DeepPrism: Channel Convolution for Lightweight Generative Models

Changqing Fu
November 26, 2023
Joint work with Laurent D. Cohen

Funded in part by the French government under management of Agence Nationale de la Recherche as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute). Supported by HPC resources from GENCI–IDRIS (Grant 2021-AD011013178).

Question

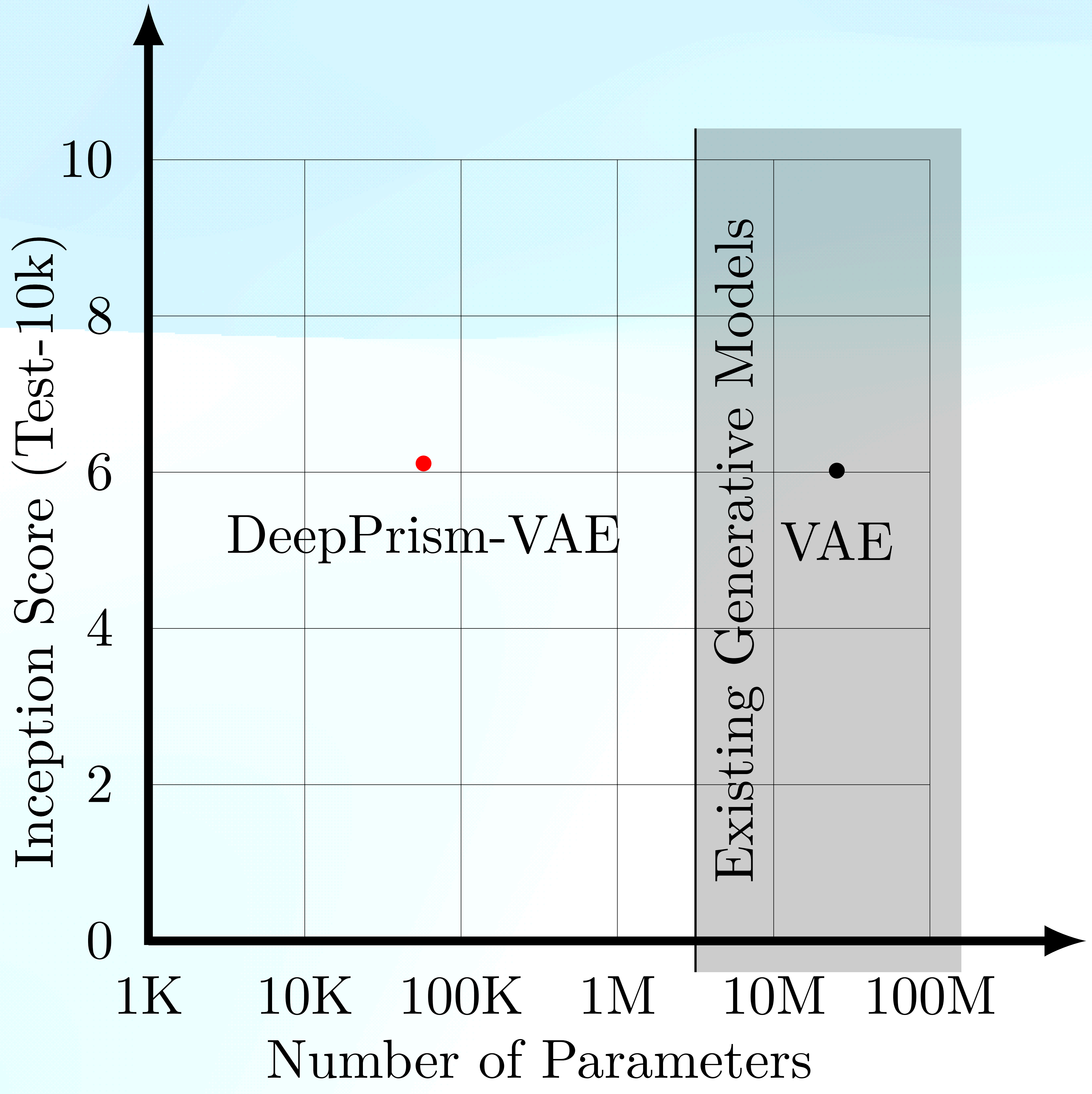
How to incorporate symmetry as a constraint in generative models for improved **parameter efficiency** and training/inference **time complexity?**

