

Mimicking Human Visual Development for Learning Robust Image Representations



Ankita Raj¹

Kaashika Prajaapat²

Tapan Kumar Gandhi¹

Chetan Arora¹

¹Indian Institute of Technology Delhi, India

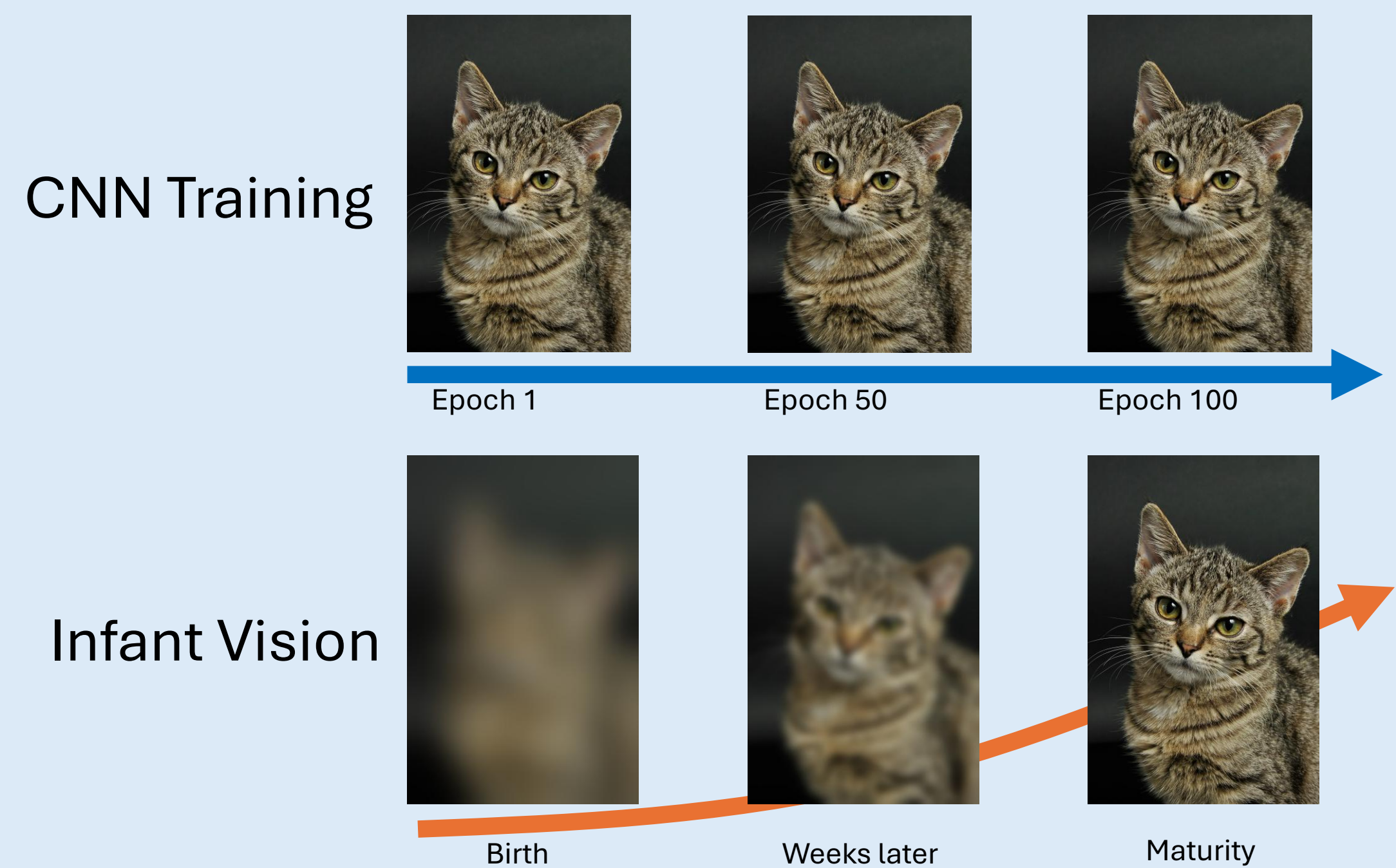
²Clinikally, India

The Problem: Brittle CNN Vision

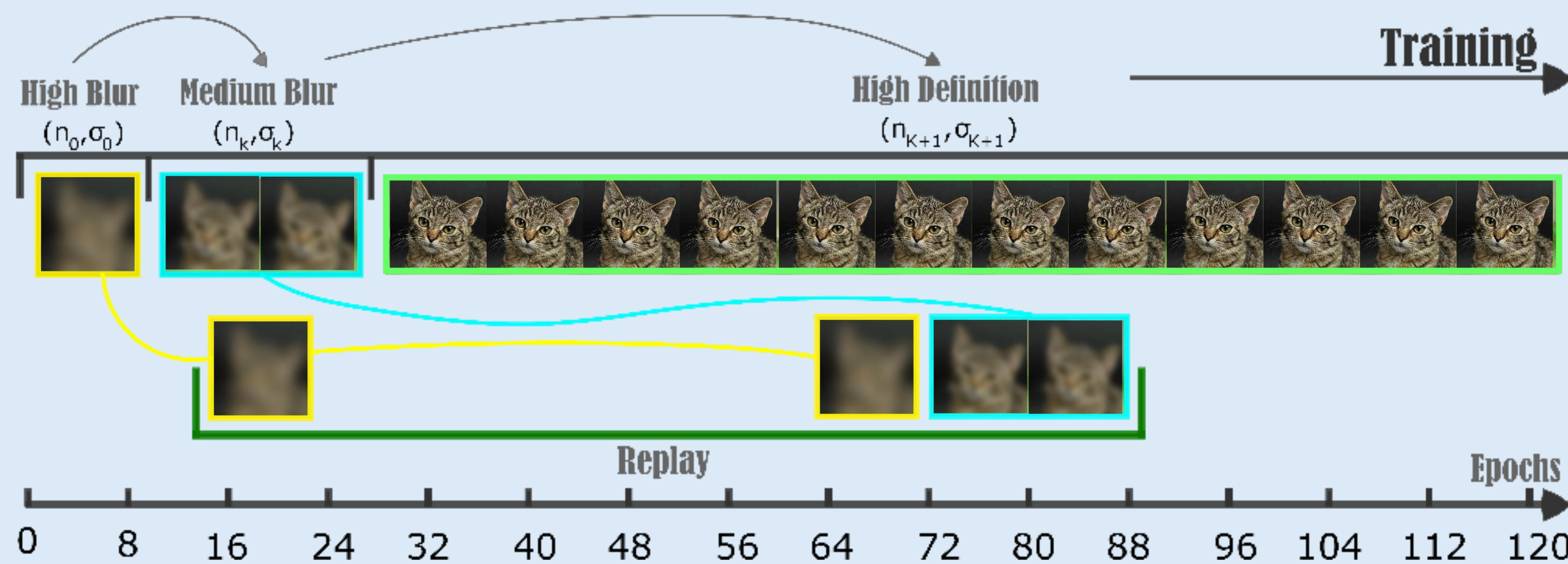


CNNs suffer a drop in accuracy when test distribution differs from training distribution – Common image corruptions or Adversarial Perturbations.

CNN vs. Human Vision



Proposed Solution: Visual Acuity Curriculum (VAC)



High Initial Blurring:

Train on highly blurred images in the initial epochs.

Gradual Refinement:

Gradually decrease blur over the epochs.

