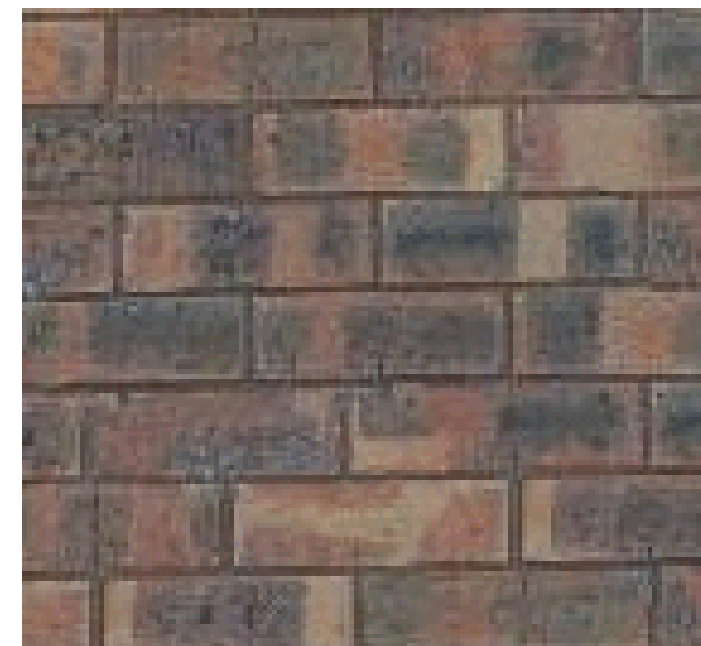
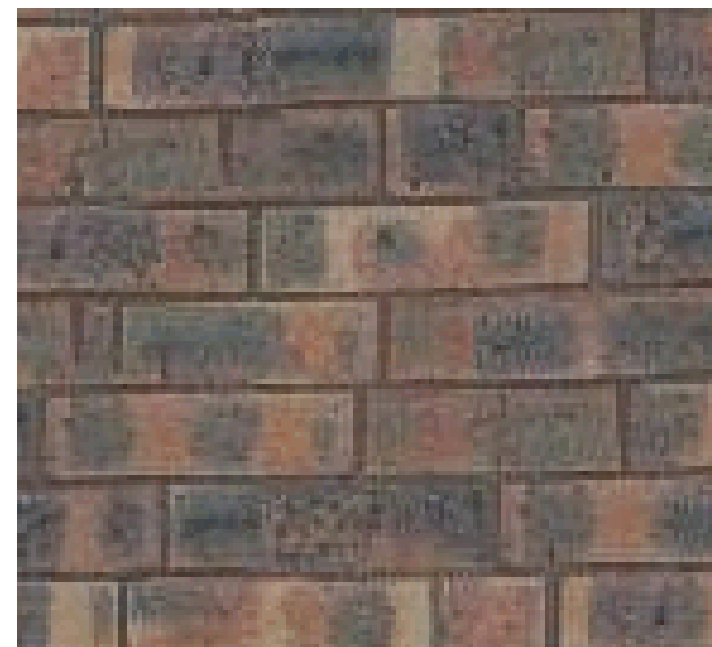


Generating Images from Noise

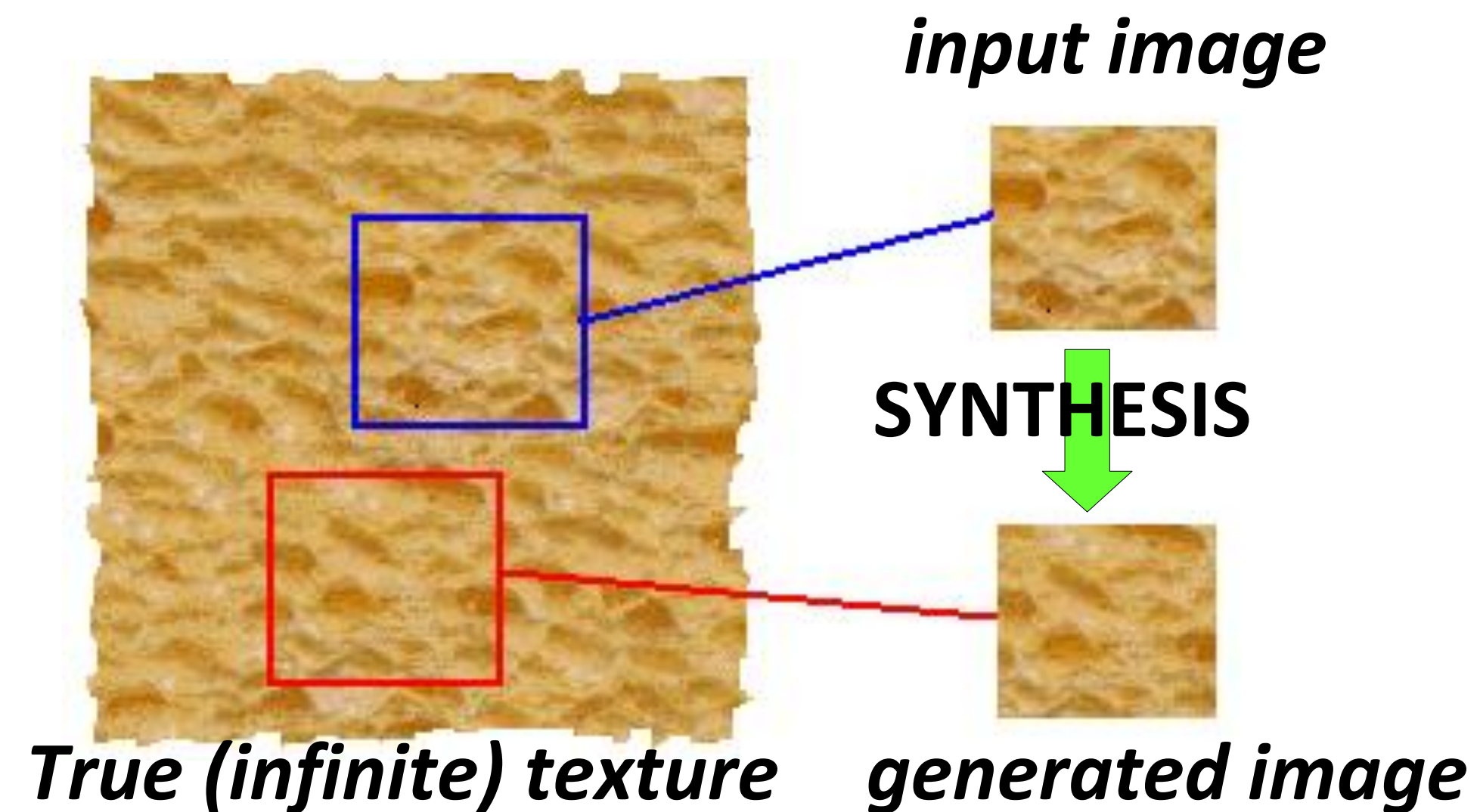


CS180: Intro to Comp. Vision and Comp. Photo
Alexei Efros, UC Berkeley, Fall 2023

Texture as samples from distribution



Parametric Texture Synthesis

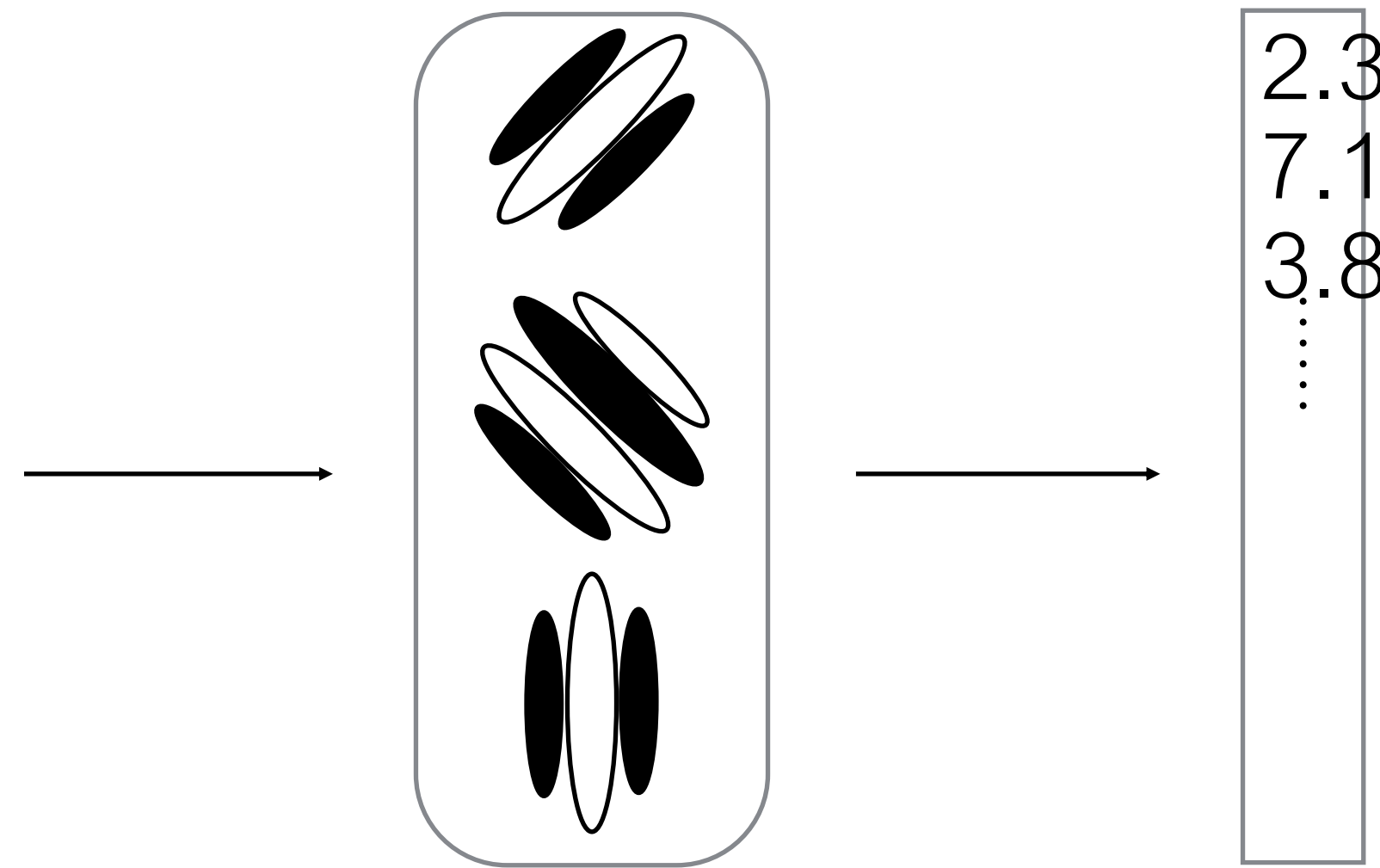
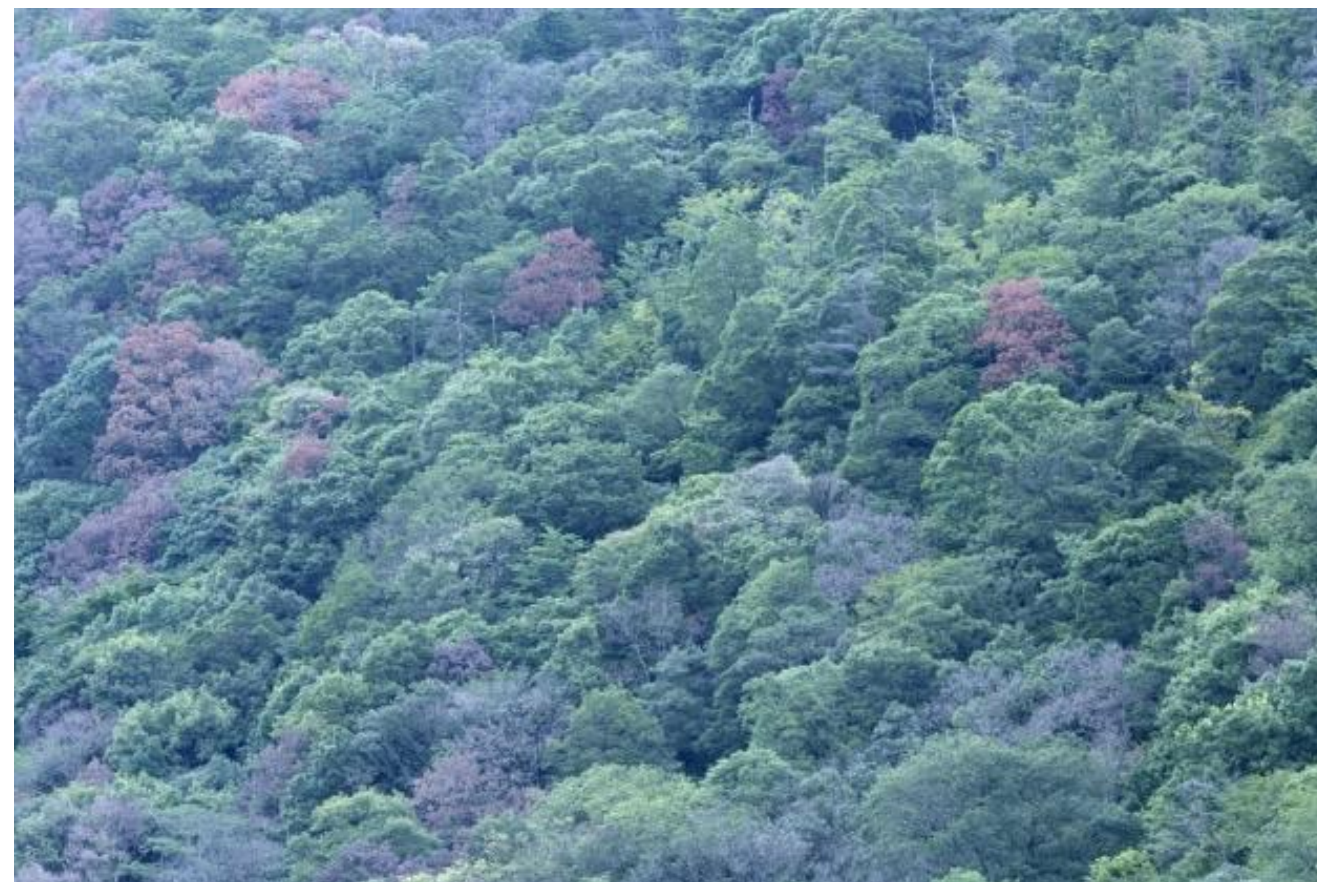


Goal: parametric **generative model** of the
“infinite texture”

Heeger & Bergen, SIGGRAPH'95



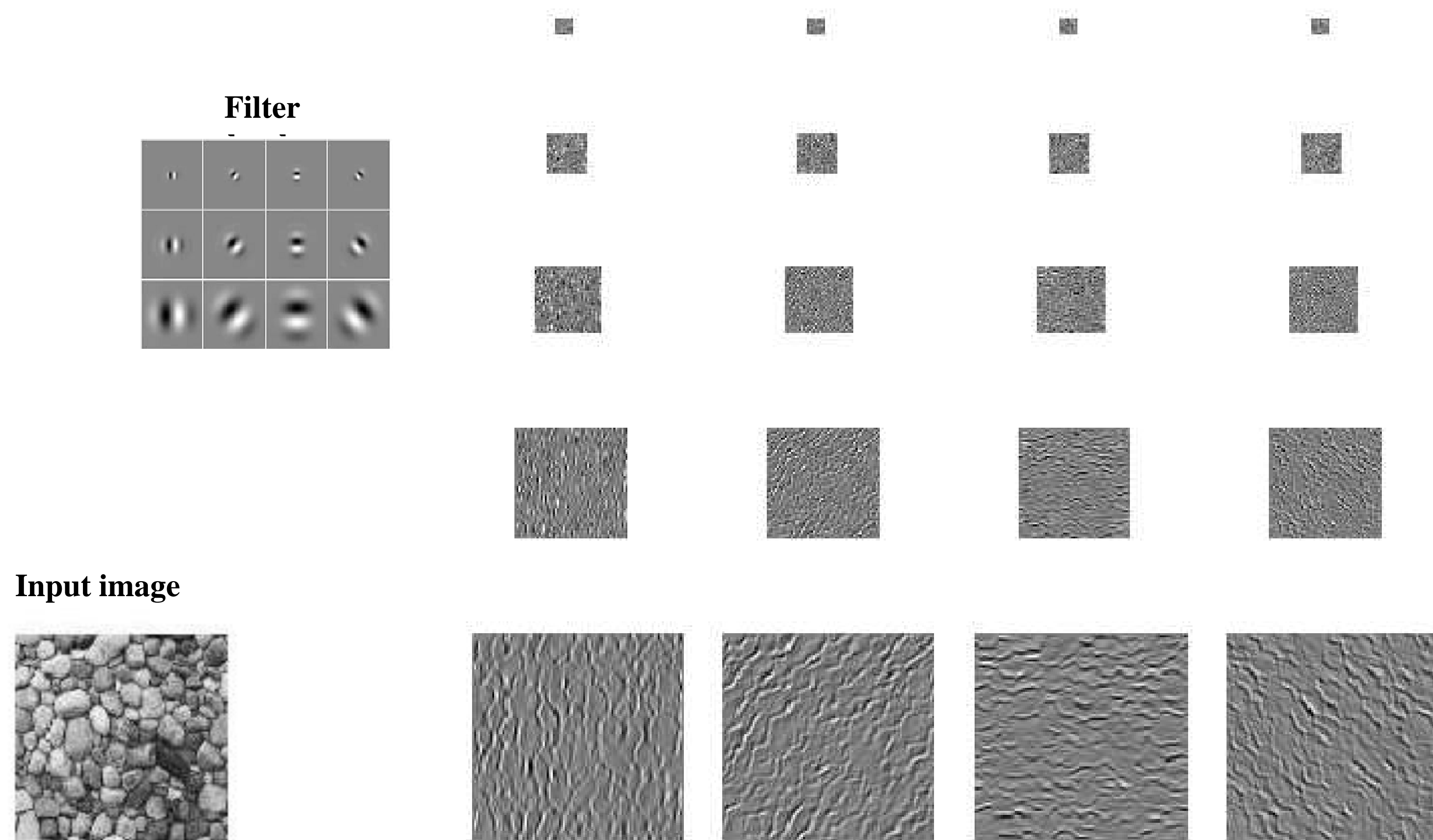
Early Vision Texture Models



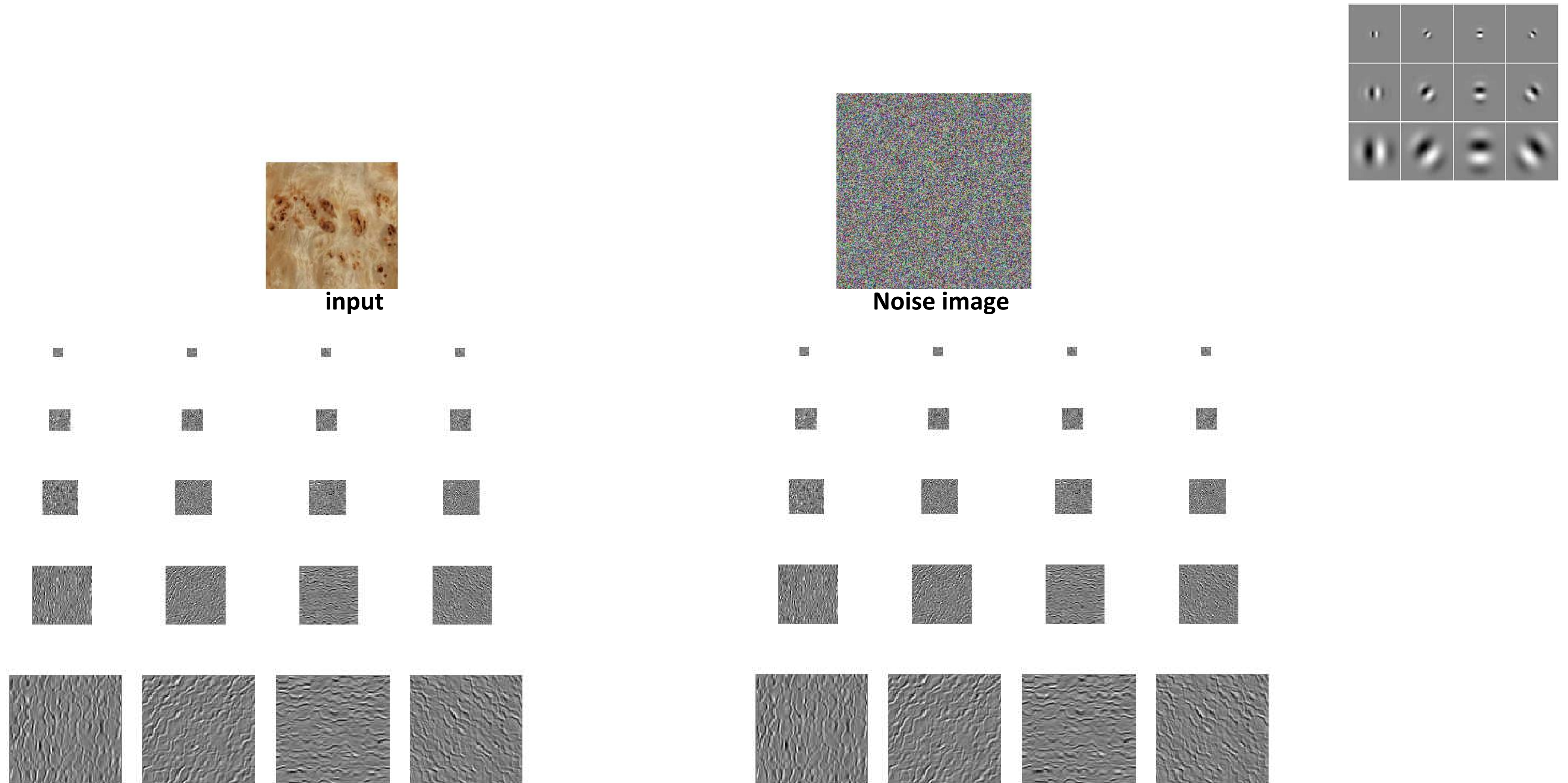
Linear filter bank

Heeger & Bergen (1995)
Portilla & Simoncelli (2000)

Multi-scale filter decomposition (steerable pyramid)



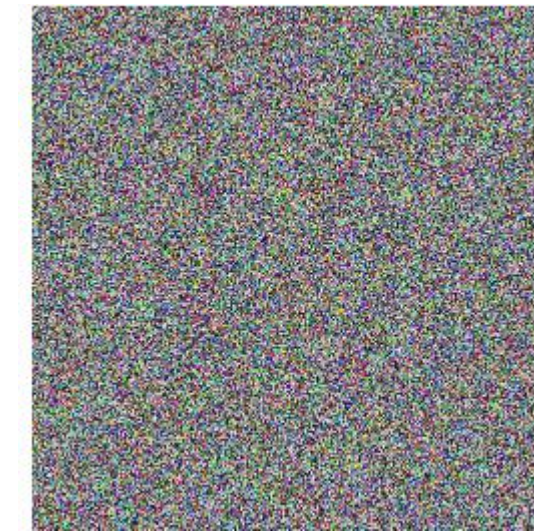
Step 1: Convolve with filterbank



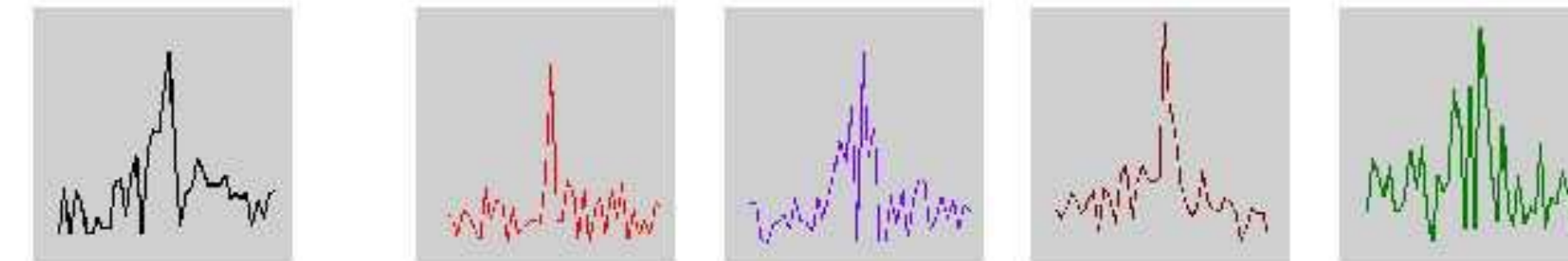
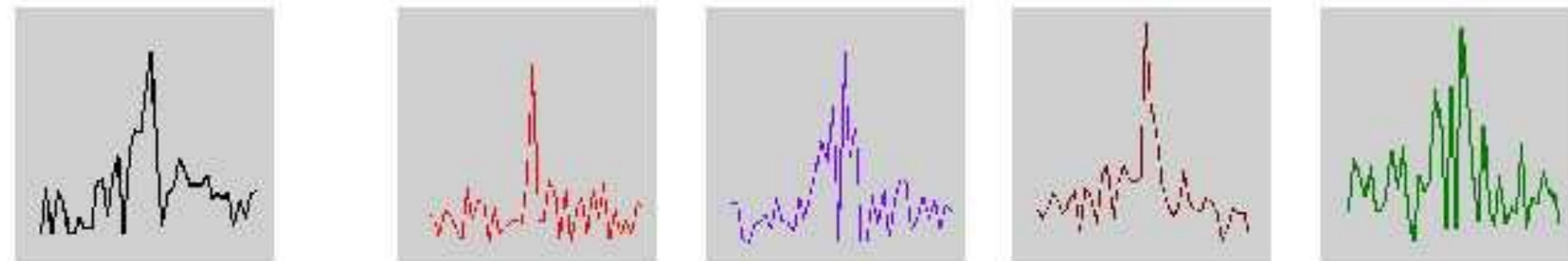
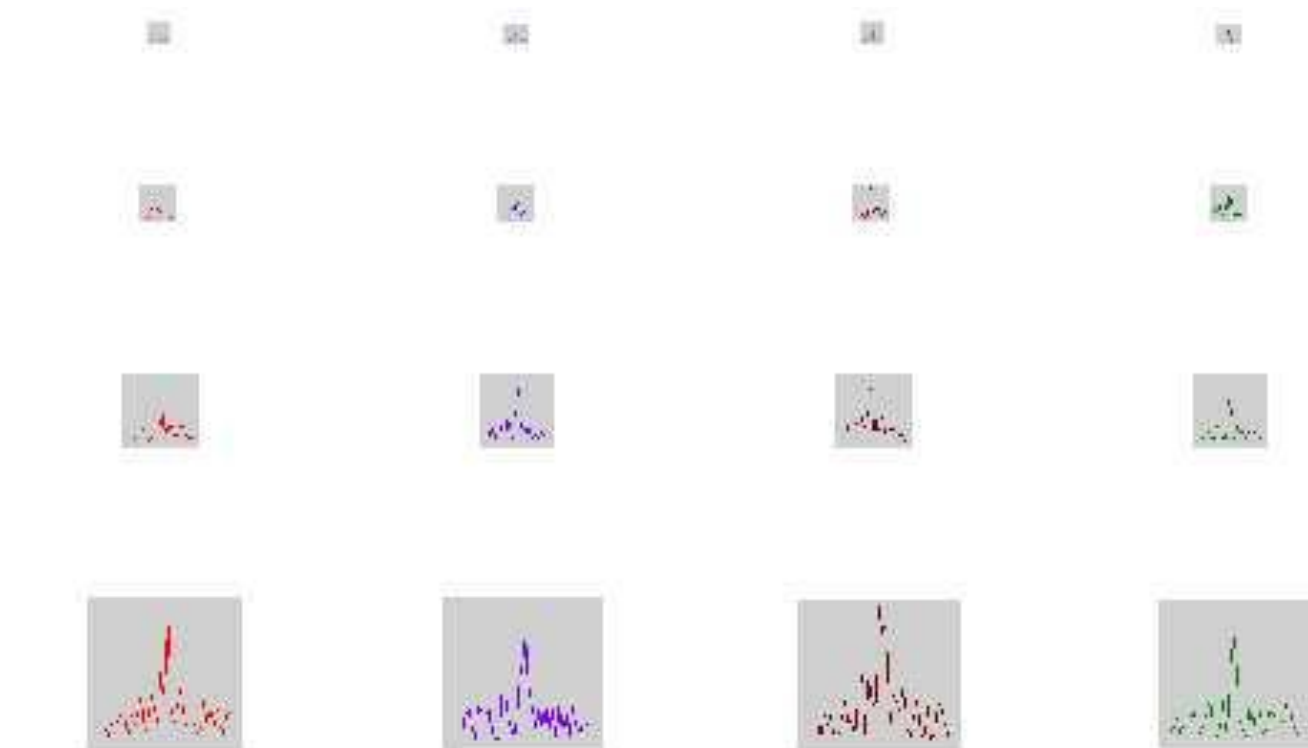
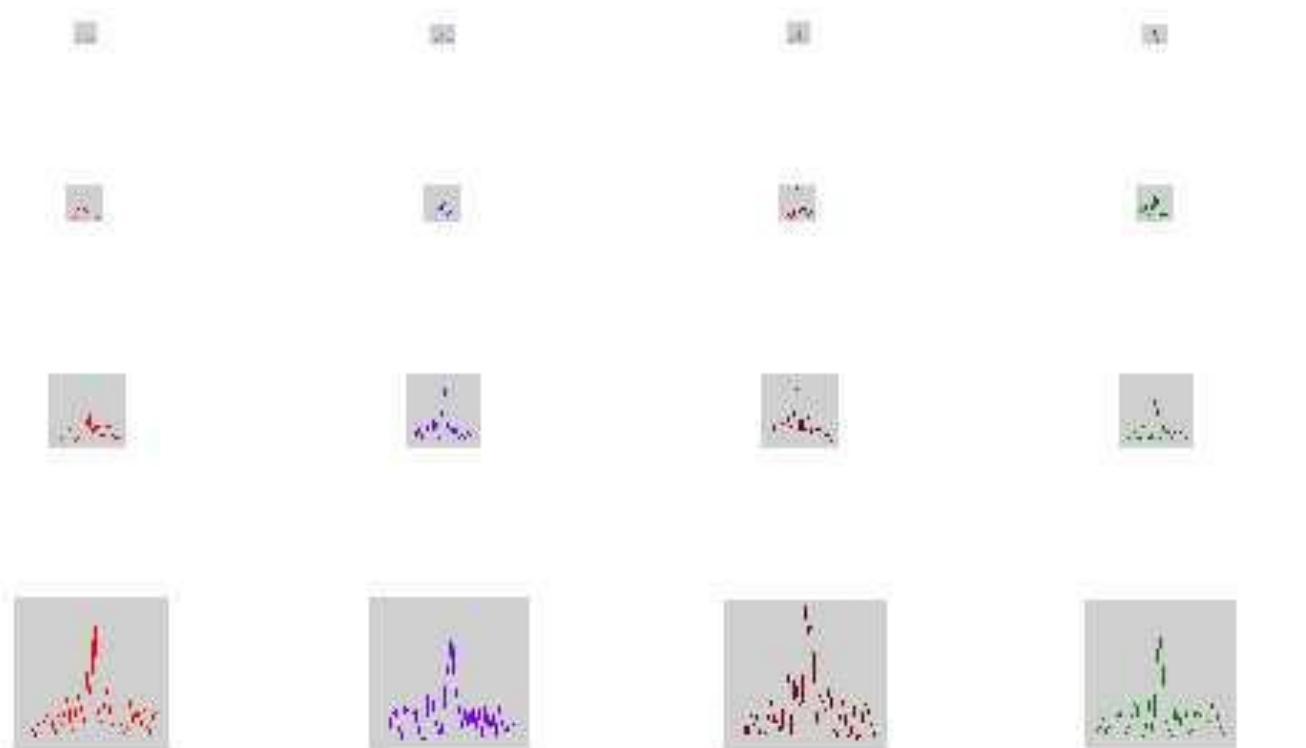
Step 2: match per-channel histograms



input



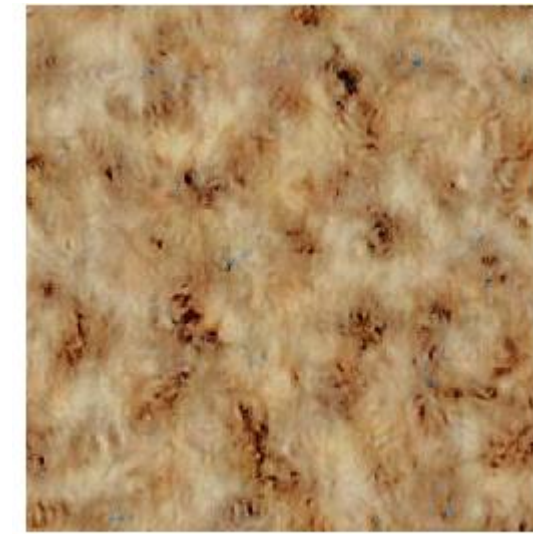
Noise image



Step 3: collapse pyramid and repeat!