Land's Symmetry

Water Flow

Solution

DeepPrism: Lightweight Convolution Kernel

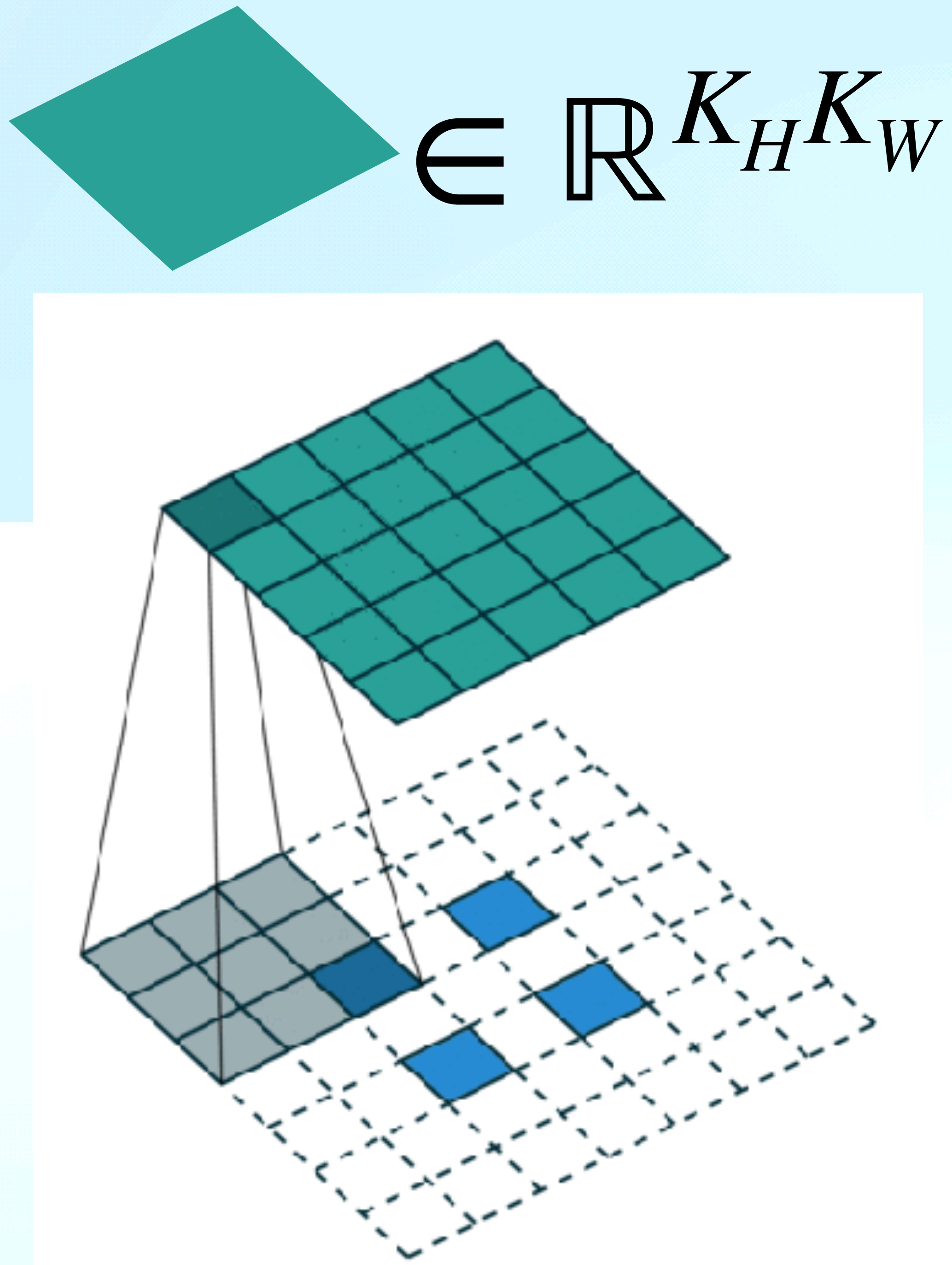


Unprecedented Lightweight Generative Model

1000x Compression!

Solution

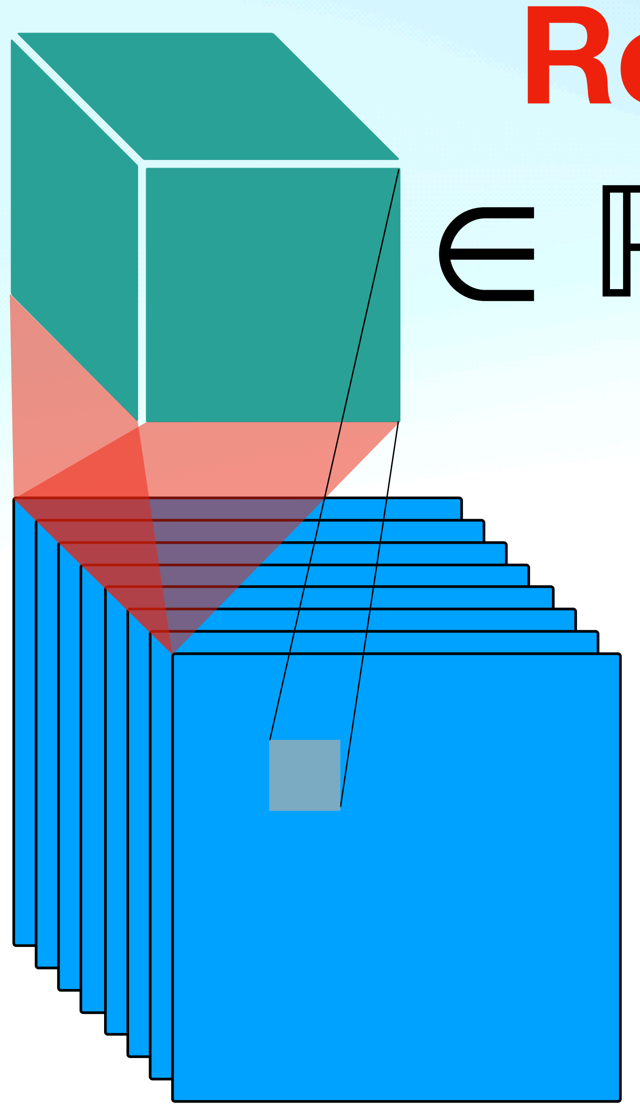
Reduce Redundancies in the Channel Dimension



