0.25	48.96	1.98 (2.12)	2.11 (3.71)	0.14	-0.01	0.82	2.34	6.61
ResNet18x0.5	61.64	1.67 (1.76)	1.76 (3.25)	0.19	-0.01	0.15	2.72	6.10
ResNet-18	69.96	1.02 (1.03)	1.01 (2.05)	0.13	0.00	0.20	2.40	4.97
ResNet-34	73.59	0.99 (1.00)	0.77 (1.67)	0.14	0.00	0.04	0.85	4.25
ResNet-50	76.39	0.98 (0.98)	0.75 (1.69)	0.09	0.00	0.06	1.24	3.95
ResNet-101	78.10	0.68 (0.69)	0.62 (1.47)	0.24	0.01	0.69	0.75	4.50
MobileNetV2-0.5	64.94	1.98 (2.00)	2.04 (3.14)	0.18	0.01	0.57	-	5.81
MobileNetV2-0.75	70.26	1.39 (1.39)	1.47 (2.56)	0.16	0.01	0.72	-	5.58
MobileNetV2-1	73.12	1.39 (1.39)	1.48 (2.43)	0.07	0.02	0.77	-	5.03
MobileNetV2-1.4	75.84	1.01 (1.02)	1.65 (2.15)	0.10	0.01	0.53	-	5.04
RegNetX-400M	70.97	1.63 (1.63)	1.42 (2.65)	0.07	0.01	0.09	-	5.70
RegNetX-800M	74.04	1.12 (1.14)	0.97 (2.00)	0.19	0.02	0.24	-	4.38
RegNetX-1.6G	76.29	0.84 (0.85)	0.79 (1.88)	0.20	0.01	0.19	-	4.15
RegNetX-3.2G	77.89	0.61 (0.62)	0.53 (1.42)	0.20	0.00	0.24	-	3.70
EfficientNet-B0	76.83	0.75 (0.76)	1.70 (3.79)	0.15	0.03	0.19	-	4.39
EfficientNet-B1	78.13	0.57 (0.58)	1.18 (2.84)	0.26	0.01	0.39	-	3.26
EfficientNet-B2	79.97	0.57 (0.58)	1.13 (2.31)	0.05	0.04	0.41	-	3.10
EfficientNet-B3	82.03	0.71 (0.72)	0.99 (1.74)	0.16	0.05	0.38	-	2.65
EfficientNet-B4	83.43	0.29 (0.30)	0.45 (0.93)	0.17	0.02	0.26	-	2.32
ViT-Tiny	75.61	1.04 (1.04)	0.99 (1.79)	0.46	0.01	0.68	-	3.21
ViT-Small	81.58	0.57 (0.58)	0.37 (1.01)	0.80	-0.01	0.80	-	2.68
Vit-Base	84.63	0.61 (0.62)	0.43 (0.74)	0.93	-0.01	1.12	-	2.89
Swin-Tiny	81.32	0.18 (0.19)	0.42 (1.76)	1.21	0.00	0.76	-	4.93
Swin-Small	83.03	0.18 (0.18)	0.23 (1.33)	1.00	0.00	0.45	-	3.51
Swin-Base	83.54	0.11 (0.30)	0.21 (1.27)	0.97	-0.01	0.55	-	3.59

The influence of SysNoise decreases as the model parameters increase

ResNet50 will lose 4% accuracy under the cumulative influence of SysNoise

The lightweight model is more affected, and the robustness improvement brought by the increase of parameters is limited

The Transformer model is affected by this noise just like the CNN models

Effect of a single SysNoise on model performance: Object Detection Task

Due to the complexity of the model, the object detection model will have more system noise than the image classification model