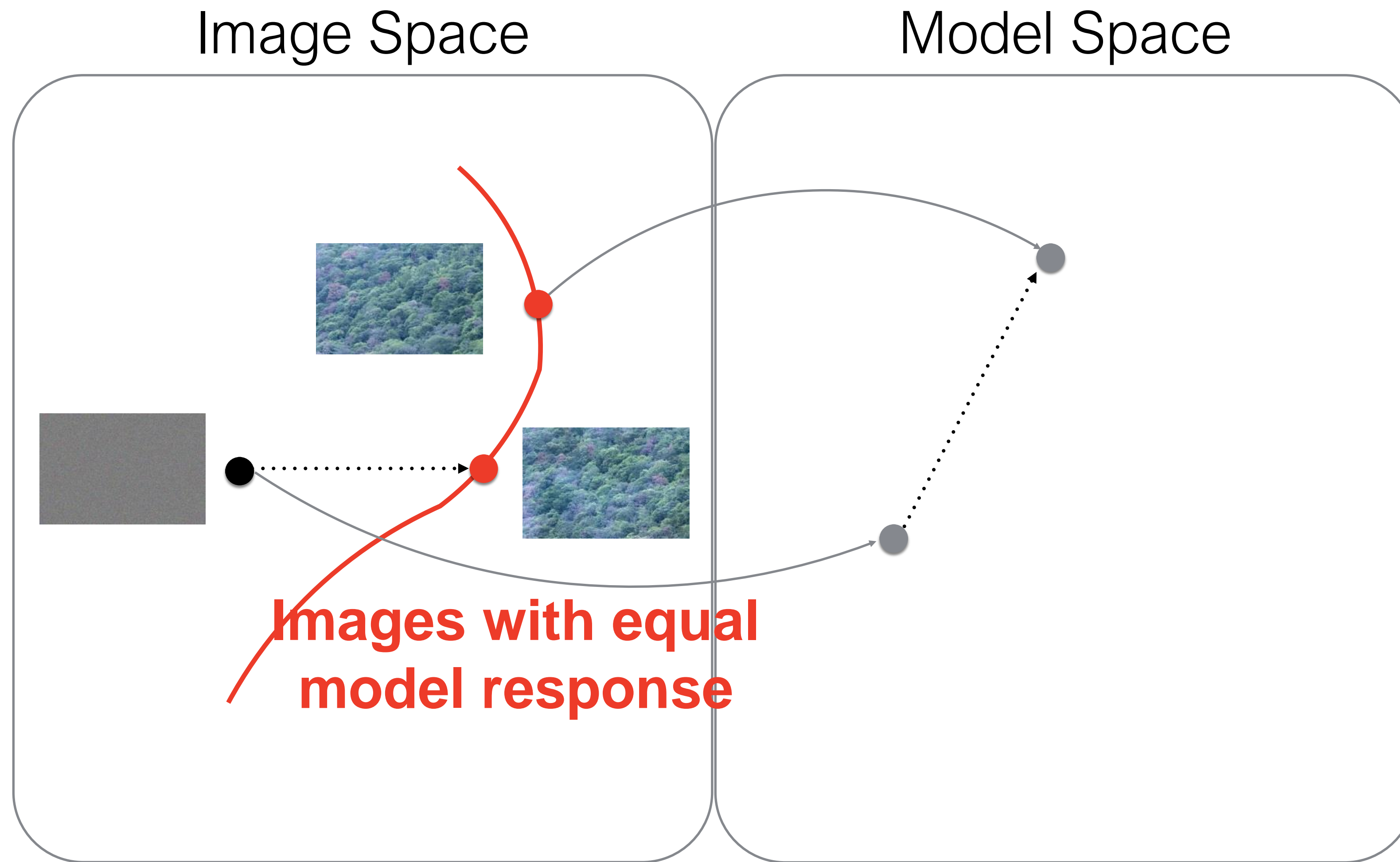


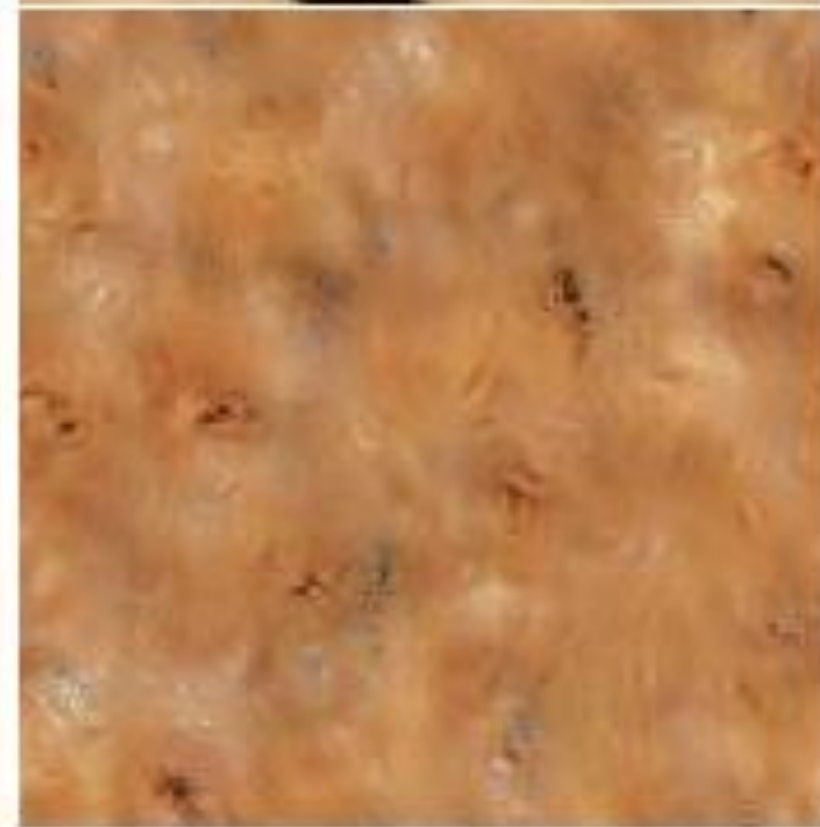
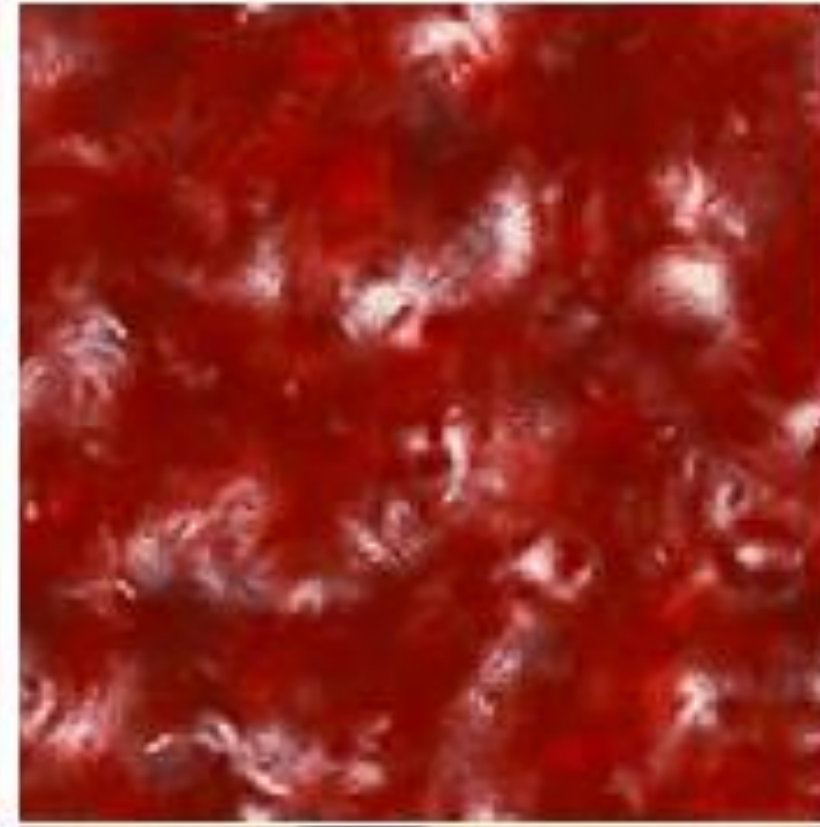
input

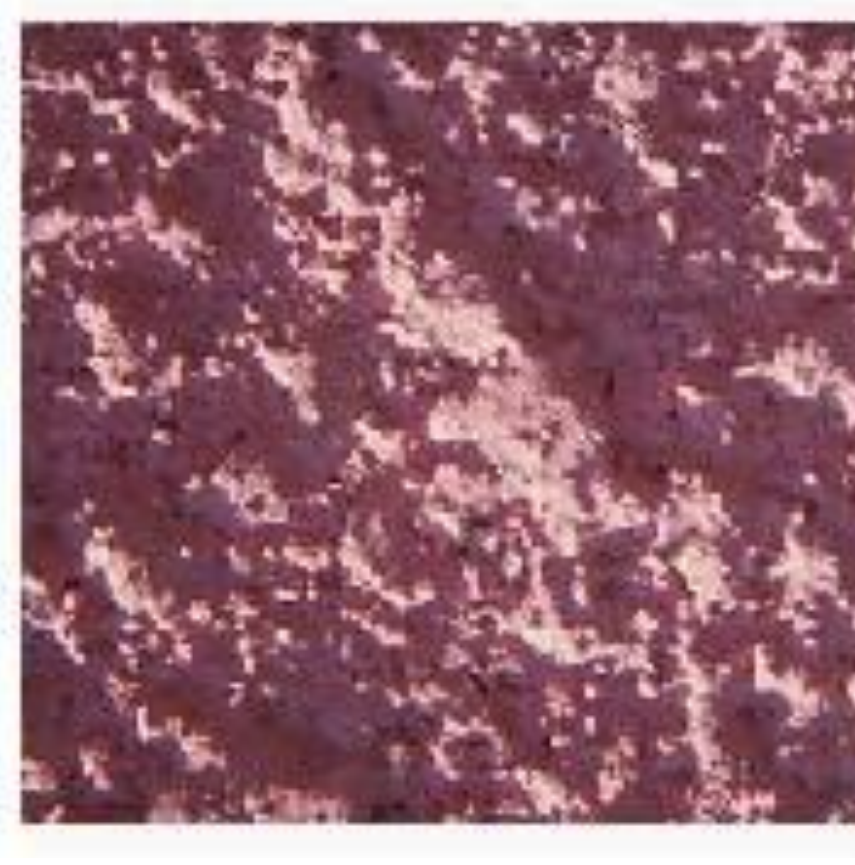
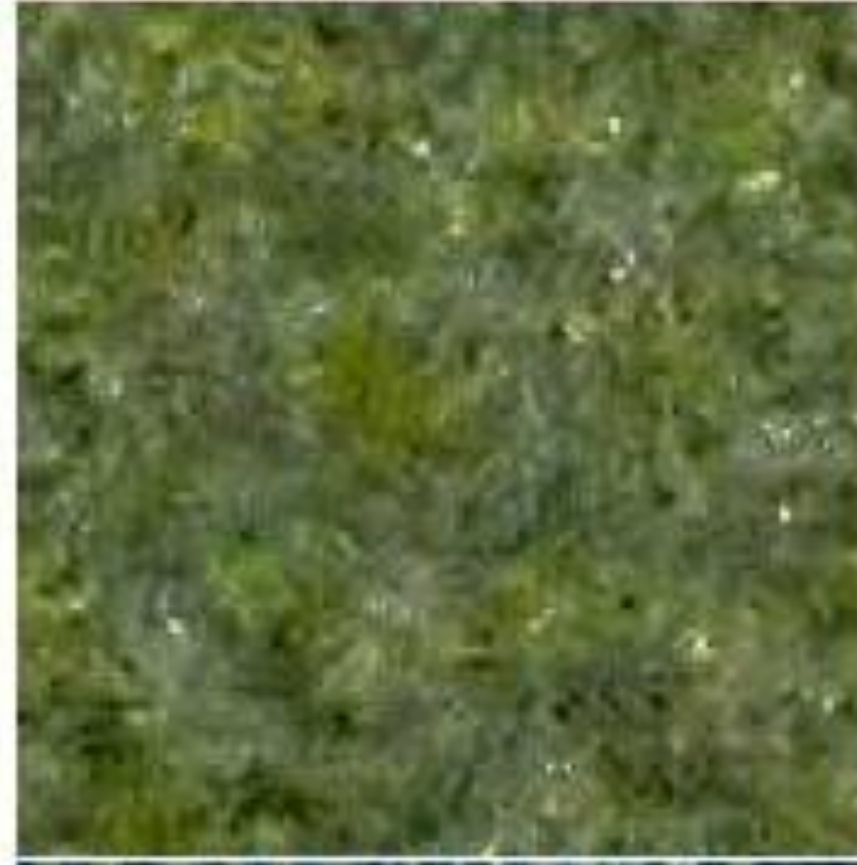
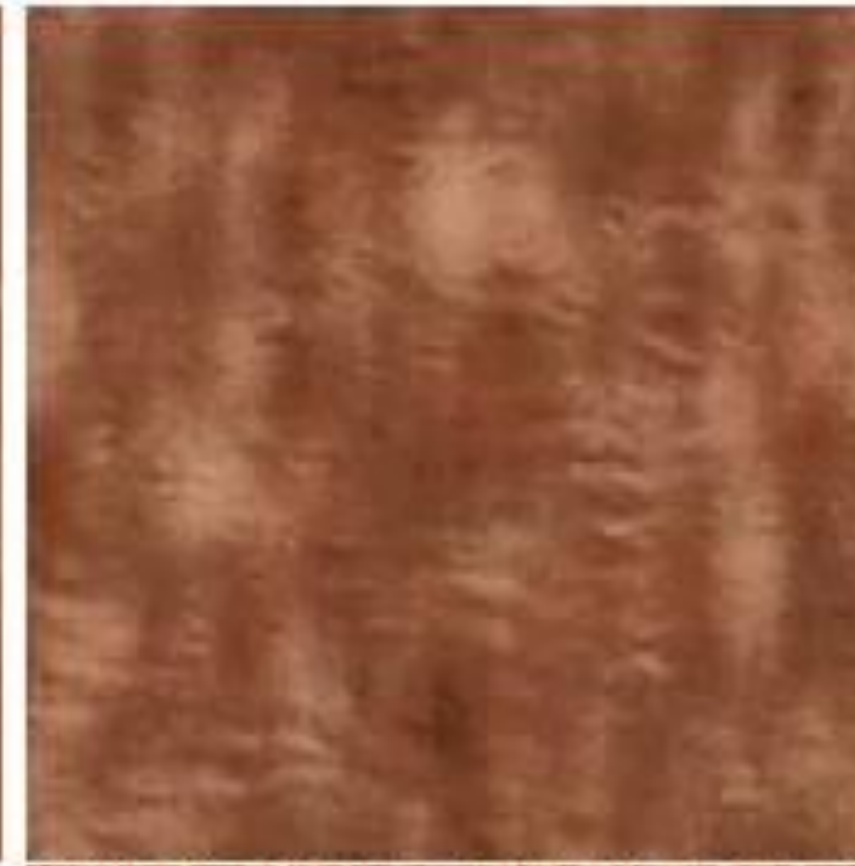
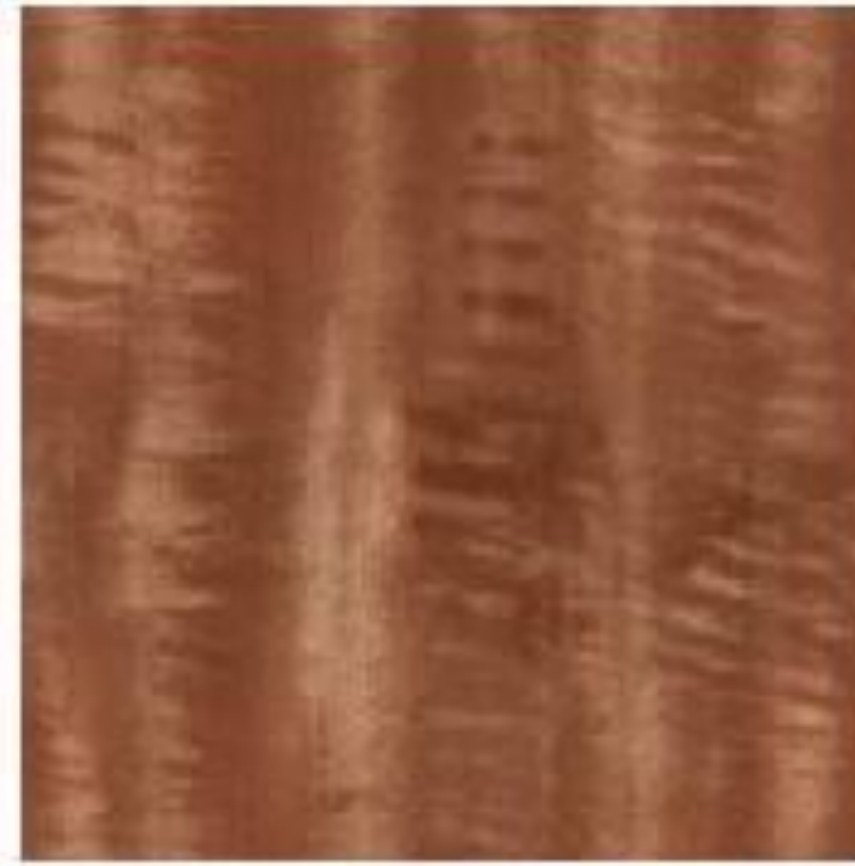


Noise image

Texture Synthesis







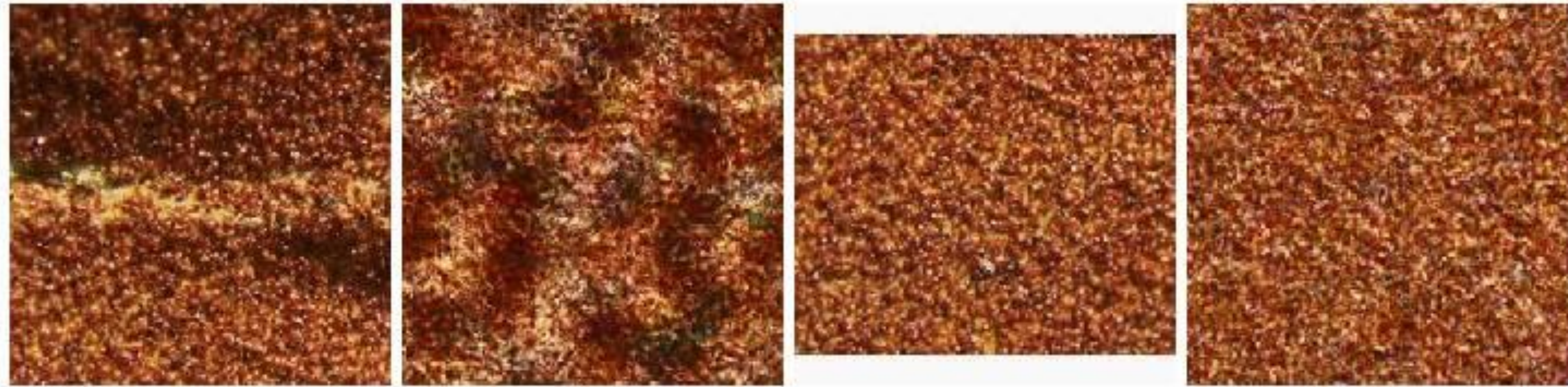


Figure 7: (Left pair) Inhomogeneous input texture produces blotchy synthetic texture. (Right pair) Homogeneous input.

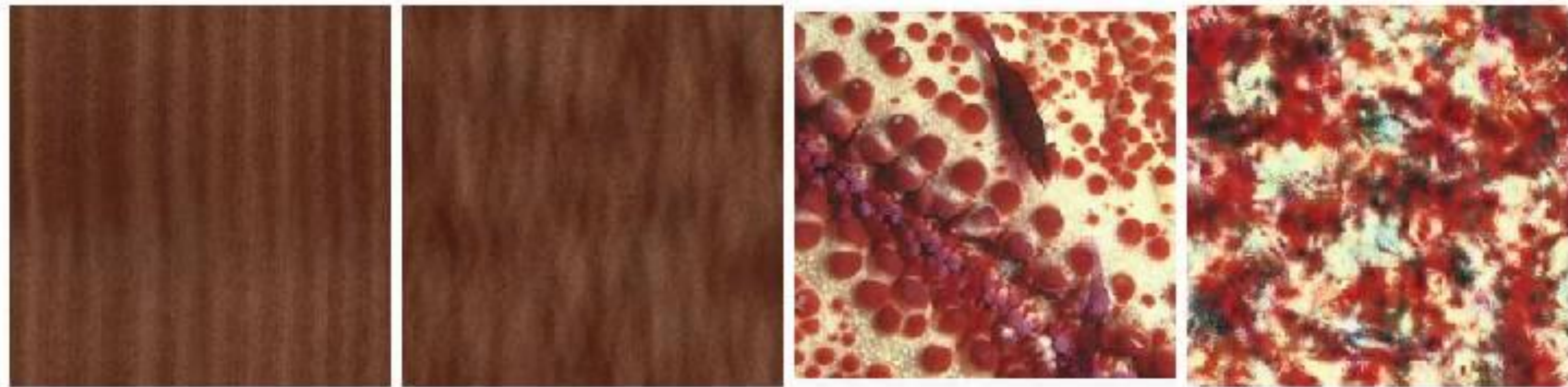


Figure 8: Examples of failures: wood grain and red coral.

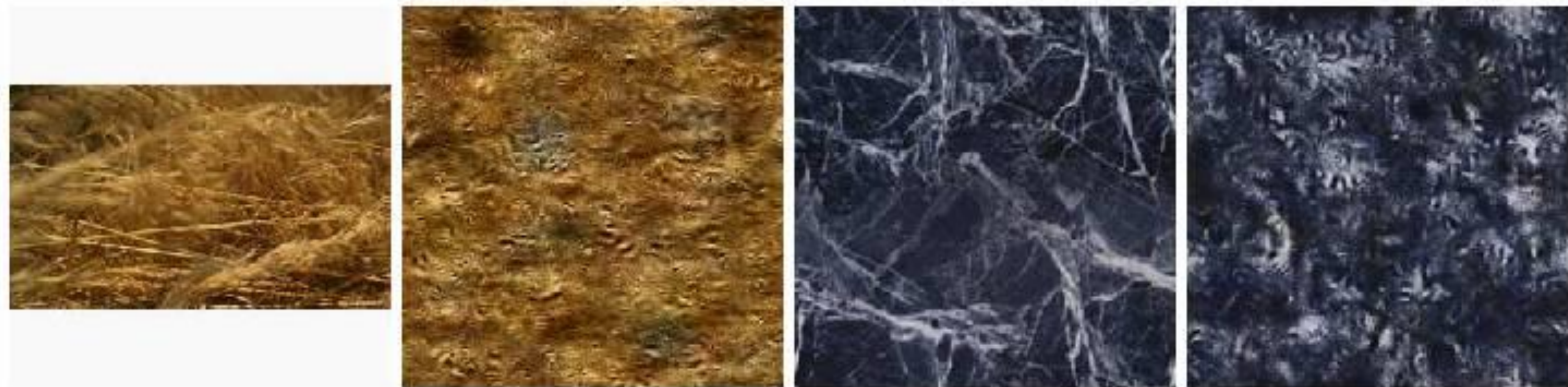


Figure 9: More failures: hay and marble.

Simoncelli & Portilla '98+

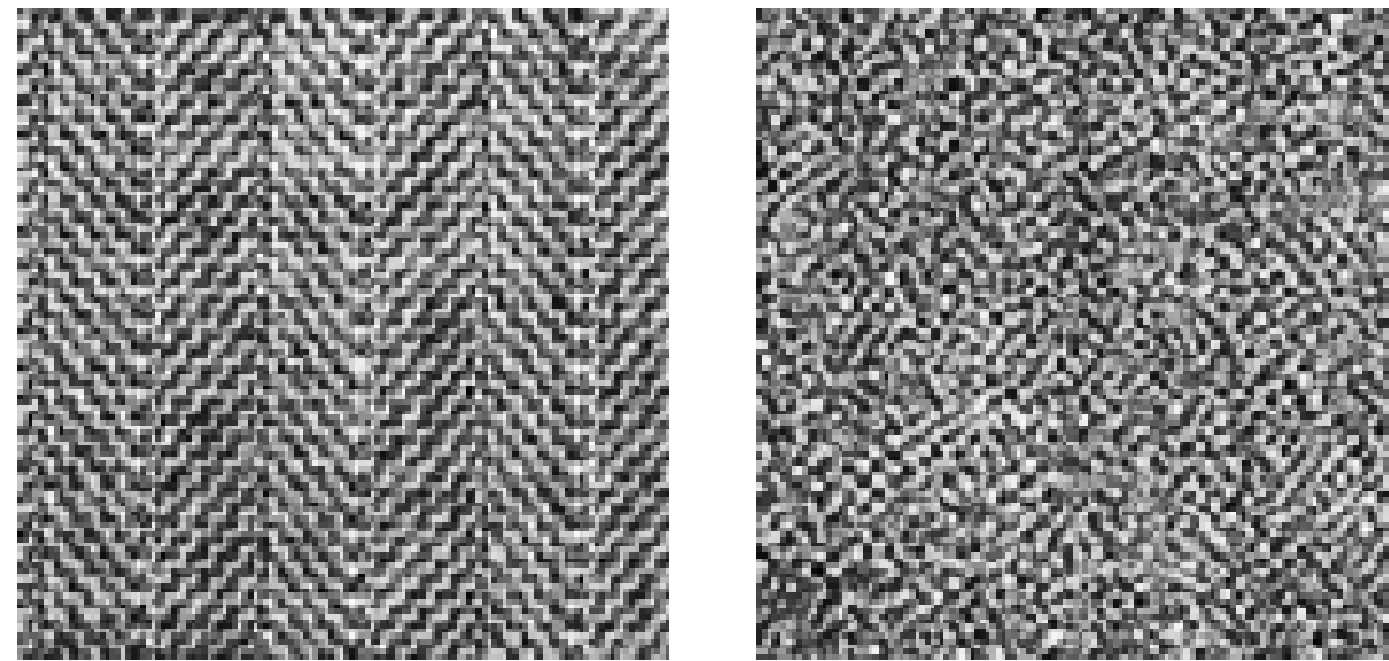
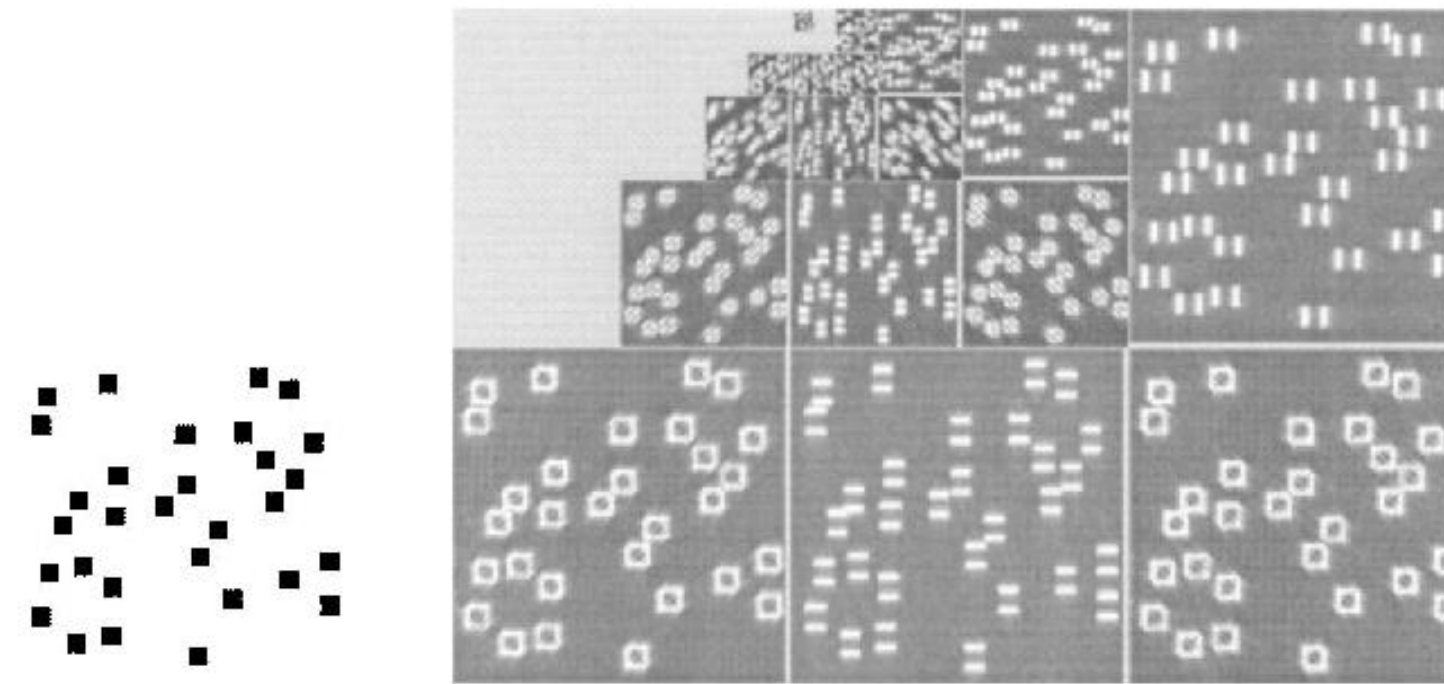


