D-GAP: Improving Out-of-Domain Robustness via Dataset-Agnostic and Gradient-Guided Augmentation in Amplitude and Pixel Spaces

Ruoqi Wang · Haitao Wang · Shaojie Guo · Qiong Luo

Email: rqwang@iaii.ac.cn

CODE

NEURAL INFORMATION PROCESSING SYSTEMS

Why domain shift matter?

Labeled source data meet unlabeled target images with shifts in background, style, or acquisition instruments.

Generic augmentation: gains vary with the domain shift.

Data-specific rules: require expert knowledge and prior analysis.

> Can the model's own response guide where to intervene?

Gradient-guided amplitude mixing

For a labeled source image x_1, measure task-loss sensitivity to its amplitude spectrum.

$$G(u,v) = \left| \frac{\partial \mathcal{L}_{\text{task}}(f(x_1), y)}{\partial A(x_1)(u,v)} \right|,$$

$$\tilde{G} = \frac{G - \mu(G)}{\sigma(G) + \varepsilon}$$

$$D = \text{clip}\big(\text{Sigmoid}(\tilde{G}), d_{\min}, d_{\max}\big)$$

Inside the square frequency window Ω_r, interpolate with the second image's amplitude.

$$A_{\text{mix}}(u,v) = \big(1 - D(u,v)\big)\, A(x_1)(u,v) + D(u,v)\, A(x_2)(u,v)$$

$$\mathcal{F}_{\text{mix}}(u,v) = A_{\text{mix}}(u,v)\, e^{-jP(x_1)(u,v)}, \hat{x}_f = \mathcal{F}^{-1}\big(\mathcal{F}_{\text{mix}}\big)$$

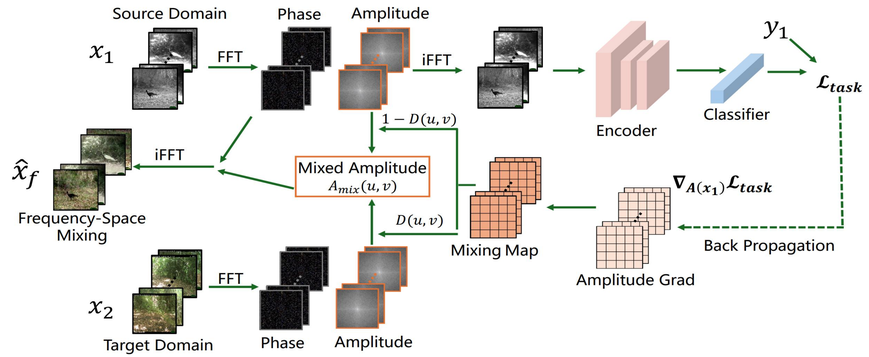
> Higher sensitivity → stronger cross-domain amplitude mixing; retain the source phase.

Training protocol

Real-world datasets: linear probing on frozen pretrained features, then joint finetuning with D-GAP.

DG benchmarks: train directly from pretrained initialization for the paper's comparisons.

Overview of gradient-guided amplitude mix



Given two images x_1 and x_2, we compute a frequency sensitivity map G(u, v) and a mixing map D(u, v) to guide amplitude mixing. The mixed amplitude is then combined with the phase of x_1 to reconstruct the augmented image.

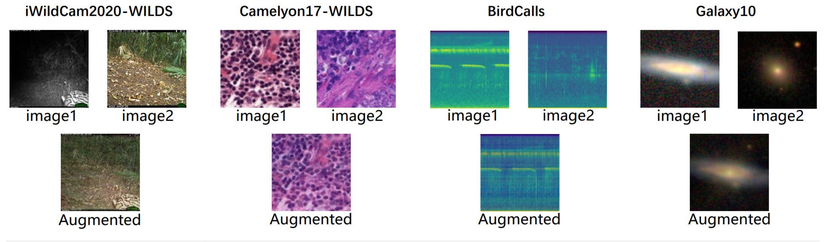
Dual-space fusion

Amplitude mixing can introduce blur or artifacts. A pixel branch adds spatial detail.

$$\hat{x}_p = (1 - \lambda_1)x_1 + \lambda_1 x_2, \;\; \hat{x} = (1 - \lambda_2)\hat{x}_f + \lambda_2 \hat{x}_p$$

Amplitude branch	+	Pixel branch	→	Augmented training image
spectral intervention		**spatial detail**		

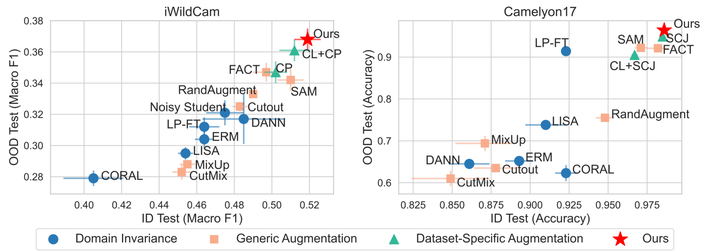
One method, four real-world shifts

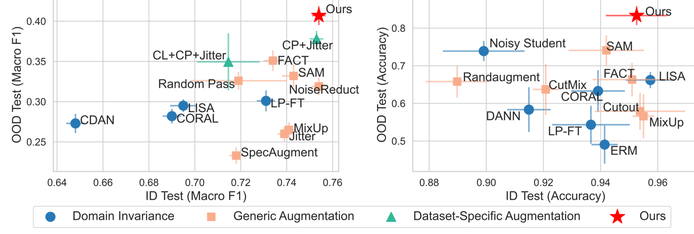


	iWildCam2020-WILDS	Camelyon17-WILDS	BirdCalls	Galaxy10
x_{obj}	animal foreground	cell morphology	bird calls	galaxy structure
$x_{d:robust}$	habitat features in background	cancer stage, tumor size and density	habitat noise	viewing angle
$x_{d:spu}$	low-level background features	stain color	microphone gain settings low-level noise	low-level noise background features
x_{noise}	xy-position	patch orientation, cell xy-positions	x-position	xy-position

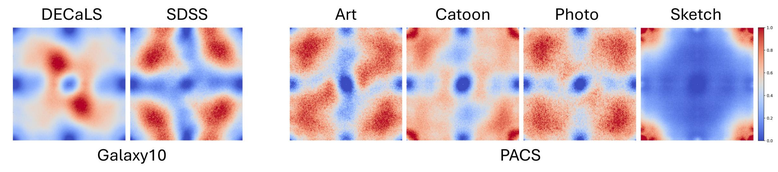
Real-world evidence: ID–OOD trade-offs

D-GAP (red star) has the highest OOD score in each original plot. Points and error bars show means and standard deviations.





Model response varies across domains:



Average amplitude sensitivity maps of ResNet50 on Galaxy10 and PACS. The method uses this response to guide mixing.

Interpretation: gradient magnitude measures local sensitivity; it does not directly identify spurious components.

Takeaways & limitation

Model-aware: task gradients set the amplitude mixing strength.
Dual-space: amplitude intervention complements pixel detail.

> Cost: additional gradient computation.
> Periodic updates and lightweight predictors are future work.