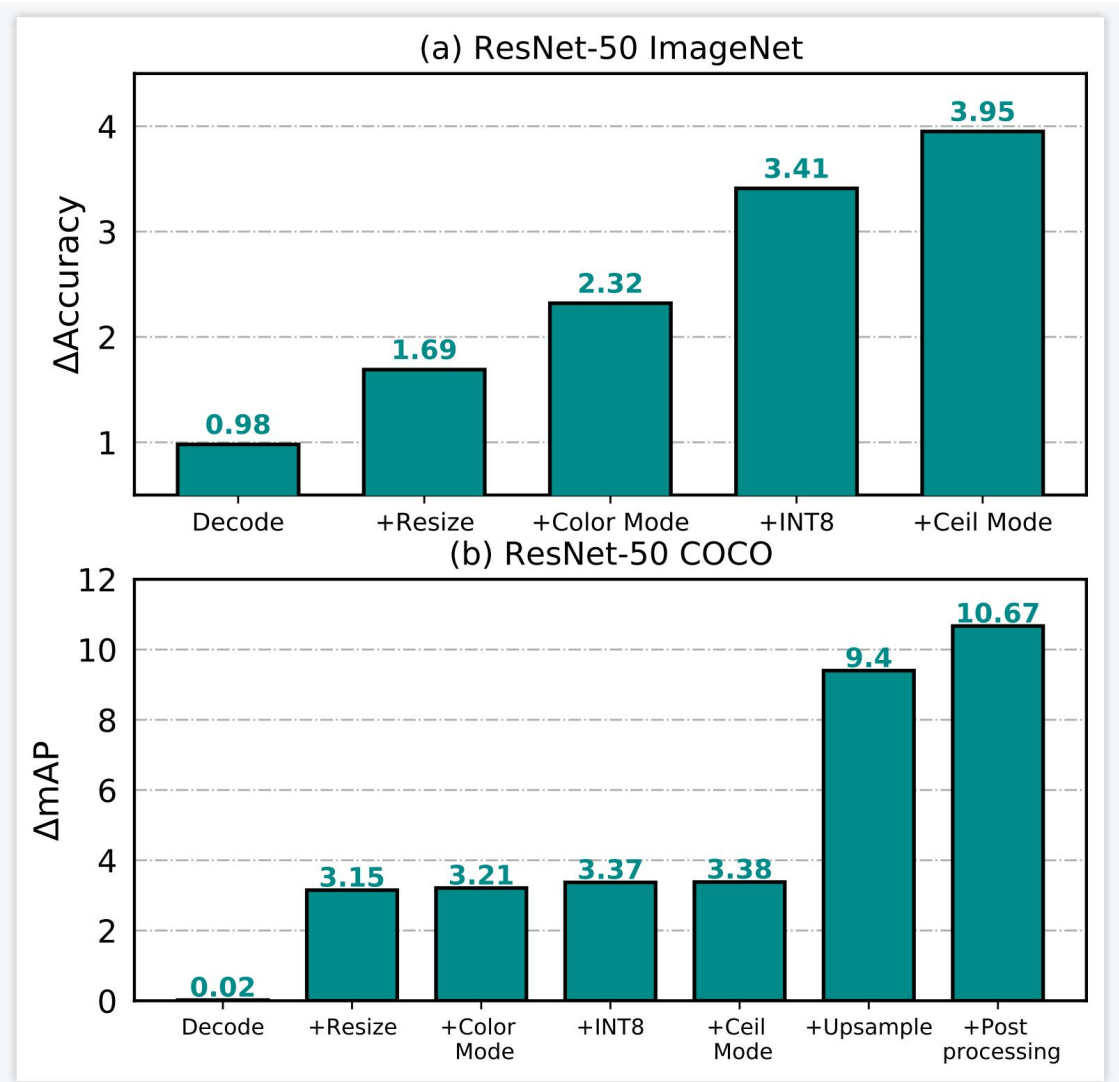
The cumulative effect of noise is more pronounced

Method	Architecture	Trained	Decode	Resize	Color Mode	Upsample	Precision INT8	Ceil Mode	Post-processing	Combined
		mAP	Δ mAP	Δ mAP	Δ mAP	Δ mAP	Δ mAP	Δ mAP	Δ mAP	Δ mAP
Faster RCNN	ResNet-34	36.76	0.02 (0.04)	0.93 (2.63)	0.25	1.28	0.06	2.50	2.29	10.25
	ResNet-50	37.36	0.02 (0.01)	1.12 (3.15)	0.10	1.66	0.10	3.14	2.39	10.67
	MobileNetV2	30.32	0.01 (0.01)	0.38 (1.14)	0.24	0.96	0.07	-	2.23	3.45
RetinaNet	ResNet-34	35.71	0.01 (0.01)	0.77 (2.20)	0.29	0.35	0.10	2.72	3.44	8.21
	ResNet-50	36.59	0.01 (0.02)	0.99 (2.78)	0.36	0.69	0.03	3.12	3.00	8.93

Compared with image classification models, object detection models are more affected by resize noise