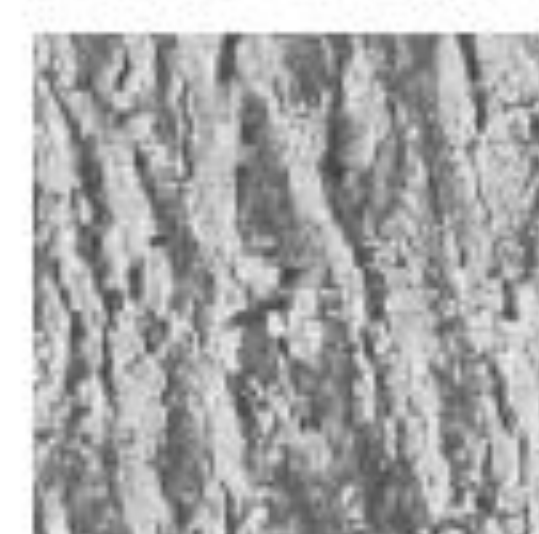
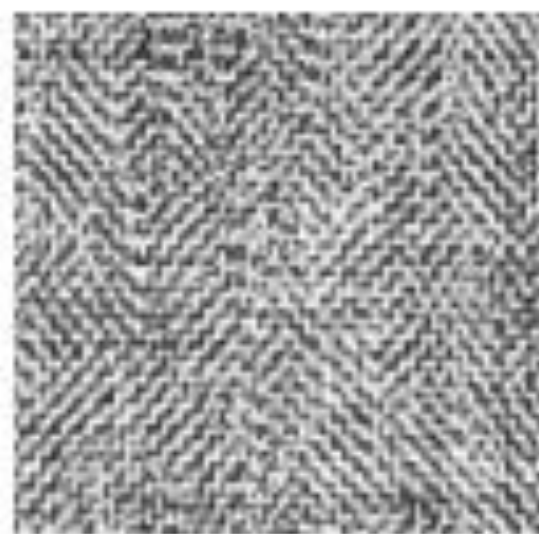
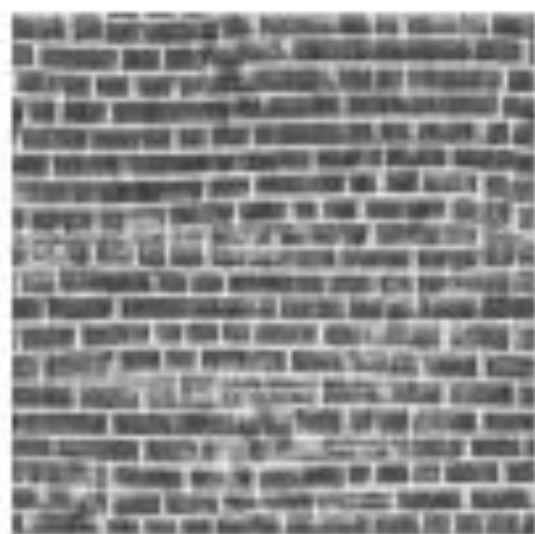
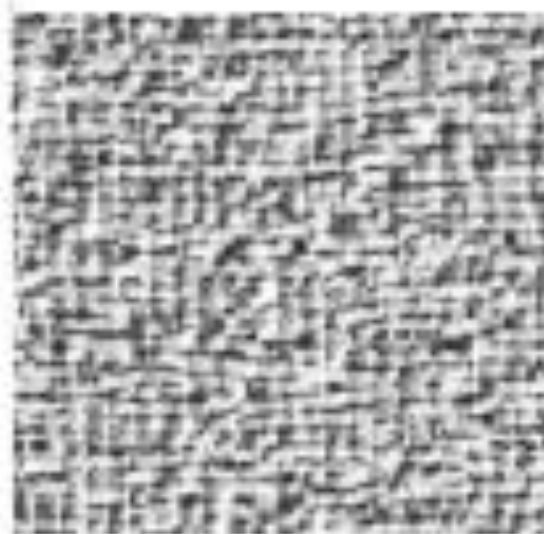
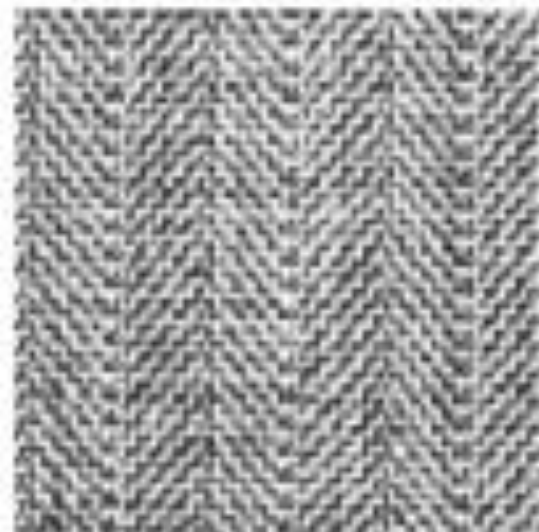
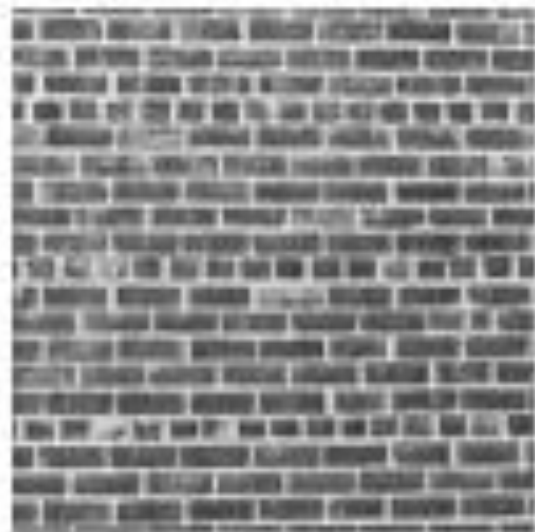
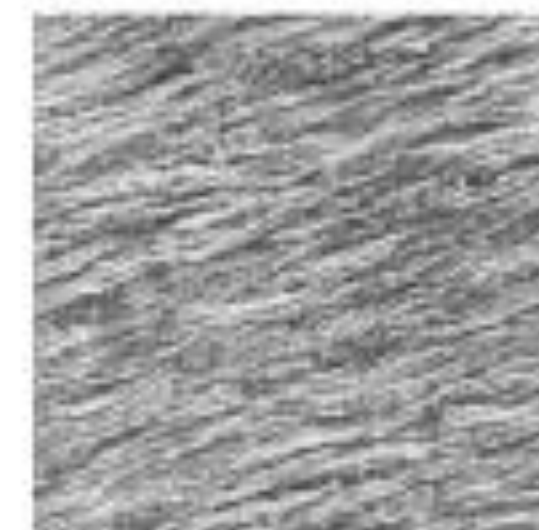
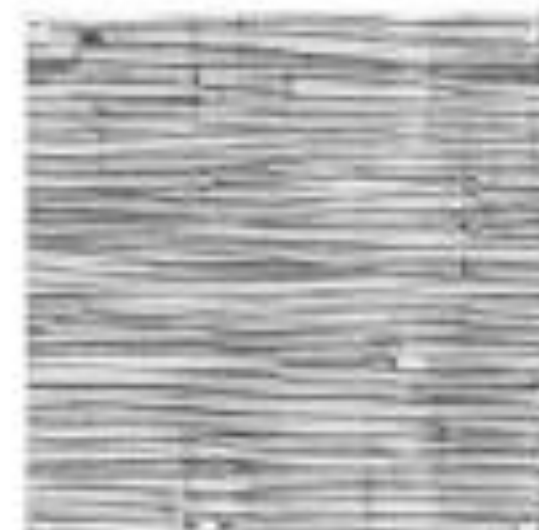
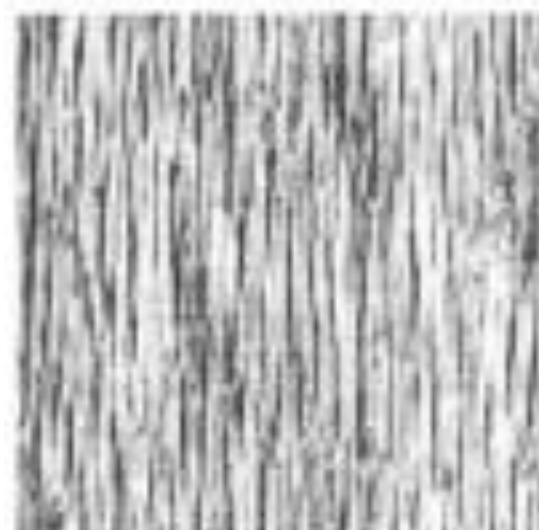
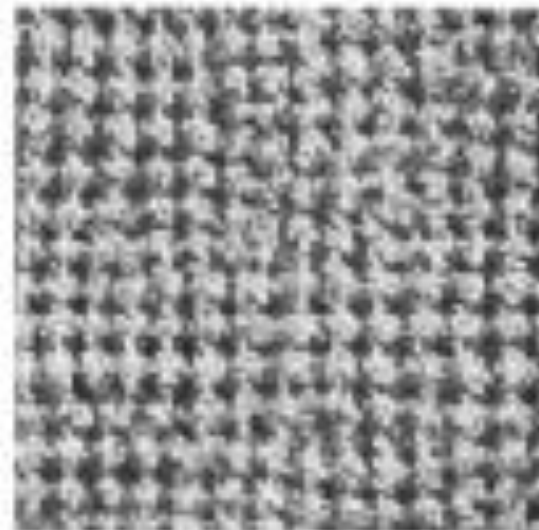
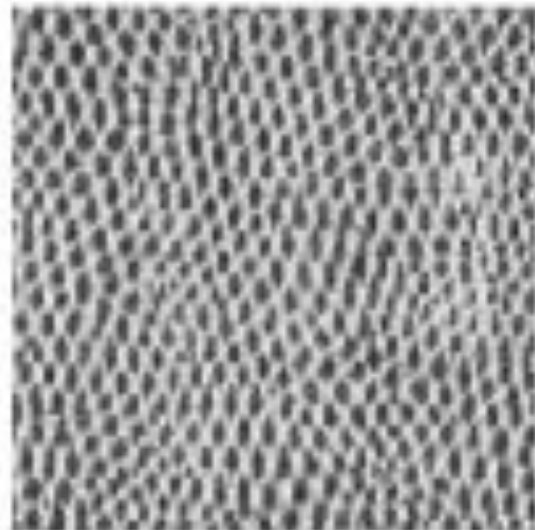
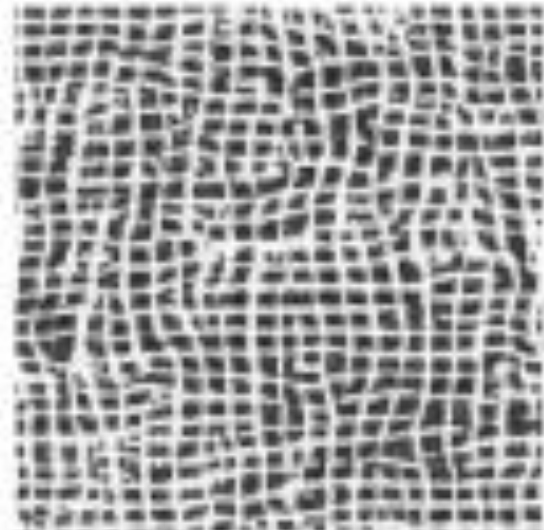
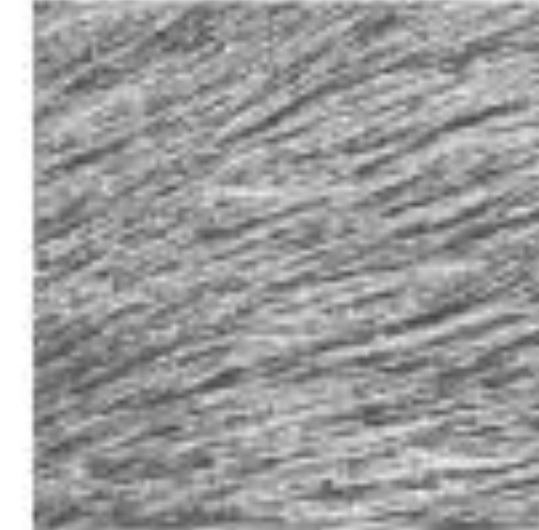
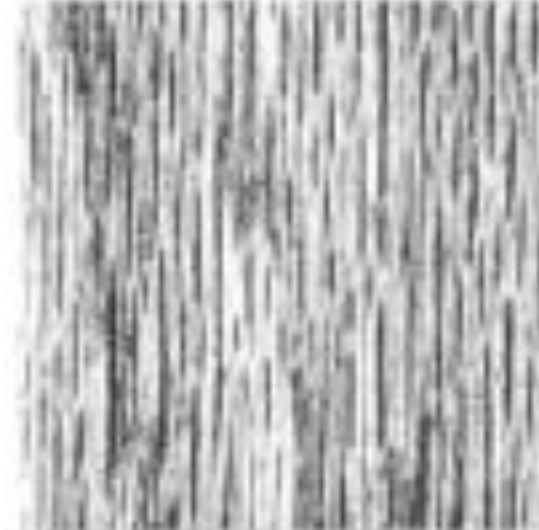
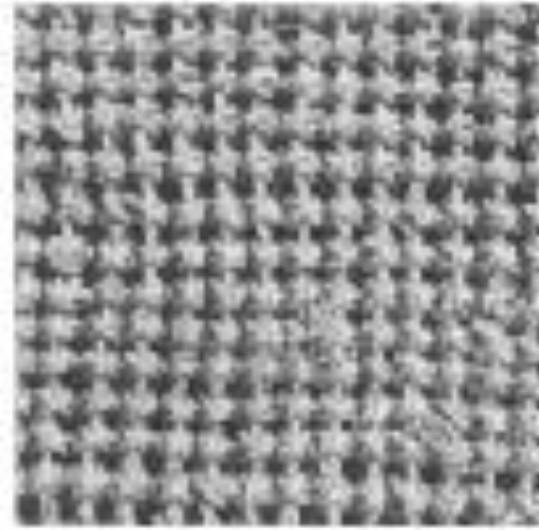
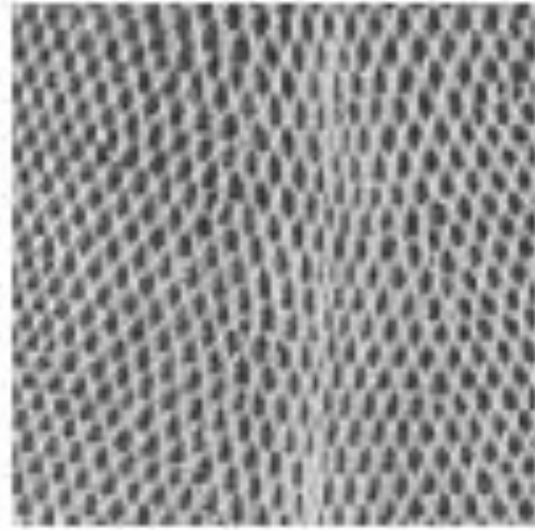
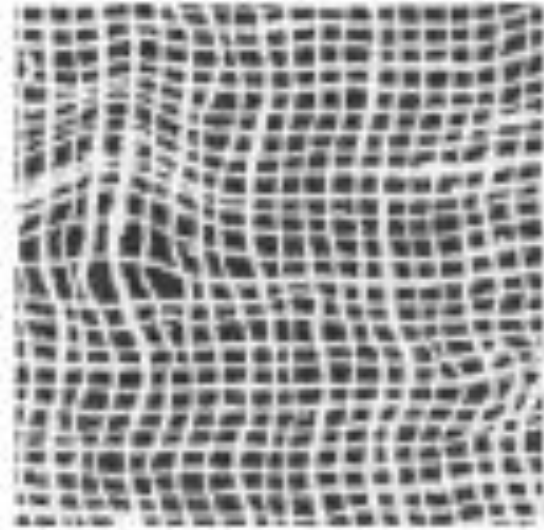
Figure 1. Textures with matching marginal statistics.

- Marginal statistics are not enough
- Neighboring filter responses are highly correlated
 - an edge at low-res will cause an edge at high-res
- Let's match 2nd order statistics too!
- J Portilla and E P Simoncelli. *A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients*. Int'l Journal of Computer Vision. 40(1):49-71, October, 2000.

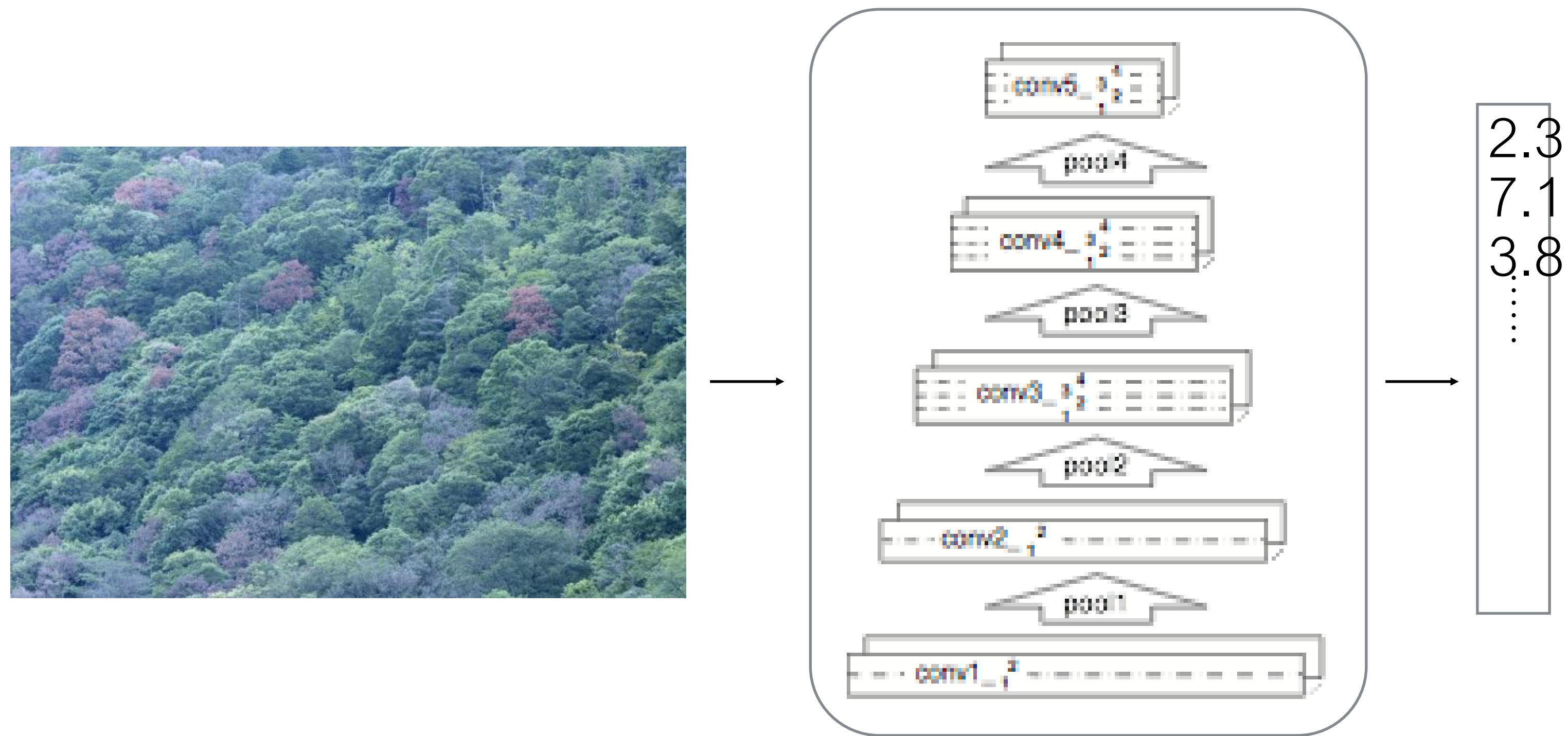
Simoncelli & Portilla '98+



- Match joint histograms of pairs of filter responses at adjacent spatial locations, orientations, and scales.
- Optimize using repeated projections onto statistical constraint surfaces