$\in \mathbb{R}^{K_H K_W}$

Single-Channel Convolution

Redundant



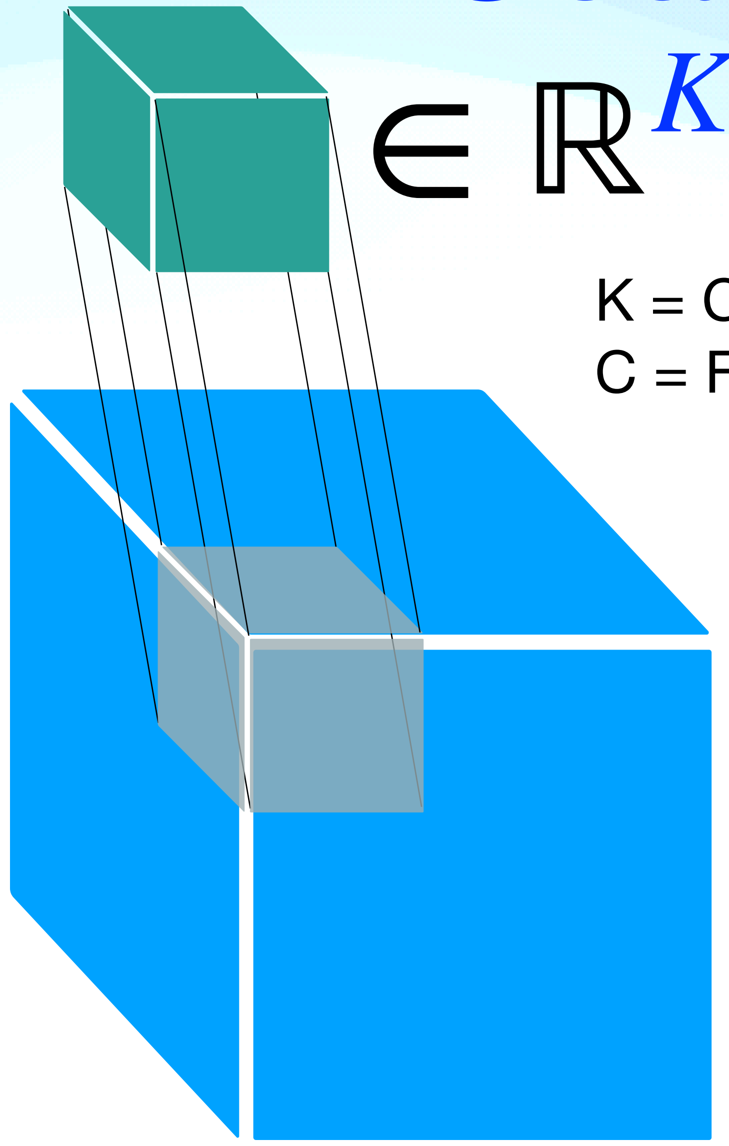
$\in \mathbb{R}^{C'CK_H K_W}$

Complexity

$O(C^2)$

Multi-Channel Convolution

Scalable



$\in \mathbb{R}^{K_C K_H K_W}$

K = Convolution Weights' Kernel Size

C = Feature Maps Channel Size

Complexity

$O(1)$

DeepPrism Convolution

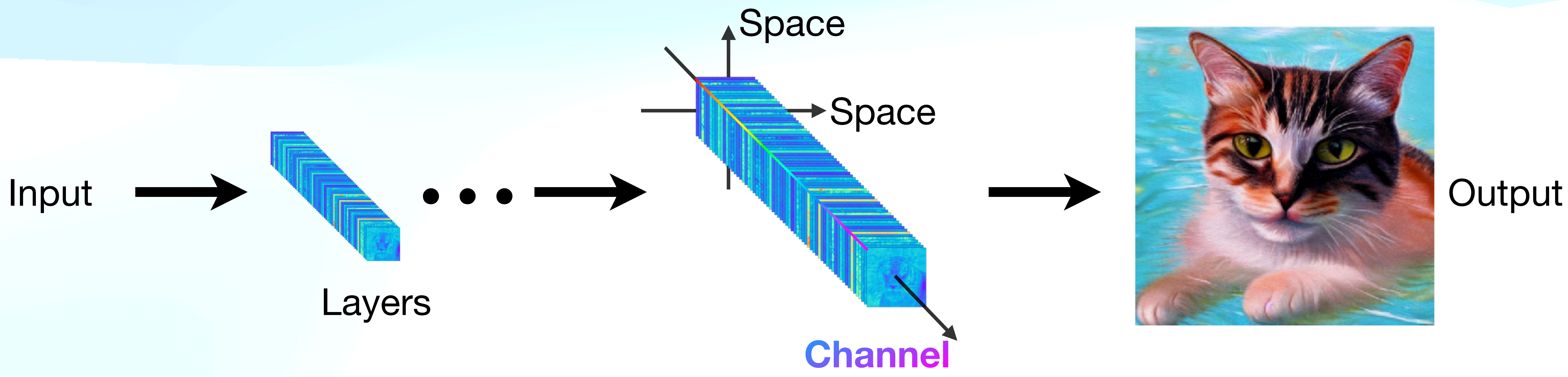
Dauphine Université Paris | PSL CEREMADE UMR CNRS 7534 PR[AI]RIE PaRis Artificial Intelligence Research InstitutE GENCI Le calcul intensif au service de la connaissance cnrs

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Motivation

Why is the channel dimension important?



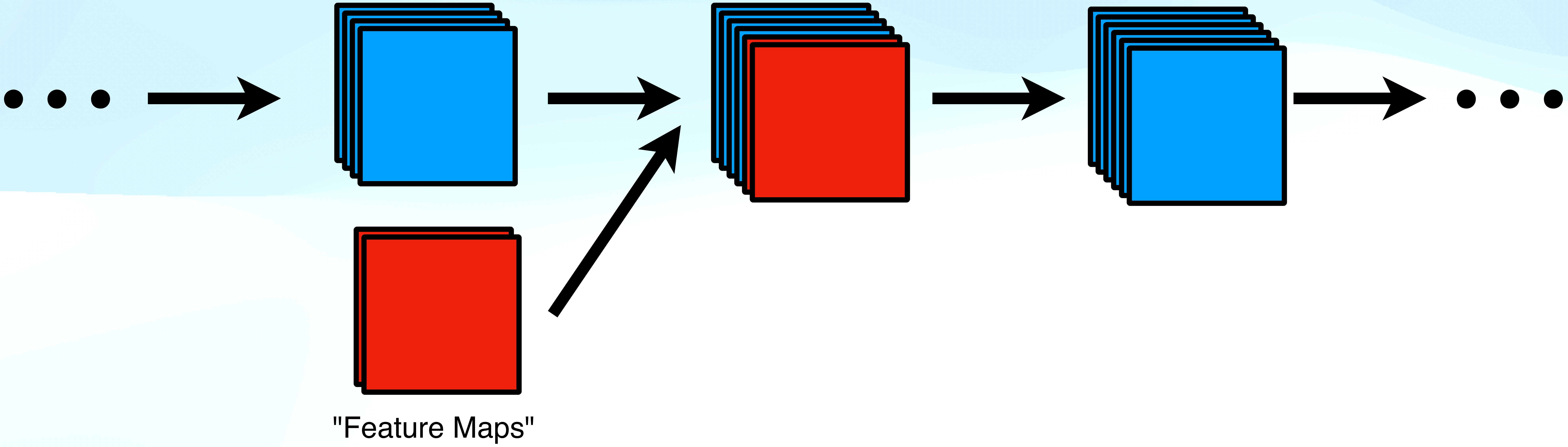
Dauphine Université Paris | PSL CEREMADE UMR CNRS 7534 PR[AI]RIE PaRis Artificial Intelligence Research InstitutE GENCI cnrs

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Controls in Generative Model