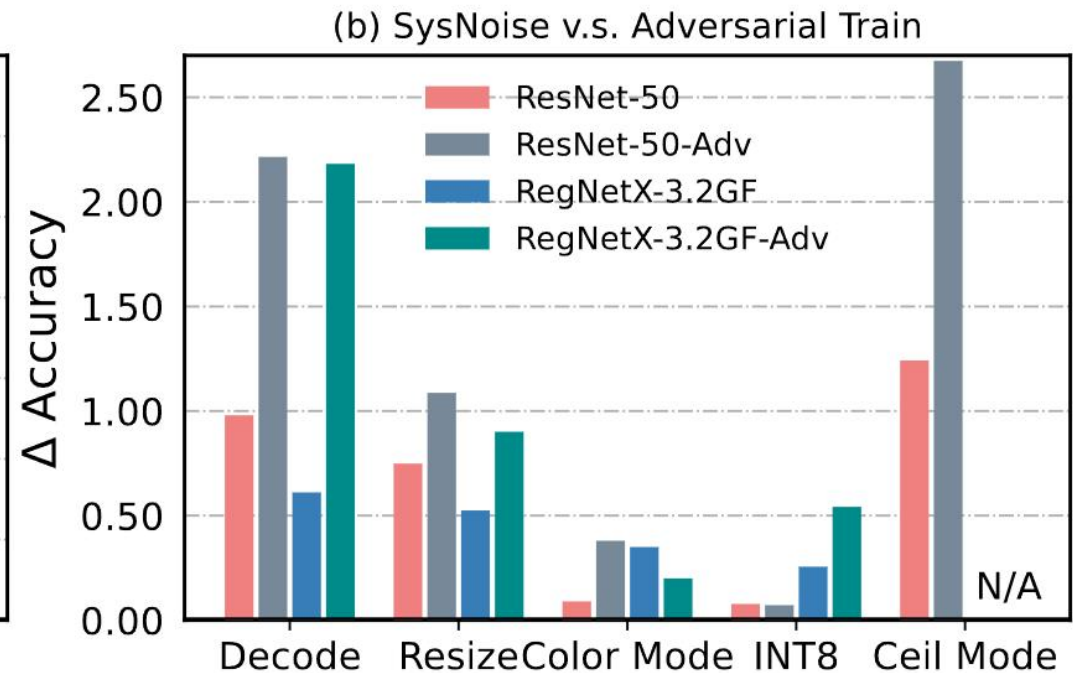
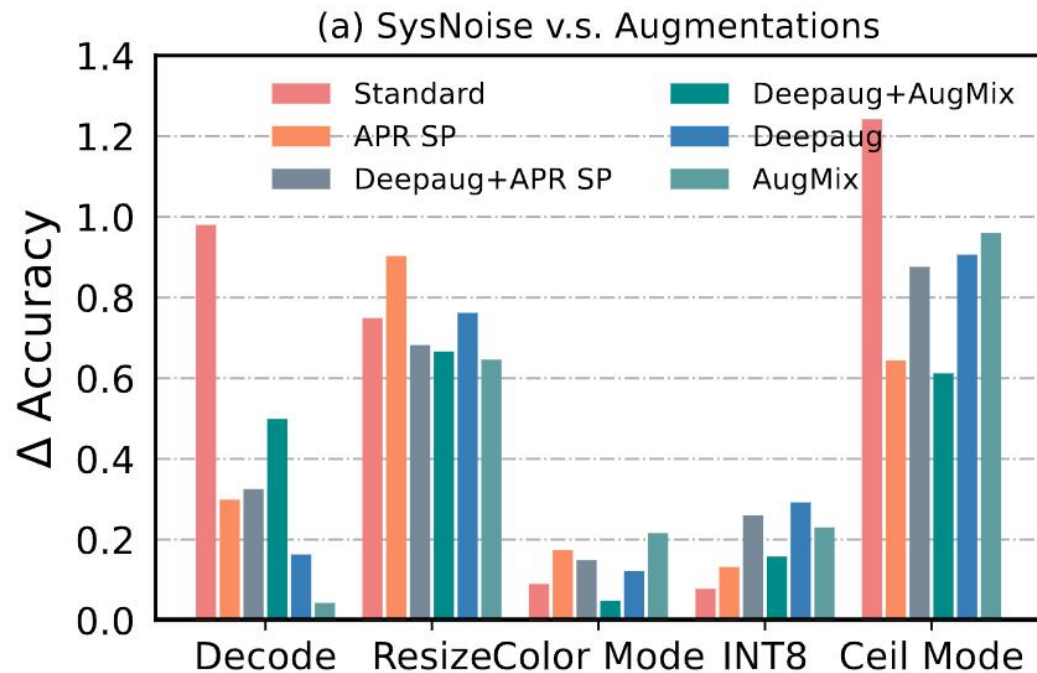
The impact of multiple SysNoise accumulations on model performance

- The cumulative effect of noise is still evident and will continue to affect model performance throughout the ML Pipeline
- Some single noises cause less damage to the model, but the combination of these noises may have a greater impact on model performance ($1+1>2$)



Possible methods against SysNoise

- Test methods against for other noise (natural noise, adversarial noise)
 - Data Augmentation
 - Adversarial Training
 - Test-time Adaptation
- The previous noise mitigation method fails on the new noise type, which proves that the data distribution of this noise is different from that of the previously proposed noise



Future Work

- Explore the effect of SysNoise on other task
- Find effective ways to enhance model robustness against system noise