Co-denoising from dummies!

Your noise, my noise; apo-noise!

Takes full responsibility for the bad memes!

!!! Disclaimer: This presentation contains memes. OMG!

Amazing idea in Marseille ;)

!!! Exact scene of us discussing the idea in Marseille!

And it seems people did it before!

Or did they?

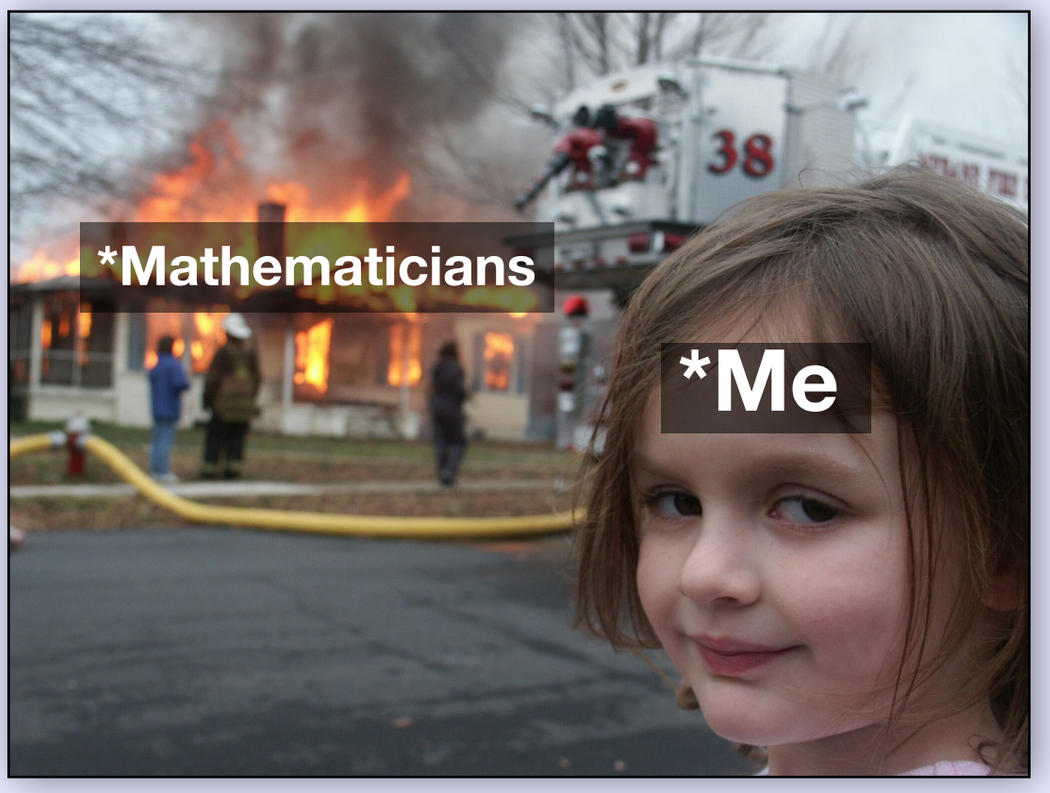
Recap: Flow Matching

Recap: Flow Matching

Flow Matching is interpolation!

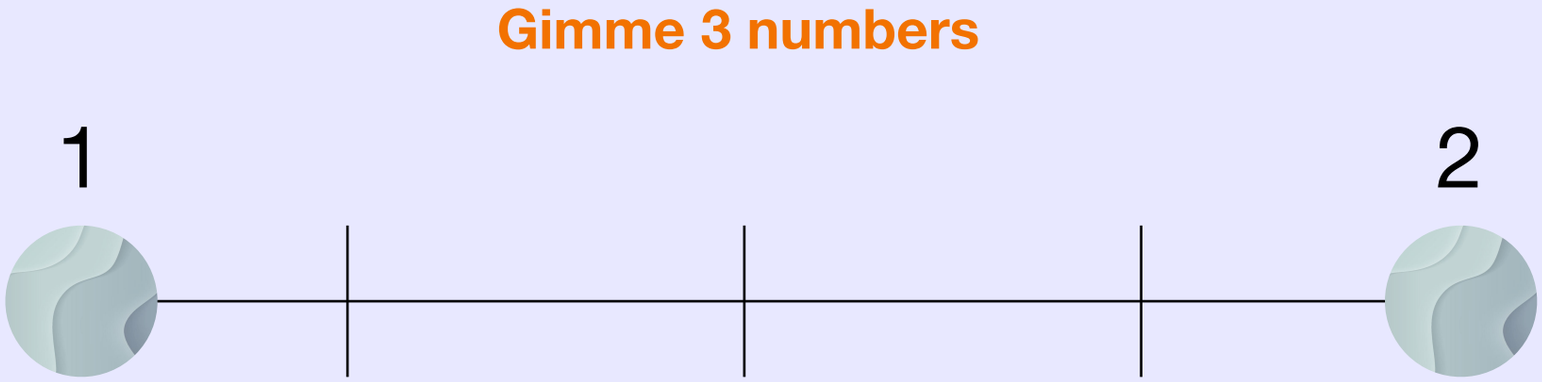
Recap: Flow Matching

Flow Matching is interpolation!



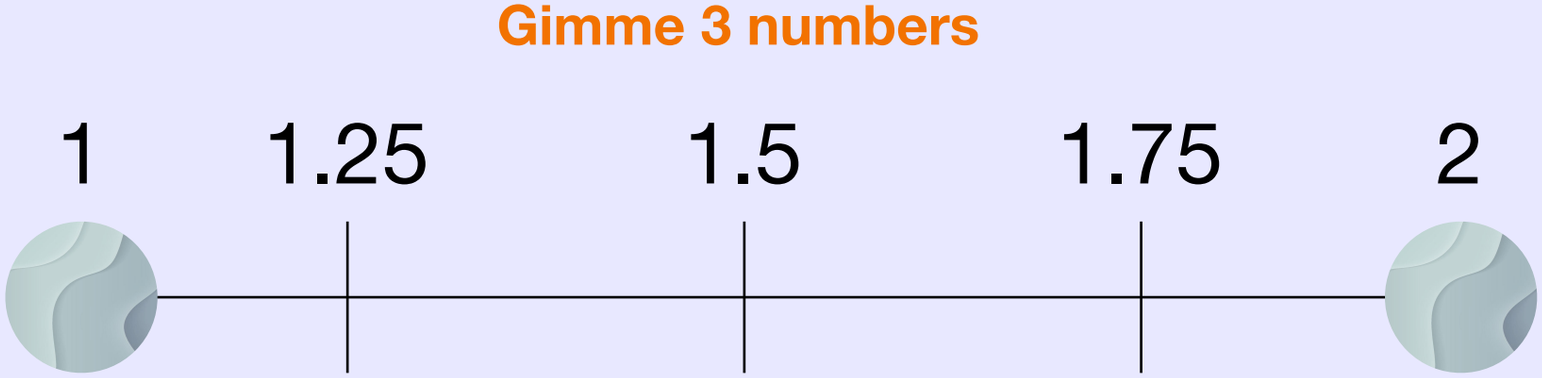
Recap: Flow Matching

Flow Matching is interpolation!



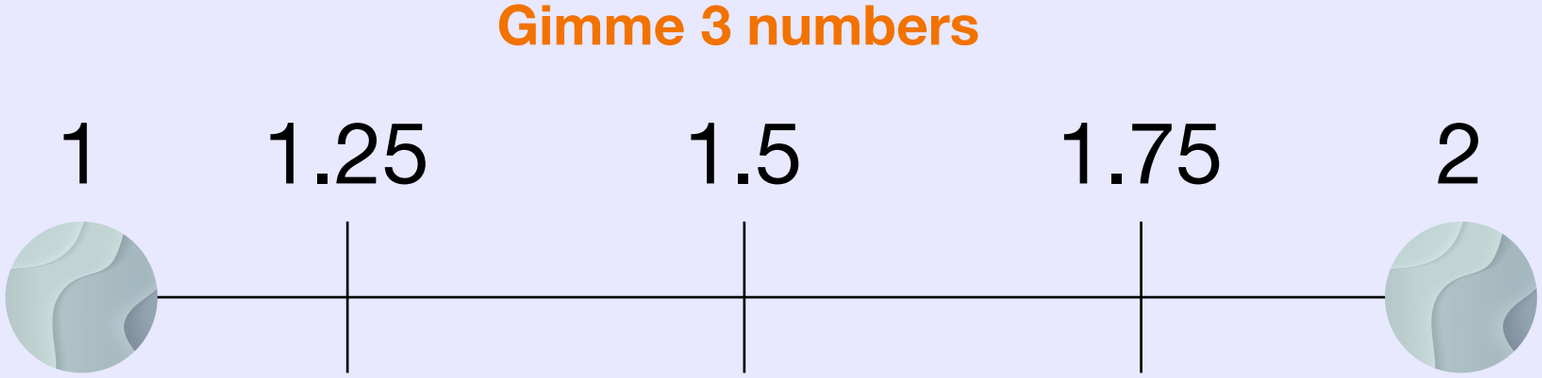
Recap: Flow Matching

Flow Matching is interpolation!



Recap: Flow Matching

Flow Matching is interpolation!

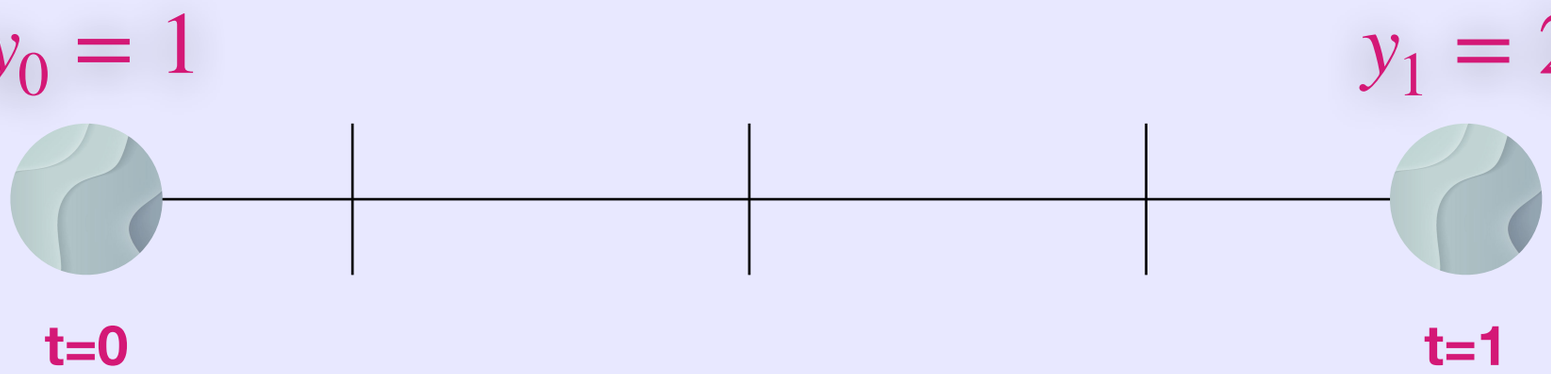


Unknowingly, we all did Flow Matching here!!

Recap: Flow Matching

Flow Matching is interpolation!

Unknowingly, we all did Flow Matching here!!



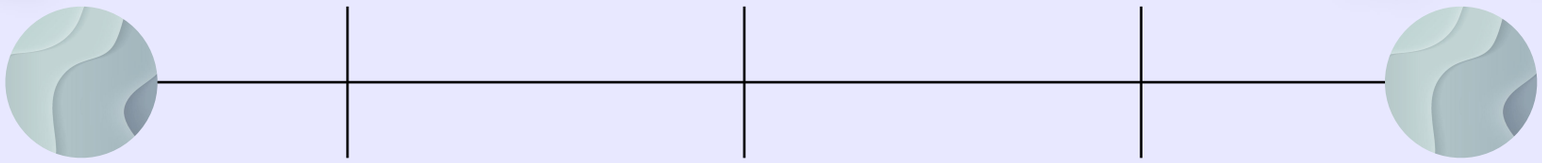
Recap: Flow Matching

Flow Matching is interpolation!

Unknowingly, we all did Flow Matching here!!

$y_0 = 1$

$y_1 = 2$



t=0

t=1

$$v = \frac{dy}{dt} = \frac{2-1}{1-0} = 1$$

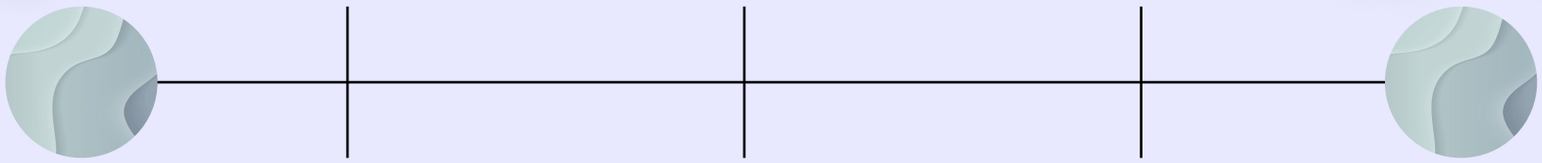
Recap: Flow Matching

Flow Matching is interpolation!

Unknowingly, we all did Flow Matching here!!

$y_0 = 1$

$y_1 = 2$



t=0

t=1

$$y_{new} = y_{old} + step \cdot v$$

$$v = \frac{dy}{dt} = 1$$

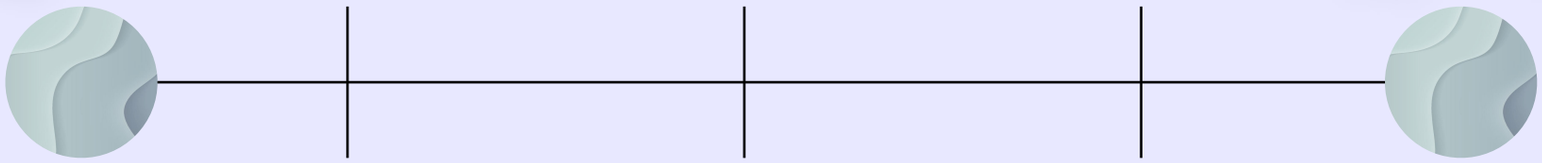
Recap: Flow Matching

Flow Matching is interpolation!

Unknowingly, we all did Flow Matching here!!

$y_0 = 1$

$y_1 = 2$



t=0

t=1

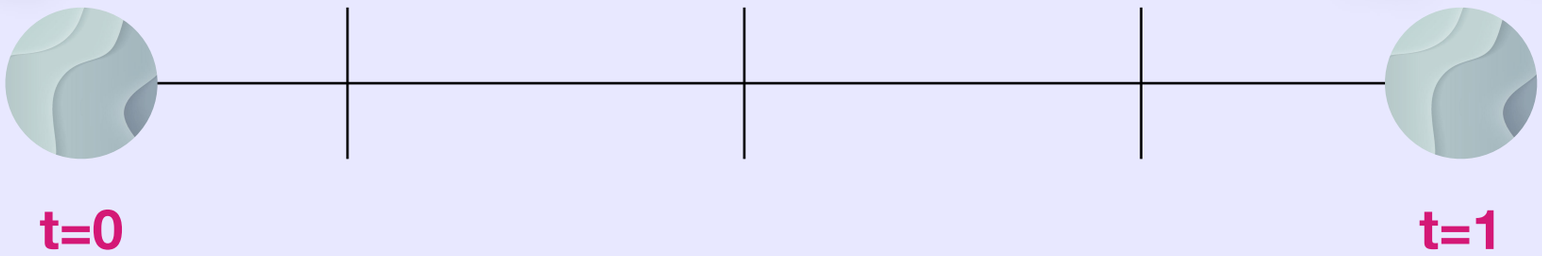
$$y_{new} = y_{old} + 0.25 \cdot v$$

$$v = \frac{dy}{dt} = 1$$

Recap: Flow Matching

Flow Matching is interpolation!

Unknowingly, we all did Flow Matching here!!



