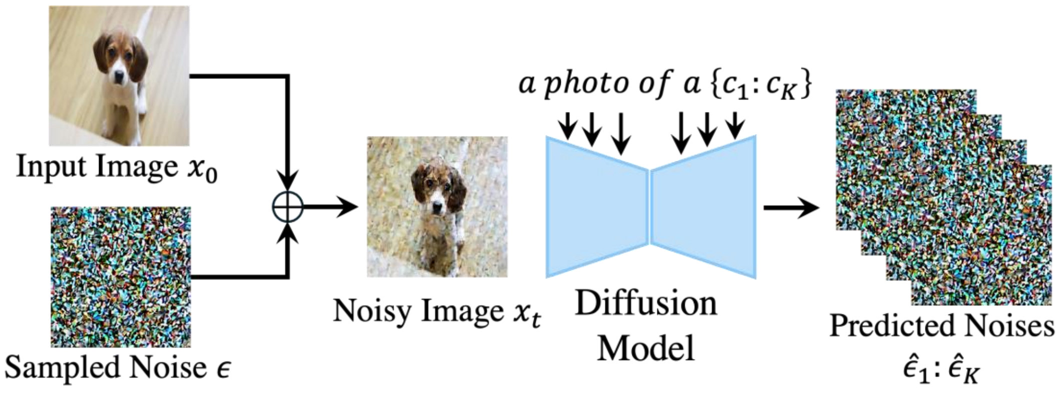
Noise Matters: Optimizing Matching Noise for Diffusion Classifiers

Yanghao Wang, Long Chen

The Hong Kong University of Science and Technology

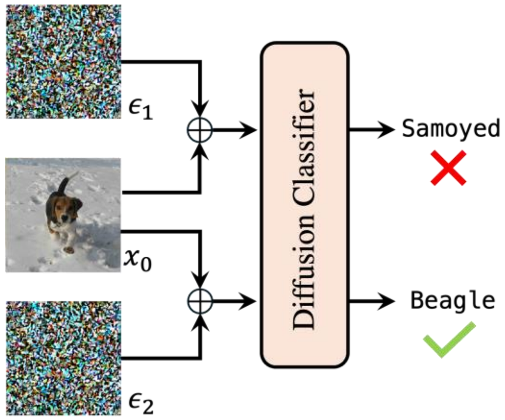
Background

- Diffusion Classifier: Predictions are based on the reconstruction performance



$$\hat{c} = \underset{c_i}{\operatorname{argmin}} ||\hat{\epsilon}_\theta(x_t, c_i, t) - \epsilon||_2^2$$

- Noise Instability: Different noises will lead to different predictions

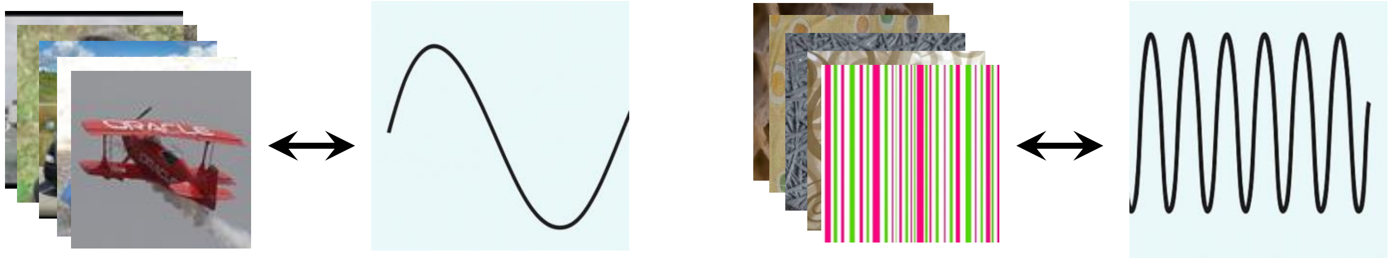


The Role of Noise in Diffusion Classifier

- The sampled noise destroys some parts of the image, and Diffusion Classifier tries to find the category that can best guide the diffusion model to reconstruct the destroyed parts.
- Thus, the "good noise" should destroy the parts that can best reflect the difference in reconstruction effect under different categories' guidance.

Motivation

- We argue that good noise should meet the following two principles
  - *Frequency Matching*: Given a dataset, the category-related signals are mainly in a specific frequency range.