Example: U-Net



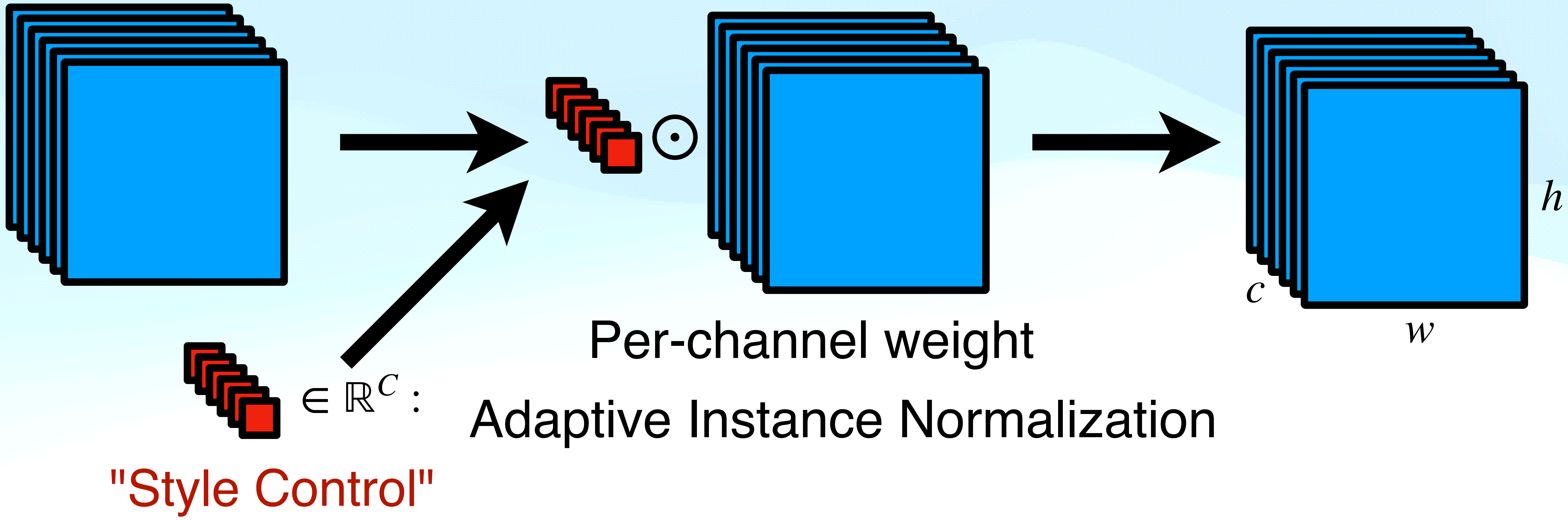
Concatenation in the Channel dimension

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Controls in Generative Model

Example: StyleGAN



$$L(x, x_1) = \lambda_{\mathrm{TV}} |\nabla x(0)| + \sum_{t \text{ first few layers}} \frac{\lambda_1}{2} |x(t) - x_1(t)|^2 + \frac{\lambda_2}{2} |\mathrm{Gram}(x(t) - x_2(t))|^2$$

Denoising Loss | Content (Perceptual) Loss | Style Loss

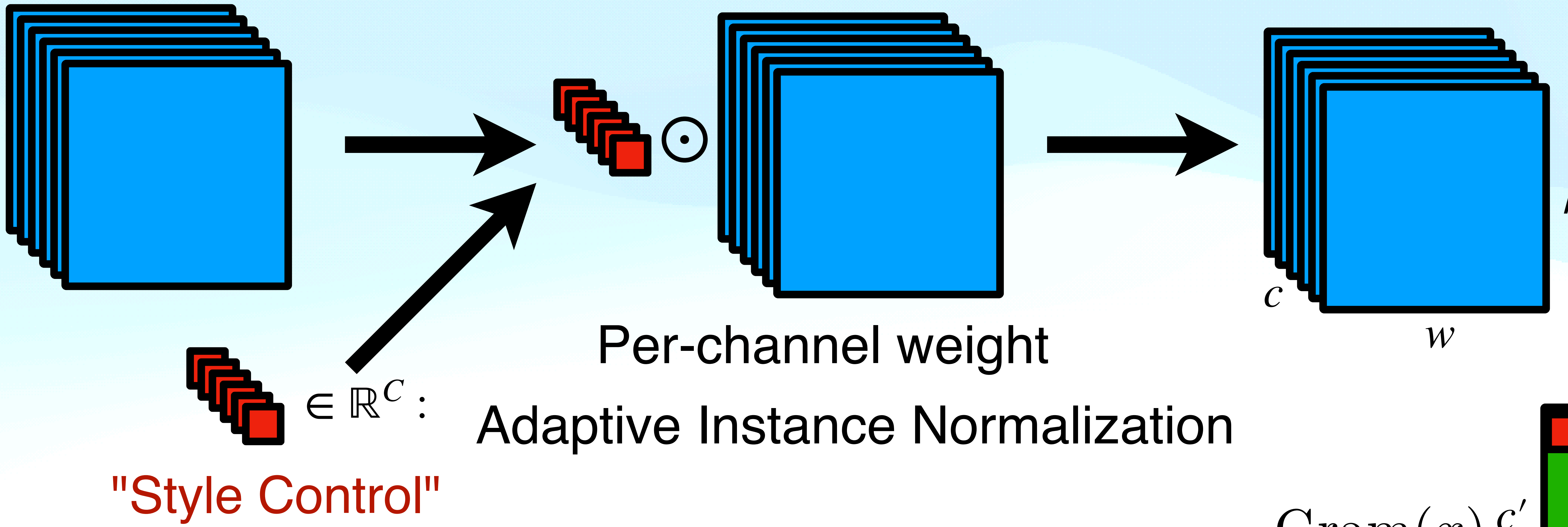
$x(t)$: t-th layer of a pre-trained classification network with input $x = x(0)$

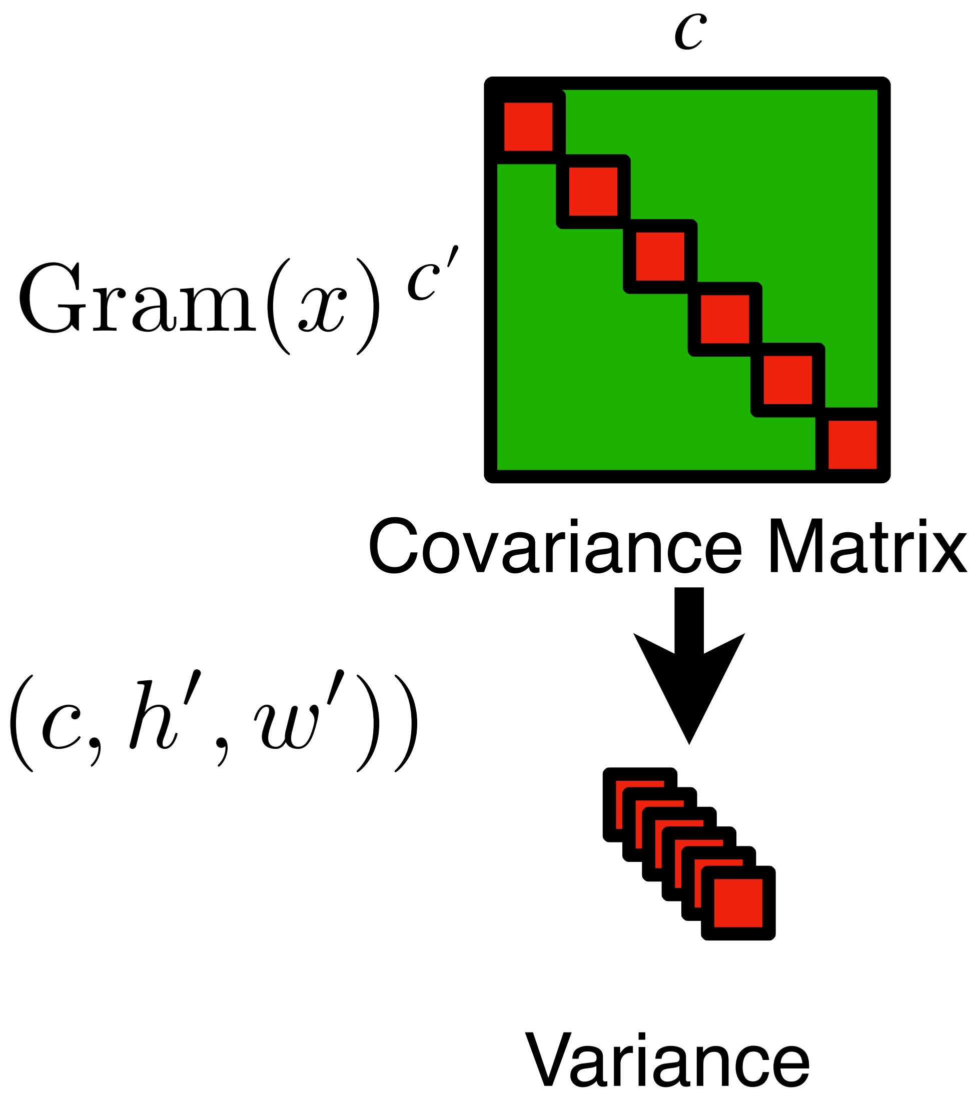
DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Controls in Generative Model