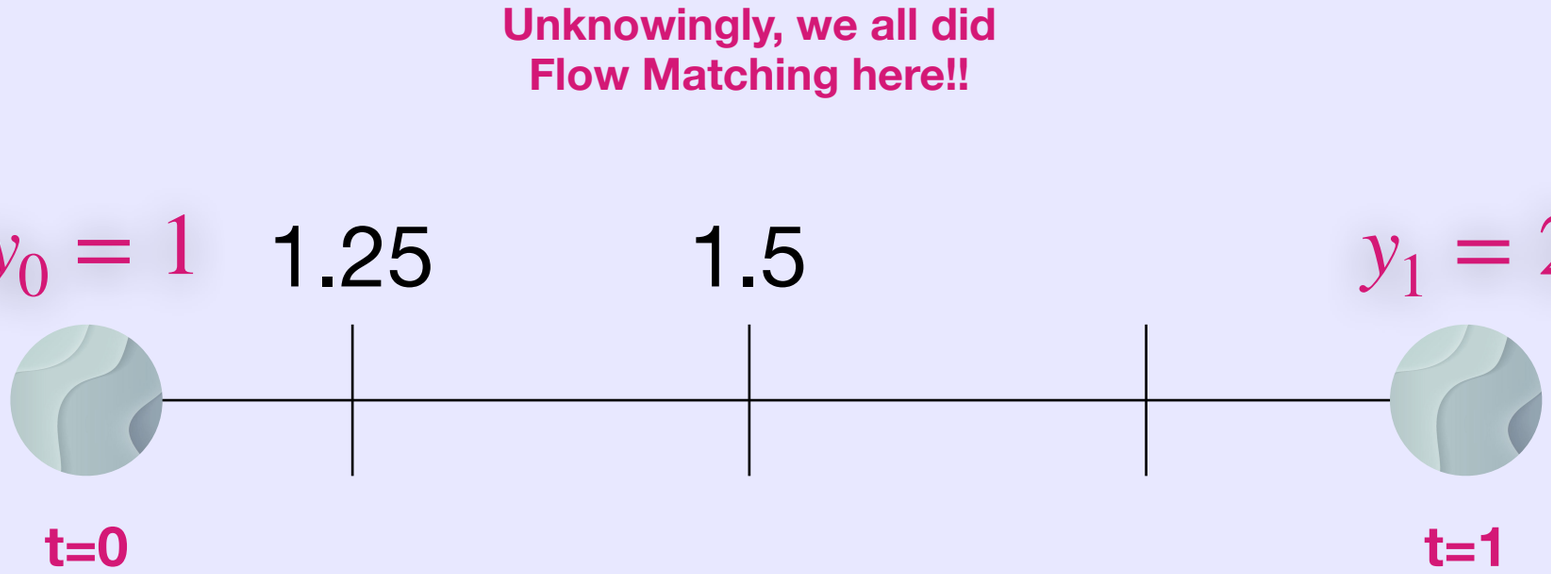
$$y_{new} = 1 + 0.25 \cdot 1 = 1.25$$

$$v = \frac{dy}{dt} = 1$$

Recap: Flow Matching

Flow Matching is interpolation!

Unknowingly, we all did Flow Matching here!!



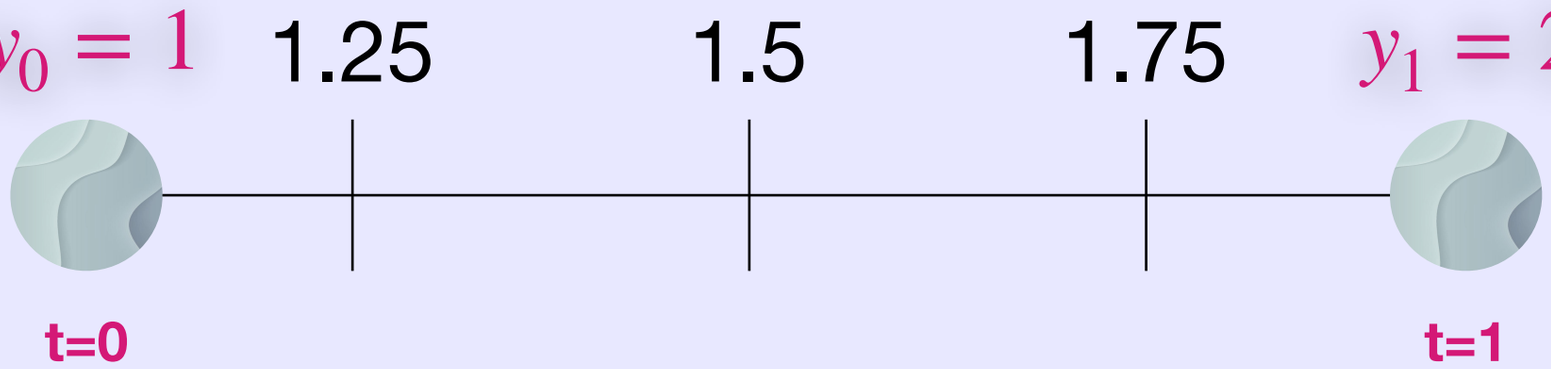
$$y_{new} = 1.25 + 0.25 \cdot 1 = 1.5$$

$$v = \frac{dy}{dt} = 1$$

Recap: Flow Matching

Flow Matching is interpolation!

Unknowingly, we all did Flow Matching here!!



$$y_{new} = 1.5 + 0.25 \cdot 1 = 1.75$$

$$v = \frac{dy}{dt} = 1$$

Recap: Flow Matching

Flow Matching is interpolation!

Source Distribution

π_0

Target Distribution

π_1

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Recap: Flow Matching

Flow Matching is interpolation!

Source Distribution

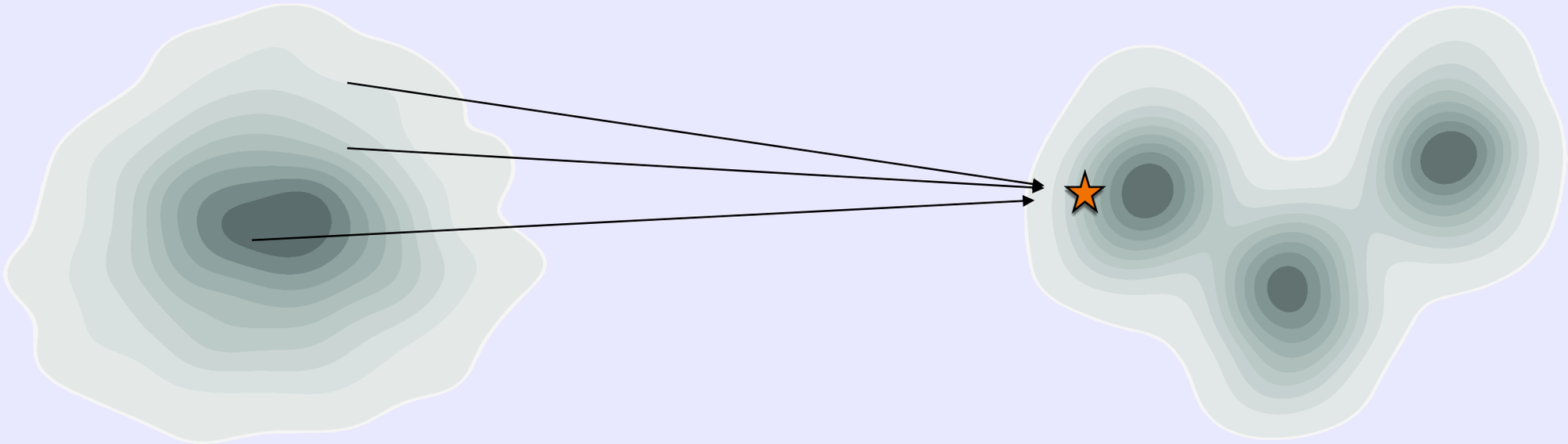
π_0

Target samples

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Recap: Flow Matching

Flow Matching is interpolation!



Source Distribution

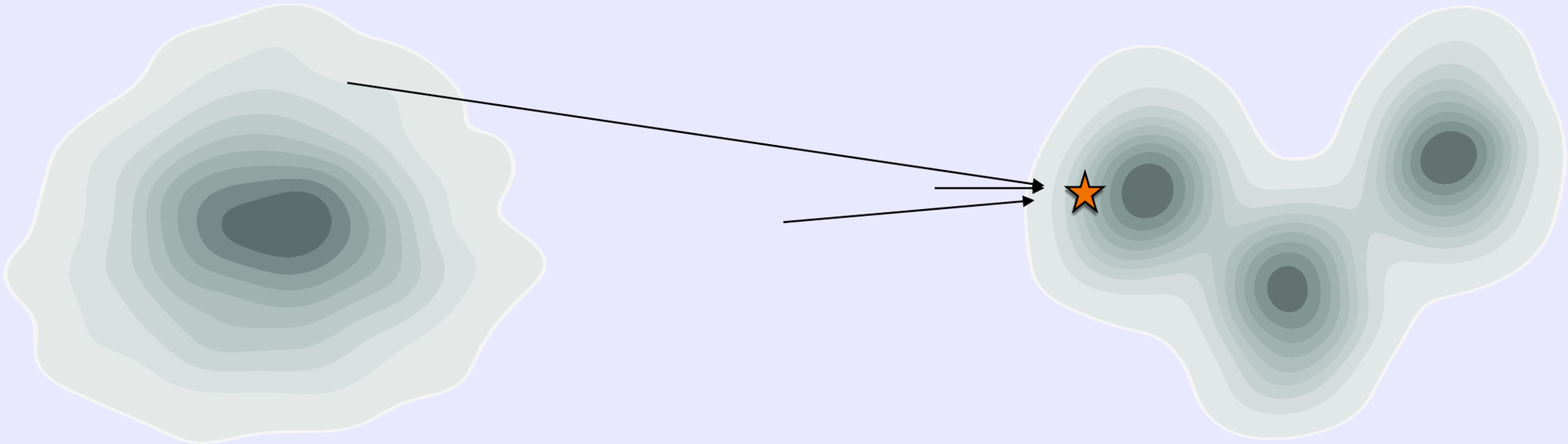
π_0

Target samples

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Recap: Flow Matching

Flow Matching is interpolation!



Source Distribution

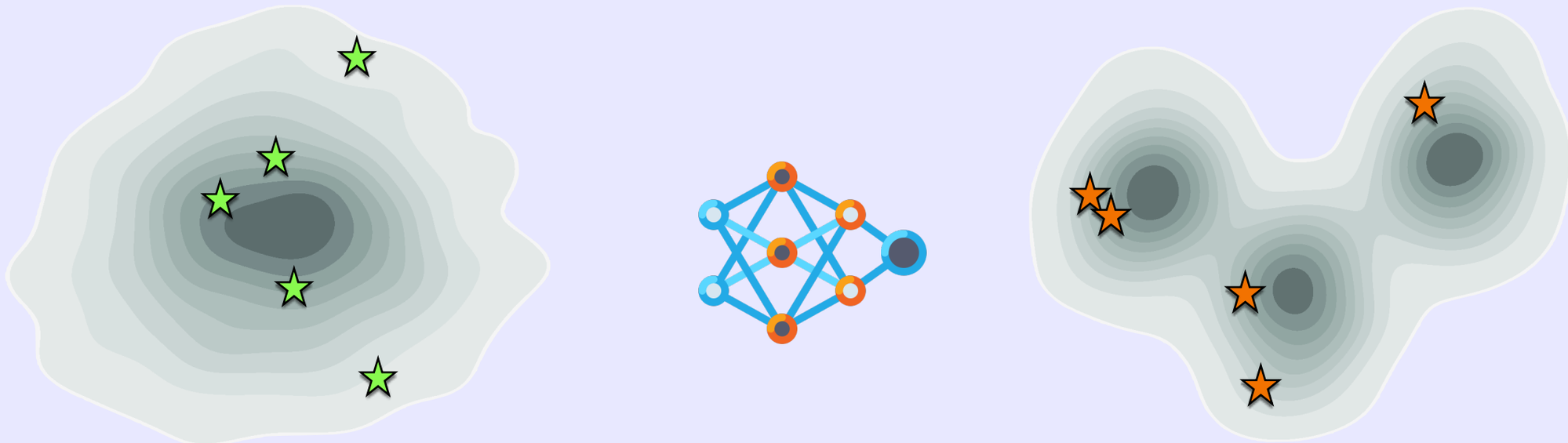
π_0

Target samples

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Recap: Flow Matching

Flow Matching is interpolation!



Source Distribution

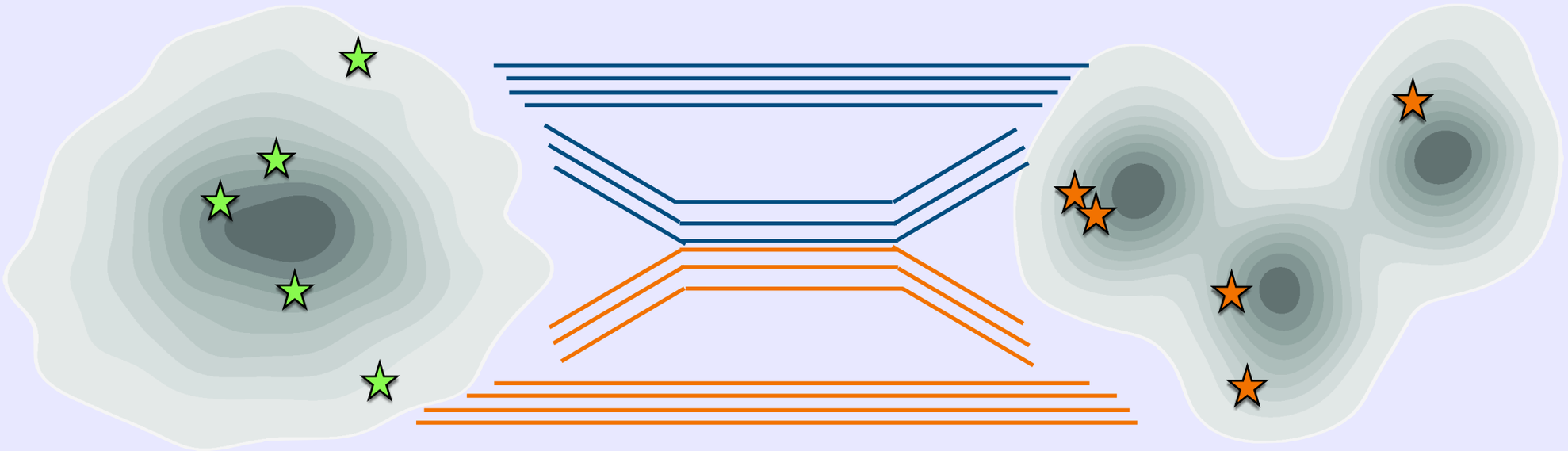
π_0

Target samples

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Recap: Flow Matching

Flow Matching is interpolation!



Source Distribution

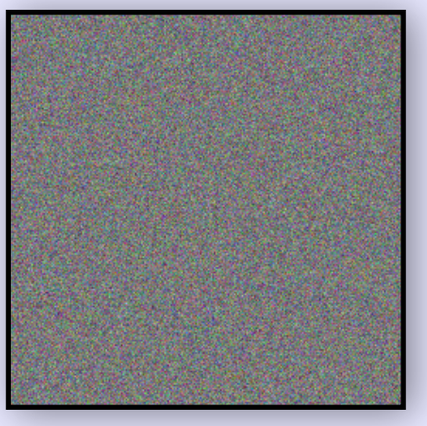
π_0

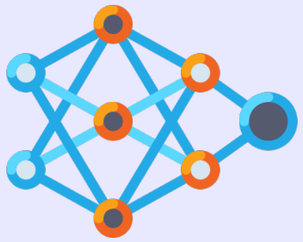
Target distribution

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image...





Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

t

Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

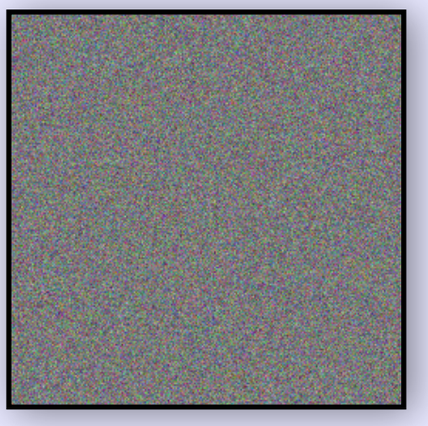
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

t

$$img_t = t \cdot img + (1 - t) \cdot noise$$



Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

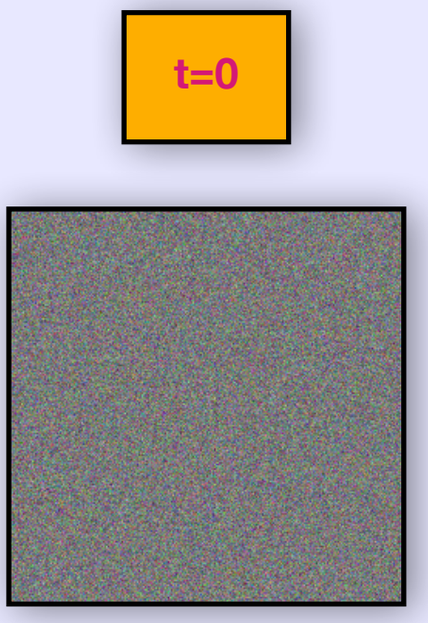
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

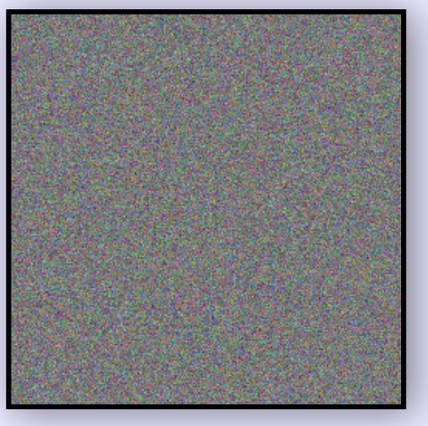
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

t=0





Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

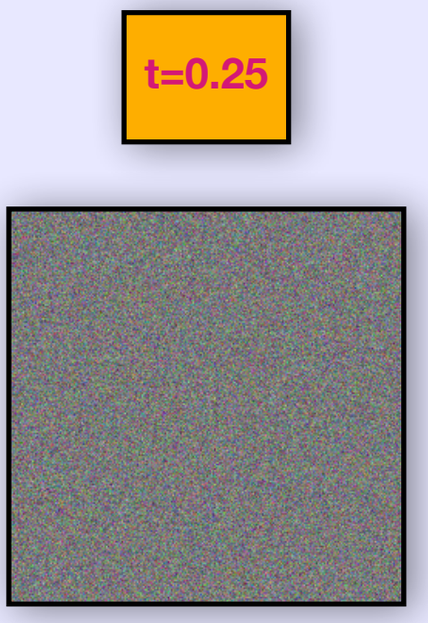
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

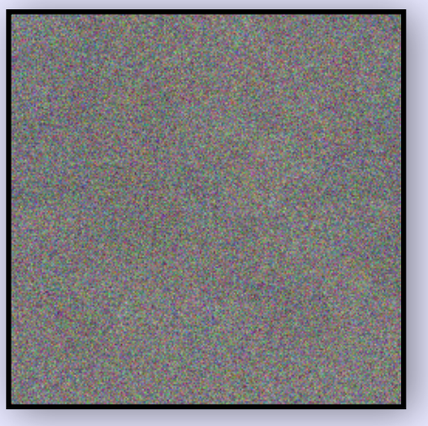
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

t=0.25





Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

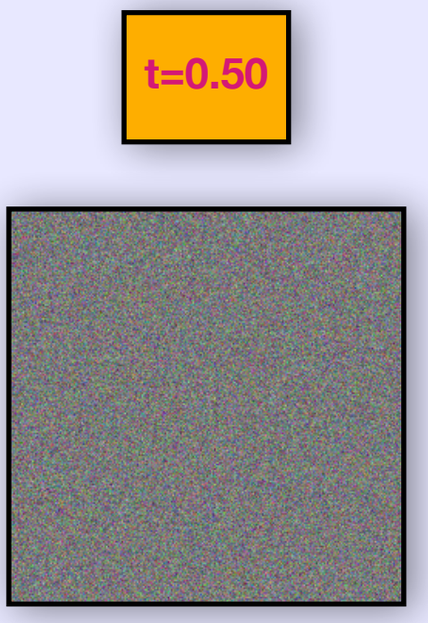
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

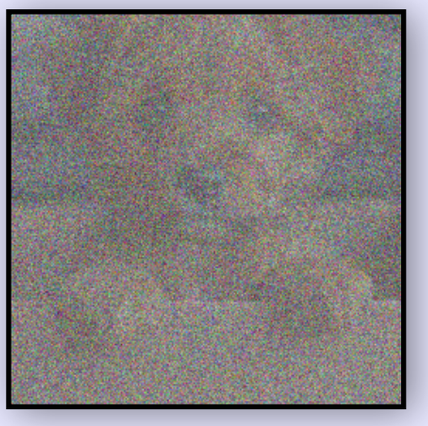
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

t=0.50





Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

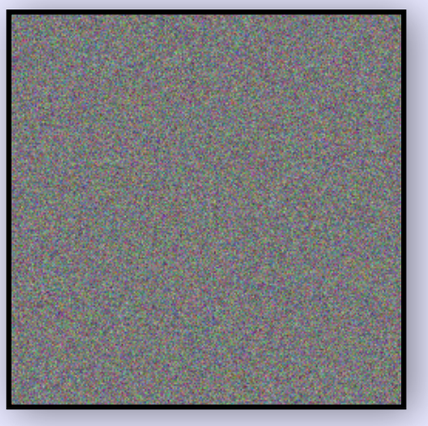
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

t=0.75



Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

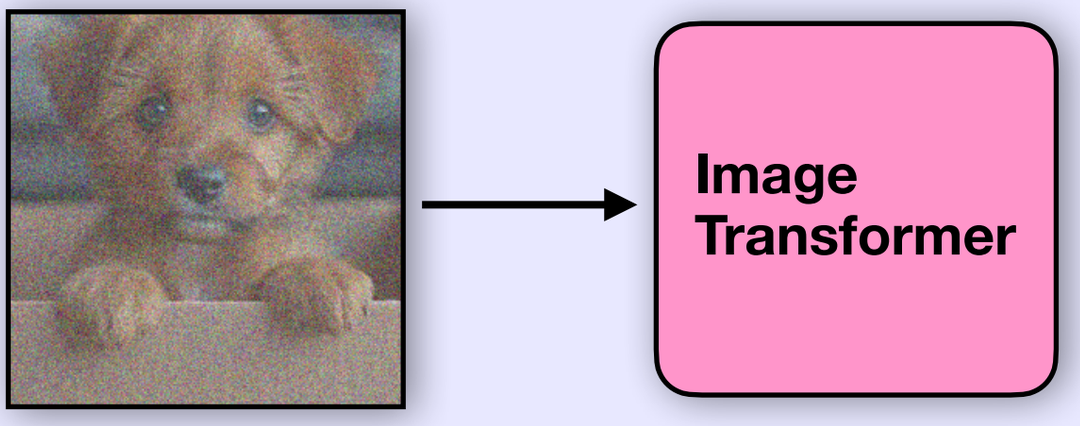
Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…



Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SIT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

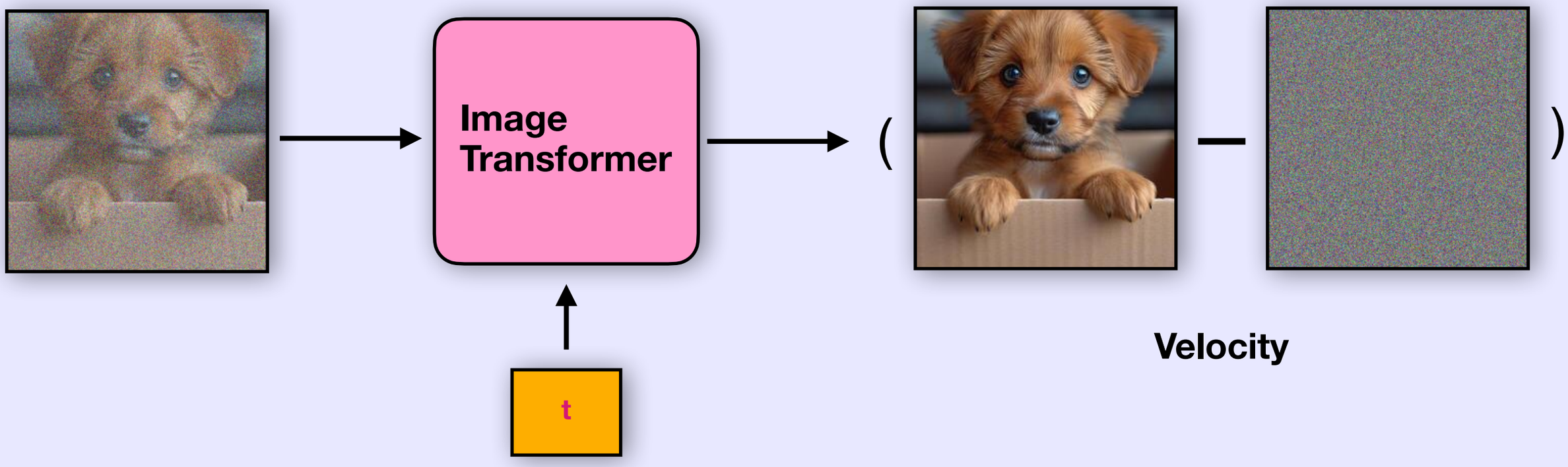
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

Training



Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

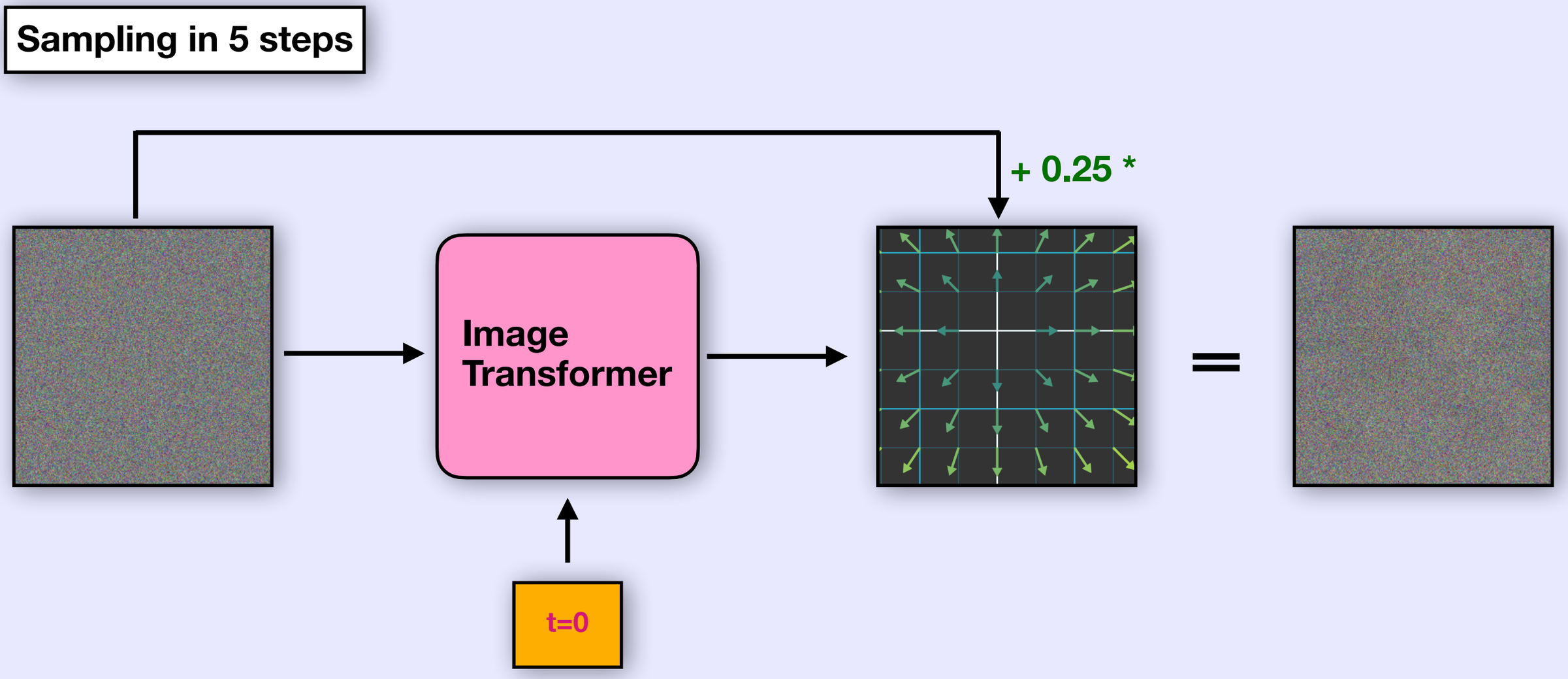
Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…



Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

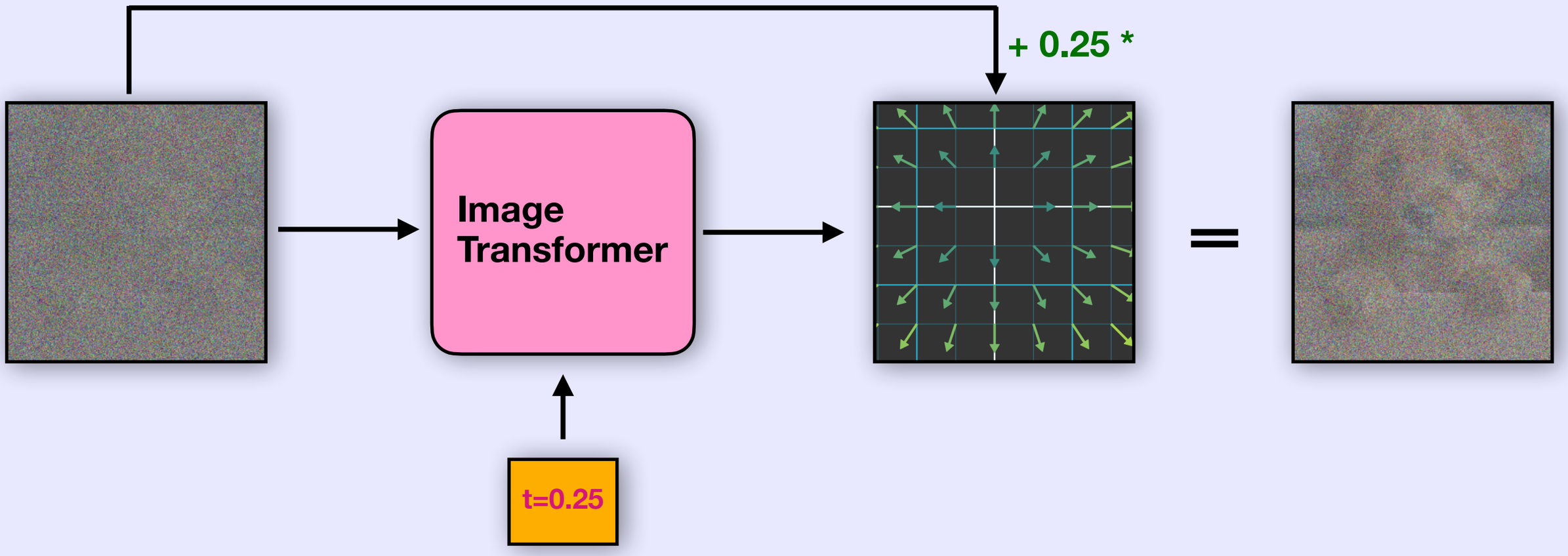
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