Convolutional Neural Network Texture Model

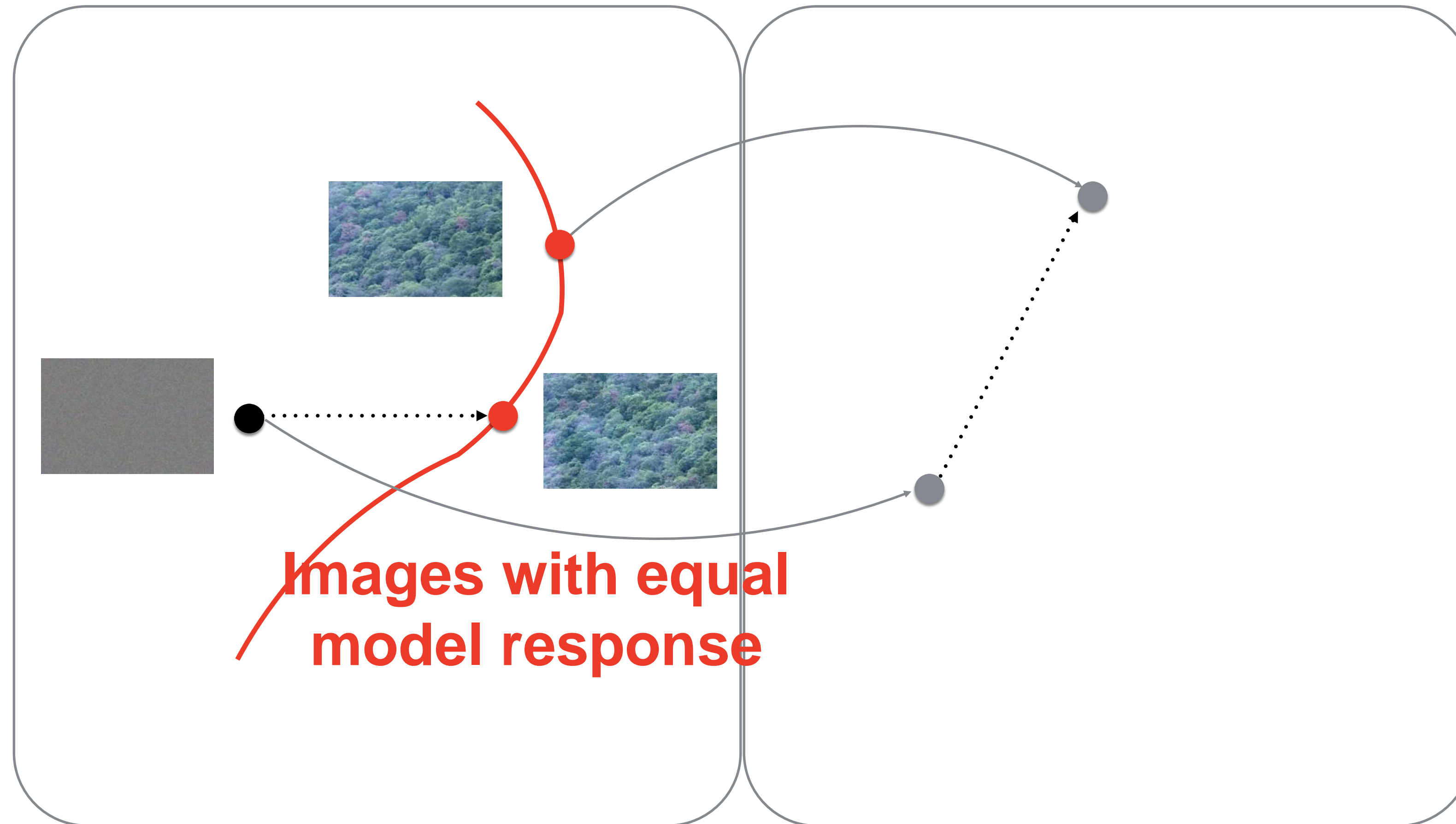


Convolutional Neural Network

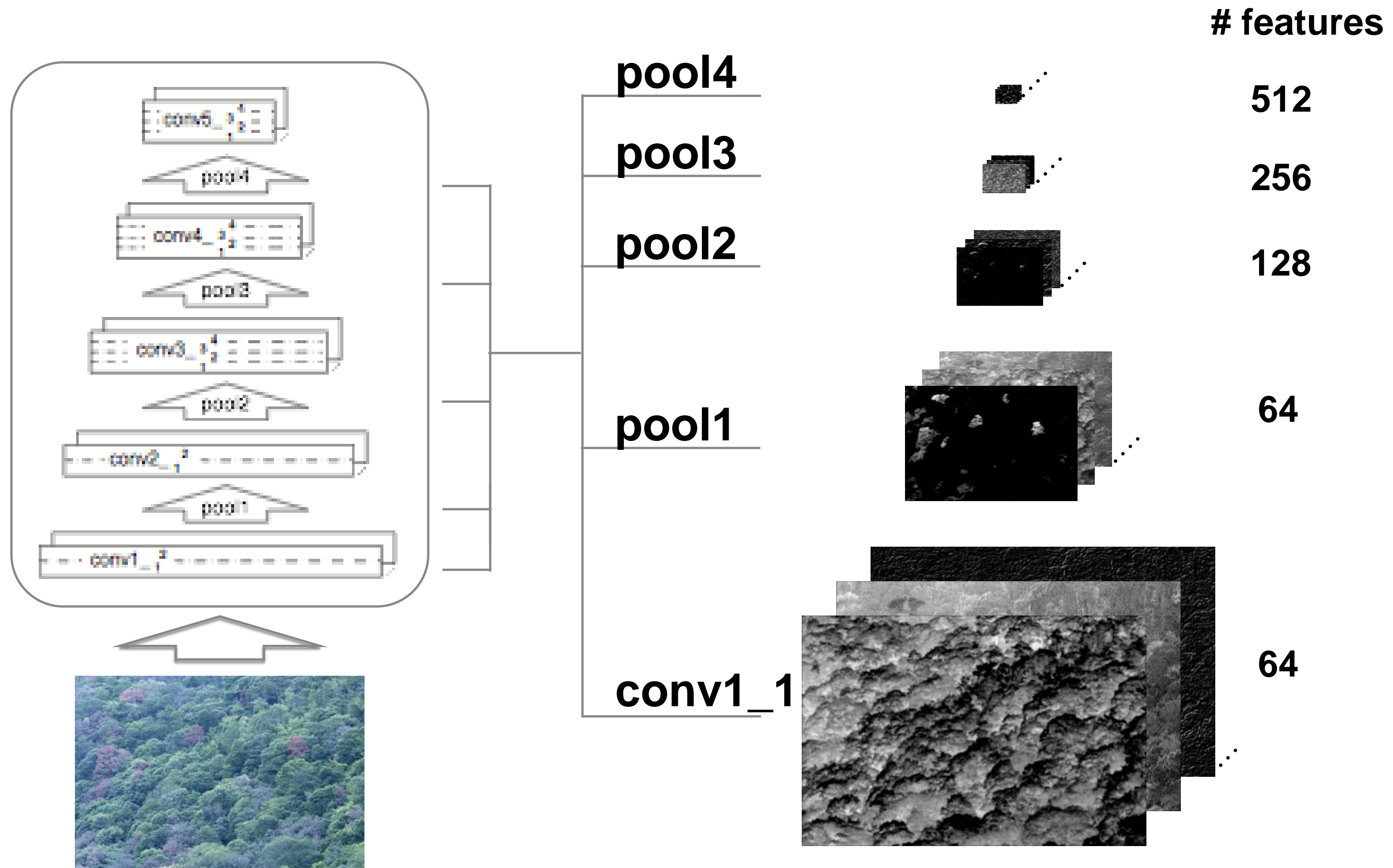
Texture Synthesis

Image Space

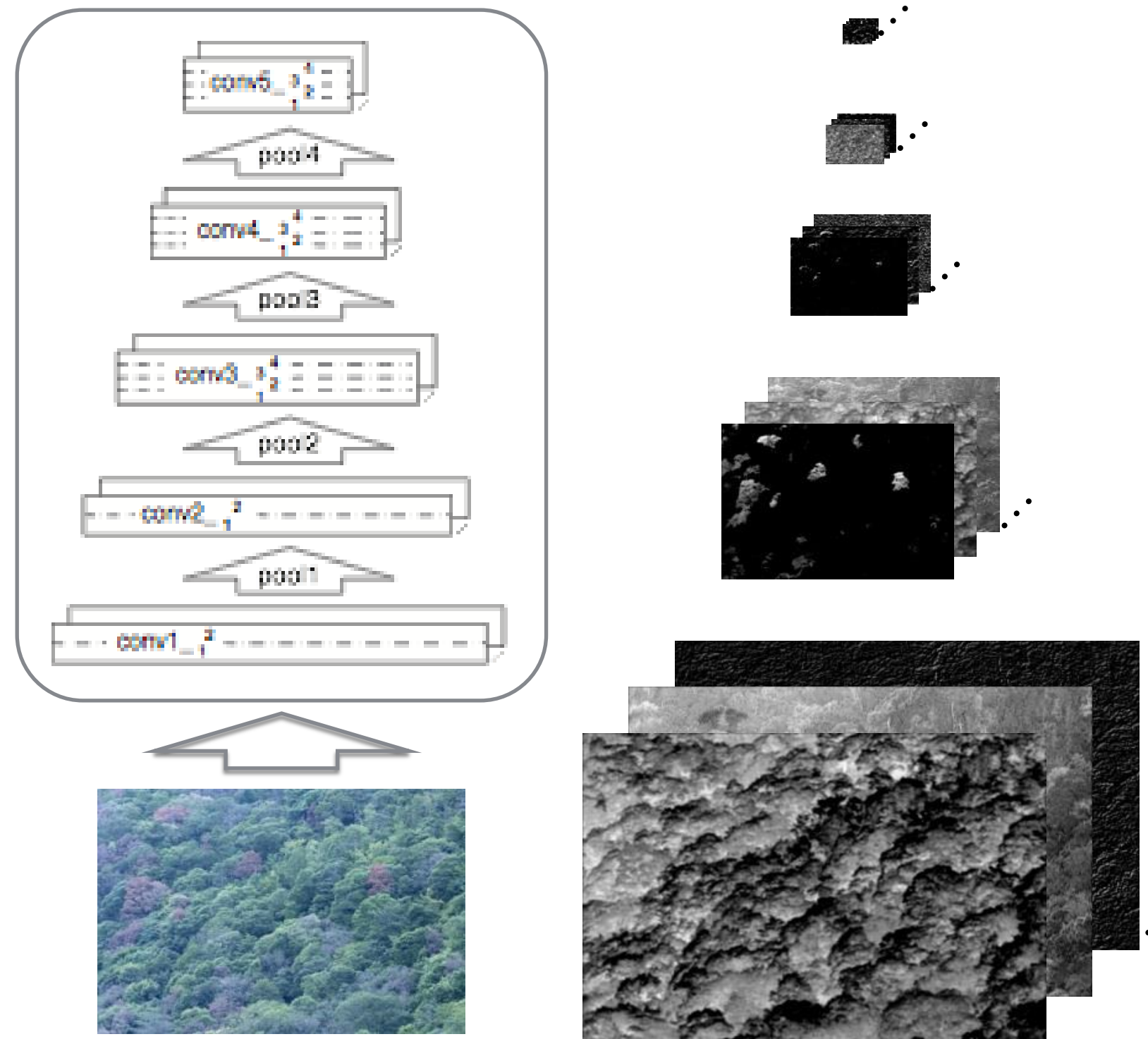
Model Space



CNN - Multiscale Filter Bank



CNN - Texture Features



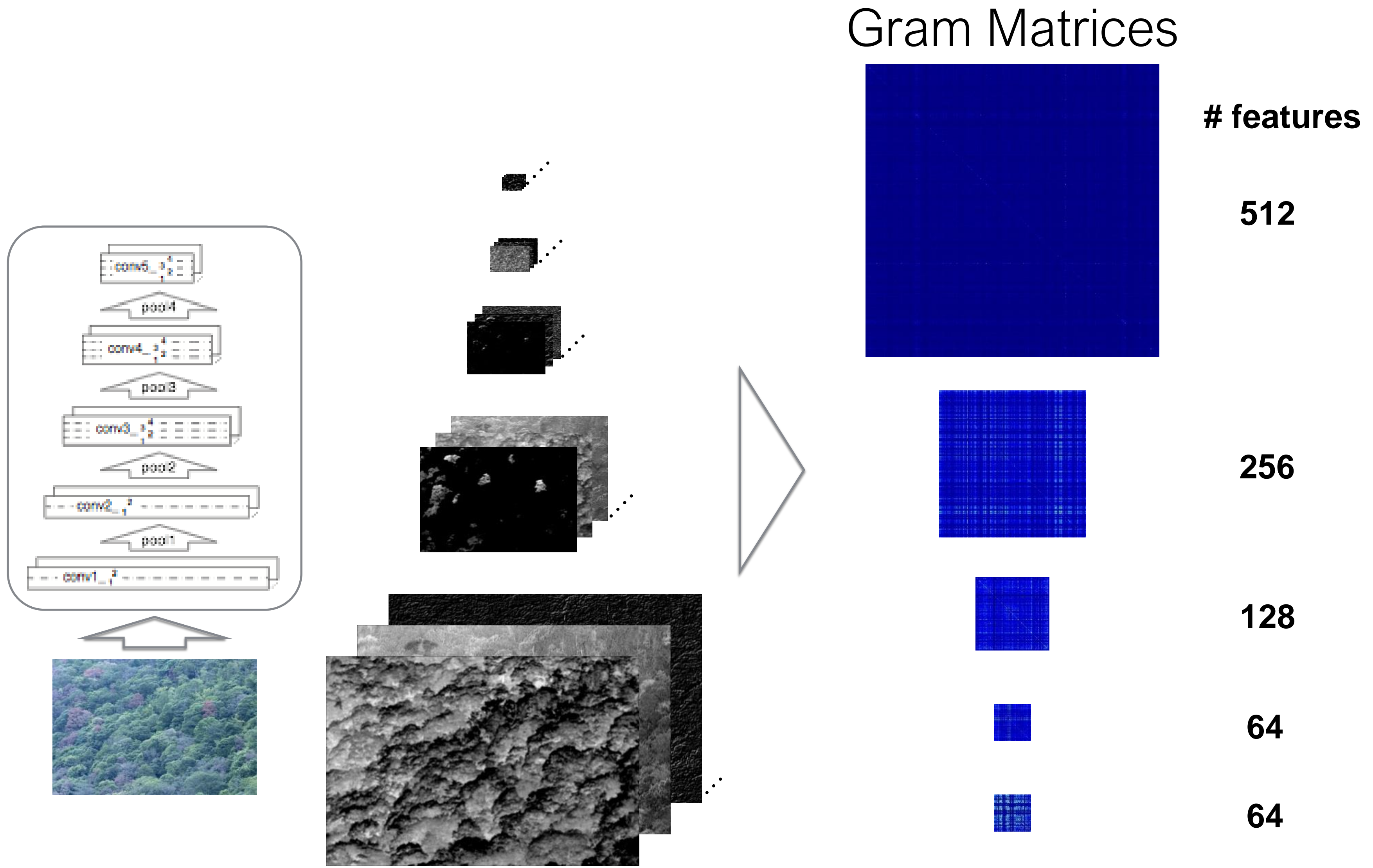
$$F = [\bar{f}_1, \bar{f}_2, \bar{f}_3, \dots, \bar{f}_N]^T$$

$$G = FF^T$$

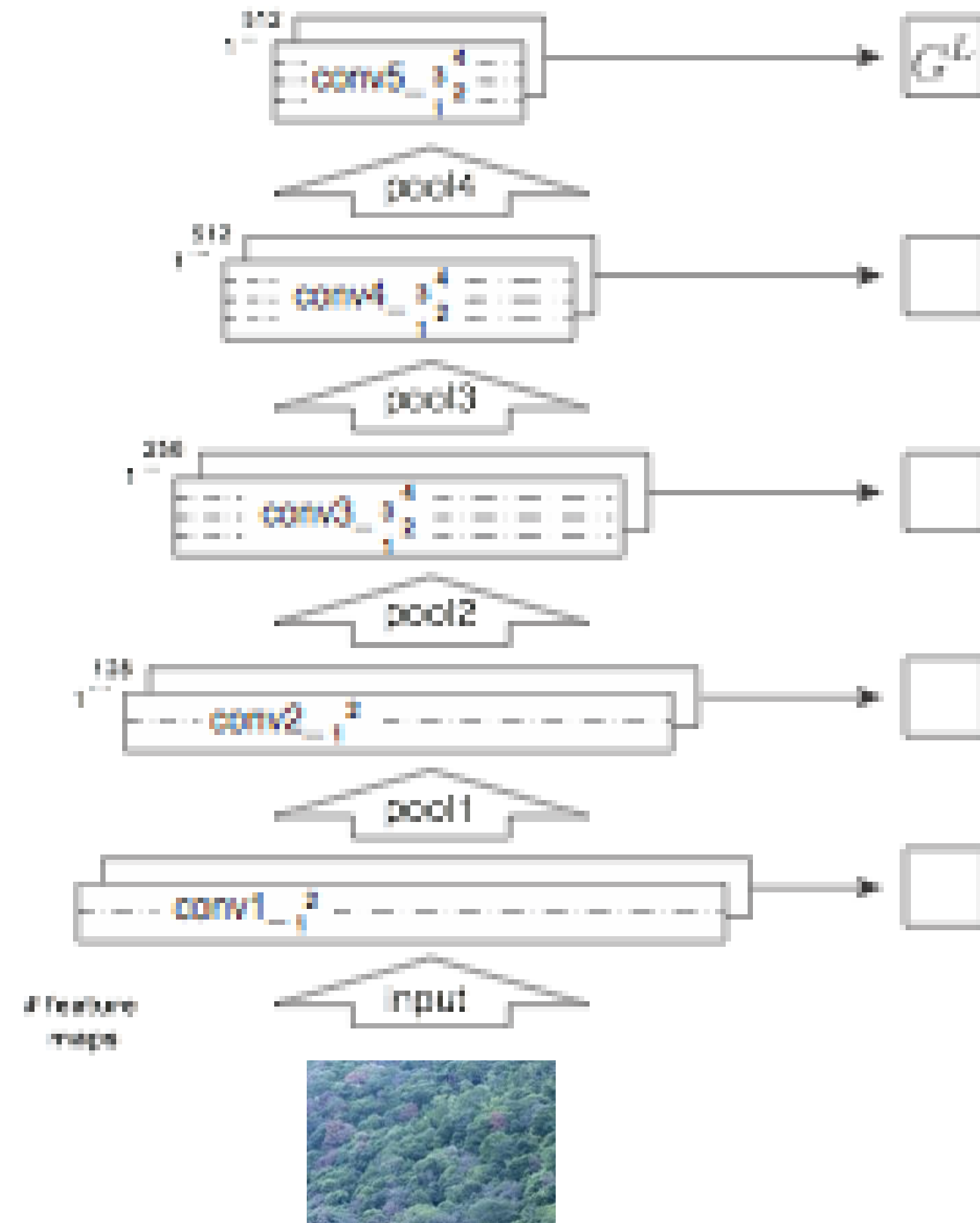
$$= \begin{pmatrix} \langle \bar{f}_1, \bar{f}_1 \rangle & \cdots & \langle \bar{f}_1, \bar{f}_N \rangle \\ \langle \bar{f}_2, \bar{f}_1 \rangle & & \vdots \\ \vdots & \ddots & \vdots \\ \langle \bar{f}_N, \bar{f}_1 \rangle & \cdots & \langle \bar{f}_N, \bar{f}_N \rangle \end{pmatrix}$$

$$\langle \bar{f}_i, \bar{f}_j \rangle = \sum_k F_{ik} F_{jk}$$

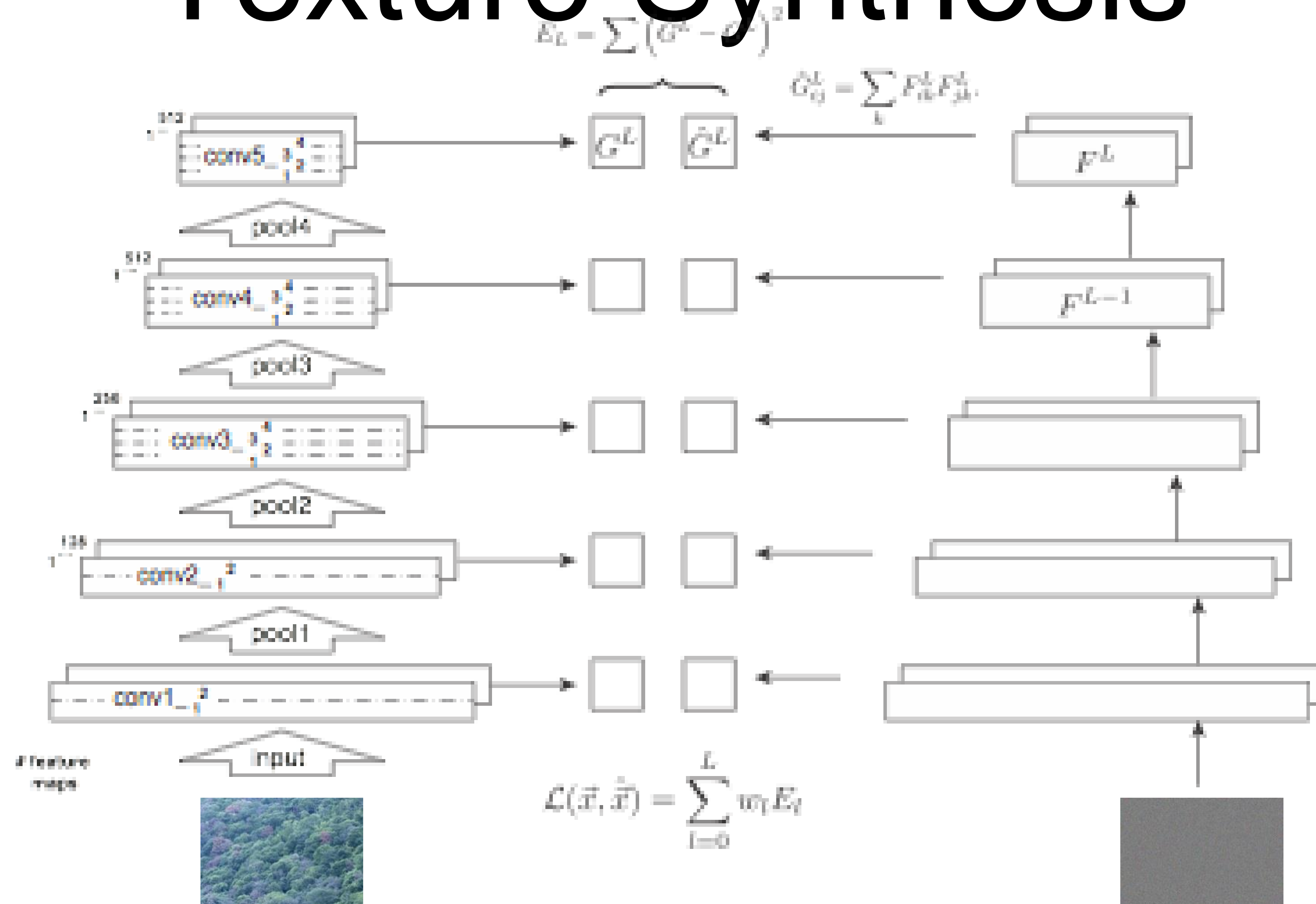
CNN-Texture Features



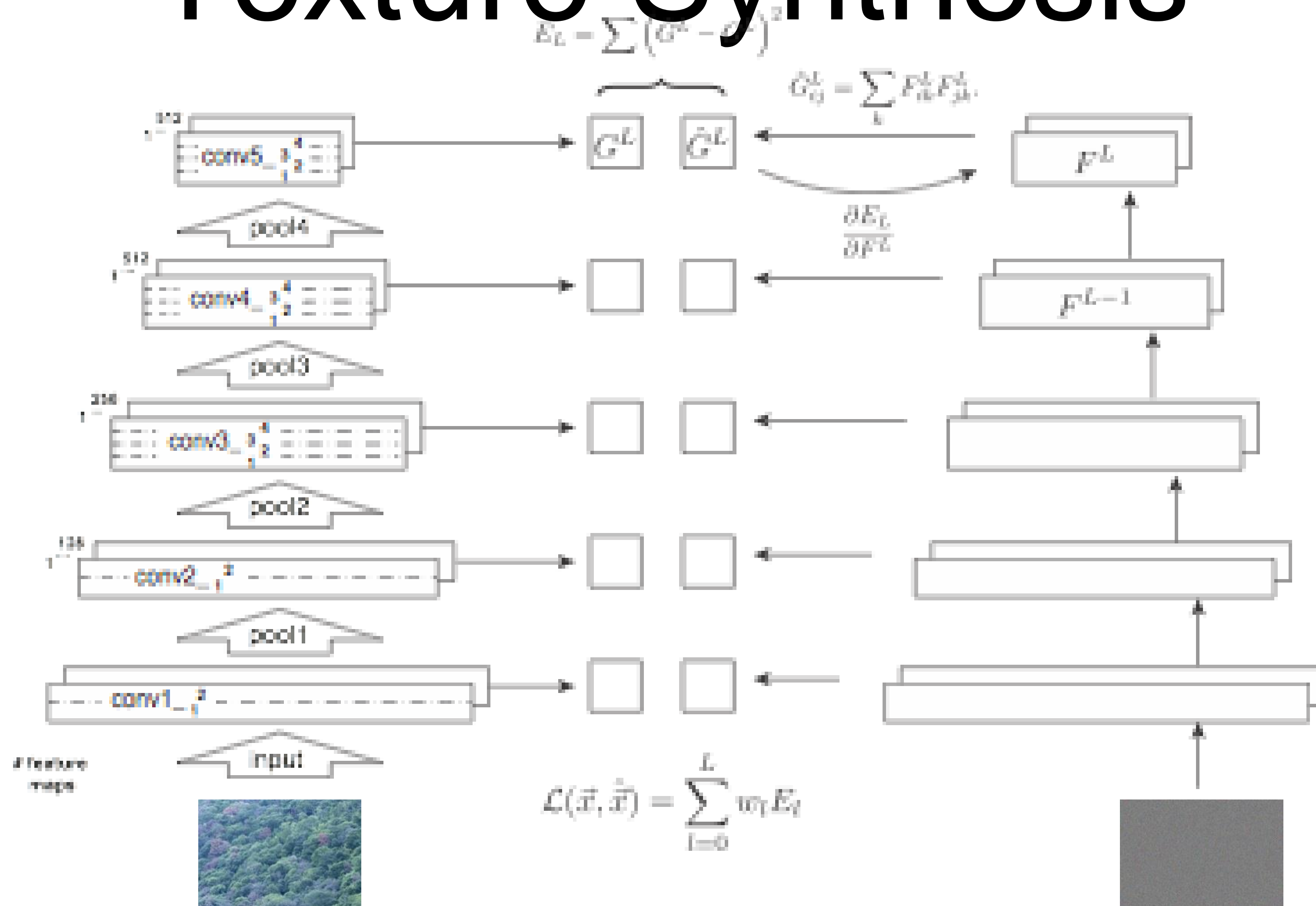
Texture Synthesis



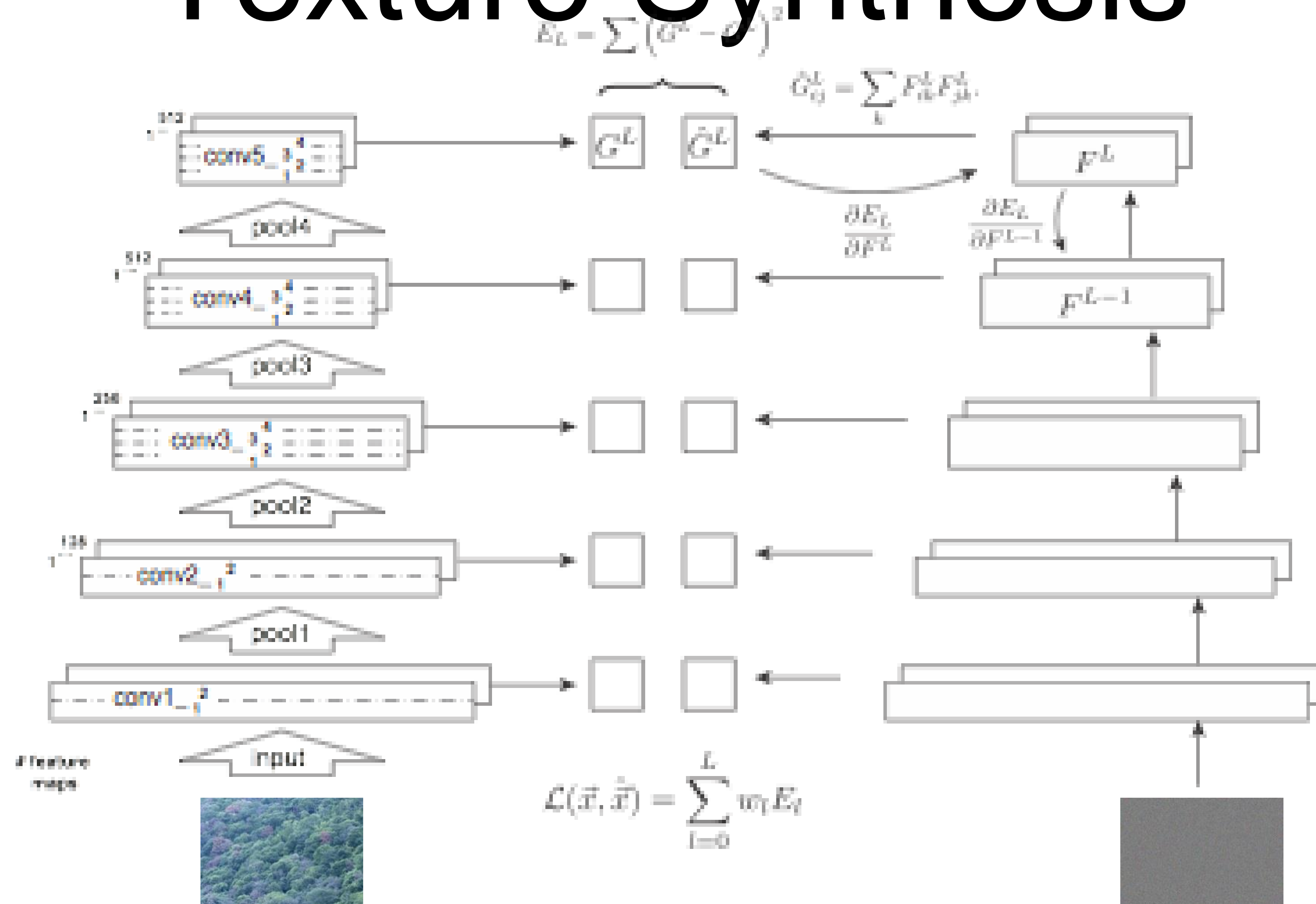
Texture Synthesis



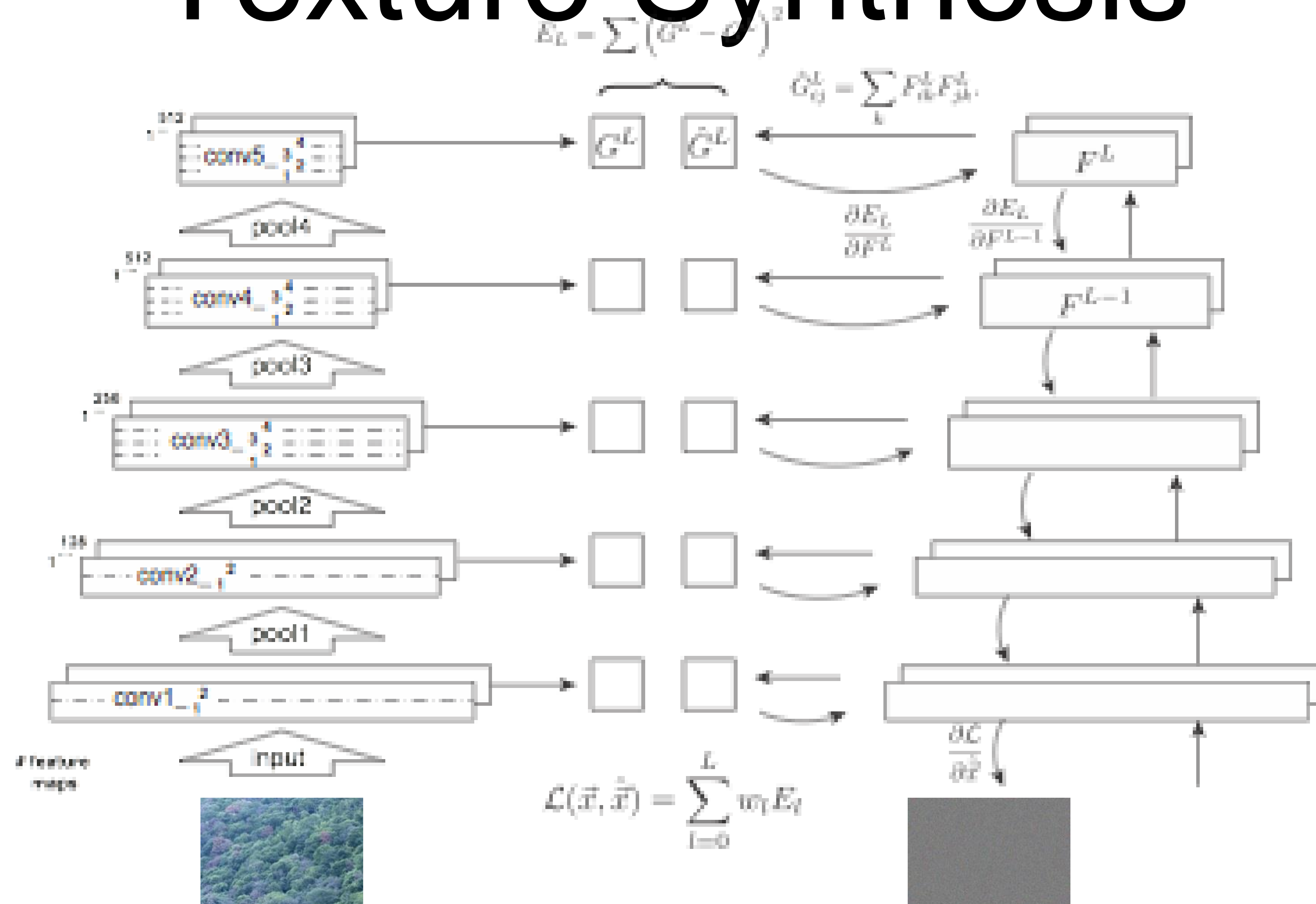
Texture Synthesis



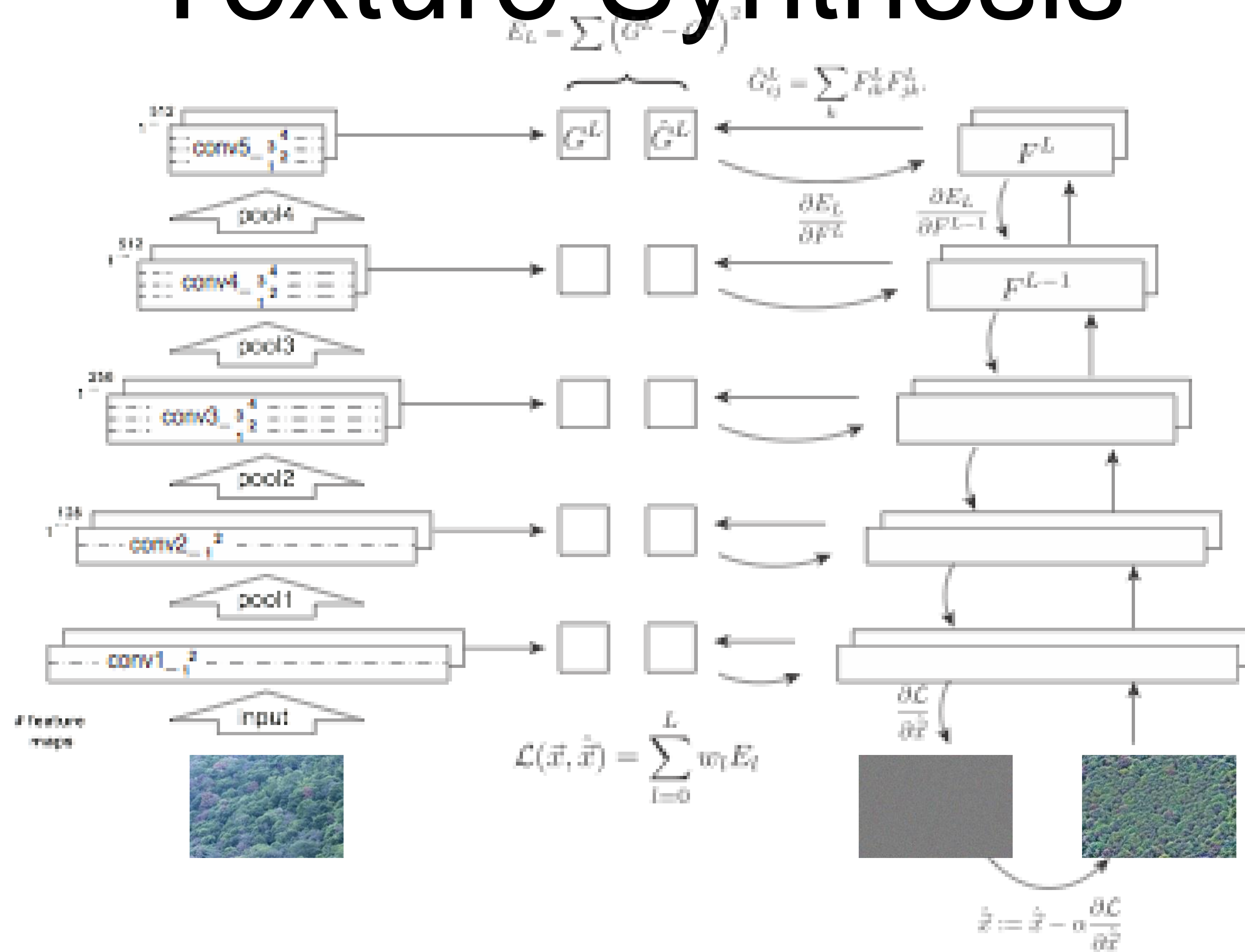
Texture Synthesis



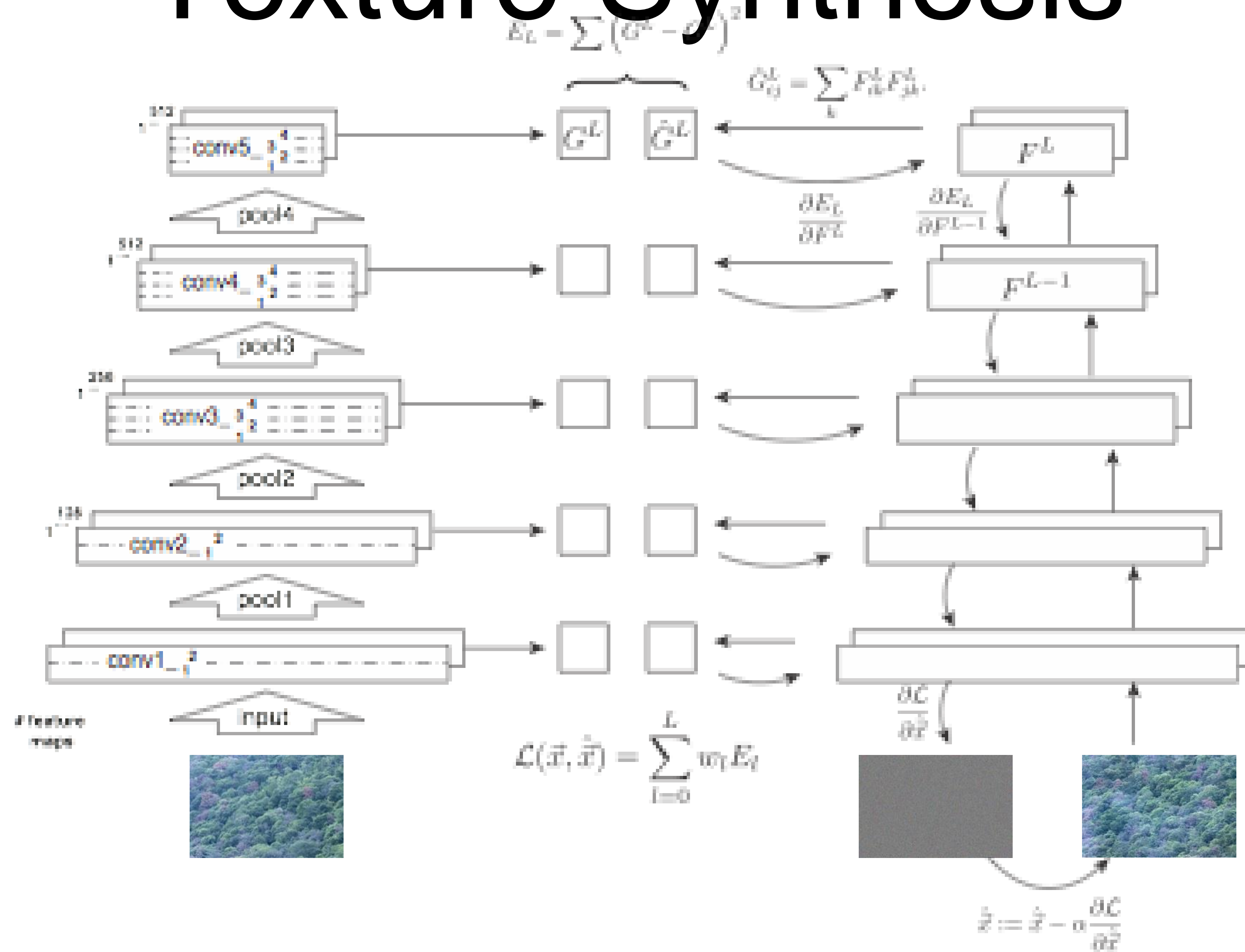
Texture Synthesis



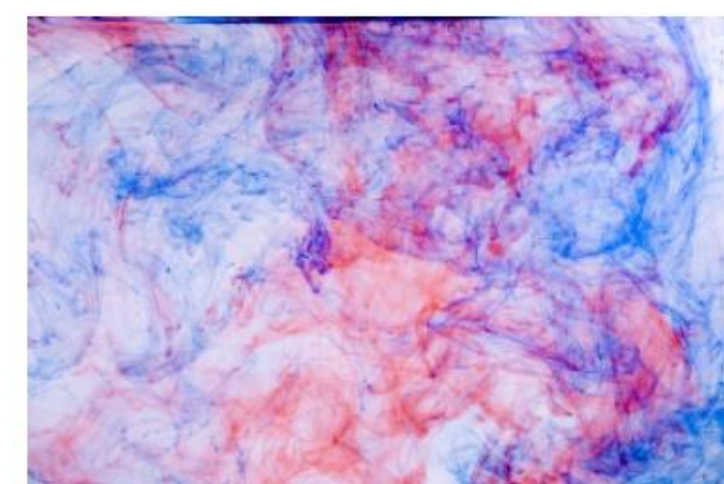
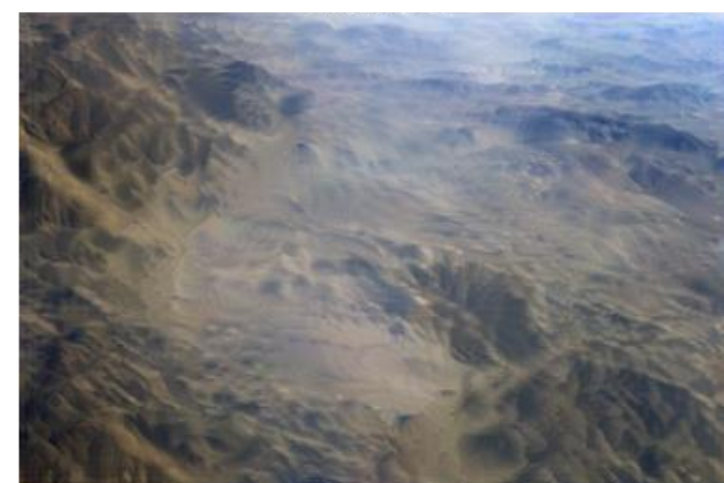
Texture Synthesis