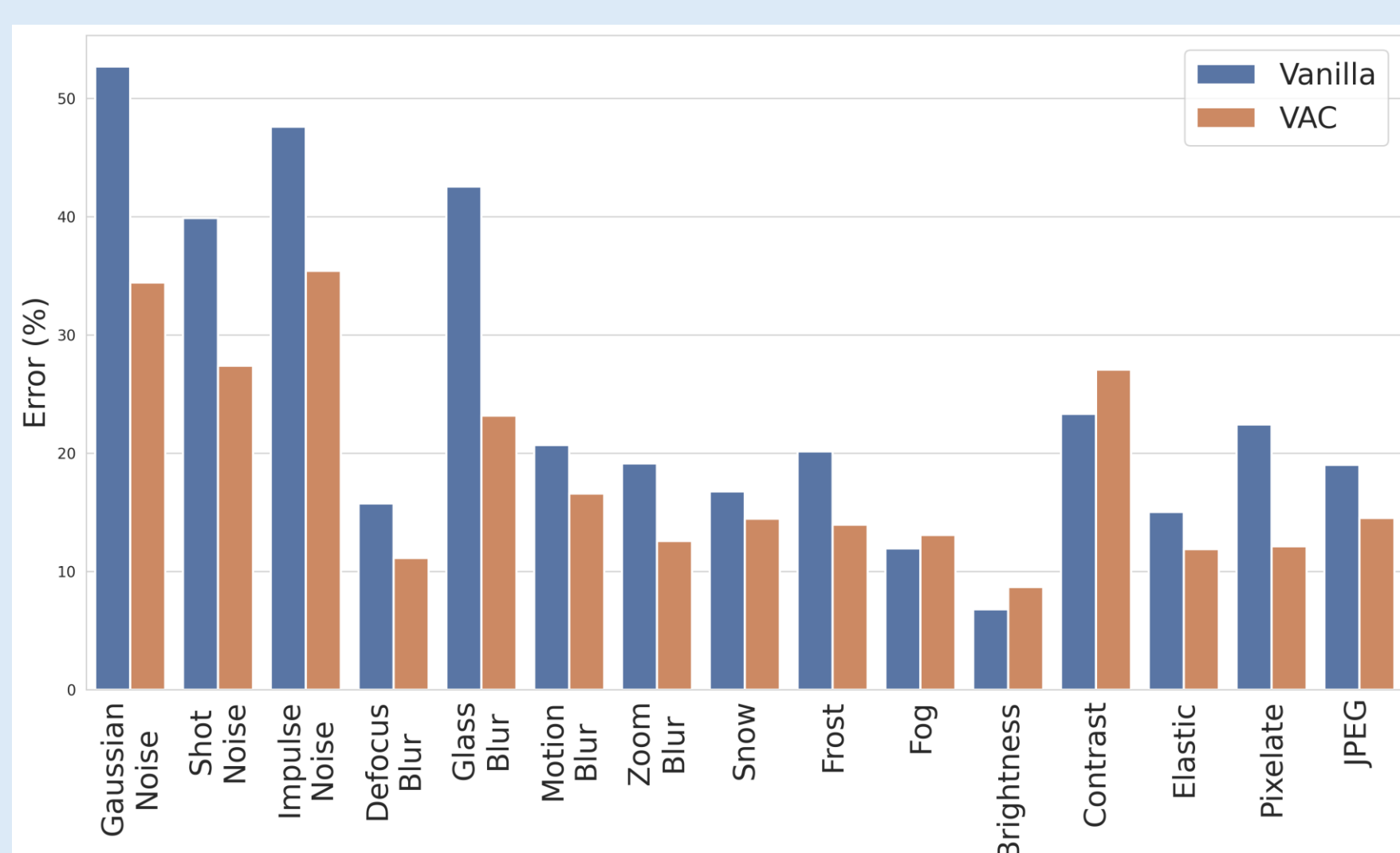
Blur Replay:

To prevent catastrophic forgetting, add a replay of previous blur levels in later epochs.

VAC improves robustness

CNNs trained with VAC are more robust to common image corruptions compared to their vanilla-trained counterparts.