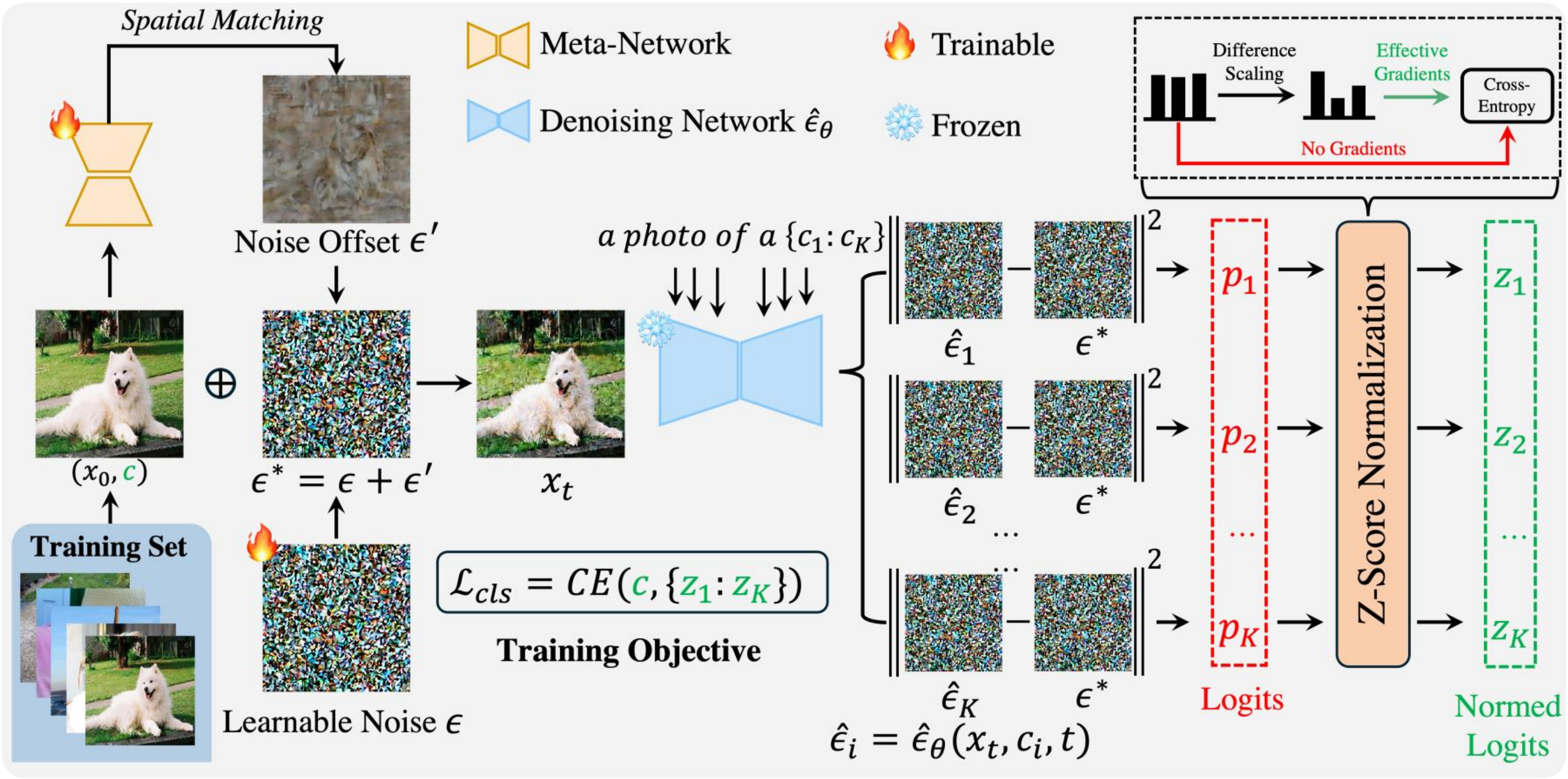
STL-10 Low Frequency DTD High Frequency

  - *Spatial Matching*: Given one image, the category-related signals are mainly in specific spatial areas.