Sampling in 5 steps



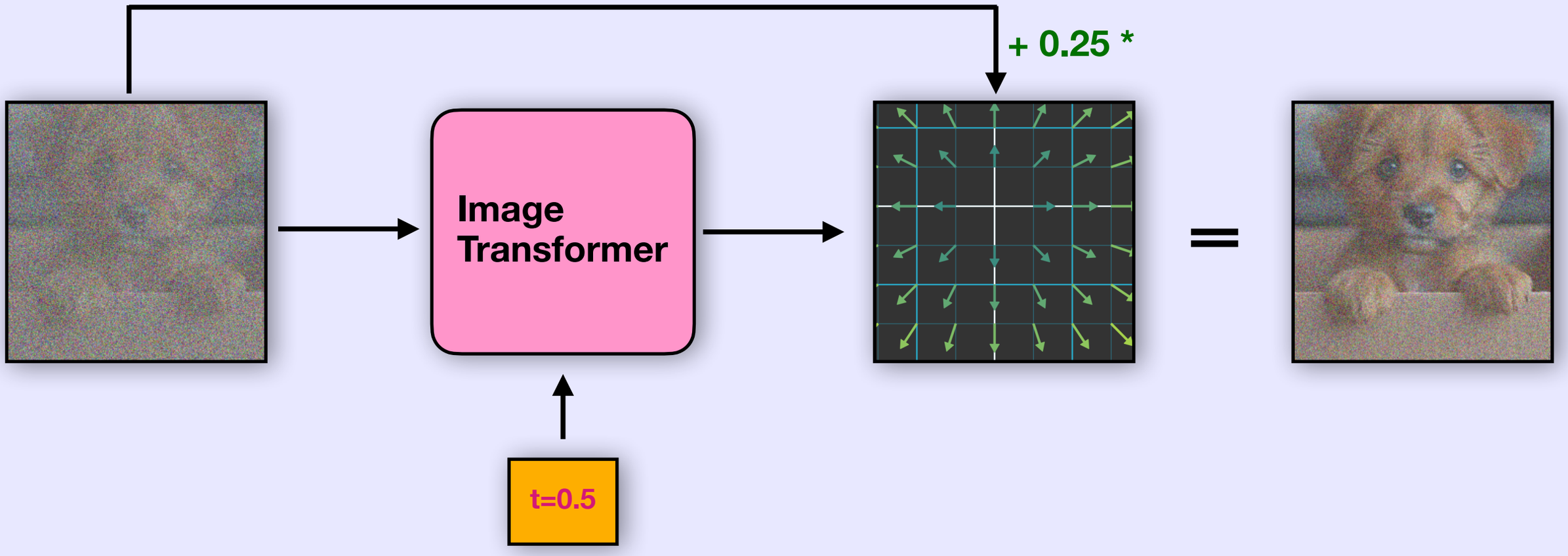
STEP: 2

Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.
Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.
Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

Sampling in 5 steps



STEP: 3

Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

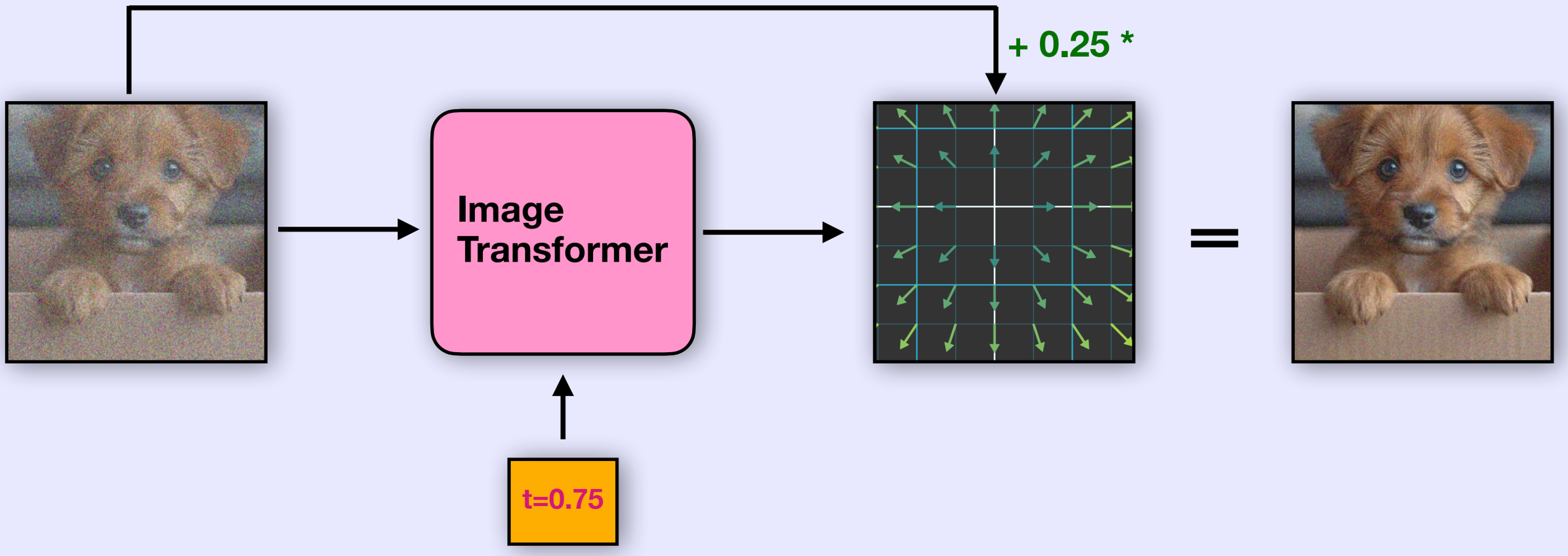
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

Sampling in 5 steps



STEP: 4

Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

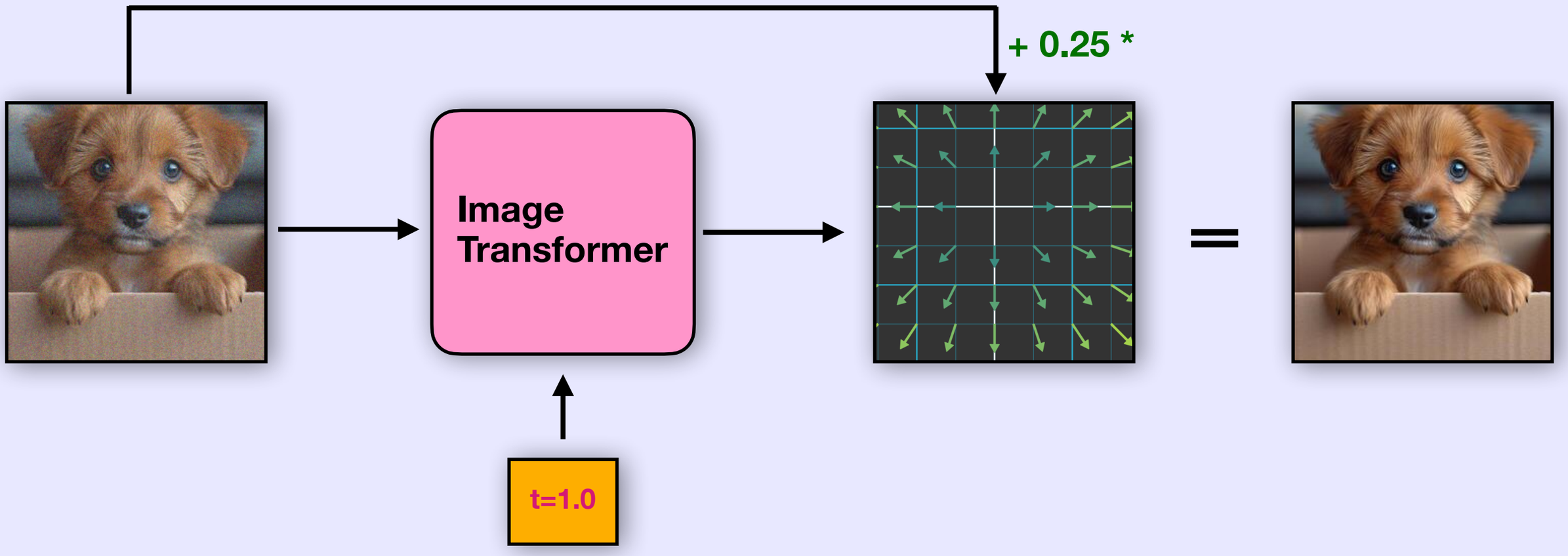
Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

Flow Matching for Image Generation

Image, image and more image…

Sampling in 5 steps



STEP: 5

Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Lipman, Yaron, et al. "Flow Matching for Generative Modeling.." ICLR 2023.

Liu, Xingchao, et al. "Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow." ICLR 2023.

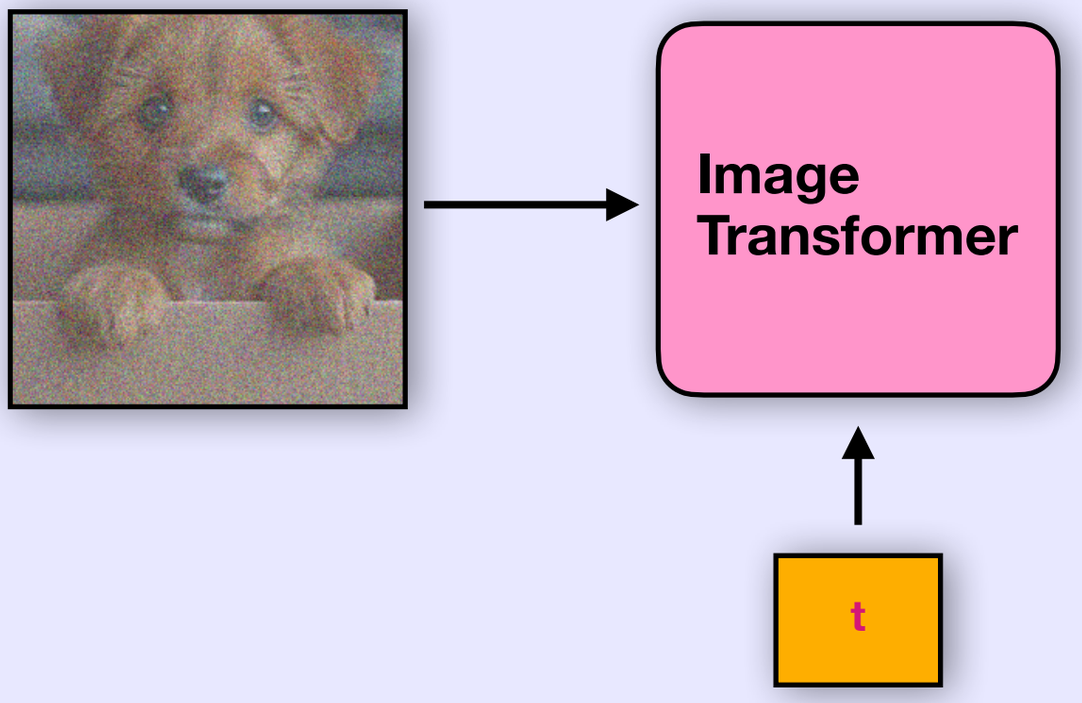
Co-denoising for Image Generation

Image, image and more image…

Can we denoise more things to improve image generation?

Co-denoising for Image Generation

I was kinda lying before!

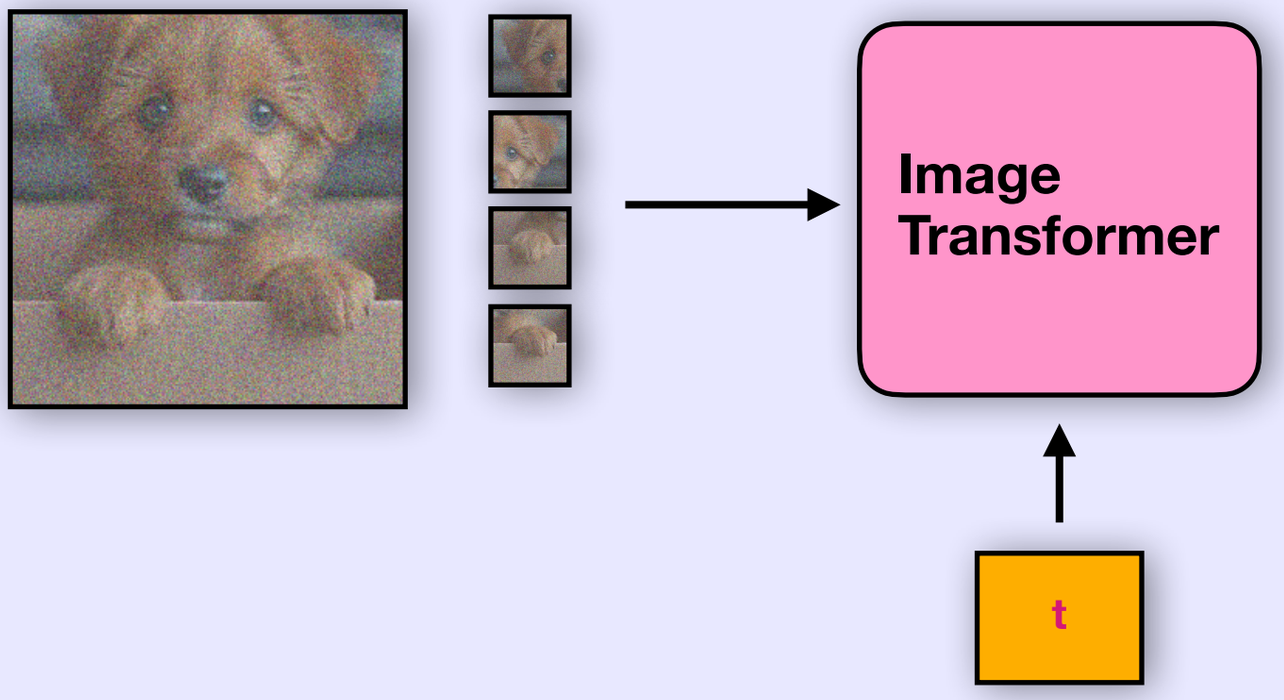


Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Co-denoising for Image Generation

There is more to it…



Li, Tianhong, et al. "Back to Basics: Let Denoising Generative Models Denoise." arXiV 2025.
Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.
Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Co-denoising for Image Generation

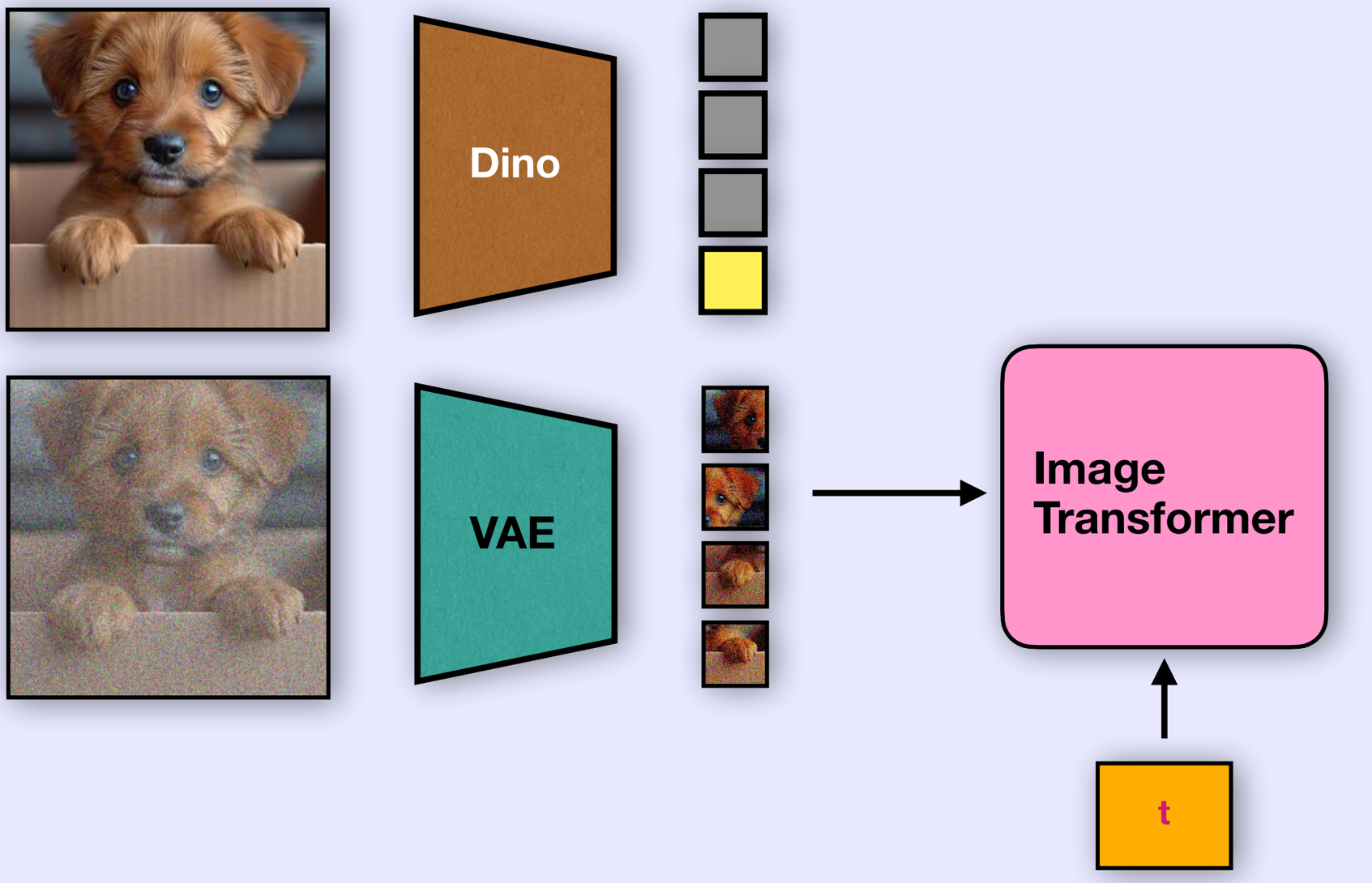
There is more to it…

Esser, Patrick, et al. "Scaling Rectified Flow Transformers for High-Resolution Image Synthesis." ICML 2024.