Texture Synthesis



Test Julesz' Conjecture



Test Julesz' Conjecture

Synthesised



Source



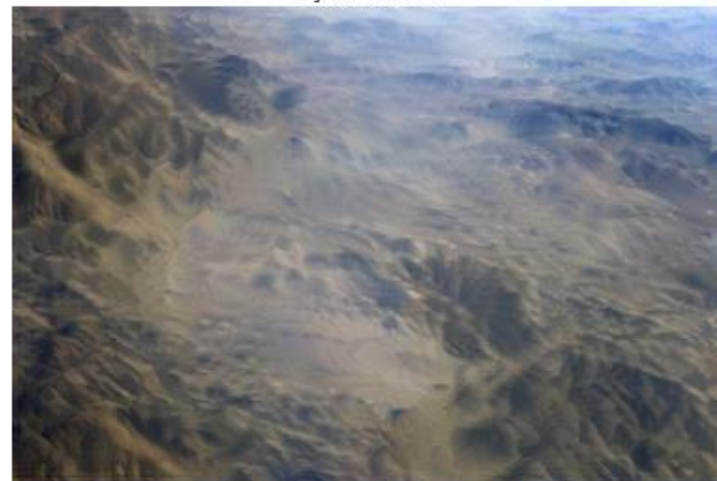
Synthesised



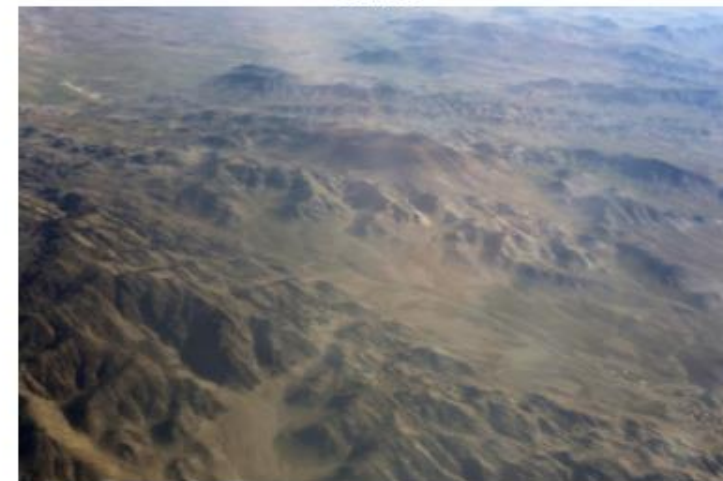
Source



Synthesised



Source



Synthesised



Source



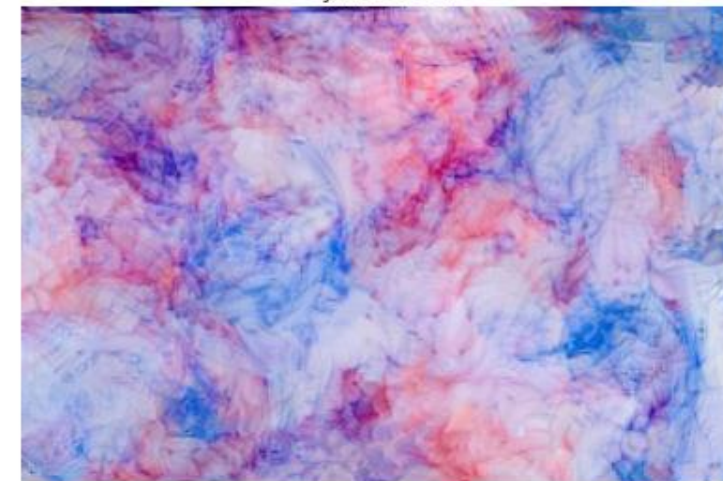
Synthesised



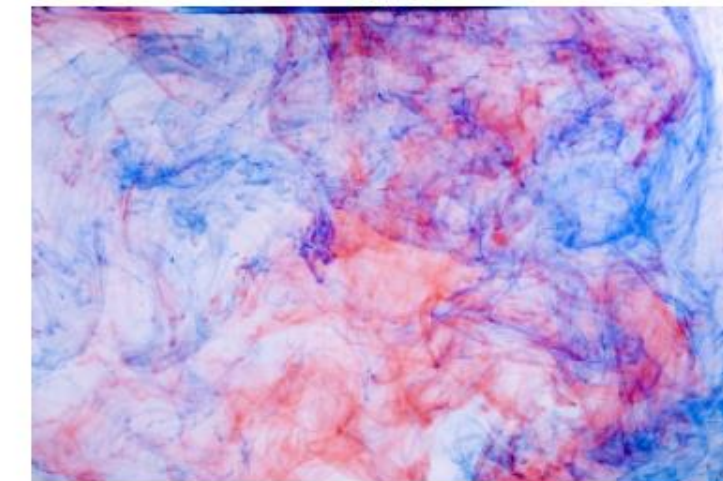
Source



Synthesised



Source



Synthesised



Source



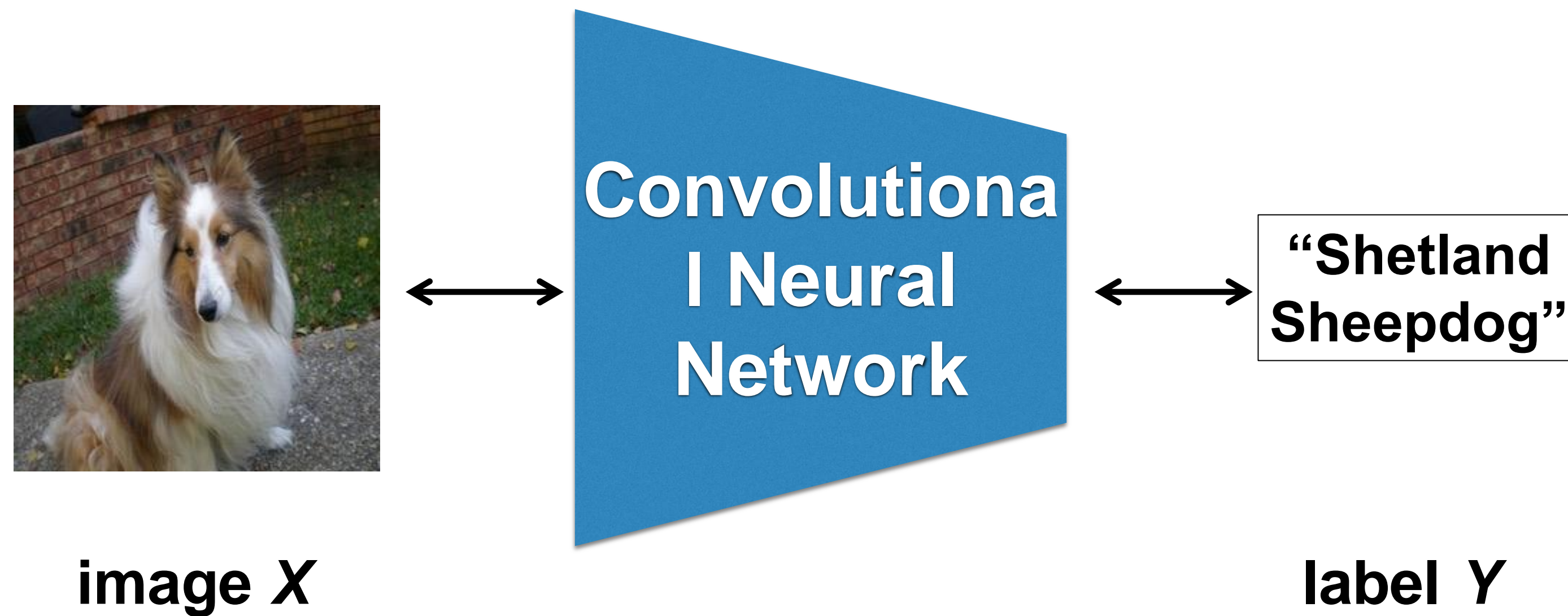
Synthesised



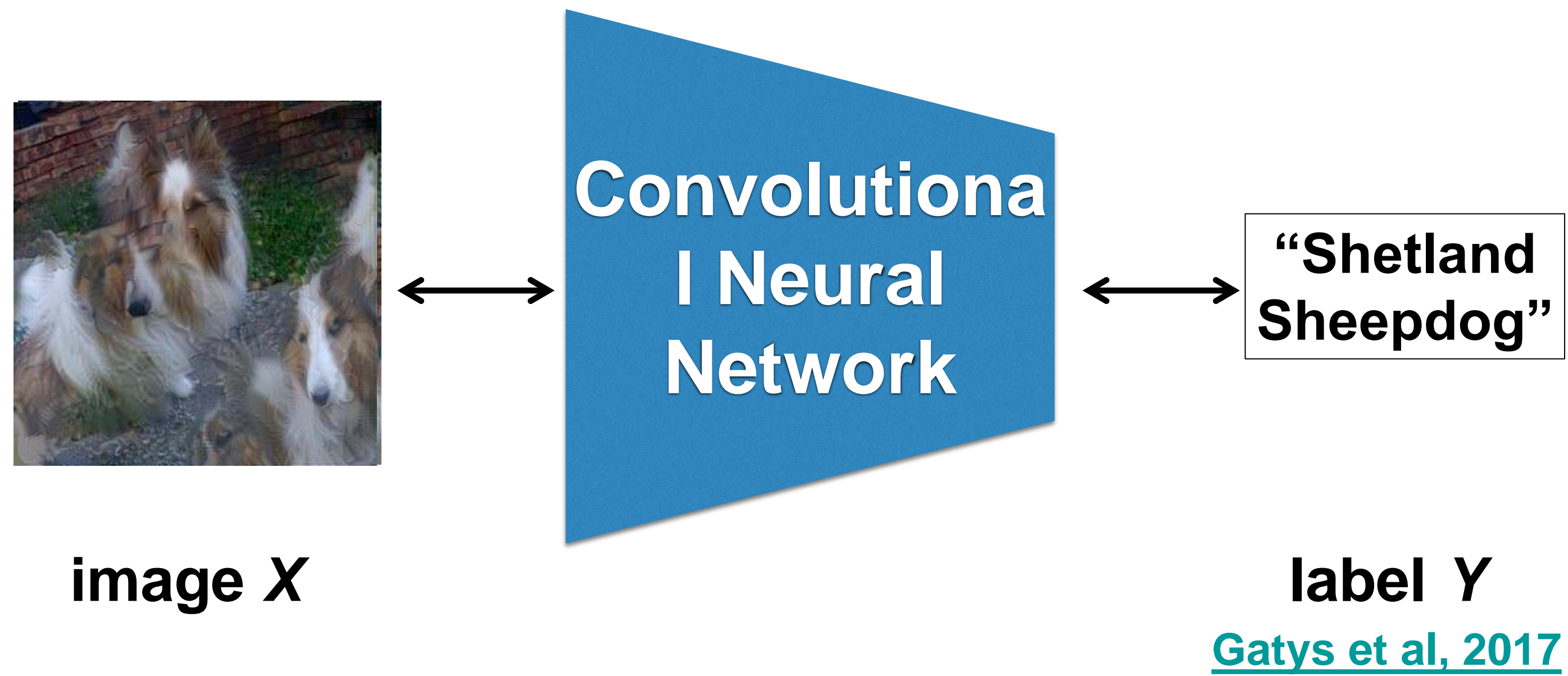
Source



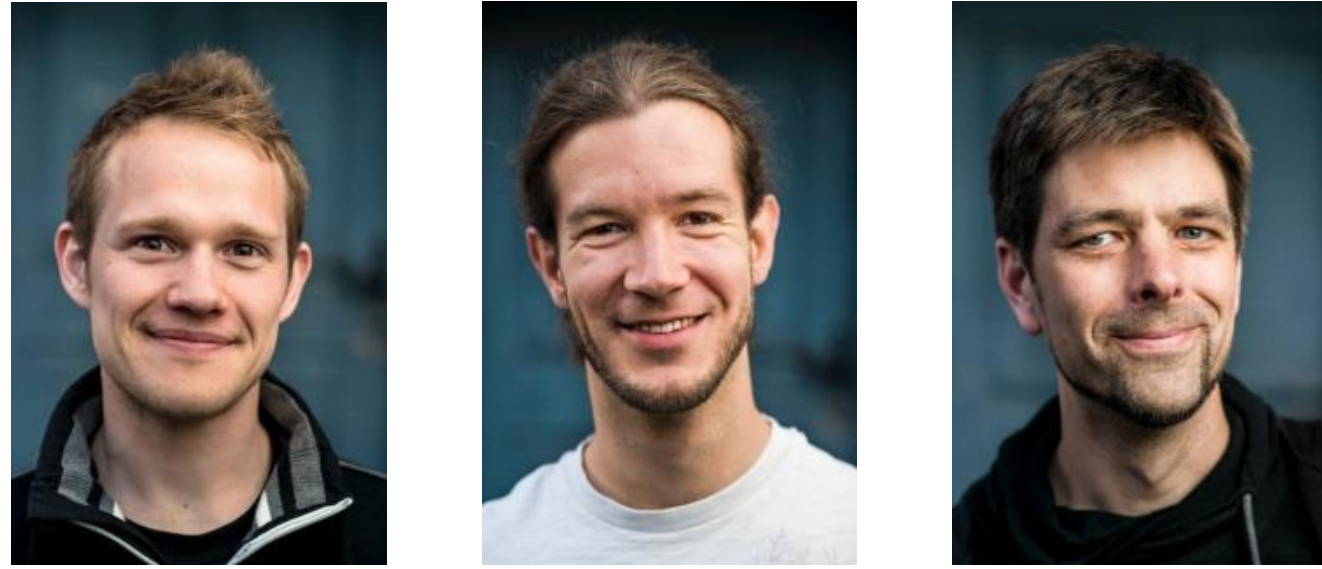
ImageNet Recognition is just Texture Recognition



ImageNet Recognition is just Texture Recognition



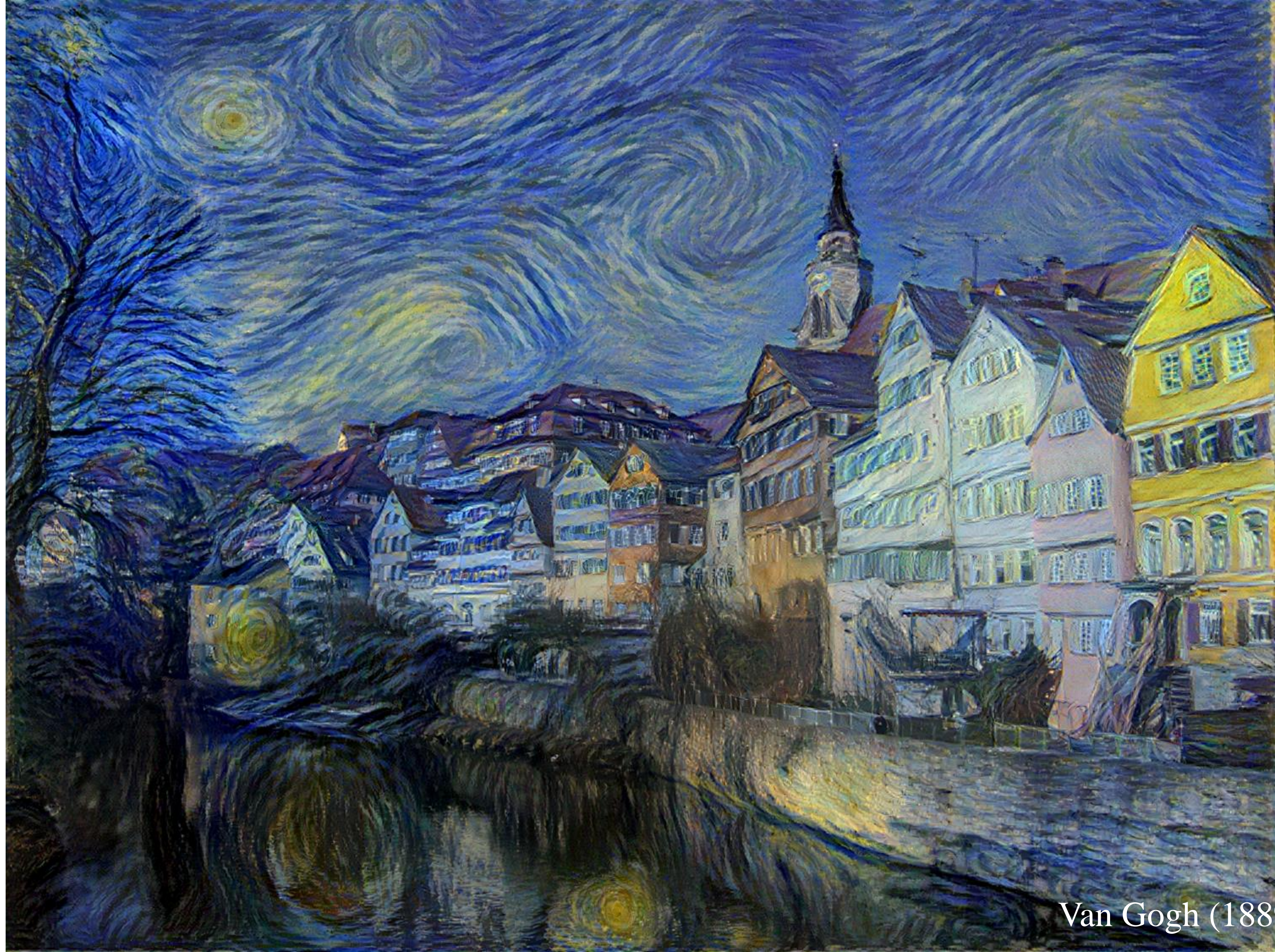
A Neural Algorithm of Artistic Style



Gatys, Ecker, Bethge (arXiv 2015)







Van Gogh (1888)



Picasso (1910)



Munch (1893)

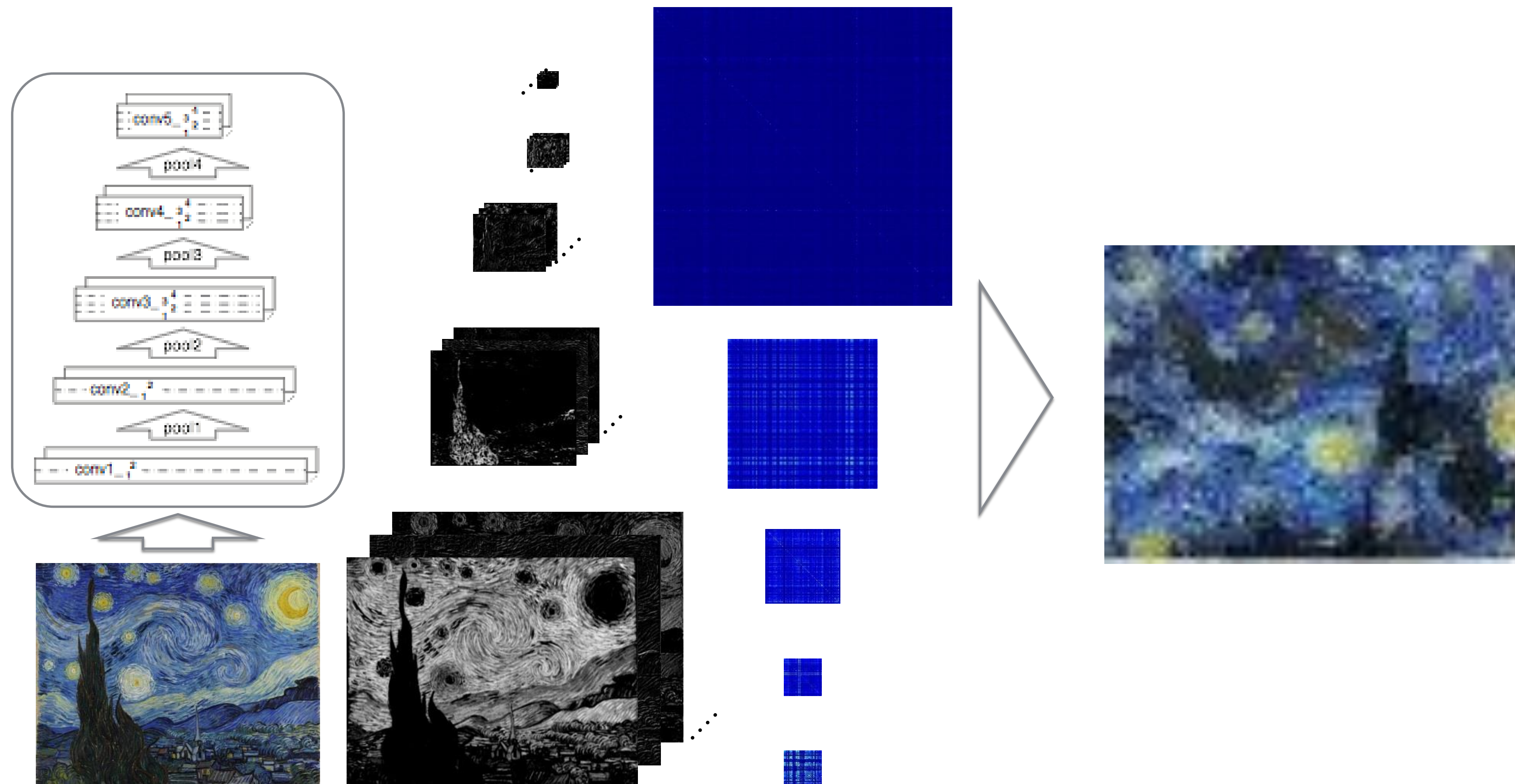


Turner (1805)



Kandinsky (1911)