Example: StyleGAN





$$\mathrm{Gram}(x)(c,c') = \frac{1}{HW}\sum_{h,w=1}^{H,W}\left(x(c,h,w) - \frac{1}{HW}\sum_{h',w'=1}^{H,W} x(c,h',w')\right)\left(x(c',h,w) - \frac{1}{HW}\sum_{h',w'=1}^{H,W} x(c,h',w')\right)$$

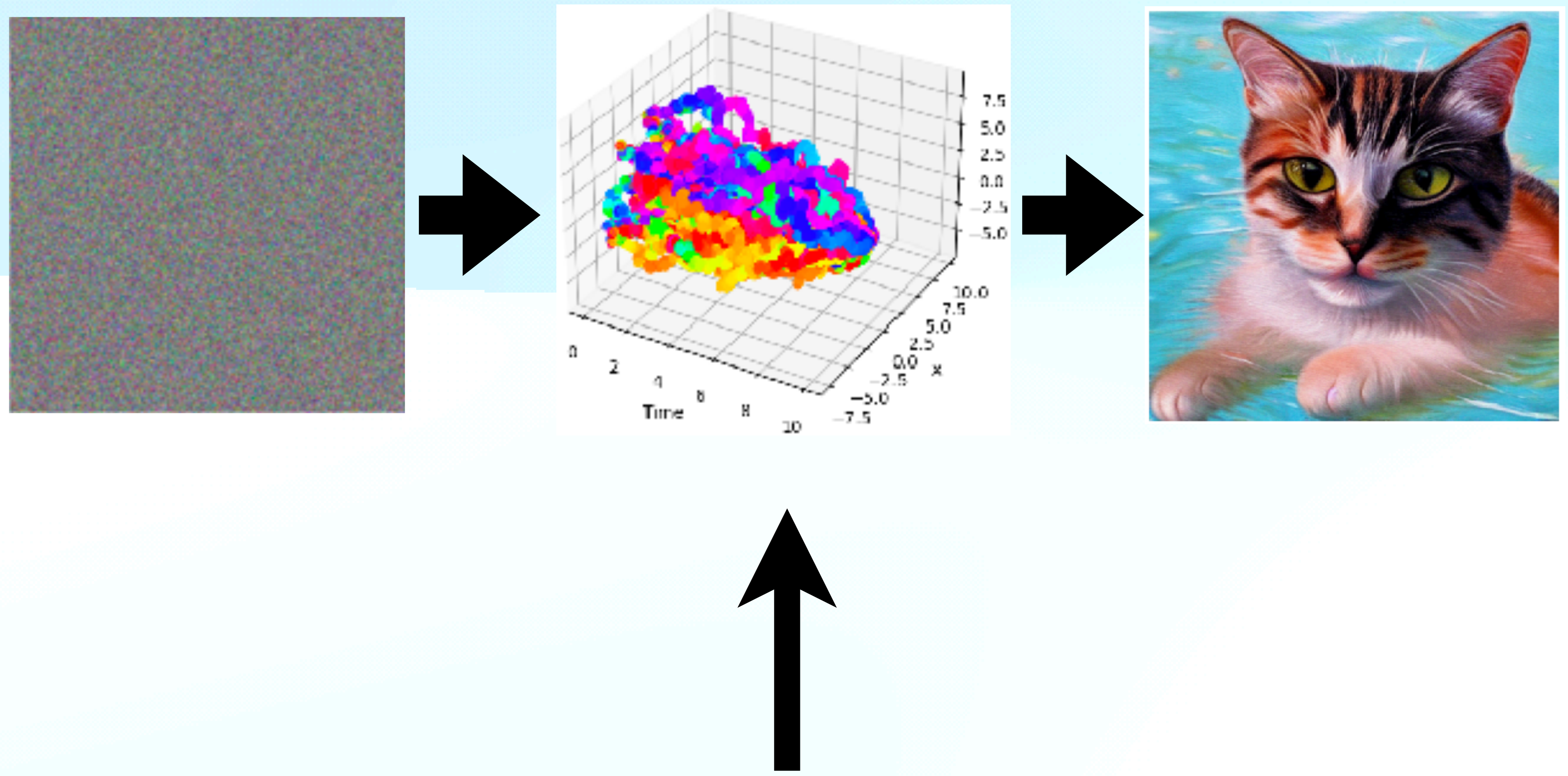
$x(c,h,w)$: channel c, height h, width w