Ma, Nanye, et al. "SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers." ECCV 2024.

Co-denoising for Image Generation

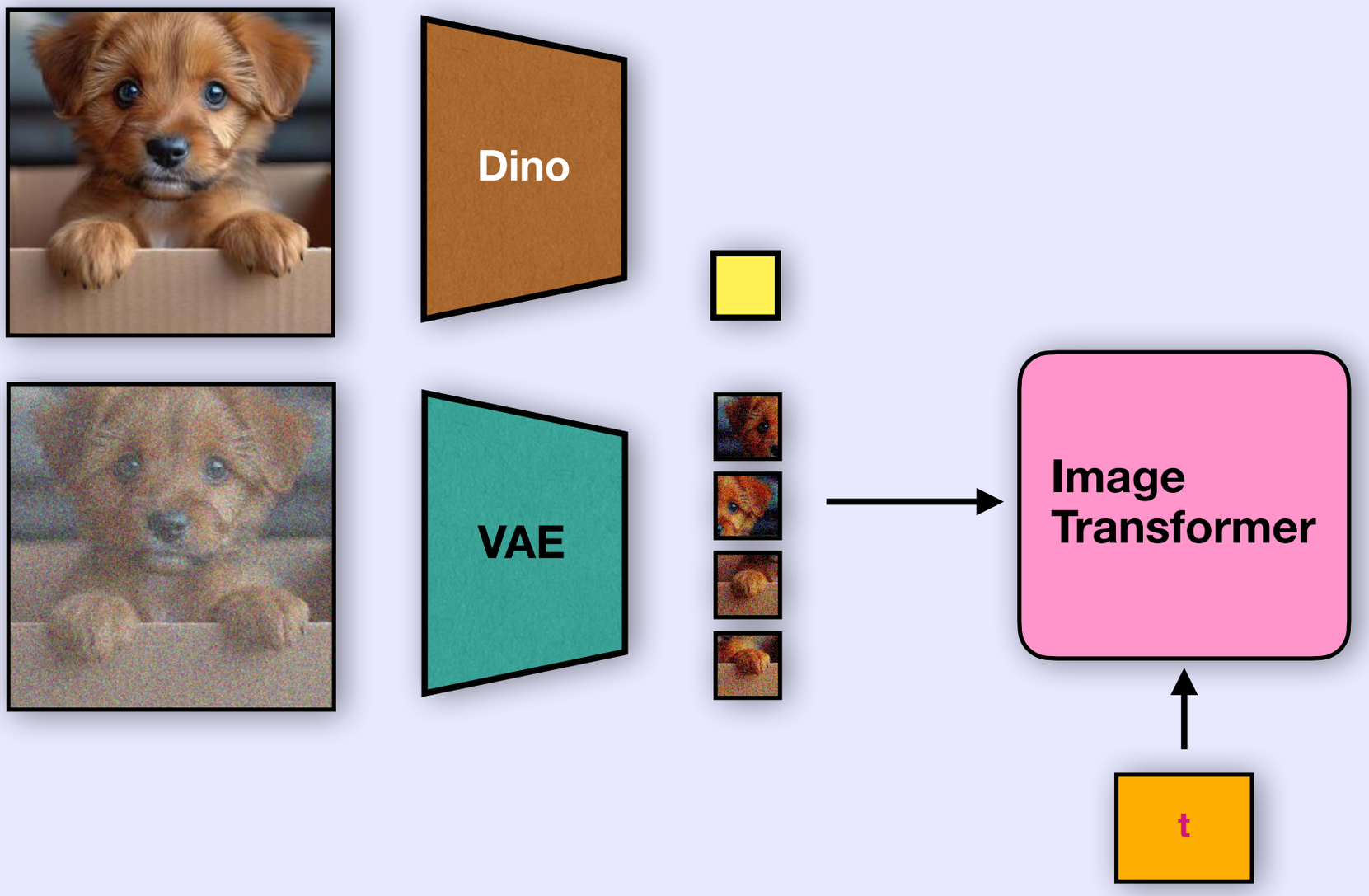
Co-denoising with Dino class token!



Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

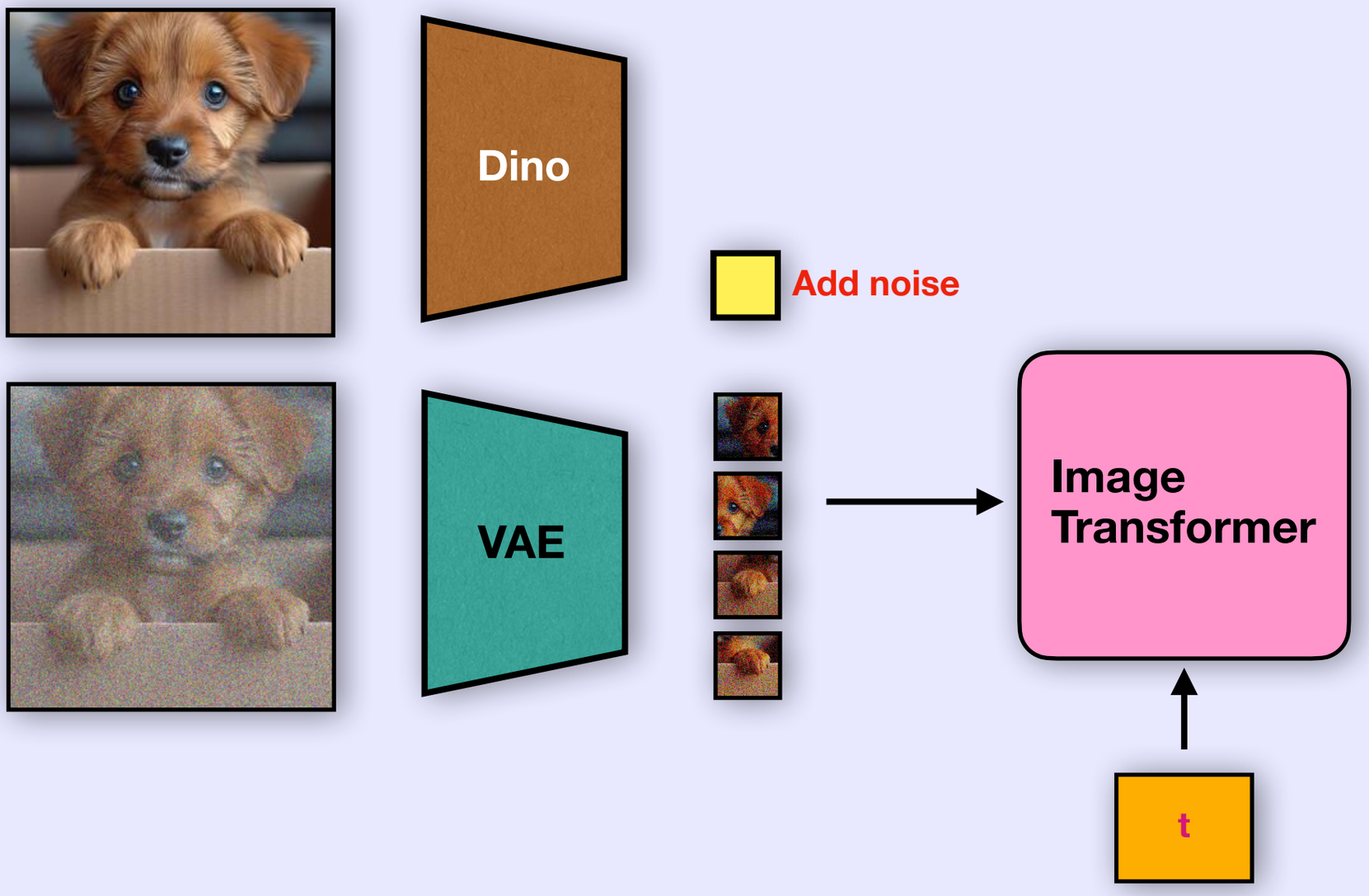
Co-denoising with Dino class token!



Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

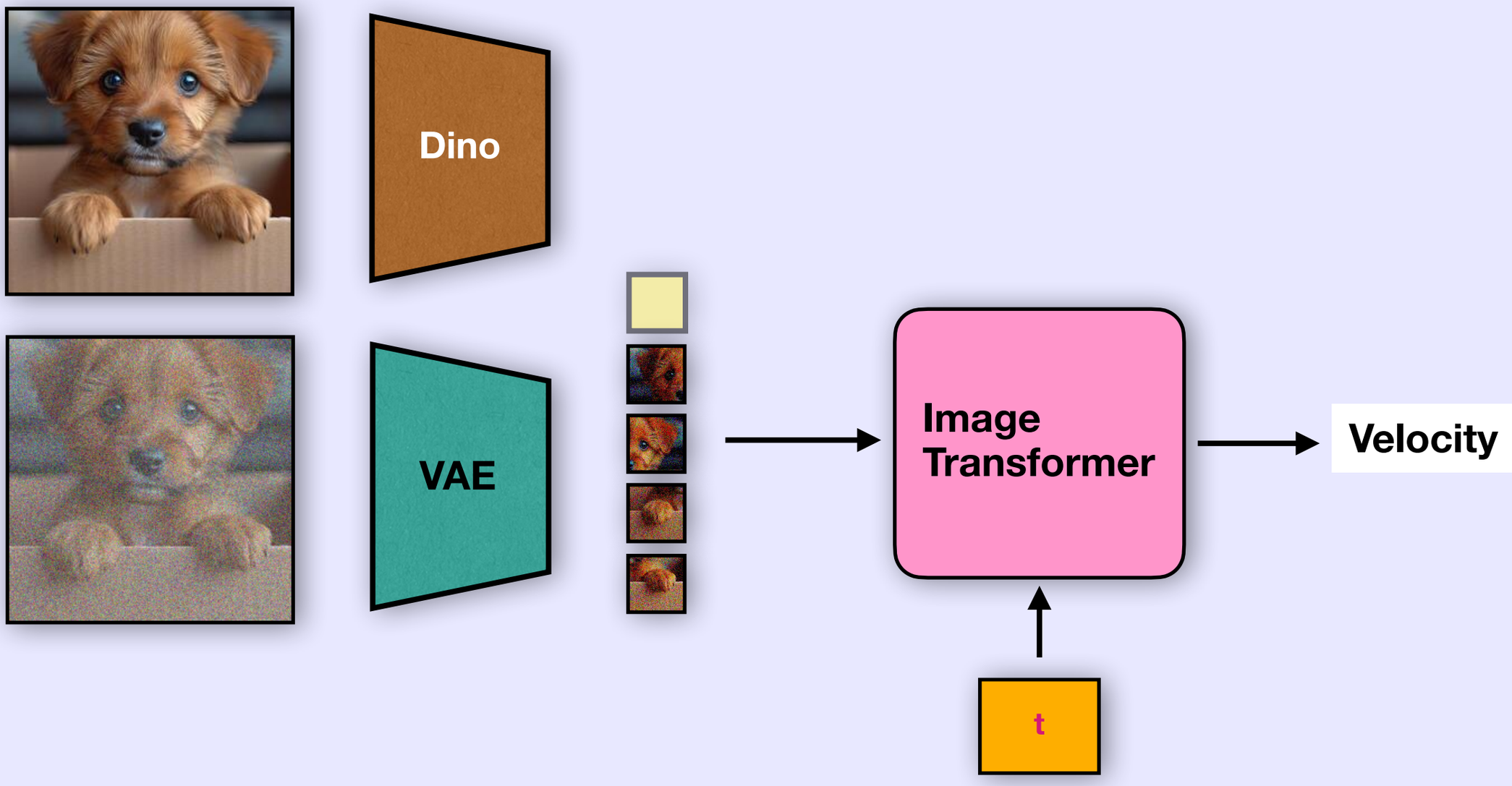
Co-denoising with Dino class token!



Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

Co-denoising with Dino class token!



Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

Co-denoising with Dino class token!

Faster Convergence

Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

Co-denoising with Dino class token!

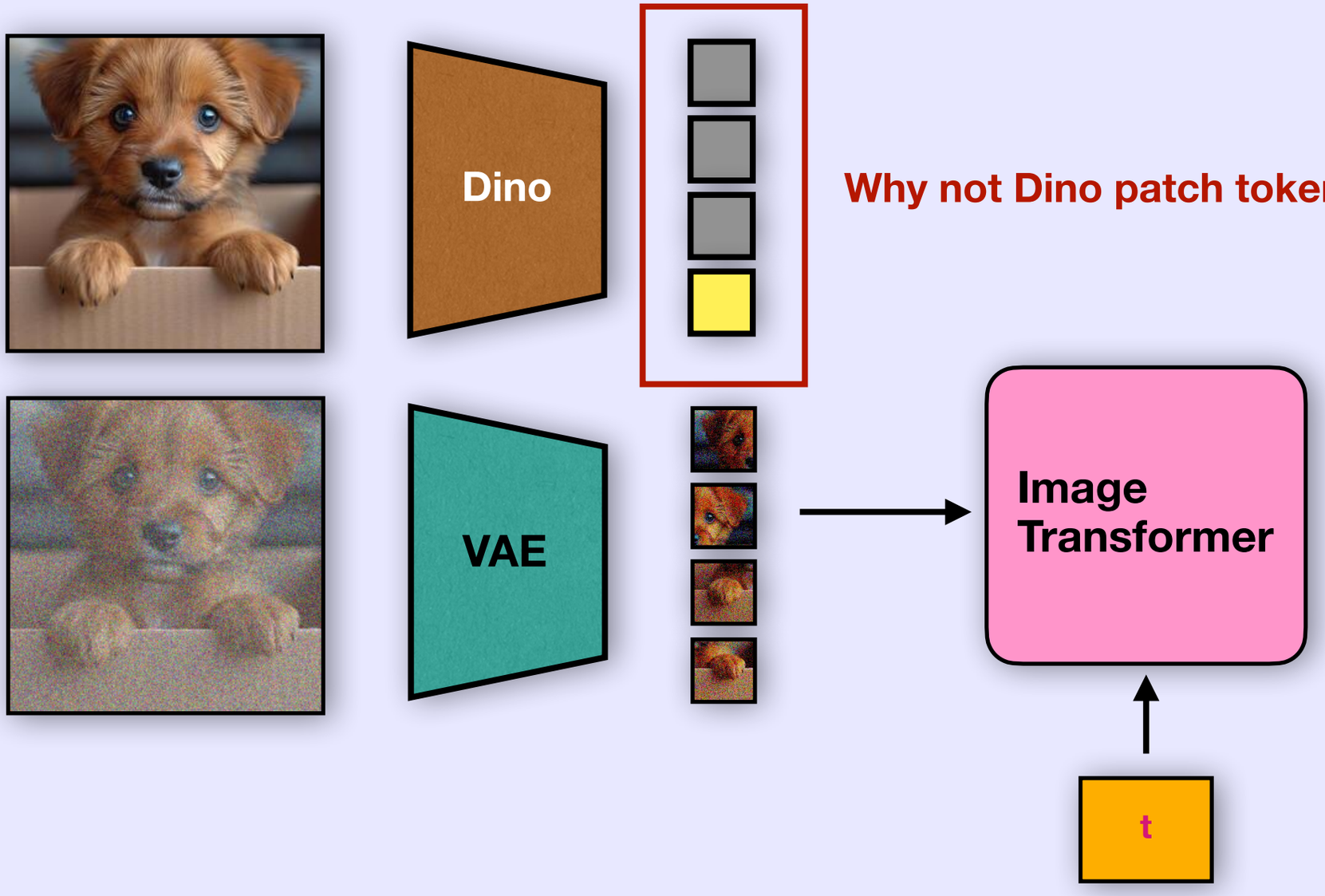
Faster Convergence

Forced to decode and understand rich representation!!!!

Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

Co-denoising with Dino class token!