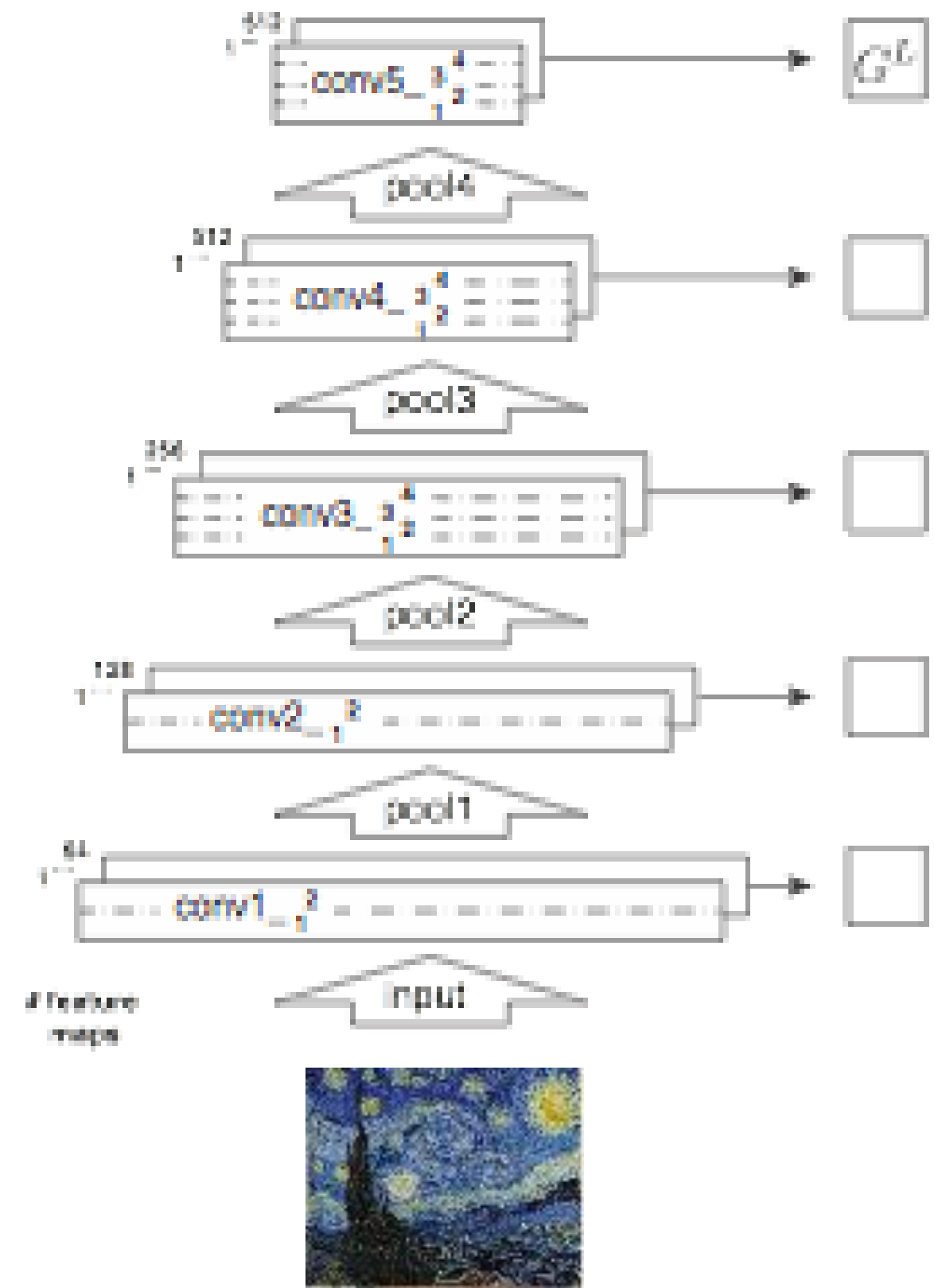
CNN - Texture Synthesis



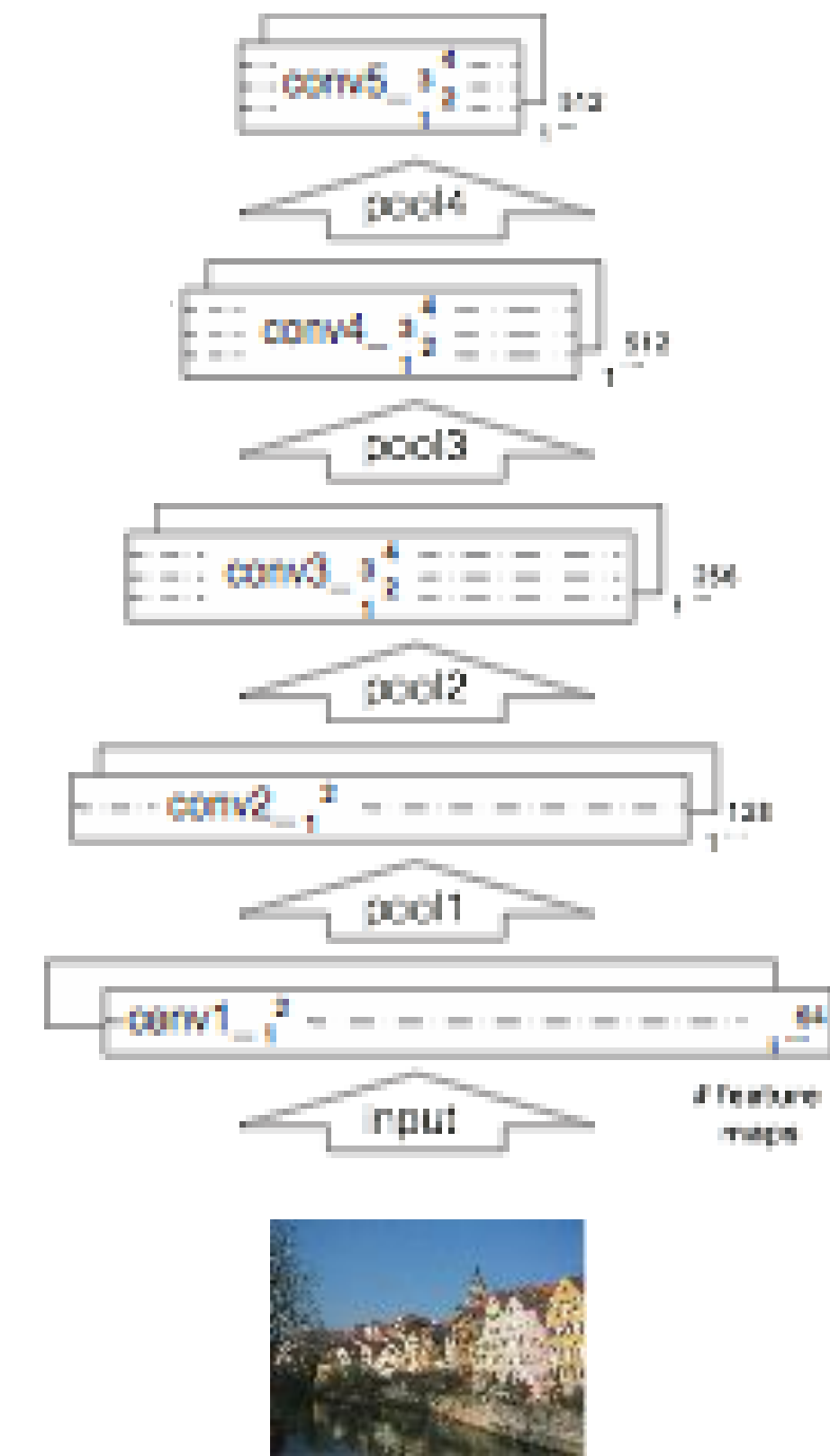
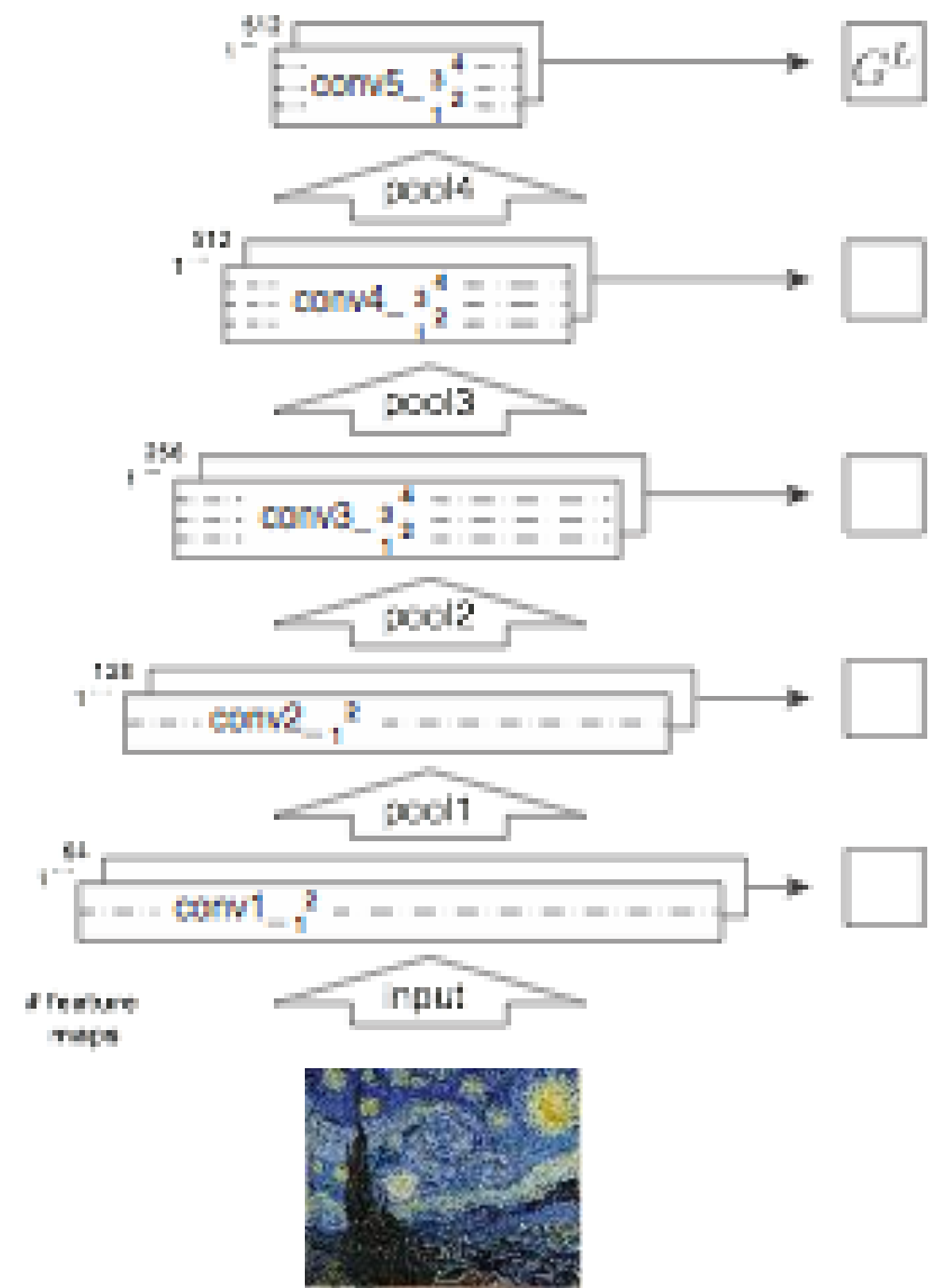
Artistic Style Transfer