DeepPrism

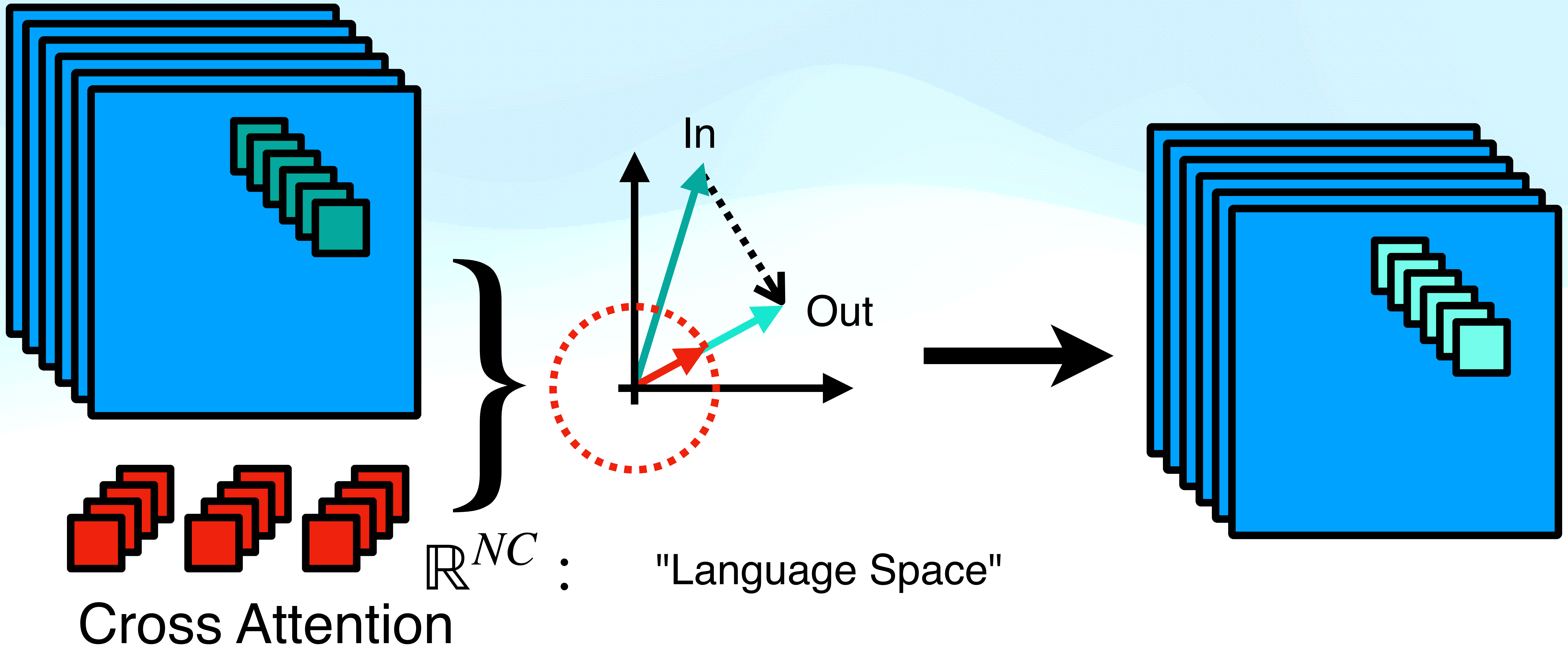
A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Controls in Generative Model

Example: Attention



Prompt = "A cute cat swimming in the pool, oil on canvas"



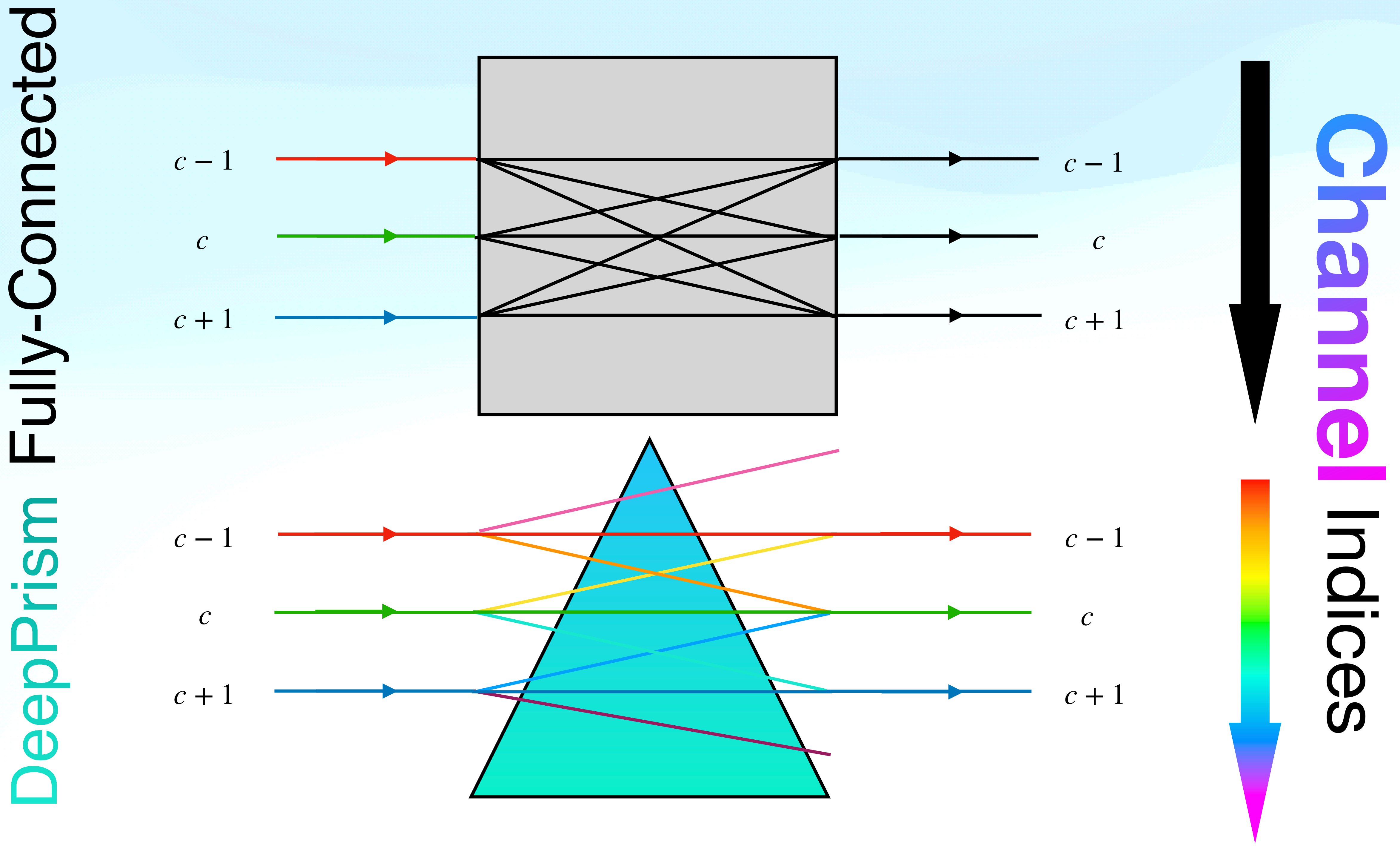
$$\mathrm{Attn}(q,k,v)_{chw} = \sum_{c',h',w'=1}^{C,H,W} \mathrm{Softmax}\left(\frac{q(c',h',w')k(c',h',w')}{\sqrt{C}}\right) v(c,h',w')$$

Channel dimension is matched with word token!