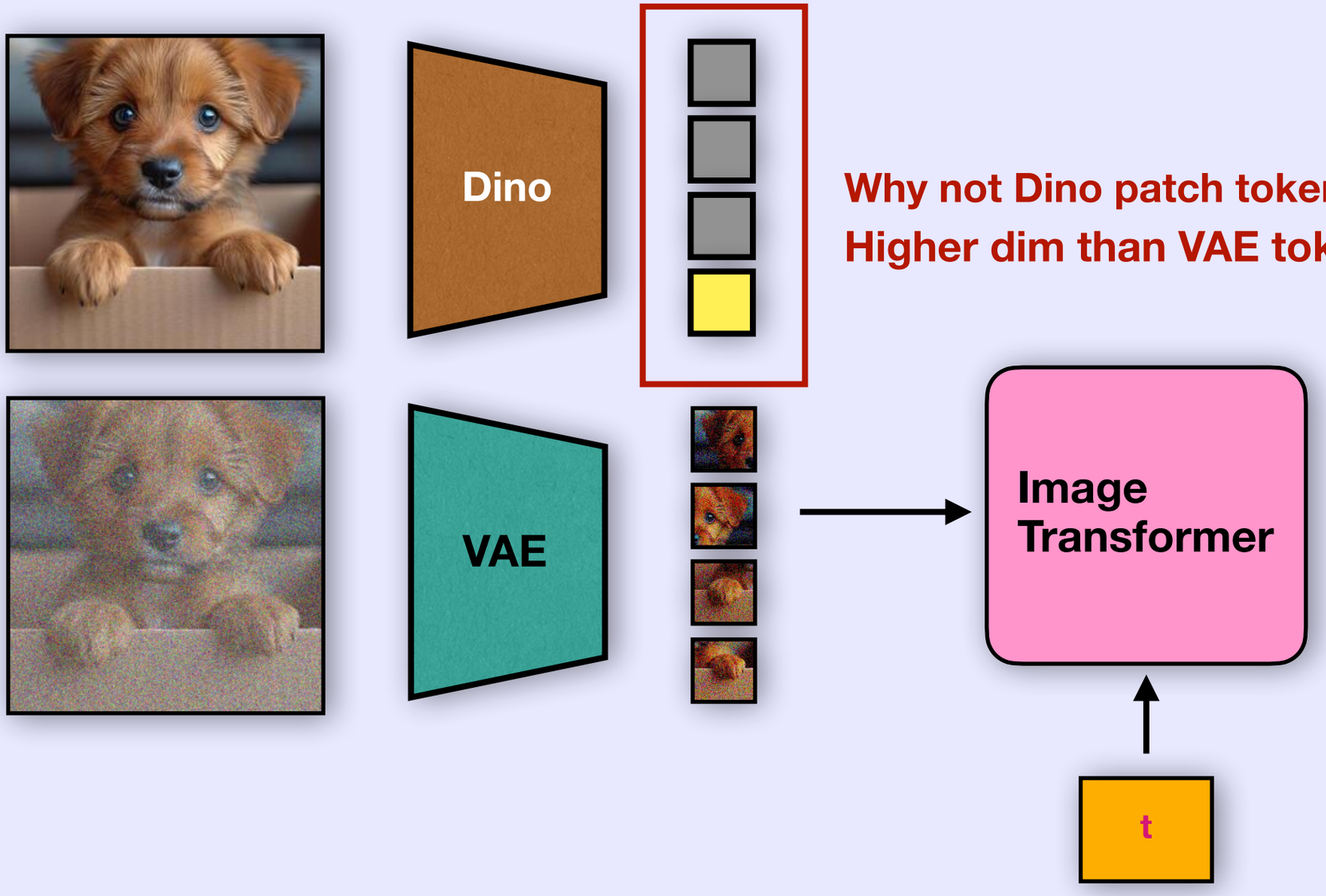


Why not Dino patch tokens?

Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

Co-denoising with Dino class token!



Why not Dino patch tokens?

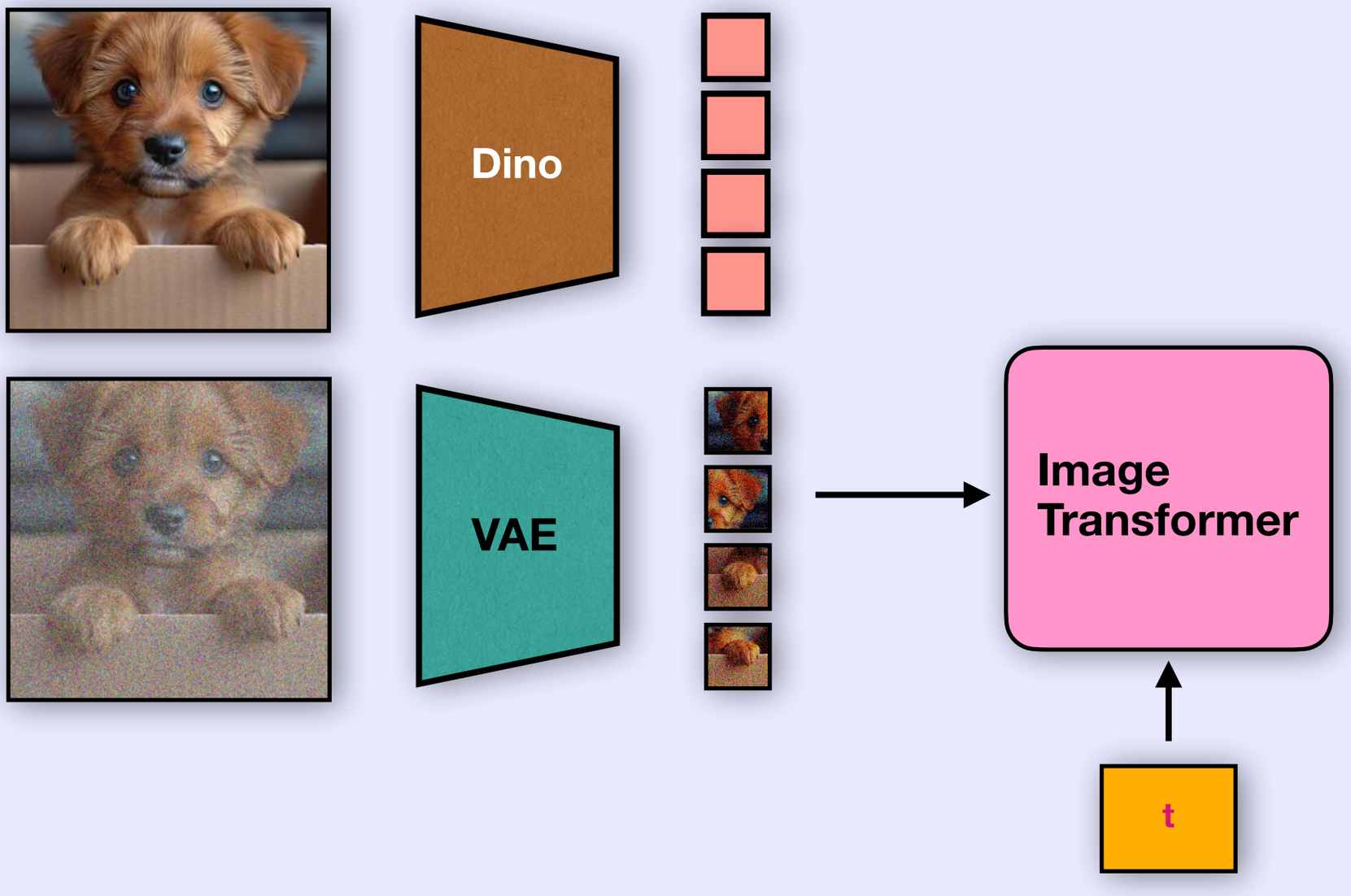
Higher dim than VAE tokens → Learns Shortcut!

Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Co-denoising for Image Generation

Co-denoising with Dino patch tokens!

Dim. Reduction w/ PCA

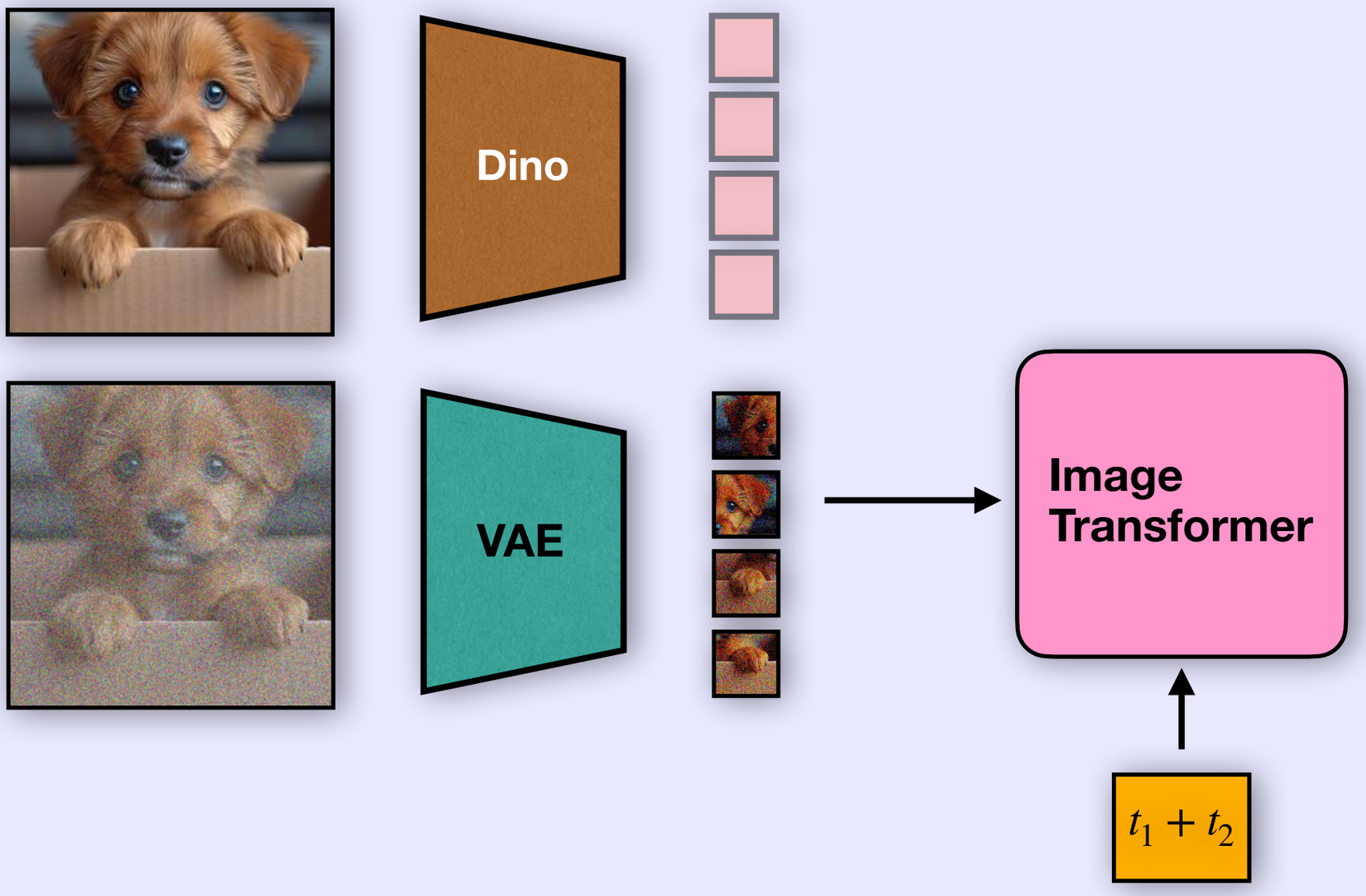


Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Co-denoising for Image Generation

Co-denoising with Dino patch tokens!

Dim. Reduction w/ PCA

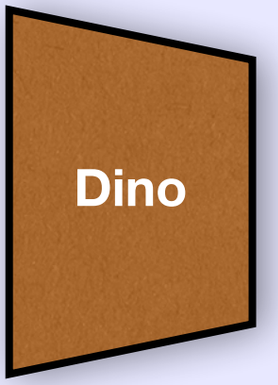


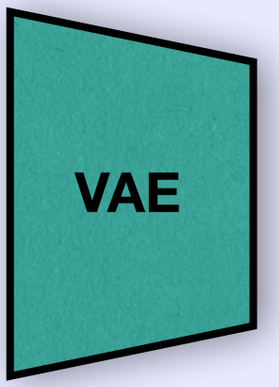
Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

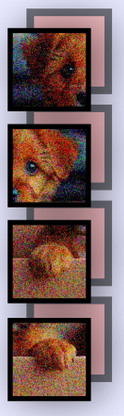
Co-denoising for Image Generation

Co-denoising with Dino patch tokens!

Dim. Reduction w/ PCA







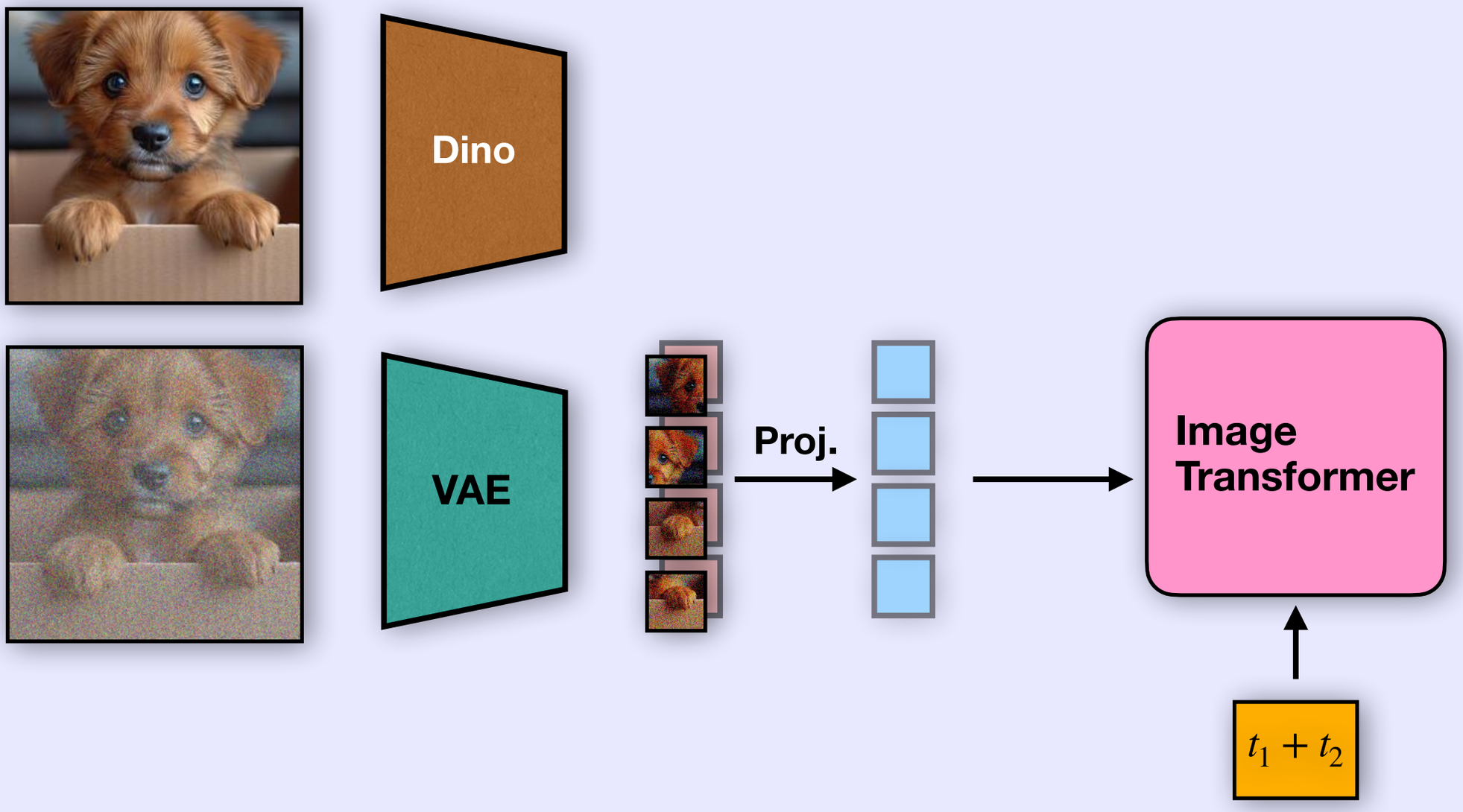
$t_1 + t_2$

Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Co-denoising for Image Generation

Co-denoising with Dino patch tokens!