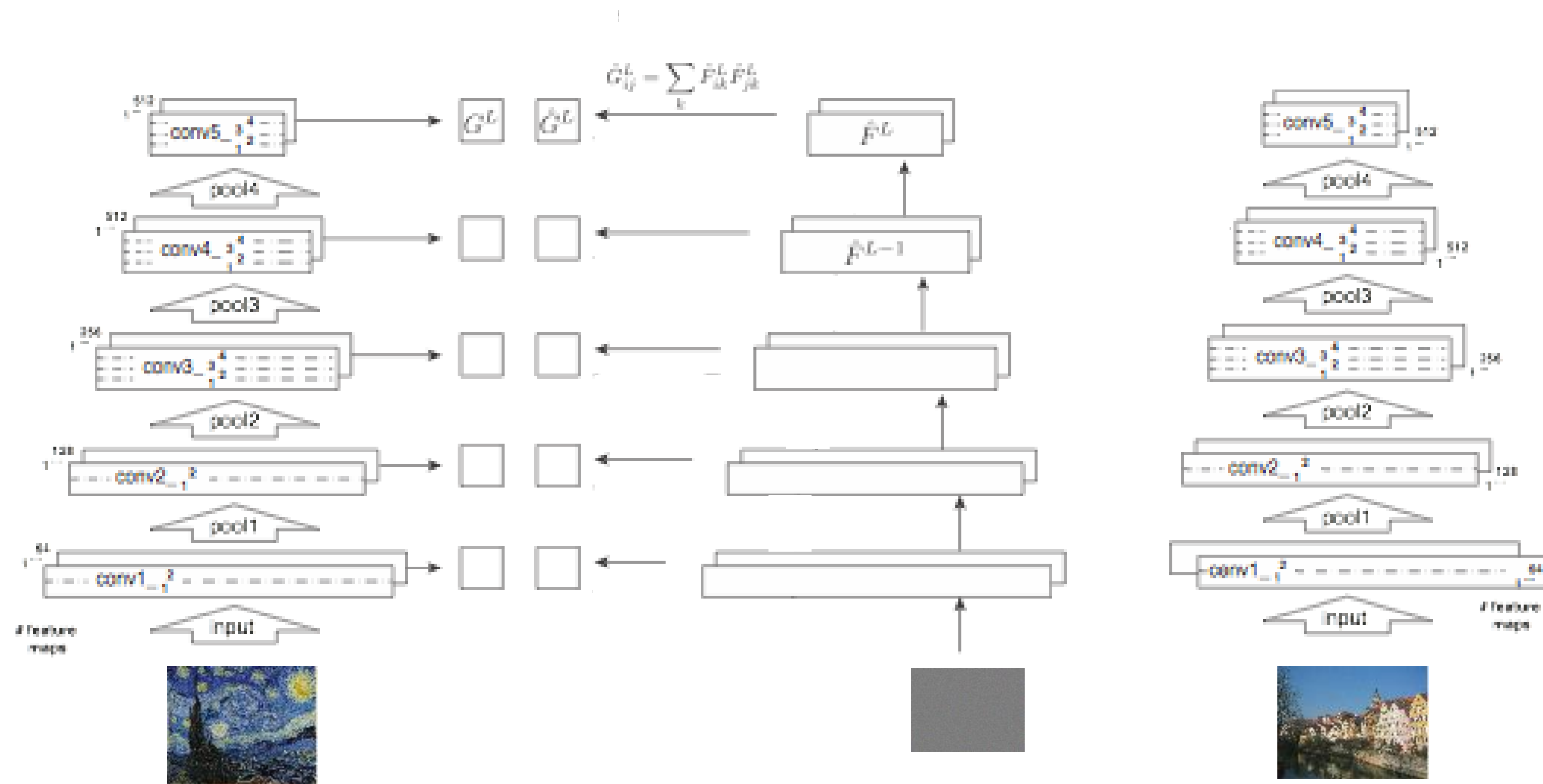
Artistic Style Transfer



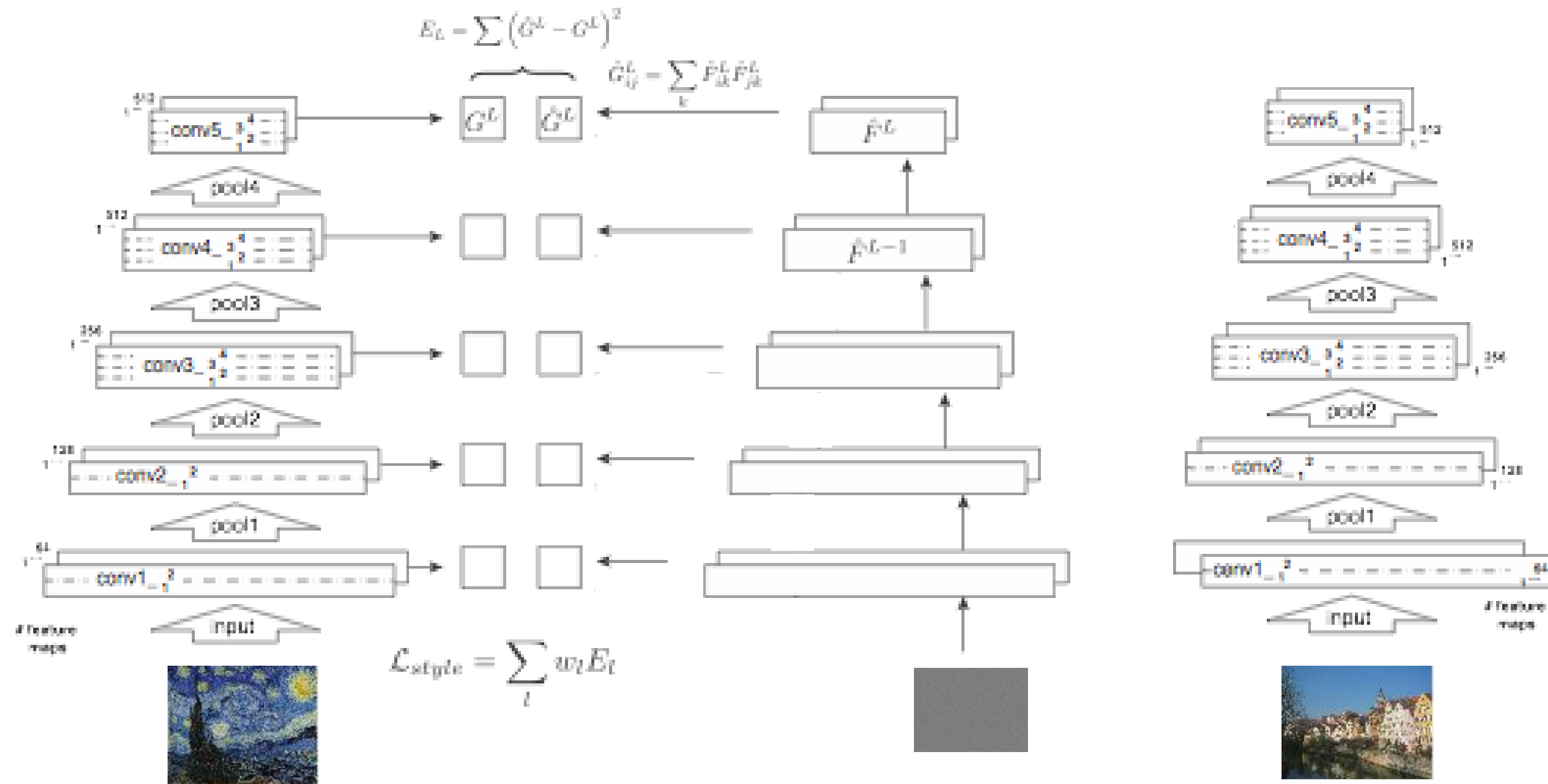
Artistic Style Transfer



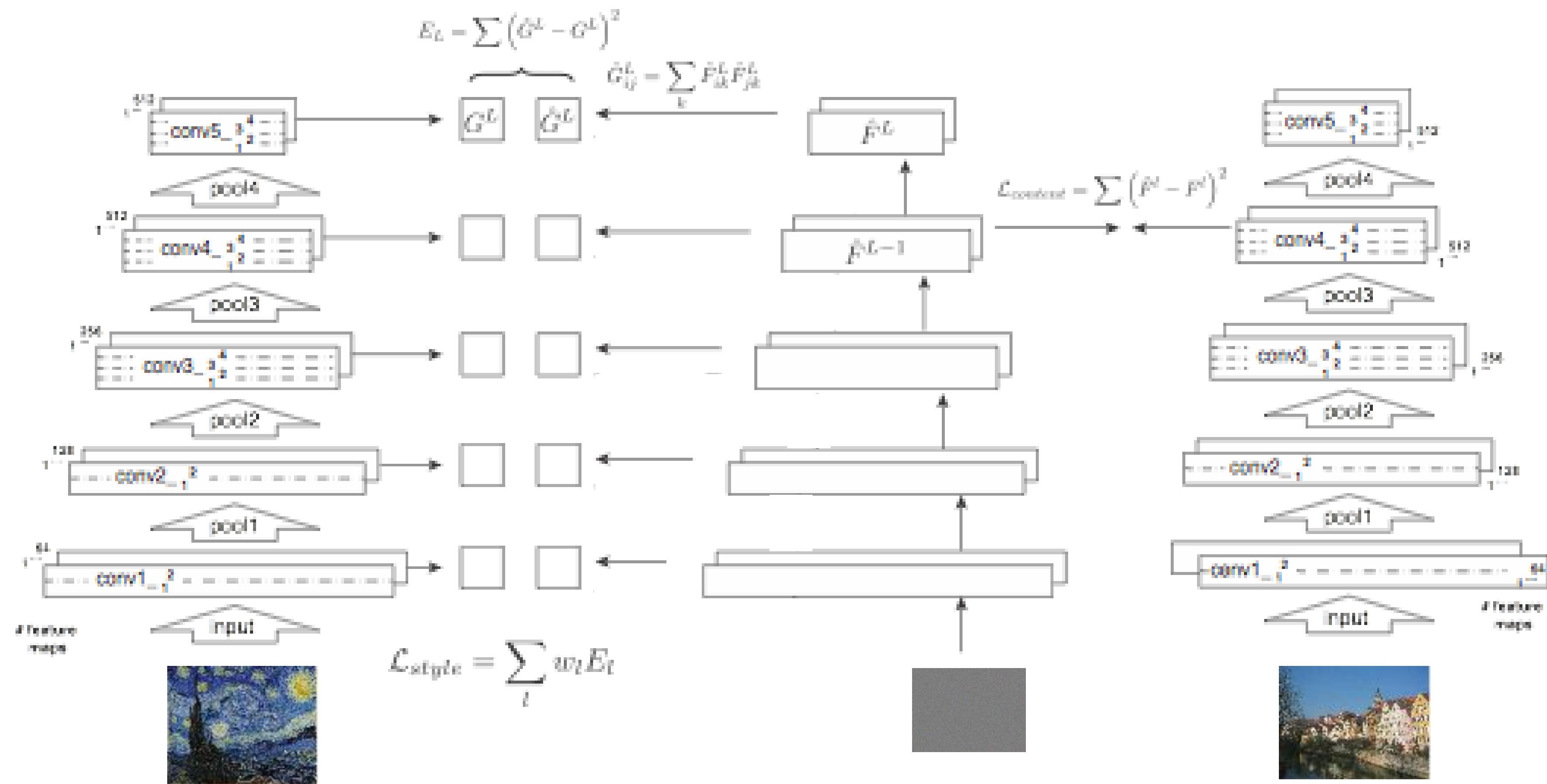
Artistic Style Transfer



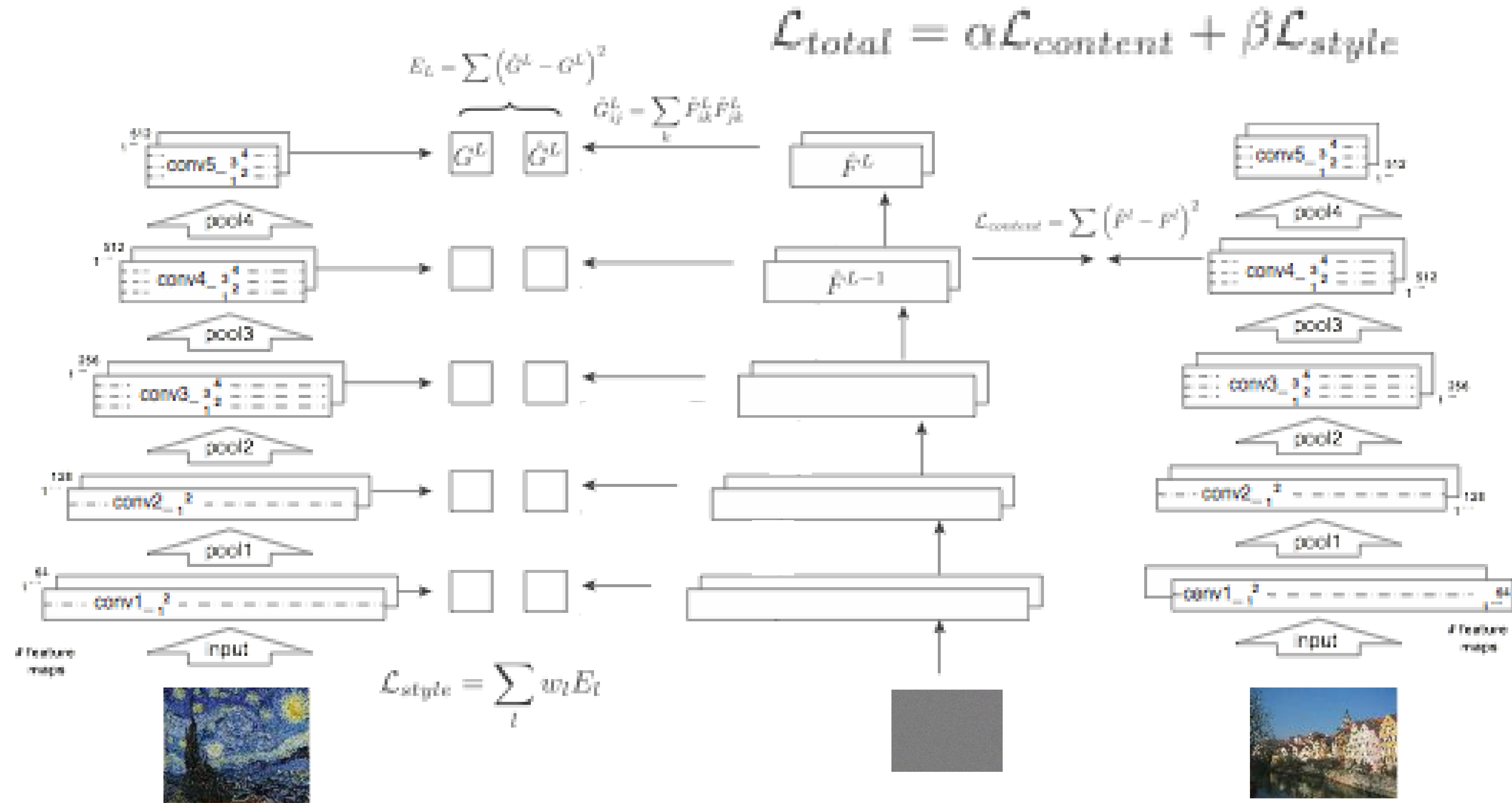
Artistic Style Transfer



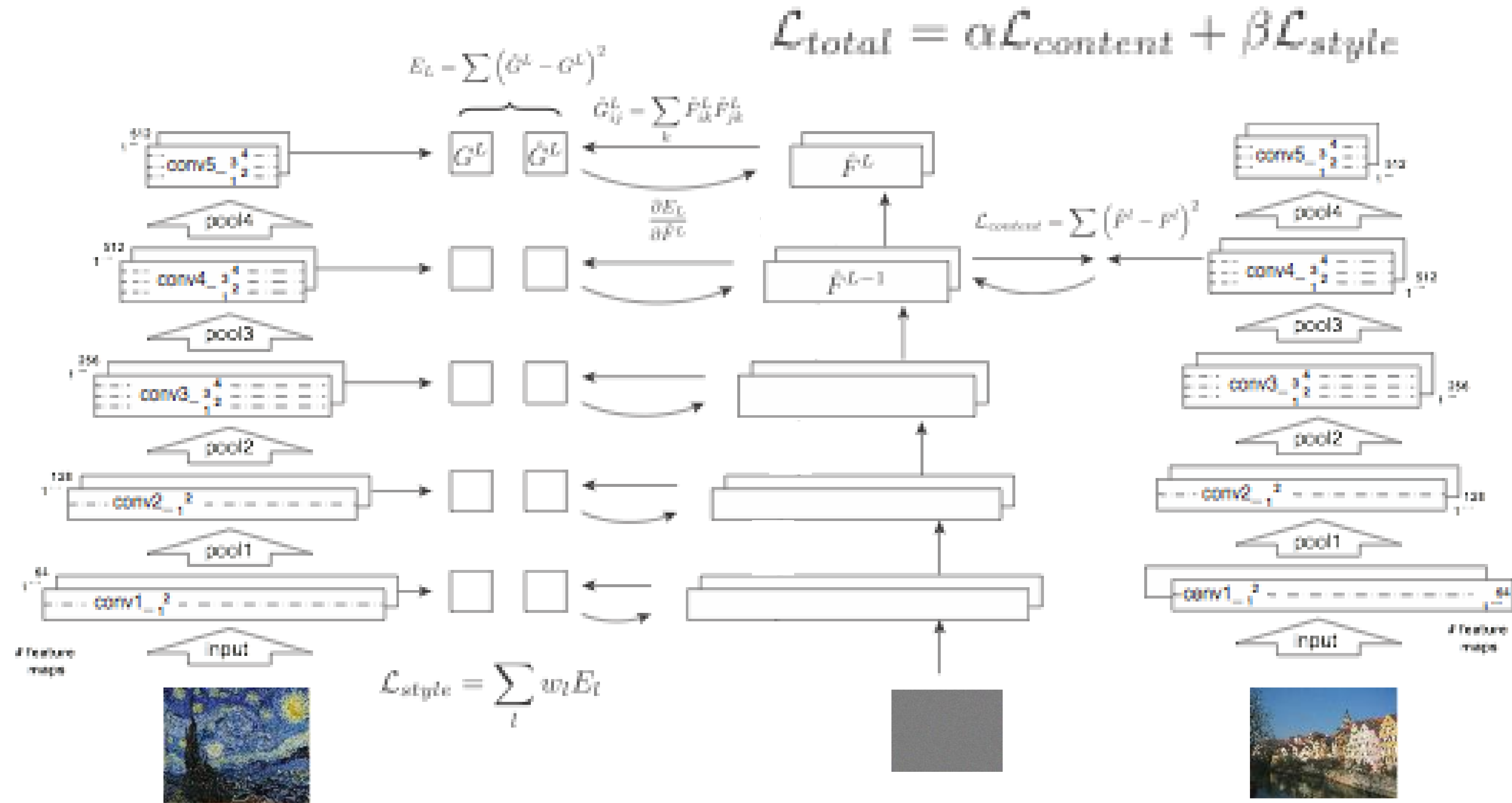
Artistic Style Transfer



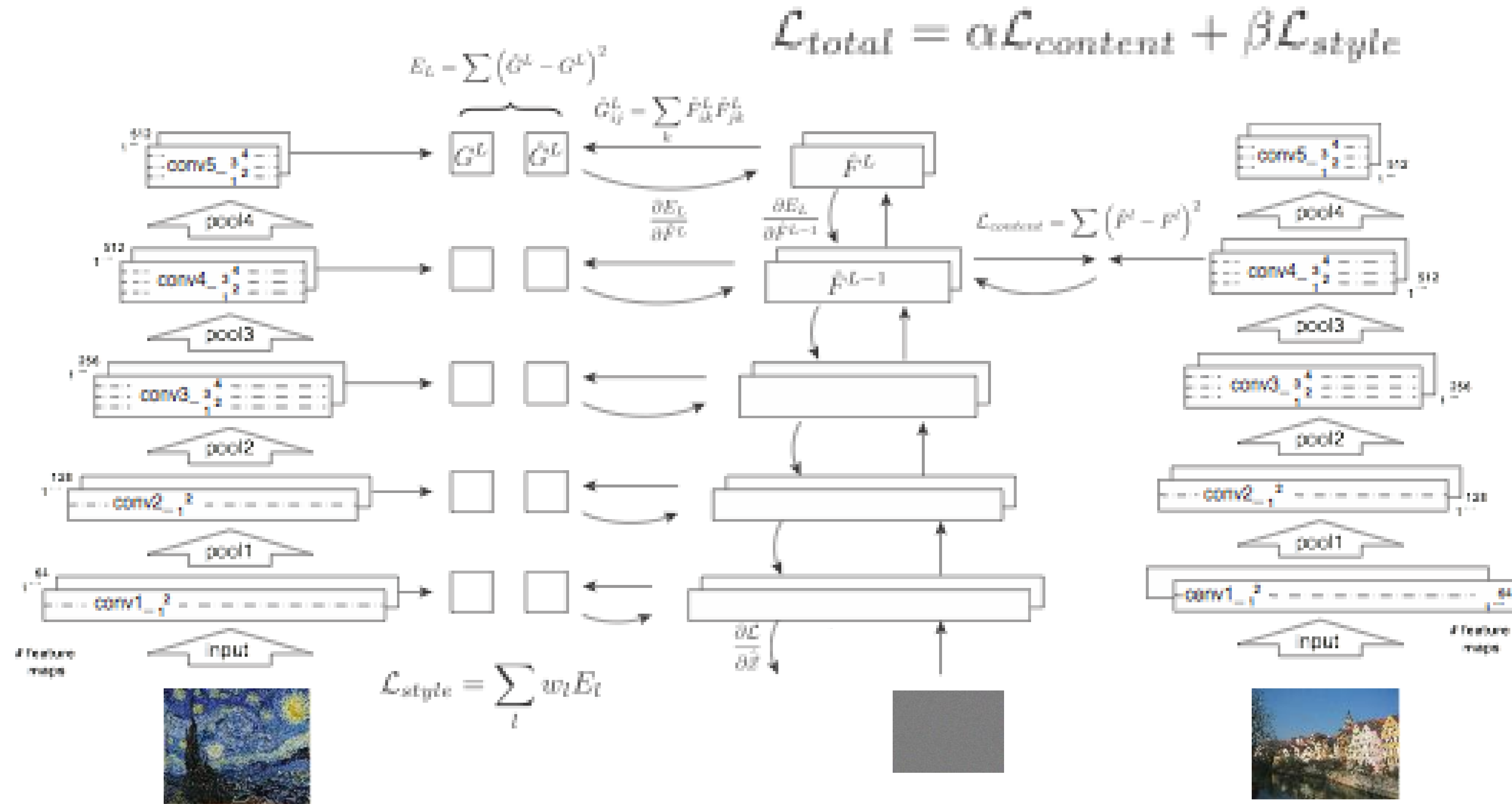
Artistic Style Transfer



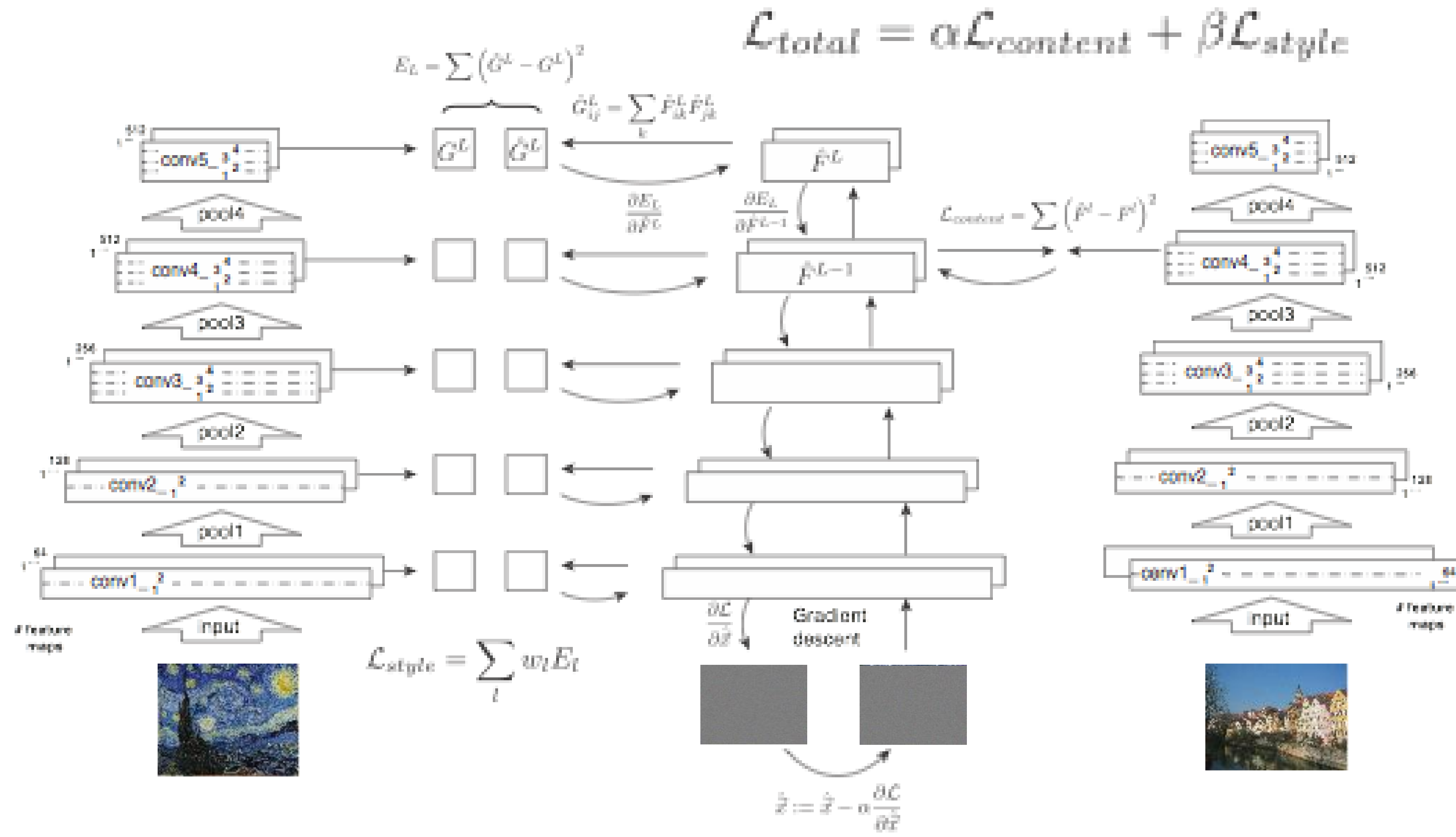
Artistic Style Transfer



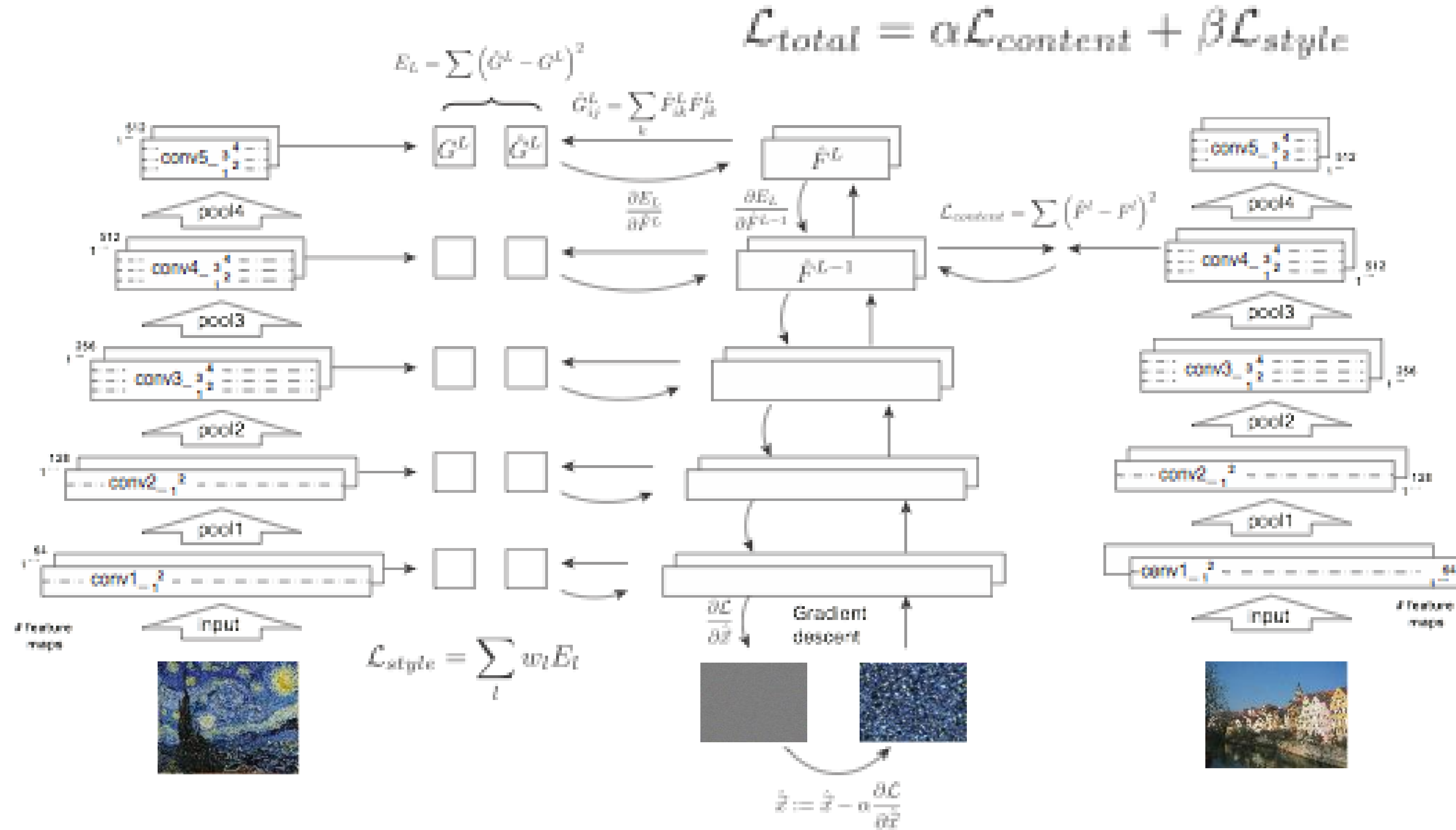
Artistic Style Transfer



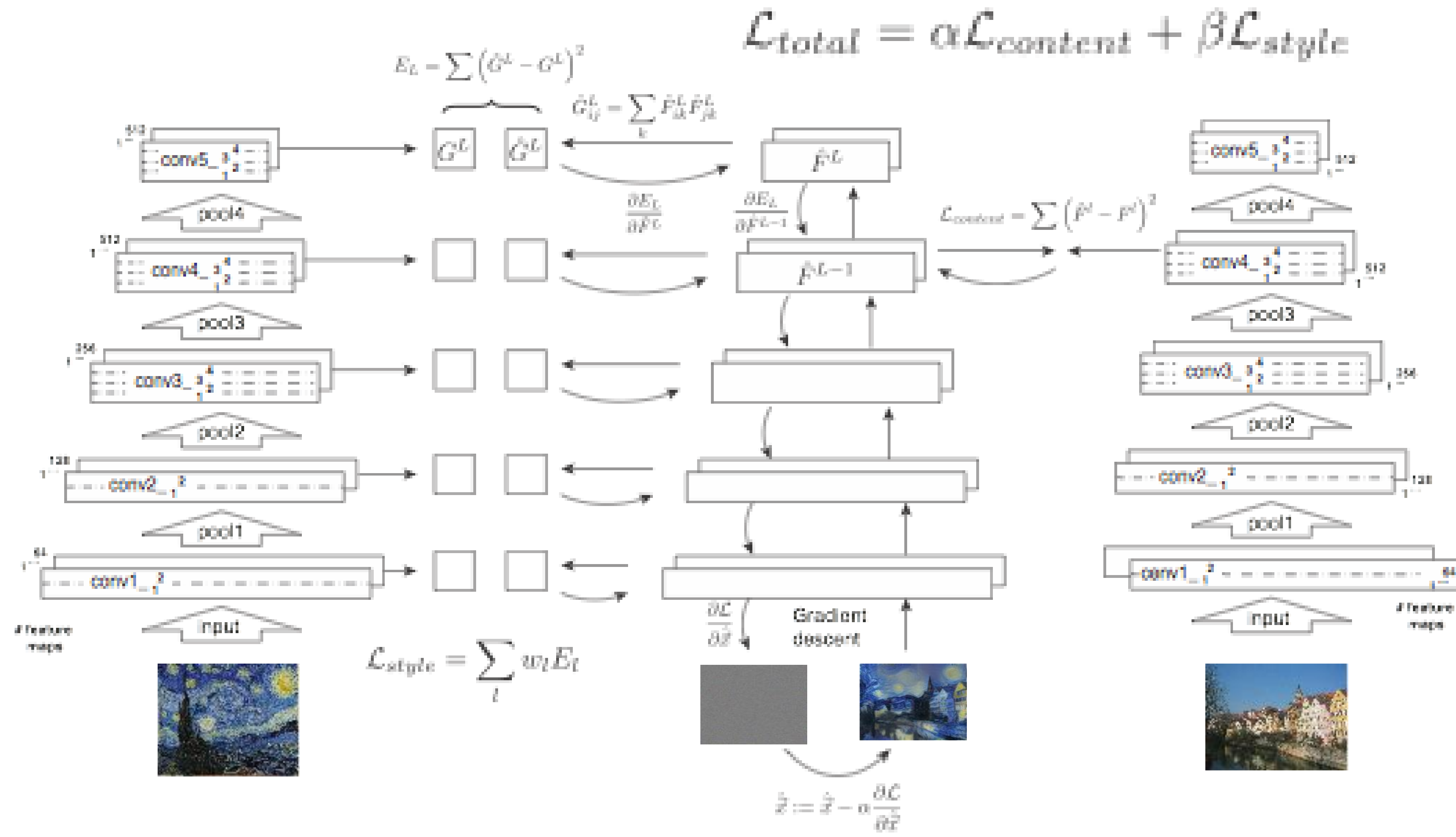
Artistic Style Transfer



Artistic Style Transfer



Artistic Style Transfer



Relative Weighting of Content and Style

$1e-4$



$1e-3$



$1e-2$



$1e-1$



Different Reconstruction Layers

Conv2_2

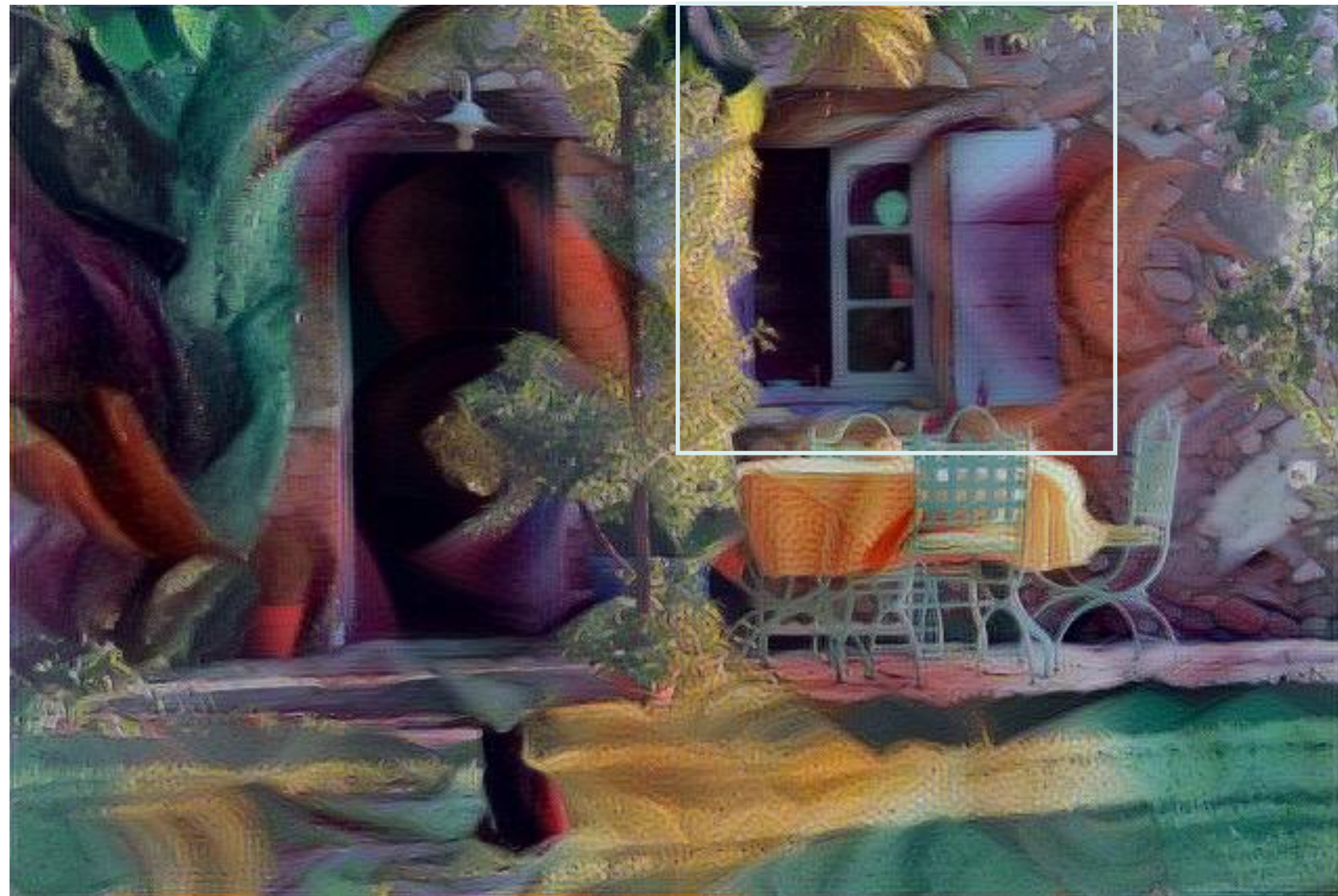


Conv4_2



Different Reconstruction Layers

Conv2_2



Conv4_2



Different Reconstruction Layers

Original



Conv2_2



Conv4_2



General Style Transfer

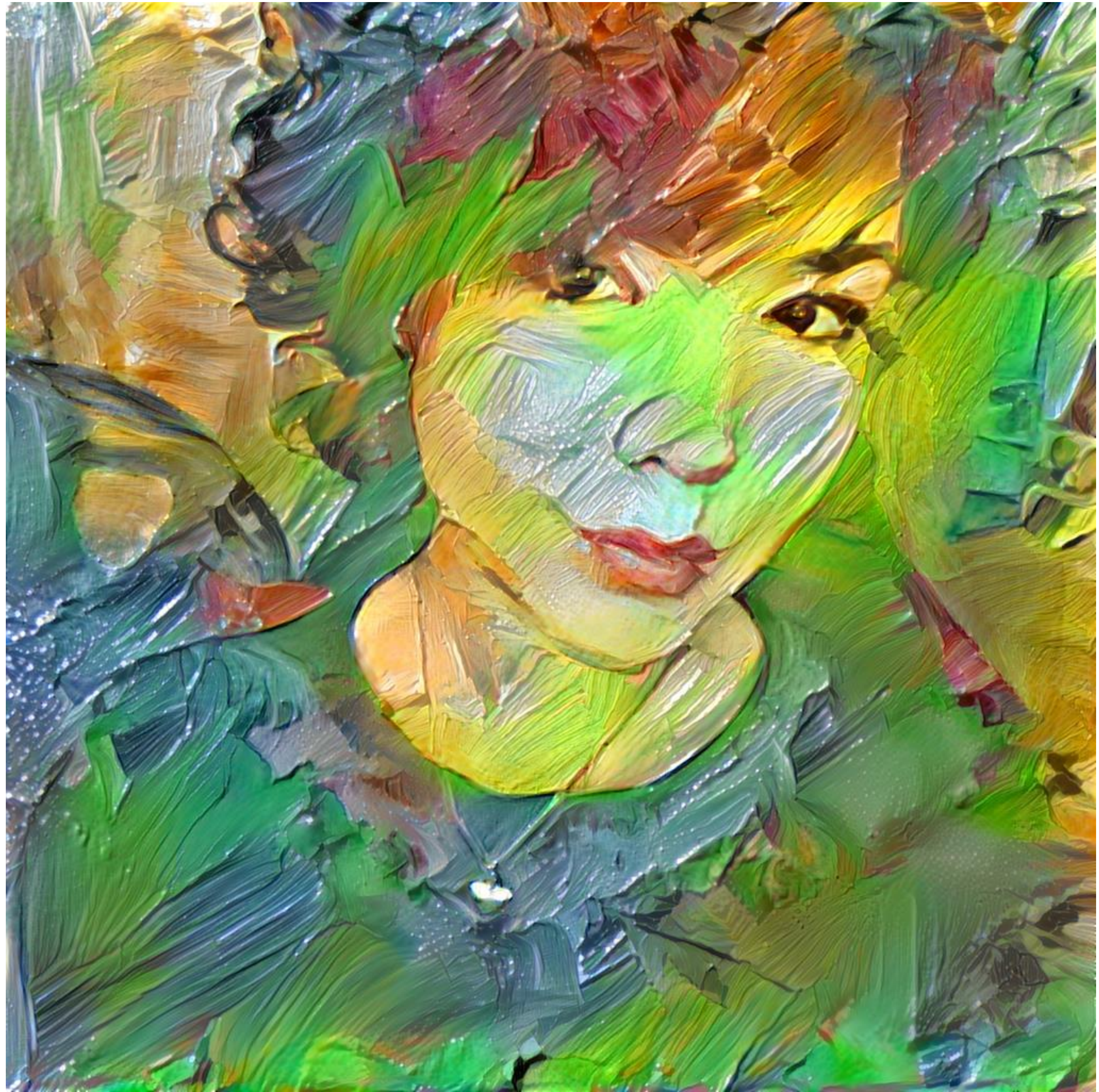


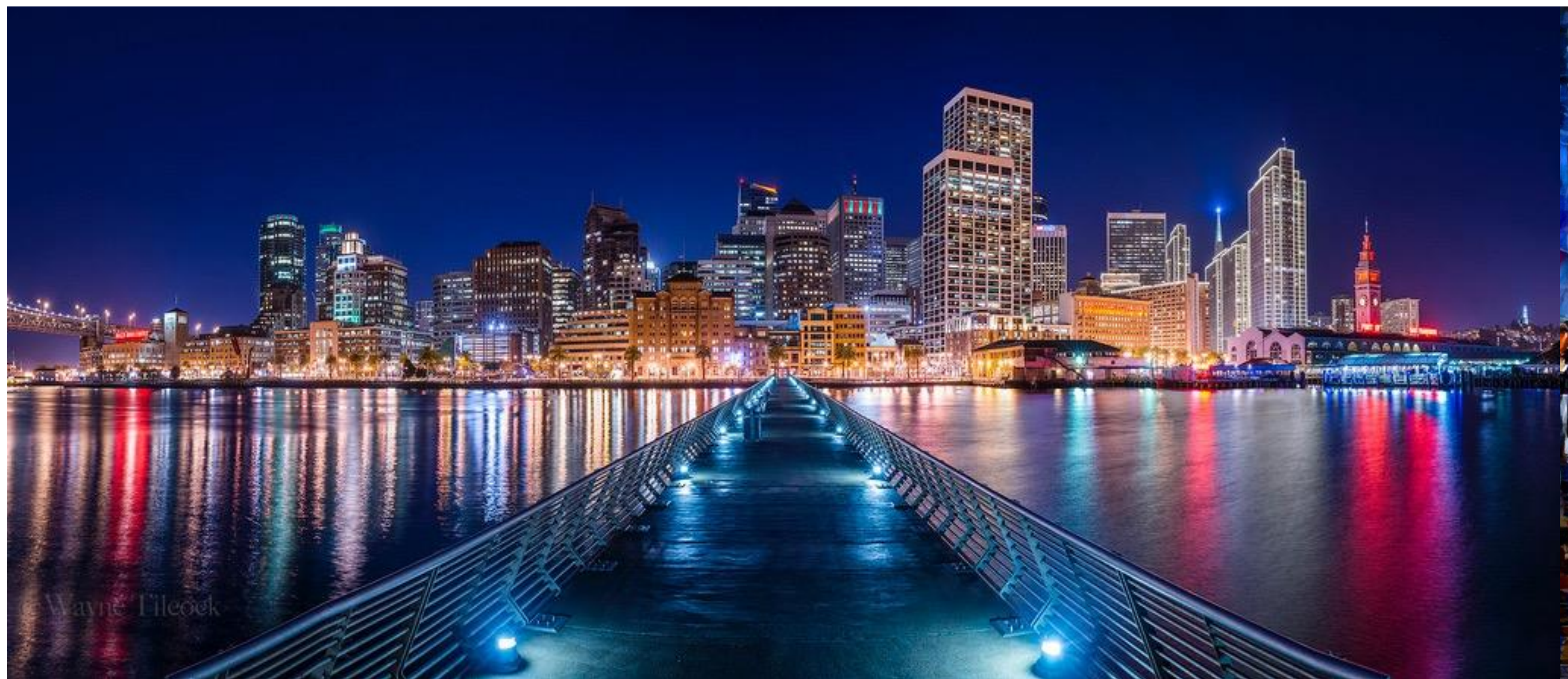
General Style Transfer



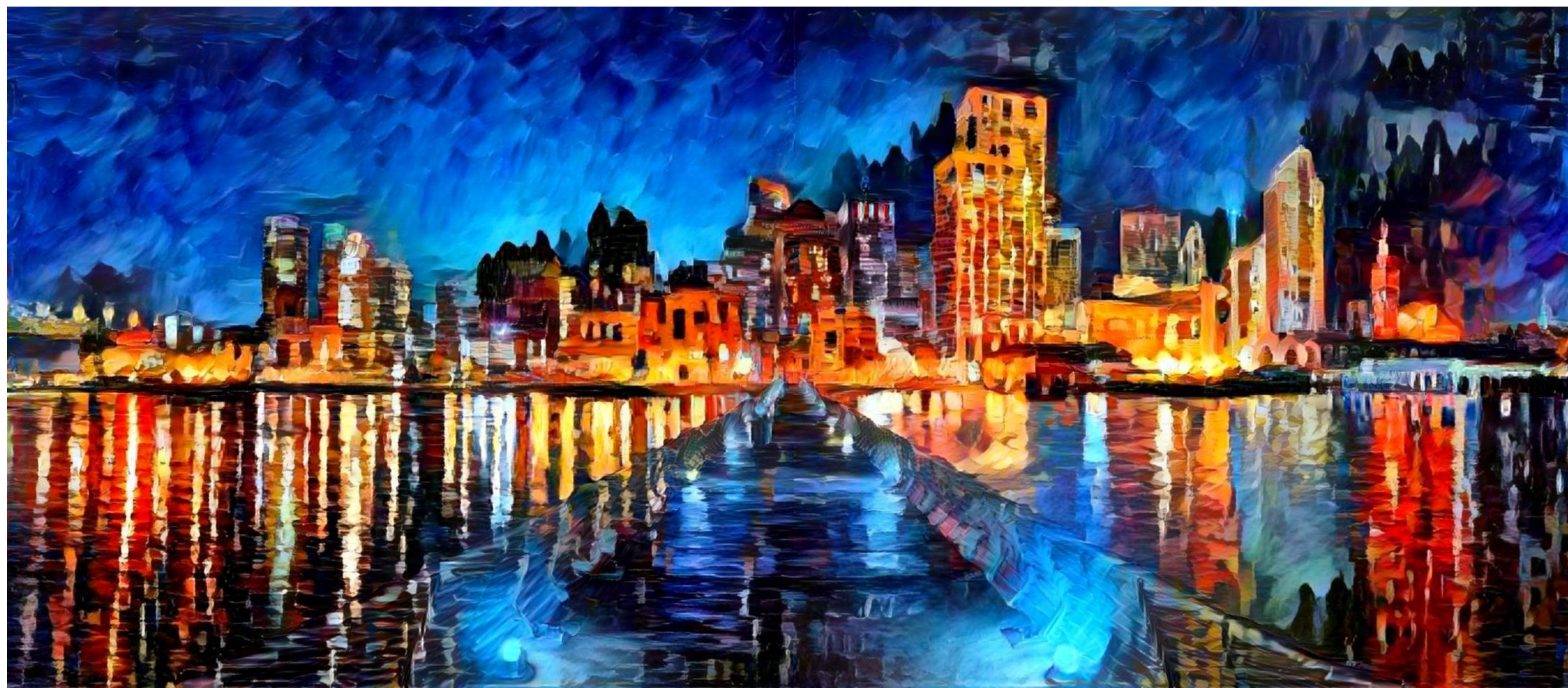




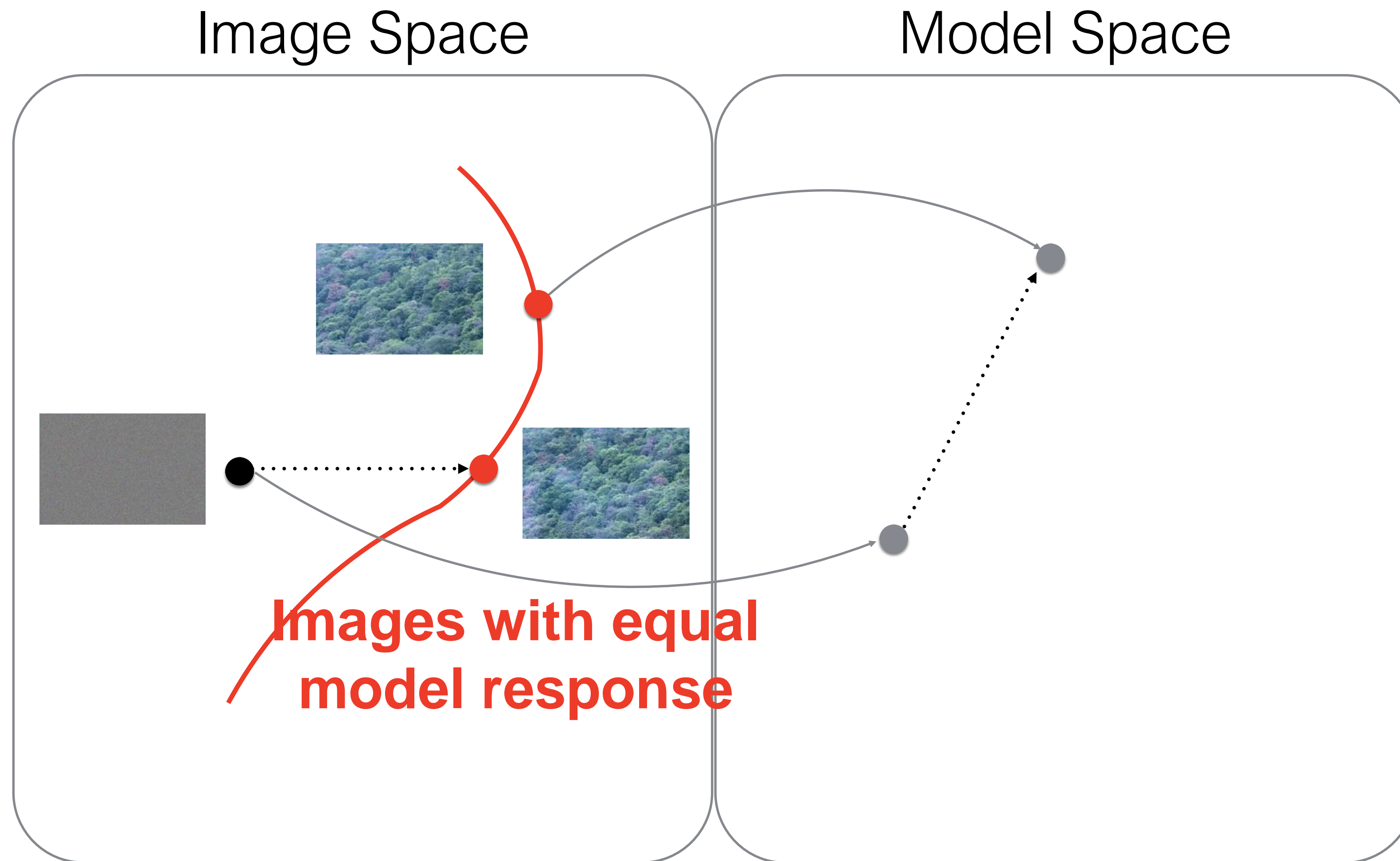




Wayne Filcock



Projecting onto the “Image Manifold”





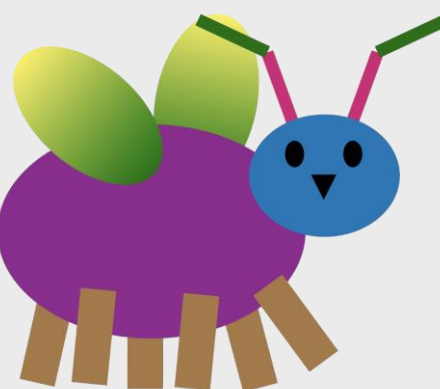
Imagen



DALL-E 2

diffusion

By Steve Seitz



**Language
Generator**

(+ pixels)

Parti

**Image
Generator**

(+ language)