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

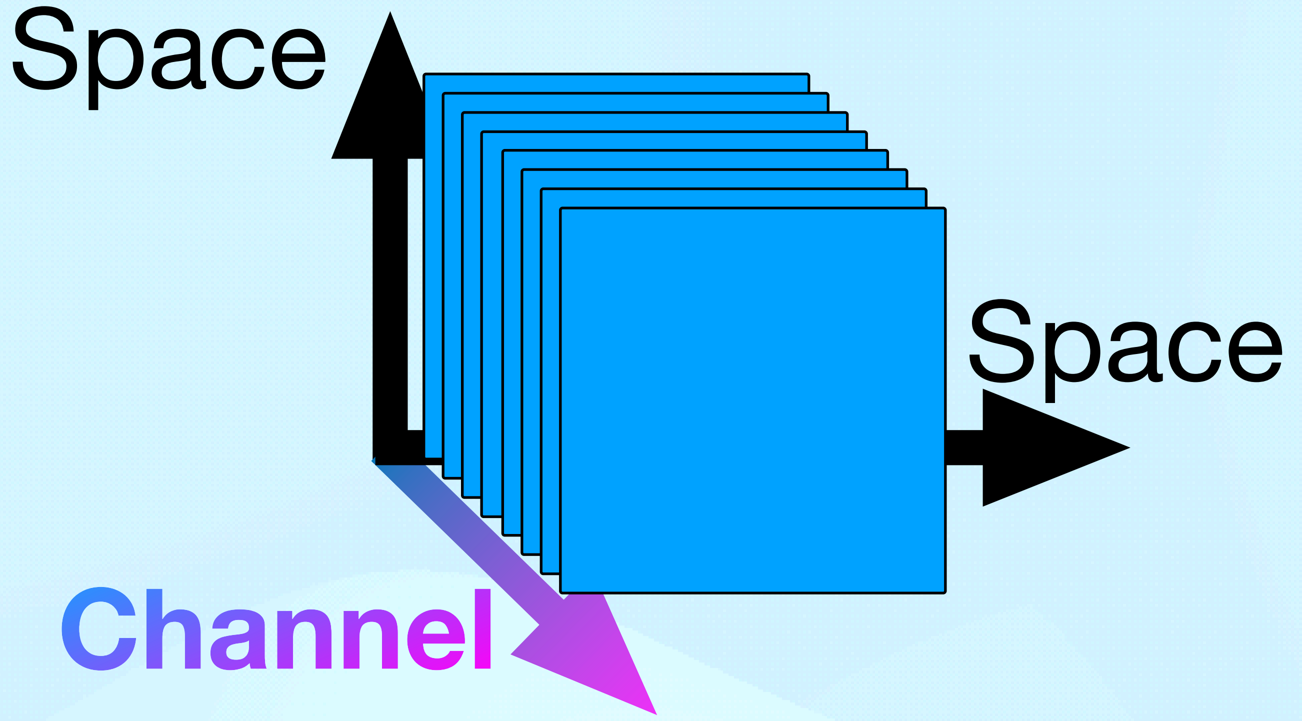
Opening the Black-Box with DeepPrism



Dauphine Université Paris | PSL CEREMADE UMR CNRS 7534 PR[AI]RIE PaRis Artificial Intelligence Research InstitutE GENCI Le calcul intensif au service de la connaissance cnrs

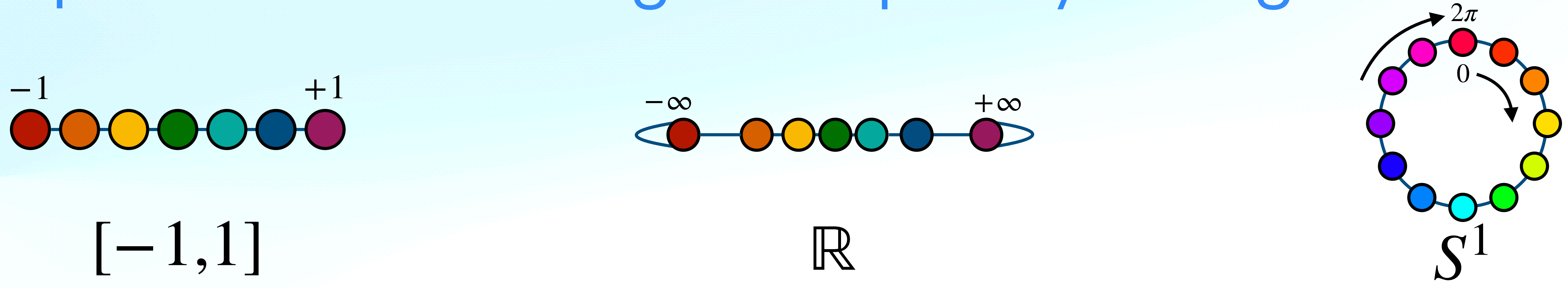
DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.



Topology in the Color Dimension

Perceptible Colors | Light Frequency | Digital Color Hue



Padding	Zero	Replicate	Circular
Topology	Interval	Line	Circle
Boundary Condition	Dirichlet	Neumann	Periodic

Implementation: From Conv2D to DeepPrism

$$\mathrm{ChannelConv}(u; w) = \mathrm{Conv1D}(w, u^\top)^\top$$

$$\mathrm{Pseudo3DConv}(u; w) = \mathrm{Conv1D}(w_c, \mathrm{Conv2D}(w_s, u)^\top)^\top$$

$$\mathrm{DeepPrism}(u; w) = \mathrm{Conv3D}(w, u)$$

where $\top$ is the transpose between channel indices and pixel locations

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Extension to Attention Layer

$$\mathrm{Attn}(q,k,v)_{chw} = \sum_{c',h',w'=1}^{C,H,W} \mathrm{Softmax}\left(\frac{q(c',h',w')k(c',h',w')}{\sqrt{C}}\right)v(c,h',w')$$

$$\mathrm{PrismAttn}(u; w_q, w_k, w_v, w_o) = \mathrm{ChannelConv}(\mathrm{Attn}(q,k,v); w_o)$$

$$q = \mathrm{ChannelConv}(u; w_q)$$

$$k = \mathrm{ChannelConv}(u; w_k)$$

$$v = \mathrm{ChannelConv}(u; w_v)$$

Time & Space Complexity

	Conv[1]	Group	Ours	Attention	MultiHeadAttn	SparseAttn [5]	Ours
Parameter	C^2	$\frac{C^2}{G}$	$\mathbf{1}$	C^2	C^2	C^2	$\mathbf{1}$
Training	BC^2N	$\frac{BC^2N}{G}$	$\boldsymbol{BCN}$	$B(CN^2+C^2N)$	$B(\frac{CN^2}{P}+C^2N)$	$B(CN^{\frac{3}{2}}+C^2N)$	$\boldsymbol{BN^2}$
Space	BCN	BCN	BCN	$B(N^2+NC)$	$B(\frac{N^2}{P}+NC)$	$B(N^{\frac{3}{2}}+CN)$	$B(N^2+NC)$
Inference	C^2N	$\frac{C^2N}{G}$	$\boldsymbol{CN}$	CN^2+C^2N	$\frac{CN^2}{P}+C^2N$	$CN^{\frac{3}{2}}+C^2N$	$\boldsymbol{N^2}$