Dim. Reduction w/ PCA

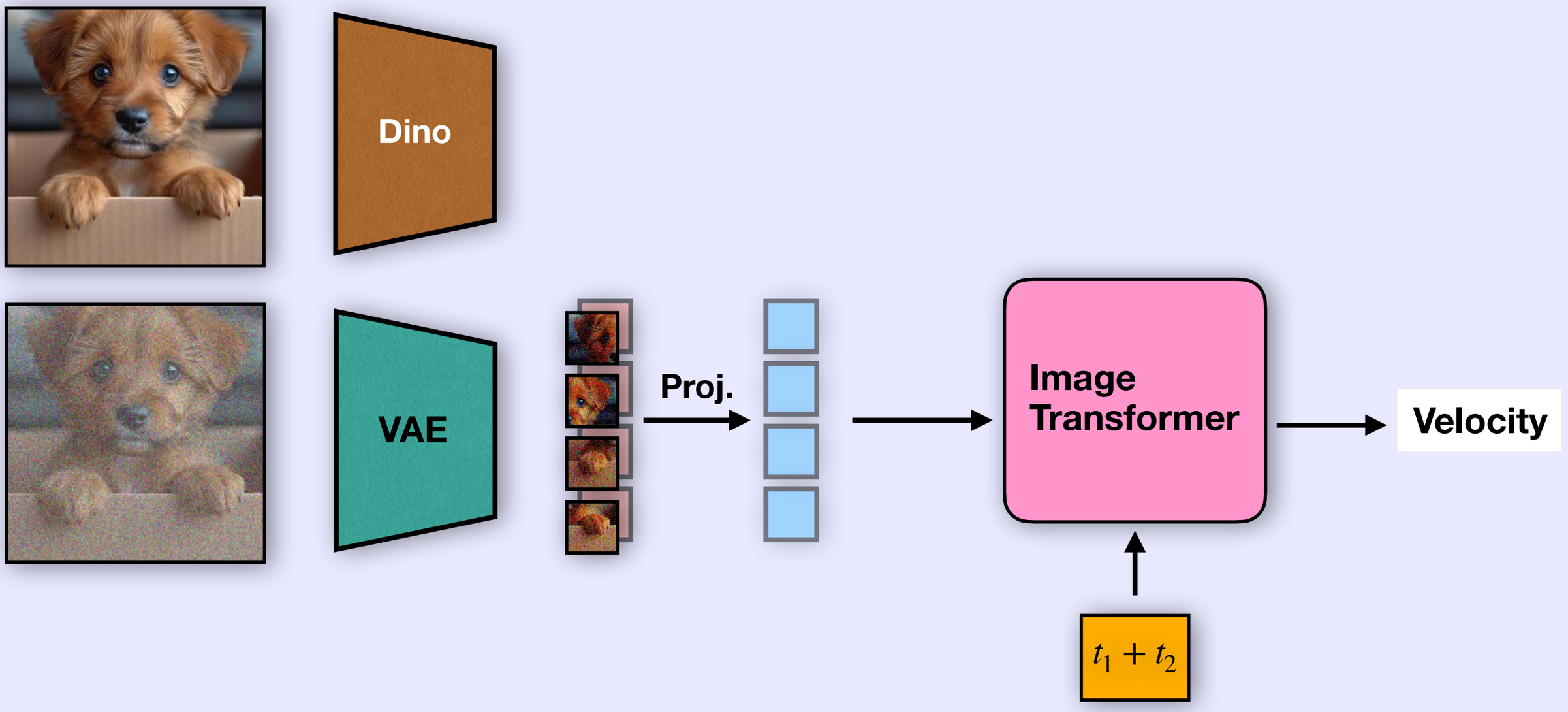


Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Co-denoising for Image Generation

Co-denoising with Dino patch tokens!

Dim. Reduction w/ PCA



Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Co-denoising for Image Generation

Co-denoising with Dino patch tokens!

Faster Convergence

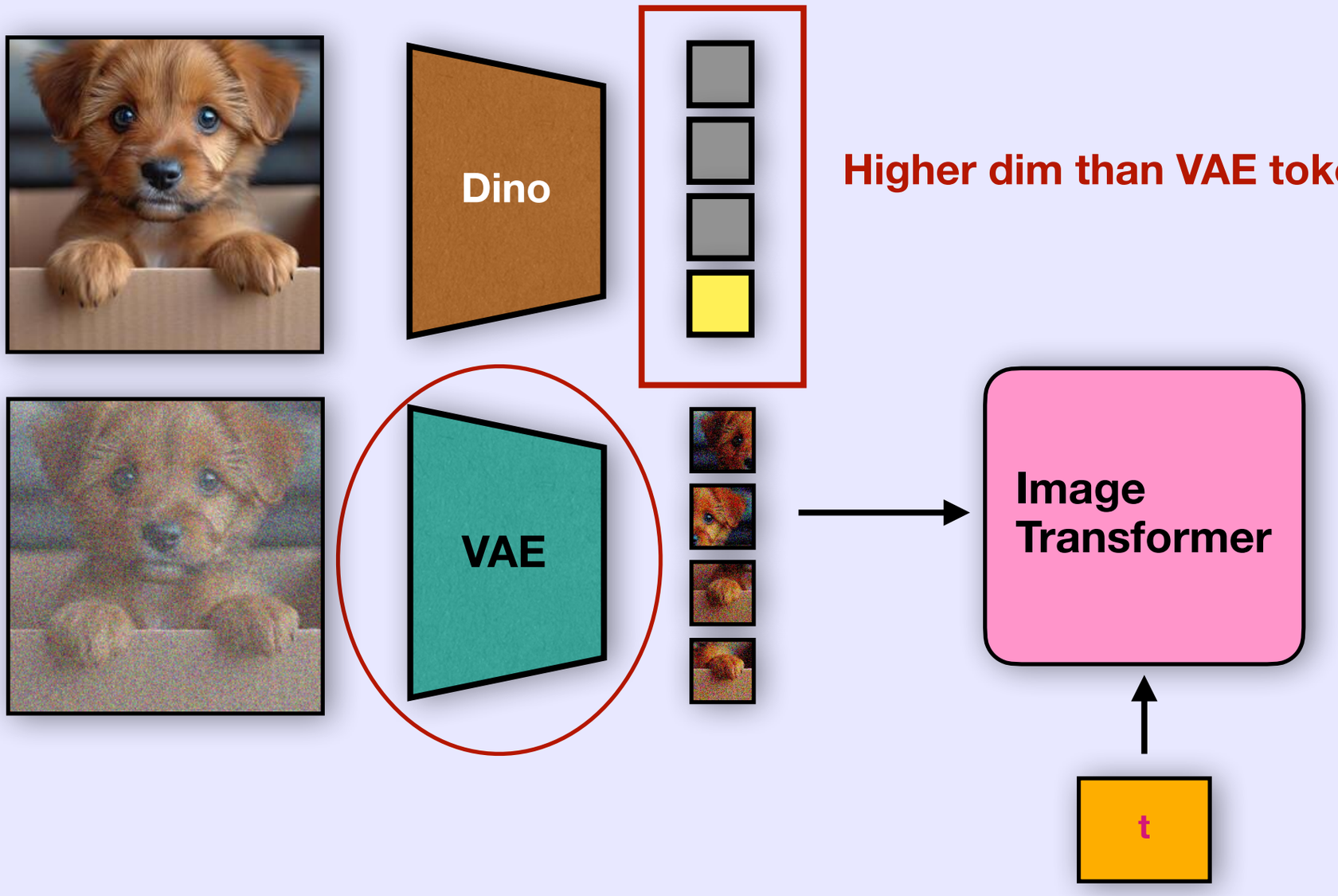
Forced to decode even better representation!!!!

Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Co-denoising for Image Generation

Co-denoising with Dino patch tokens!

Why PCA?

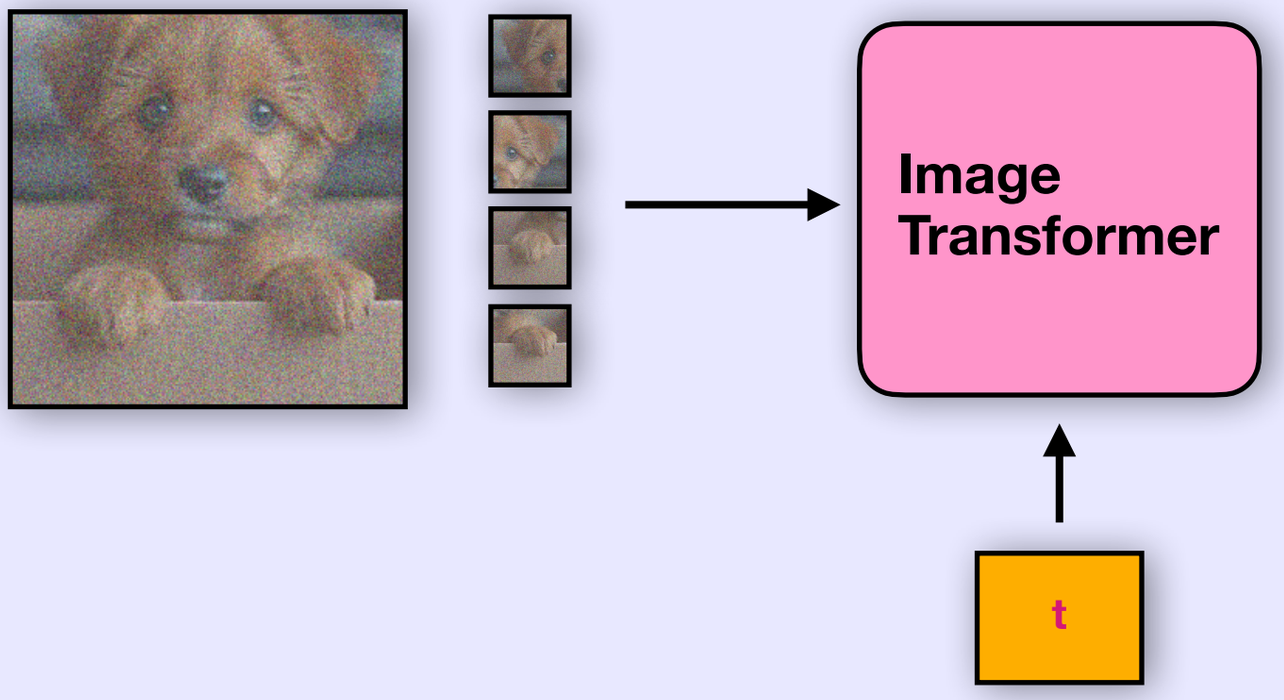


Higher dim than VAE tokens → Learns Shortcut!

Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Co-denoising for Image Generation

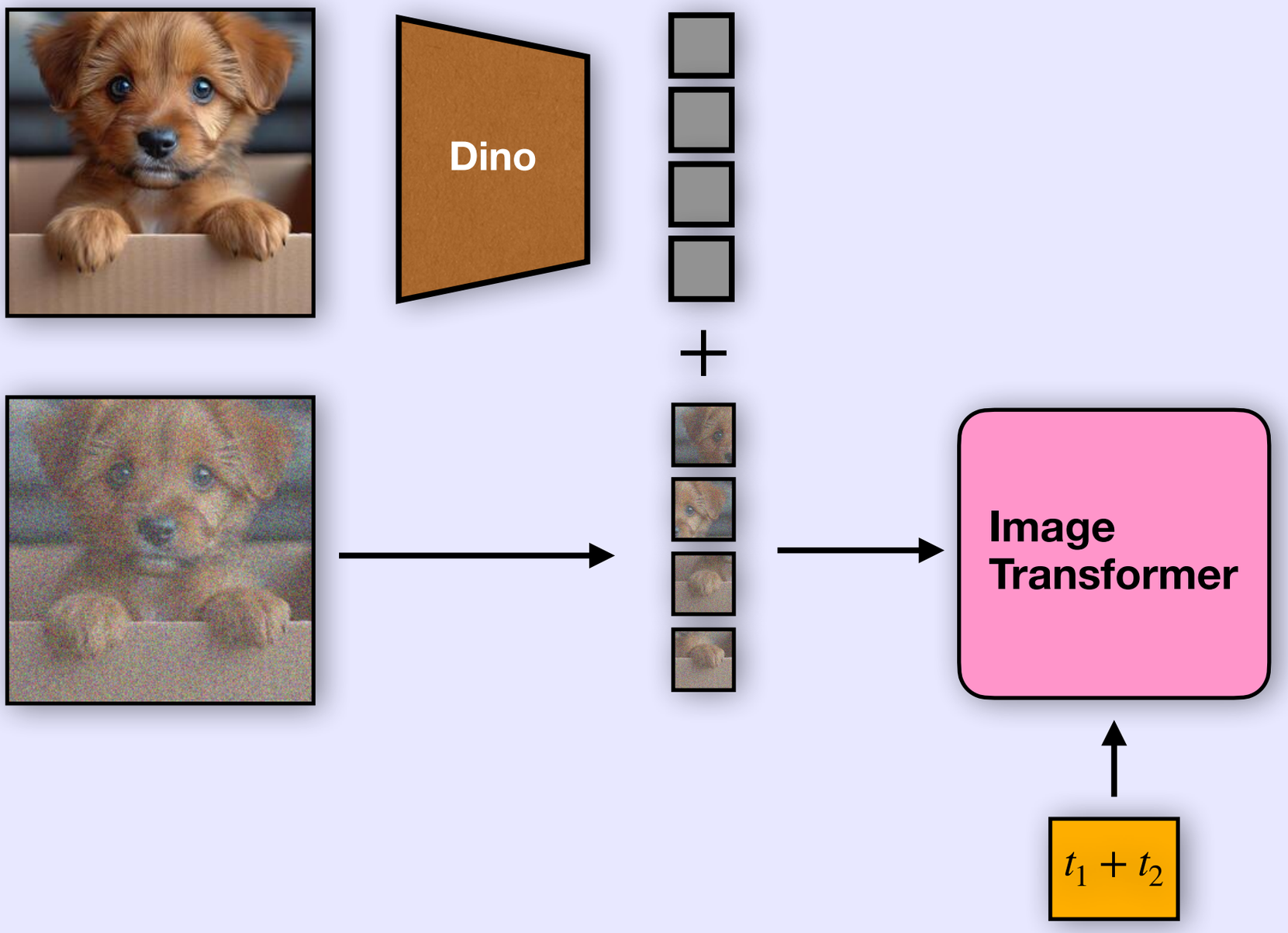
Pixel Models are a thing!!! So no more VAE!



Li, Tianhong, et al. "Back to Basics: Let Denoising Generative Models Denoise." arXiV 2025.

Co-denoising for Image Generation

Pixel Models + Raw Dino!!