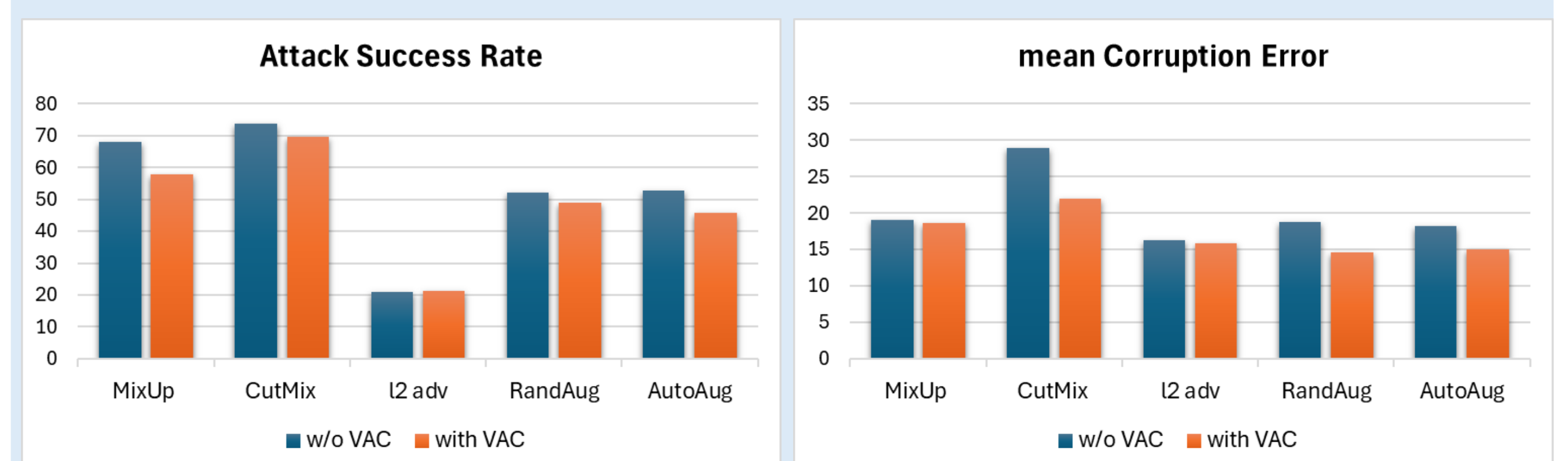
Dataset	Architecture	Vanilla training		VAC	
		Clean error (↓)	mCE (↓)	Clean error (↓)	mCE (↓)
CIFAR-10	All-Conv	6.13	31.19	6.77	27.83
	ResNet-18	5.24	25.03	6.92	18.78
	ResNet-56	6.55	28.67	10.37	20.37
	MobileNet-v2	14.66	33.03	16.90	29.60
	ConvNeXt	7.07	17.00	7.66	16.39
ImageNet-100	ResNet-18	13.12	53.66	15.24	49.23
	ResNet-50	10.96	51.86	15.98	49.03



Robustness gains are achieved across corruptions that the model was never trained on.

- VAC provides better corruption robustness compared to other curriculum and generalization-focused methods.
- VAC is not the same as the popular Gaussian blur augmentation, which lacks a temporal structure. VAC achieves lower corruption error (17.58) compared to blur-based augmentation (18.02).

VAC complements existing augmentations



A Unified Hypothesis of Early-Stage Blurring

Critical Learning Period Hypothesis

- Sensory deprivation in early days causes permanent visual impairment in animals
- Blurring in early epochs causes drop in accuracy of CNNs

Low Visual Acuity Hypothesis

- Low visual acuity at birth promotes extended spatial processing in humans
- Blurring in early epochs is helpful for face recognition in CNNs

Our Hypothesis

Blurring in early epochs improves robustness of CNNs to out-of-distribution corruptions, with a marginal drop in in-domain accuracy.