Image Optimized Noise Image Optimized Noise

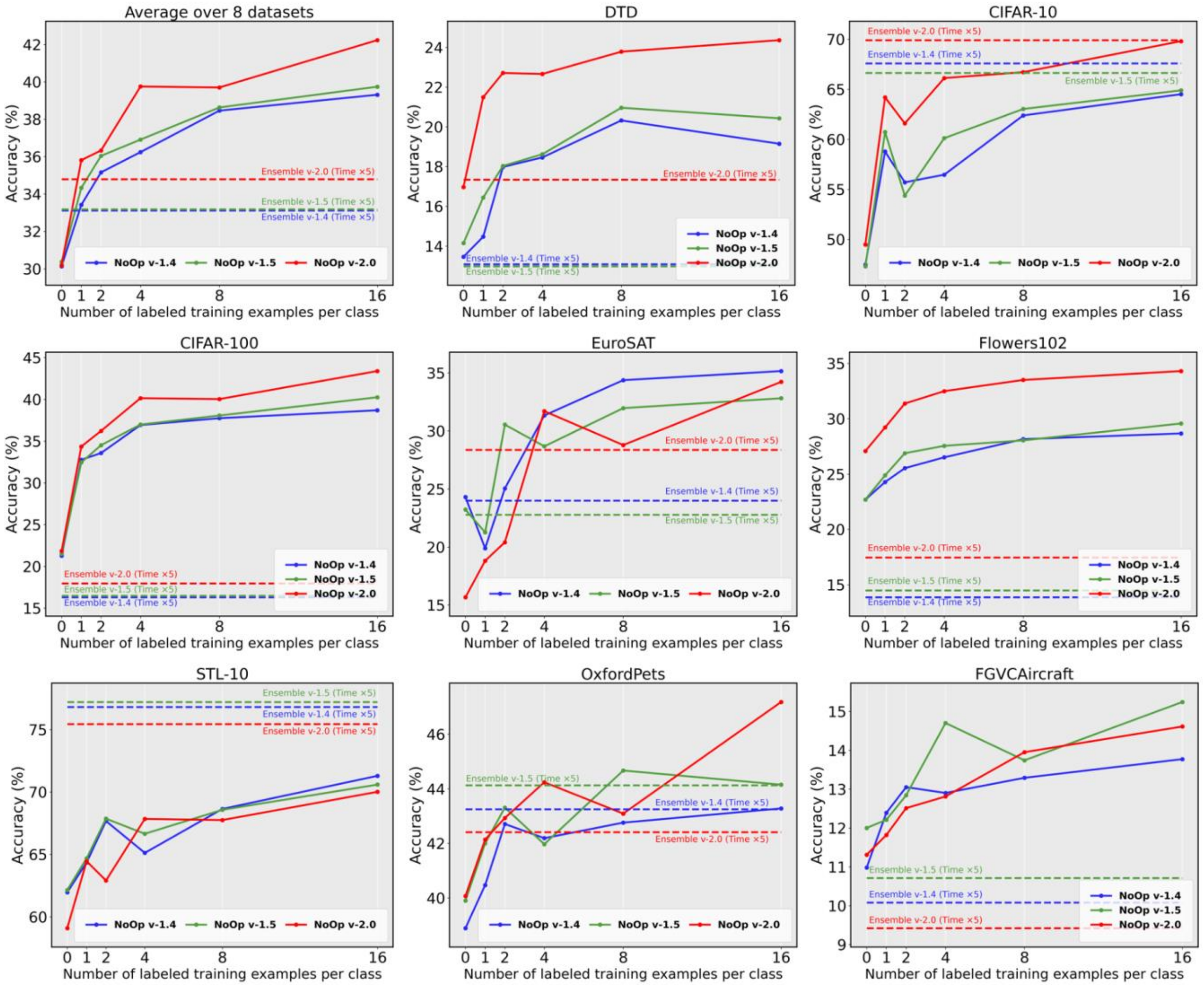
Solution

- NoOp: Noise optimization based on two Principles



Few-shot Learning

- Stable Improvement for Few-shot Classification



Generalization of "Good Noise"

- Cross-dataset Generalization

	Source	Target								
	ImageNet	DTD	CIFAR-10	CIFAR-100	EuroSAT	Flowers102	STL-10	OxforfPets	FGVCAircraft	*Average*
Ensemble (Time ×5)	25.94	17.34	**69.91**	17.96	28.37	17.47	**75.44**	42.41	9.42	34.79
NoOp	**26.34**	**21.70**	63.26	**29.36**	**29.31**	**29.48**	71.66	**45.90**	**10.56**	**37.65**
Δ	+0.40	+4.36	-6.65	+11.40	+0.94	+12.01	-3.78	+3.49	+1.14	+2.86

- The noise optimization can learn some generalized knowledge that is beneficial to the classification, i.e., how to destroy the target part of the image.

Orthogonal to the Prompt Learning

- Comparison with Prompt Learning

Method	ISIC-2019					FGVCAircraft				
	1	2	4	8	16	1	2	4	8	16
Zero-shot DC	13.25	13.25	13.25	13.25	13.25	11.31	11.31	11.31	11.31	11.31
TiF Learner	17.25	13.89	19.76	19.91	17.53	15.60	17.04	16.98	19.47	21.03
NoOp	**18.41**	**20.80**	**23.72**	18.41	20.82	11.82	12.51	12.81	13.95	14.61
NoOp + TiF	18.32	19.23	14.45	**29.29**	**21.59**	**15.99**	**18.54**	**19.59**	**22.44**	**25.74**

- Overall, both prompt optimization (TiF) and noise optimization are effective few-shot learners.
- Moreover, their effects are complementary.
- NoOp is a new few-shot learner, which has a different effect and mechanism from the current prompt optimization.

Conclusion

- Today’s diffusion models (include VP/VE/Flow Matching SDEs) rely on progressive corruption (i.e., the random noise).

- Since the random noise leads to instability in both perception and generation tasks, the role of noise, its impact, and how to optimize it are underexplored in many tasks.

Noise Matters: Optimizing Matching Noise for Diffusion Classifiers

Yanghao Wang, Long Chen

The Hong Kong University of Science and Technology

ywangtg@connect.ust.hk *longchen@ust.hk*

WeChat

GitHub

Thanks!