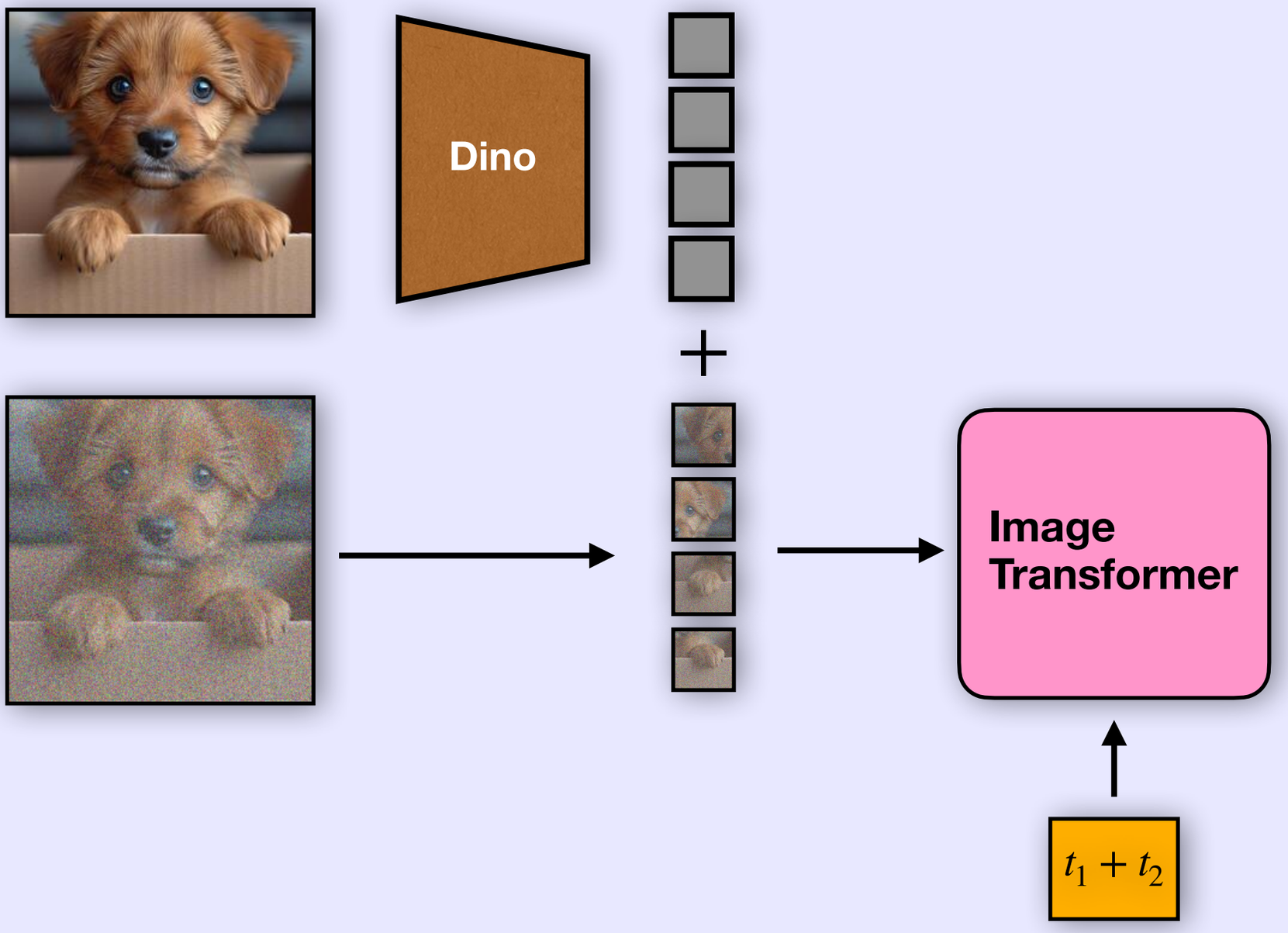


Baade, Alan, et al. "Latent Forcing: Reordering the Diffusion Trajectory for Pixel-Space Image Generation." arXiV 2025.

Co-denoising for Image Generation

Pixel Models + Raw Dino!!

How to get 2 velocities?

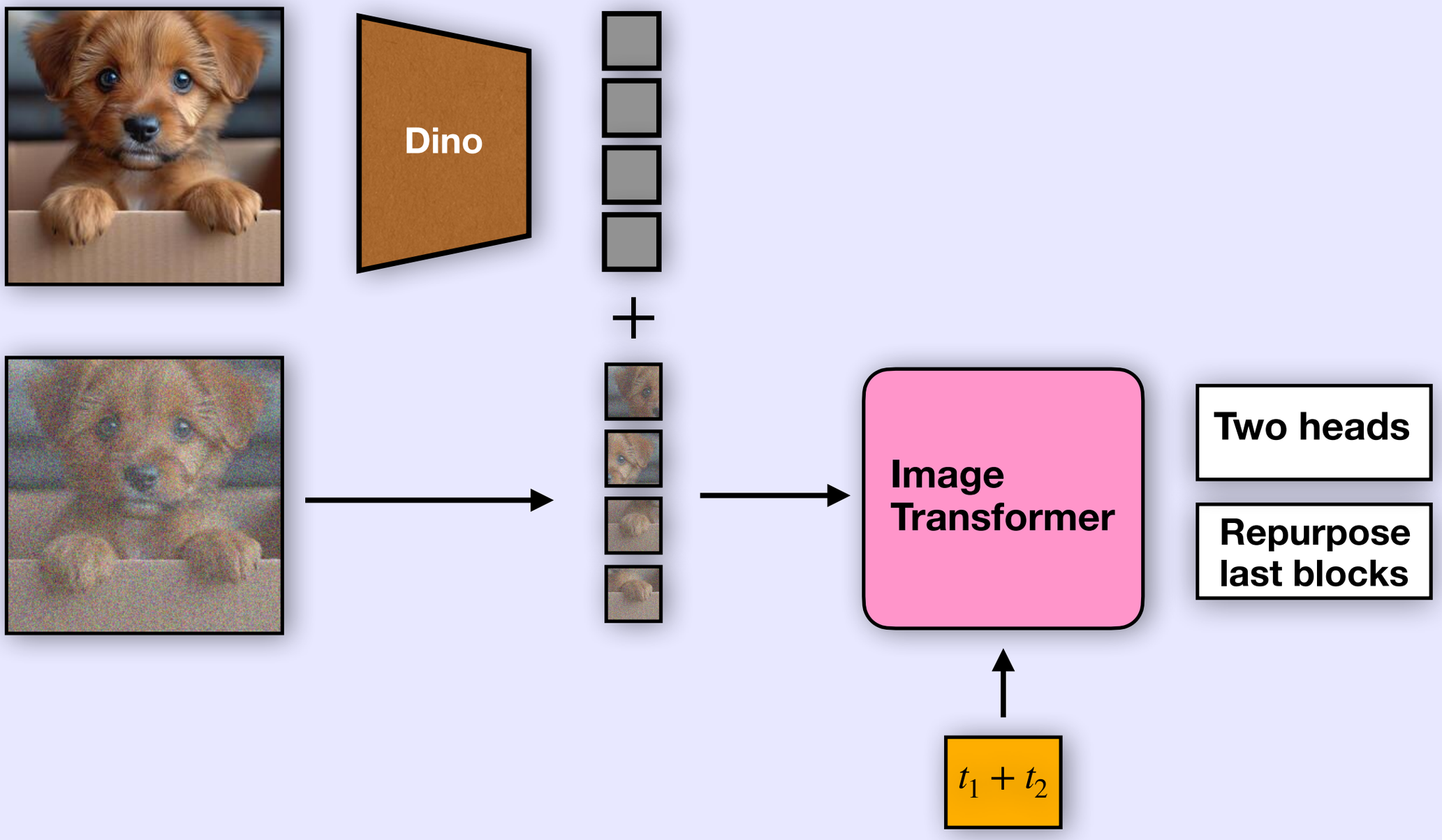


Baade, Alan, et al. "Latent Forcing: Reordering the Diffusion Trajectory for Pixel-Space Image Generation." arXiV 2025.

Co-denoising for Image Generation

Pixel Models + Raw Dino!!

How to get 2 velocities?



Baade, Alan, et al. "Latent Forcing: Reordering the Diffusion Trajectory for Pixel-Space Image Generation." arXiV 2025.

Co-denoising for Image Generation

Pixel Models + Raw Dino!!

Cool Insight: Denoise Dino features faster than Image features!

Baade, Alan, et al. "Latent Forcing: Reordering the Diffusion Trajectory for Pixel-Space Image Generation." arXiV 2025.

Co-denoising for Image Generation

All previous models!

Cool Insight: An intristic guidance to image generation!

Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Baade, Alan, et al. "Latent Forcing: Reordering the Diffusion Trajectory for Pixel-Space Image Generation." arXiV 2025.

Co-denoising for Image Generation

All previous models!

RANT: Does **denoising DINO features** make any sense?

Wu, Ge, et al. "Representation Entanglement for Generation: Training Diffusion Transformers Is Much Easier Than You Think." ECCV 2024.

Kouzelis, Theodoros, et al. "Boosting Generative Image Modeling via Joint Image-Feature Synthesis." ICLR 2025.

Baade, Alan, et al. "Latent Forcing: Reordering the Diffusion Trajectory for Pixel-Space Image Generation." arXiV 2025.

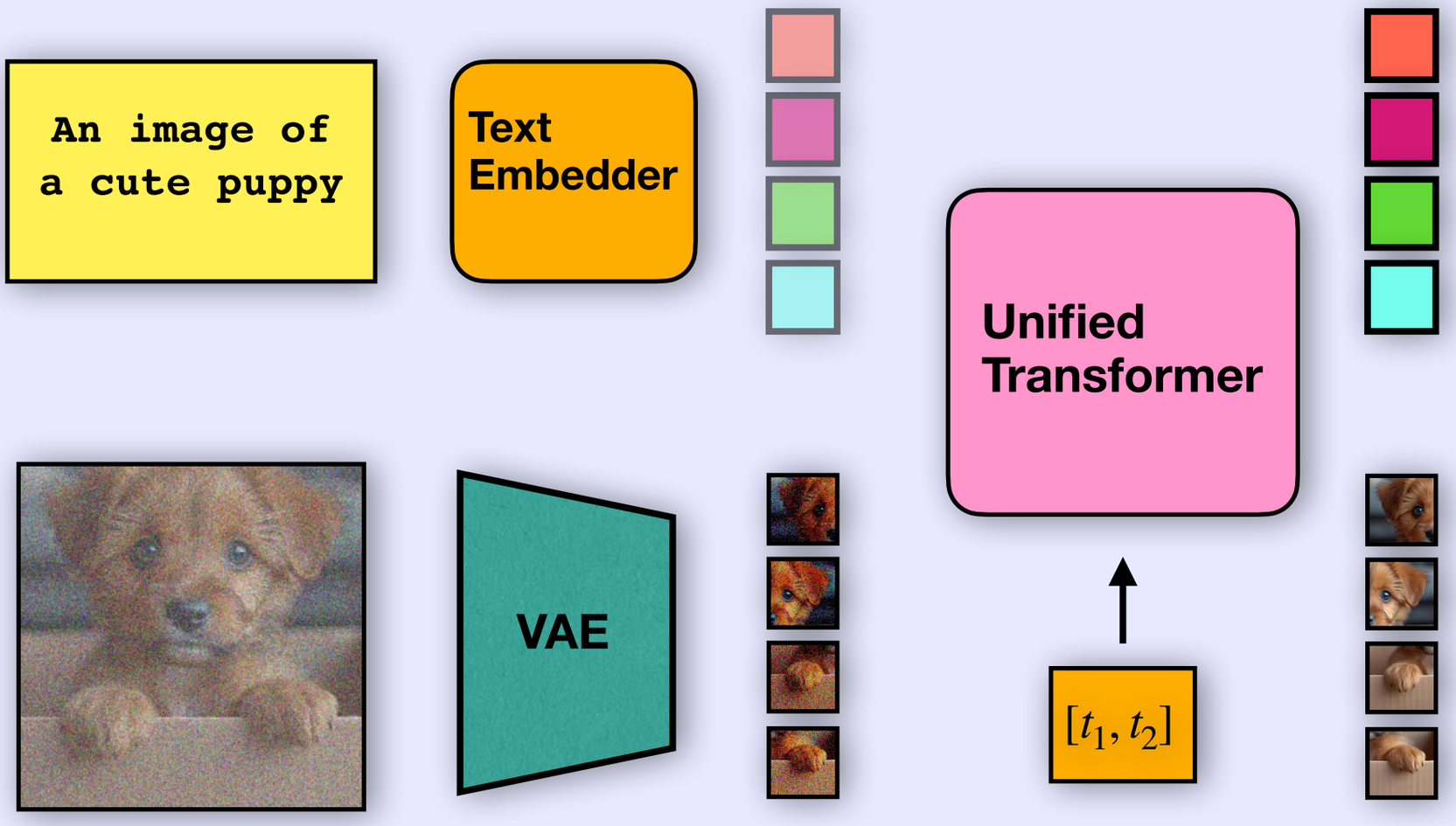
Co-denoising!

Denoise something meaningful!

Bao, Fan, et al. "One Transformer Fits All Distributions in Multi-Modal Diffusion at Scale." CVPR 2023.

Co-denoising!

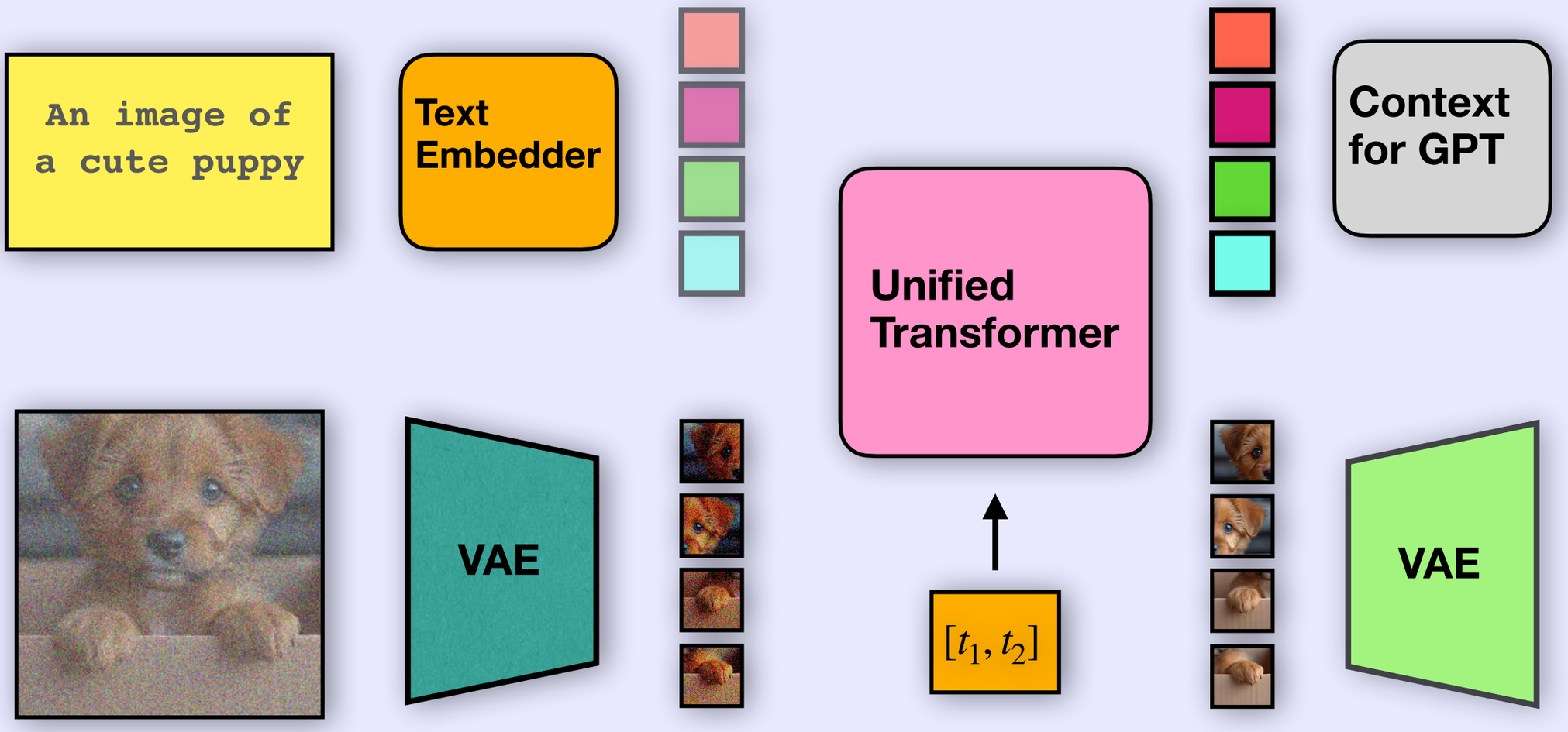
Denoise text and image together!!!!



Bao, Fan, et al. "One Transformer Fits All Distributions in Multi-Modal Diffusion at Scale." CVPR 2023.

Co-denoising!

Denoise text and image together!!!!



Bao, Fan, et al. "One Transformer Fits All Distributions in Multi-Modal Diffusion at Scale." CVPR 2023.

Co-denoising!

Denoise text and image together!!!!

Cool thing: Learn **conditional and marginal** in one framework!

Bao, Fan, et al. "One Transformer Fits All Distributions in Multi-Modal Diffusion at Scale." CVPR 2023.

Co-denoising!

Denoise text and image together!!!!

A guess: Two **t** with two different modalities; slower convergence!

Bao, Fan, et al. "One Transformer Fits All Distributions in Multi-Modal Diffusion at Scale." CVPR 2023.

This was the amazing idea in Marseille ;)

But with something else as an application ;)

Thank you!