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

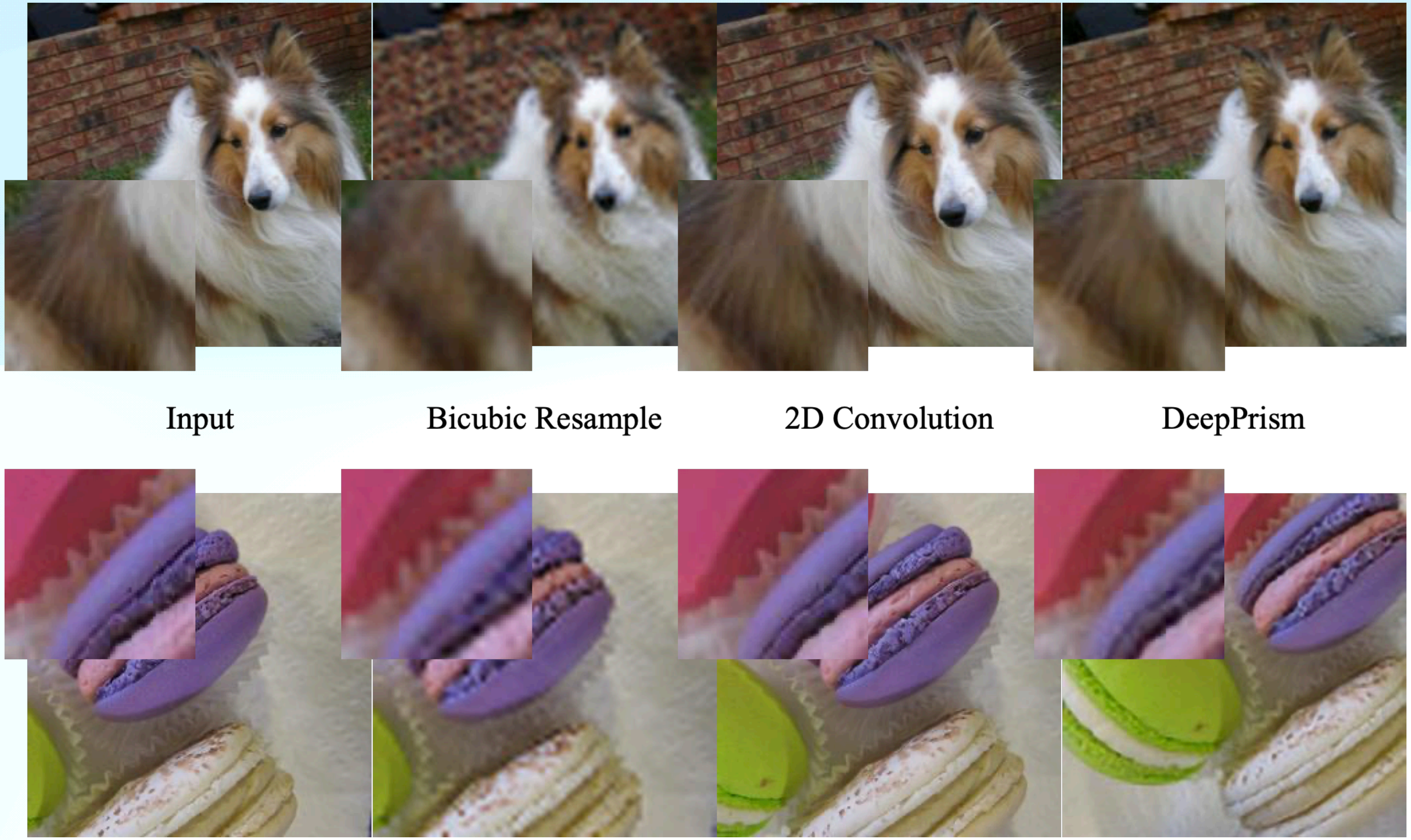
Performance on VAE on CIFAR10

Method	PSNR (↑)	IS (↑)	KID (↓)	PS (↑)	SSIM (↑)	Parameter (↓)
Baseline	18.8 ± 2.4	5.99 ± 0.12	0.0146 ± 0.0011	2.07 ± 0.49	0.609 ± 0.106	25.3M
Ours	24.1 ± 2.5	6.08 ± 0.15	0.0179 ± 0.0011	1.03 ± 0.27	0.850 ± 0.053	**57.8K**

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Results on VAE on ImageNet



DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Losses and Regularizers

$$L(E,G,D,w) = \min_{E,G,Z} \max_{D} \mathbb{E}_{u\sim\mathcal{D}}[\frac{1}{2}\|u - G(E(u))\|^2$$ Reconstruction Loss

$$+\frac{\lambda_1}{2}\sum_{\ell} w_\ell \|V_\ell(u) - V_\ell(G(E(u)))\|^2$$ Perceptual Loss

$$+\frac{\lambda_2}{2}(\|\overline{E(x)} - q_Z(E(u))\|_2^2 + \|E(x) - \overline{q_Z(E(u))}\|_2^2)$$ VQ Loss

$$+\lambda_2 H\big(\mathcal{N}(E(u), E'(u)^2\mathbf{I}) \,|\, \mathcal{N}(0,\mathbf{I})\big)$$ VAE Loss

$$+\lambda_3 D(u)_+ - D(G(E(u))_+]$$ GAN Loss $\mathcal{W}_1(\mathcal{D}, (G\circ E)^\sharp\mathcal{D})$

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Results on Single-Image Diffusion Model

Input

SinDDM

Ours (1000x smaller)

DeepPrism

A symmetry constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.

Results on Latent Diffusion Model

Dauphine UNIVERSITÉ PARIS | PSL CEREMADE UMR CNRS 7534 PR[AI]RIE PaRis Artificial Intelligence Research InstitutE GENCI Le calcul intensif au service de la connaissance cnrs

Conclusion

DeepPrism is a **symmetry** constraint on generative models for improved **parameter efficiency** and training/inference **time complexity**.