Imagen

Generate 100 images



Generate 100 images

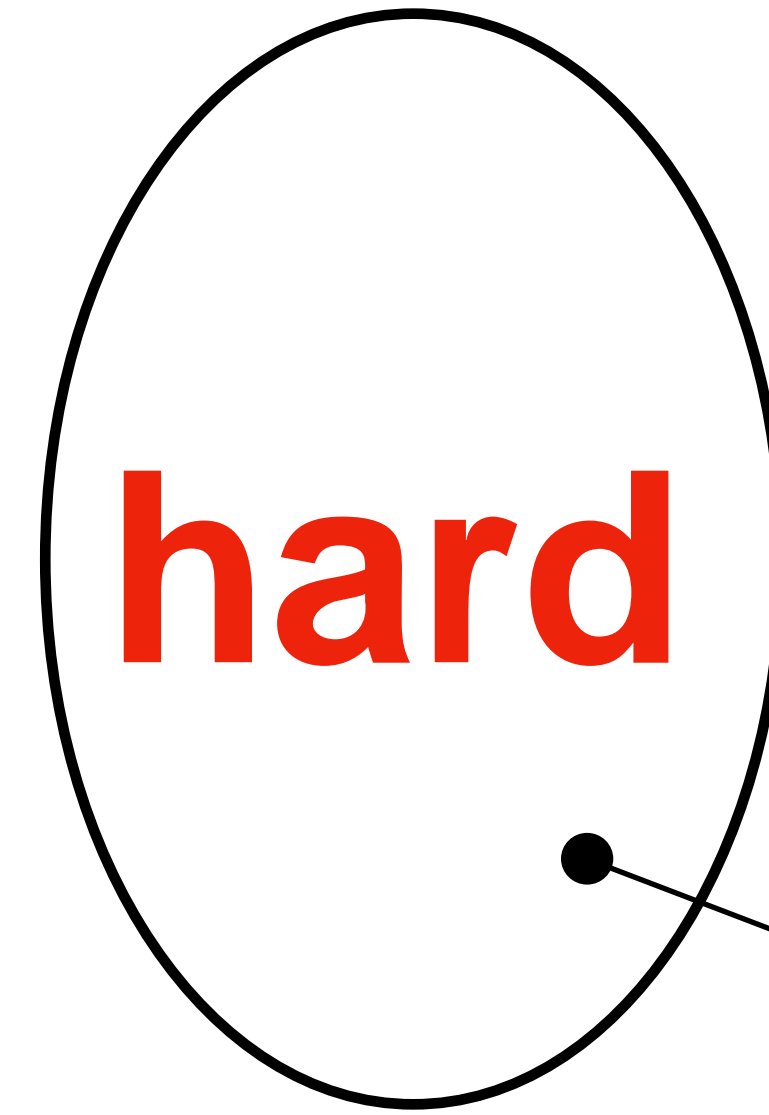


Generate 100 images of raspberries



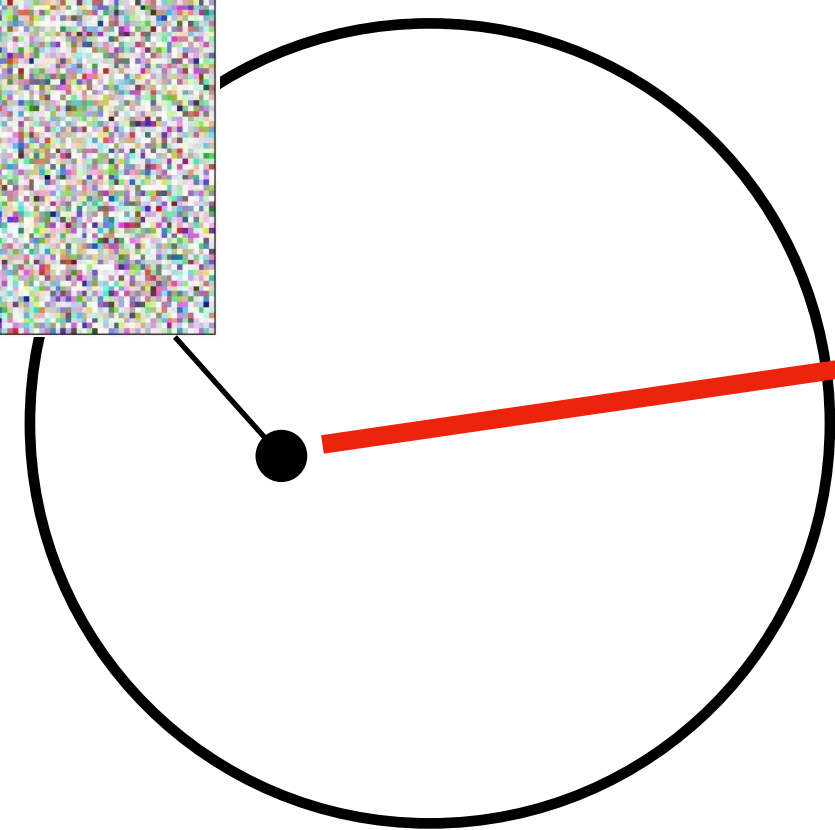
easy

random
images



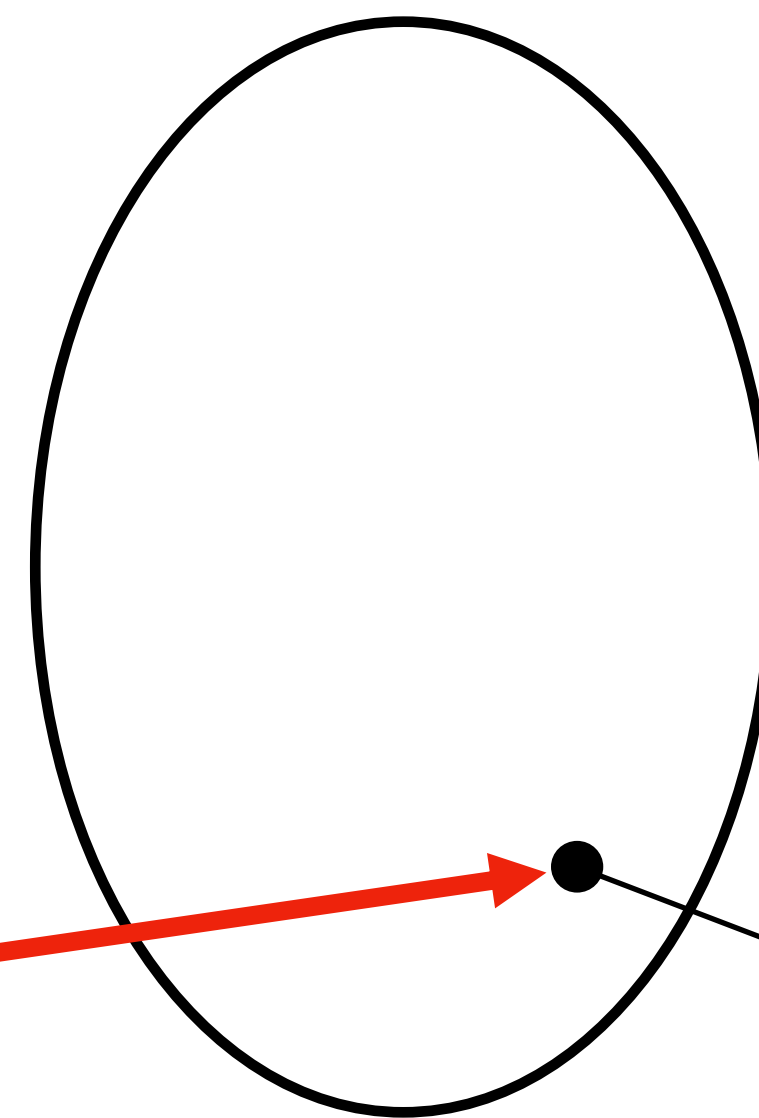
raspberry
images





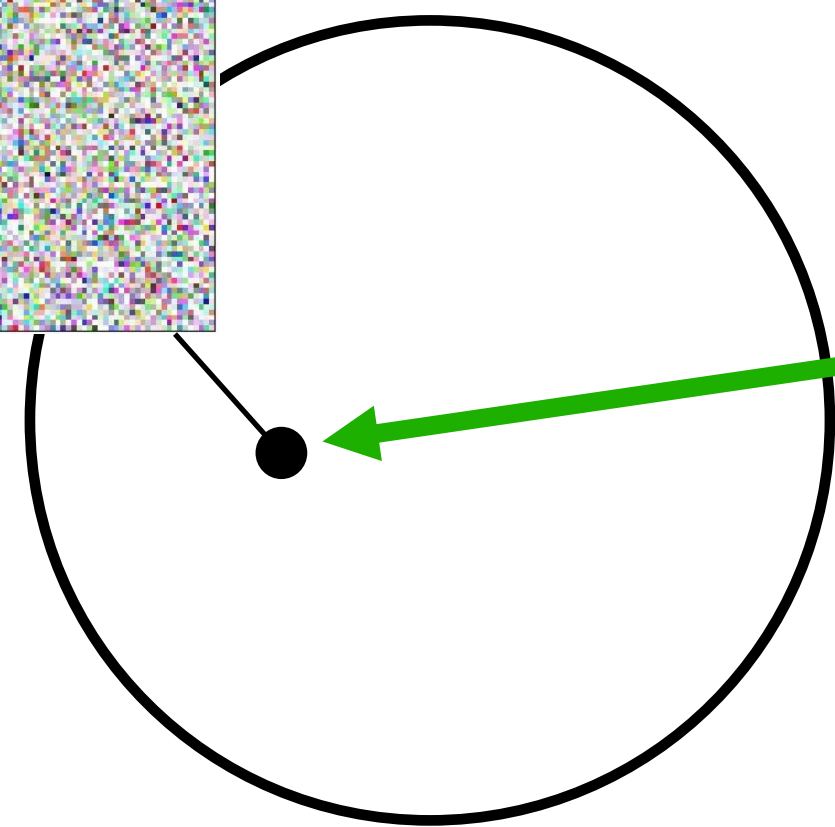
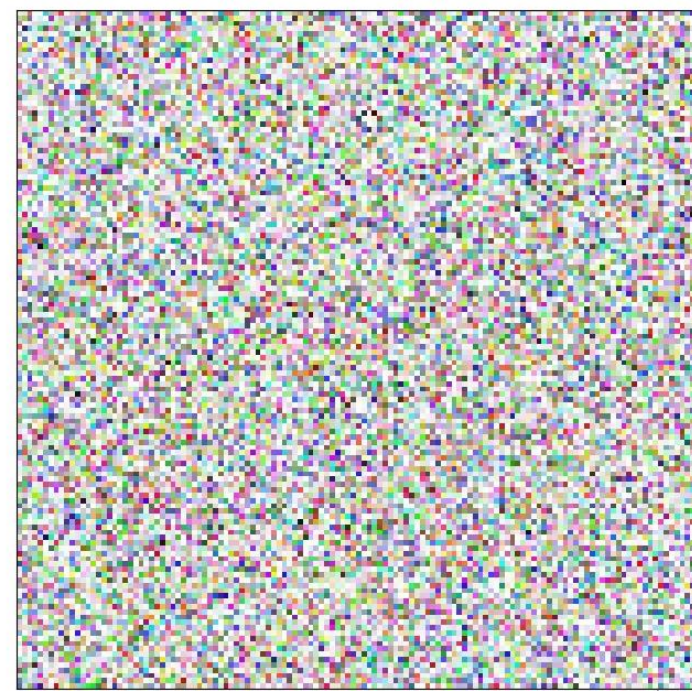
random
images

hard



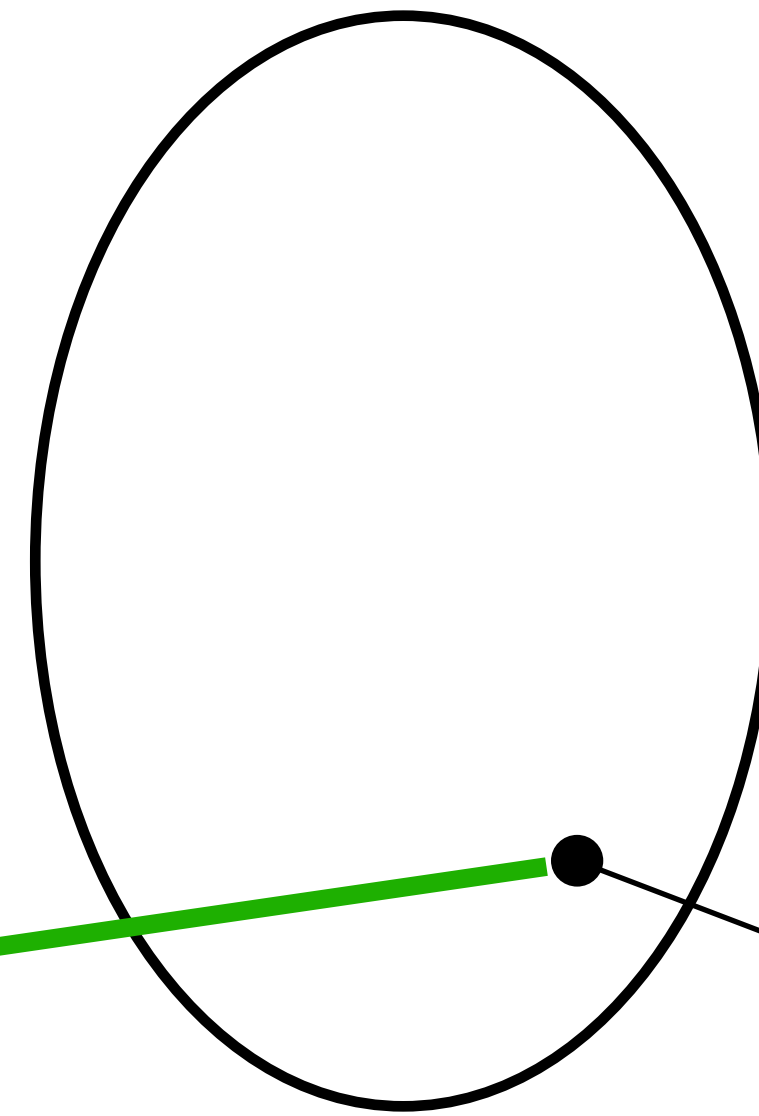
raspberry
images





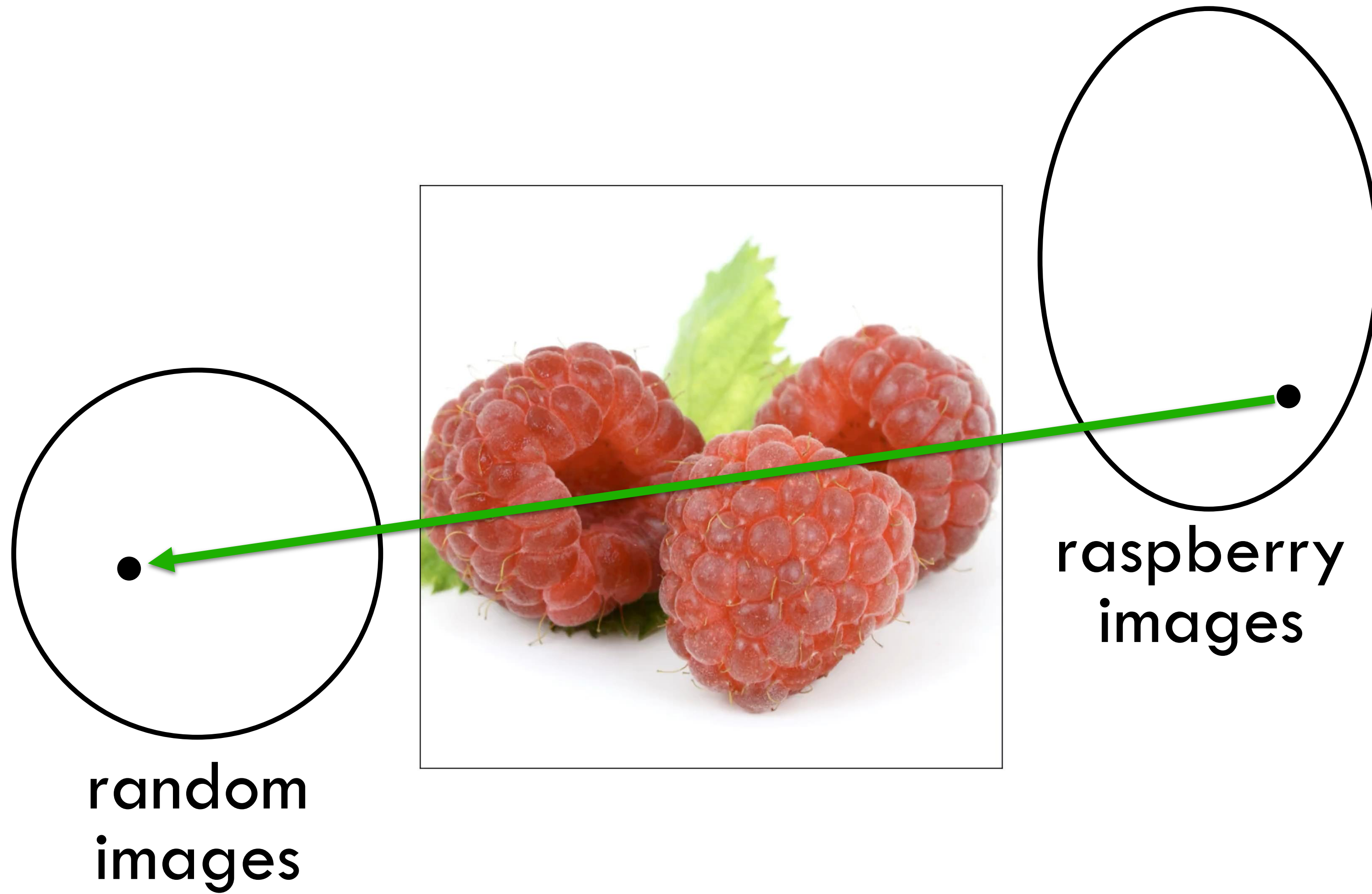
random
images

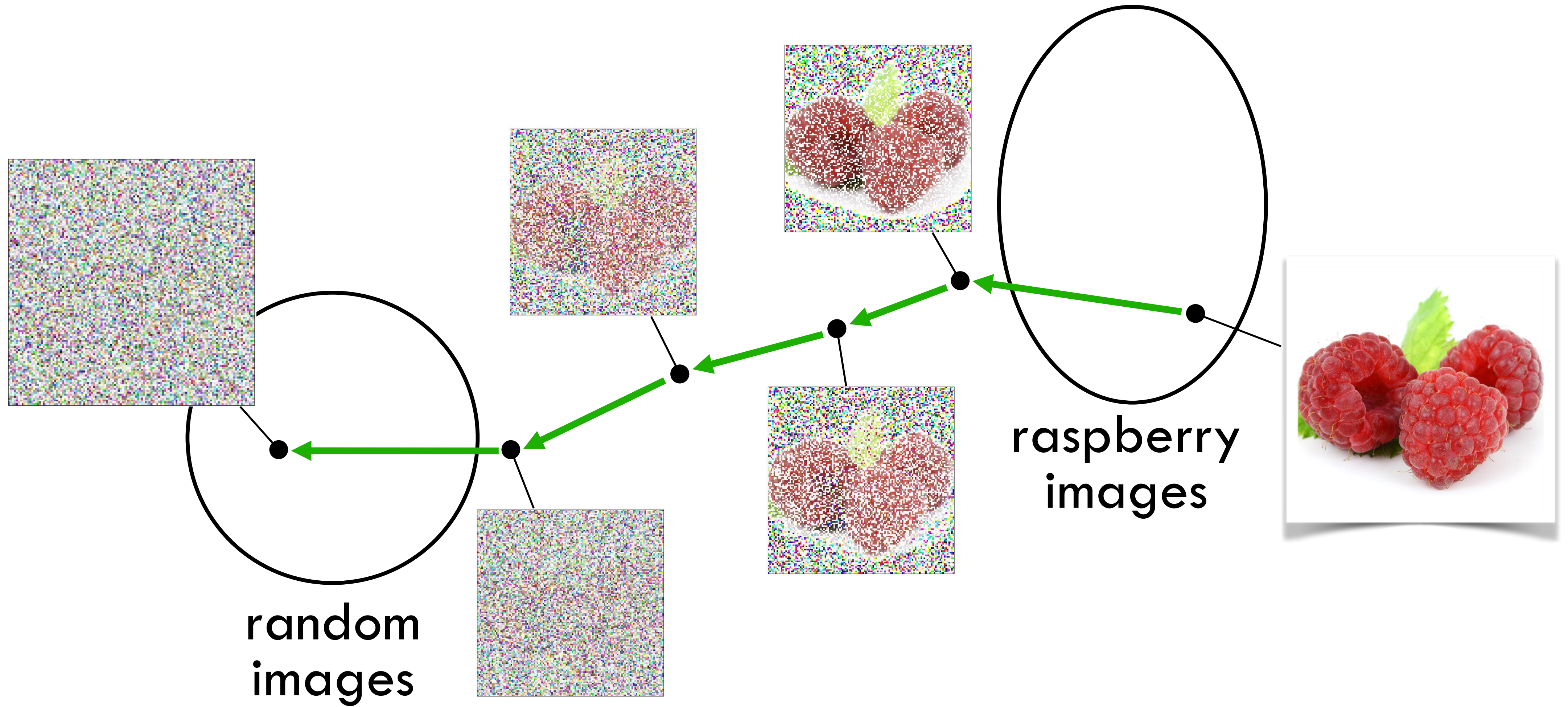
easy

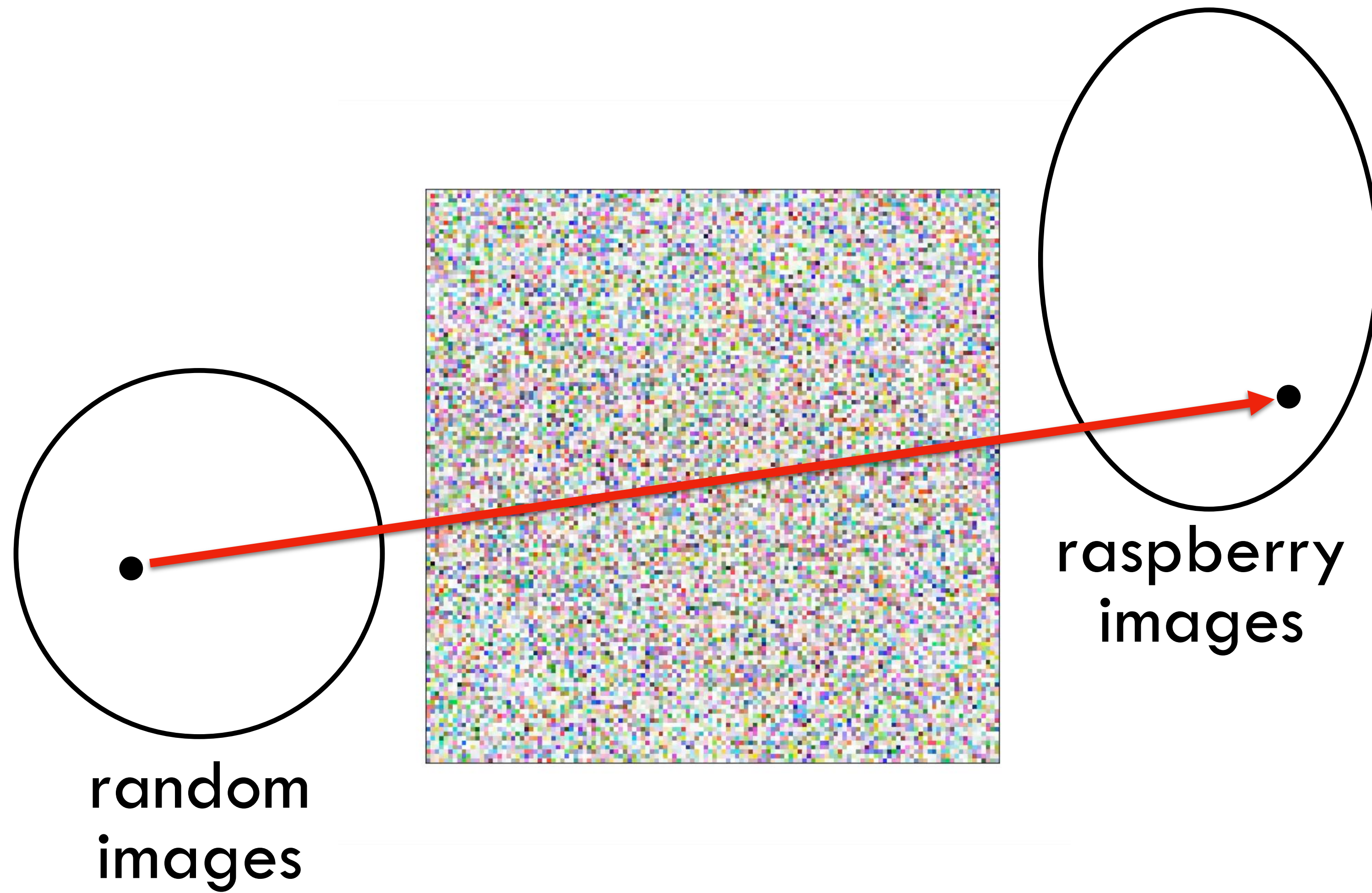


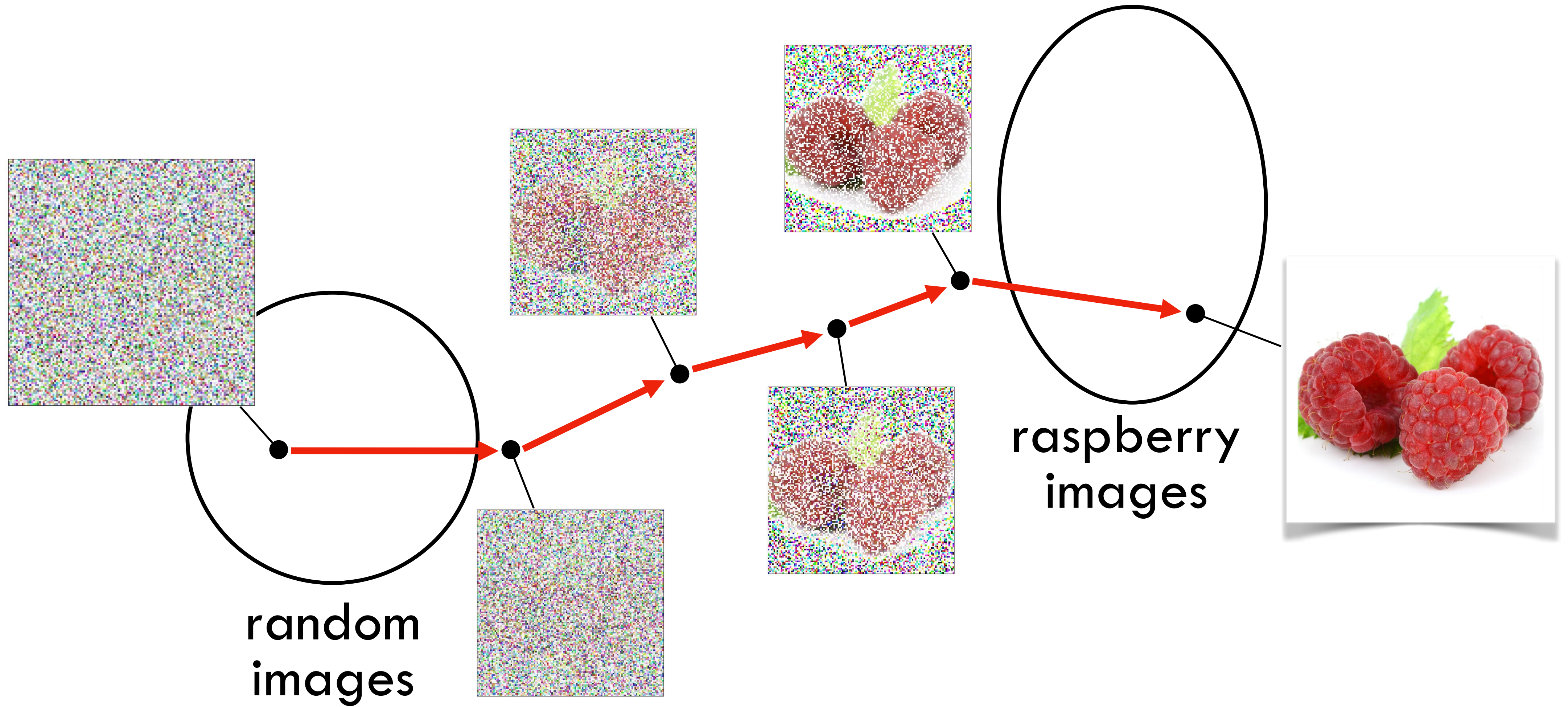
raspberry
images

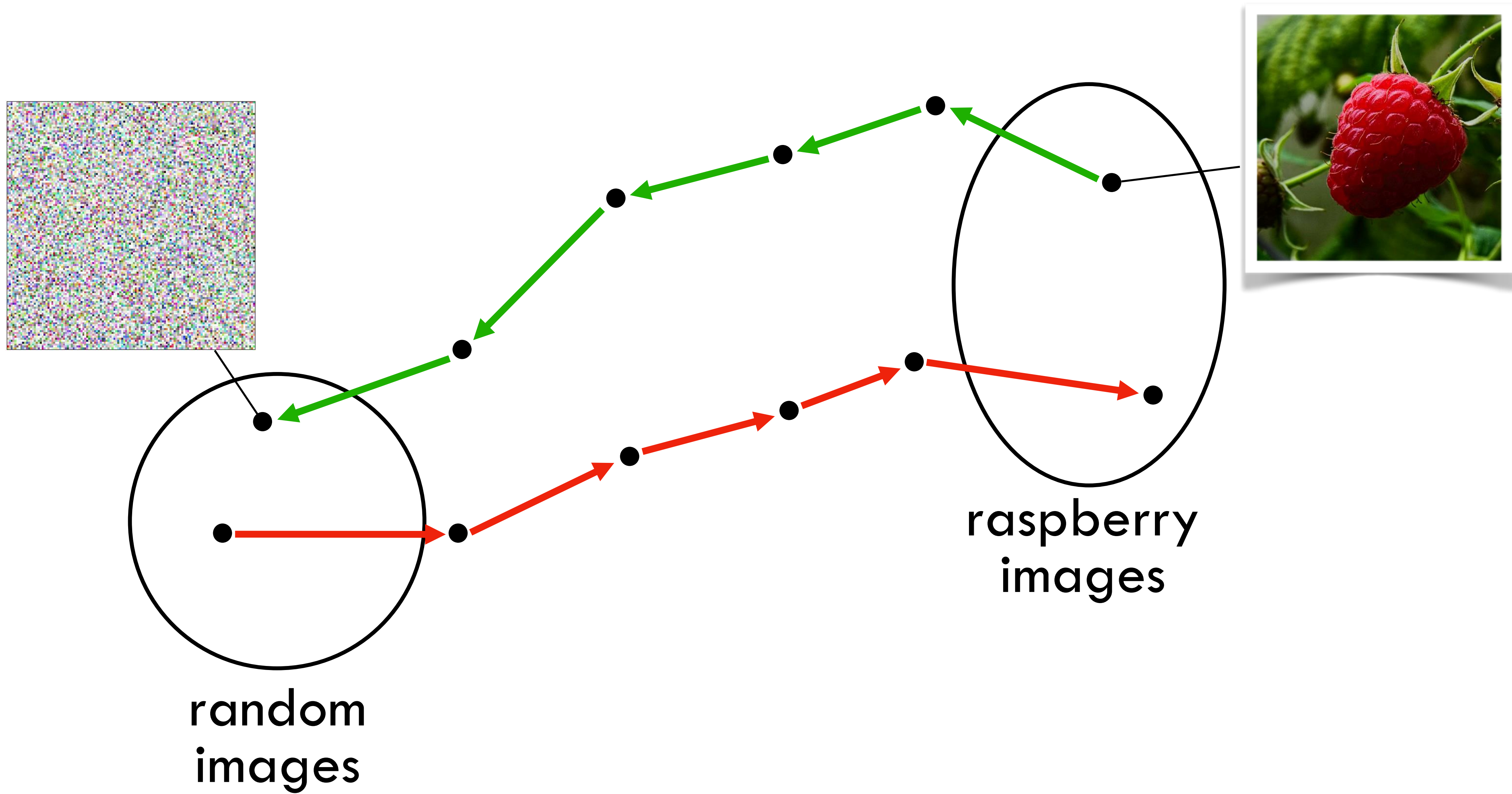


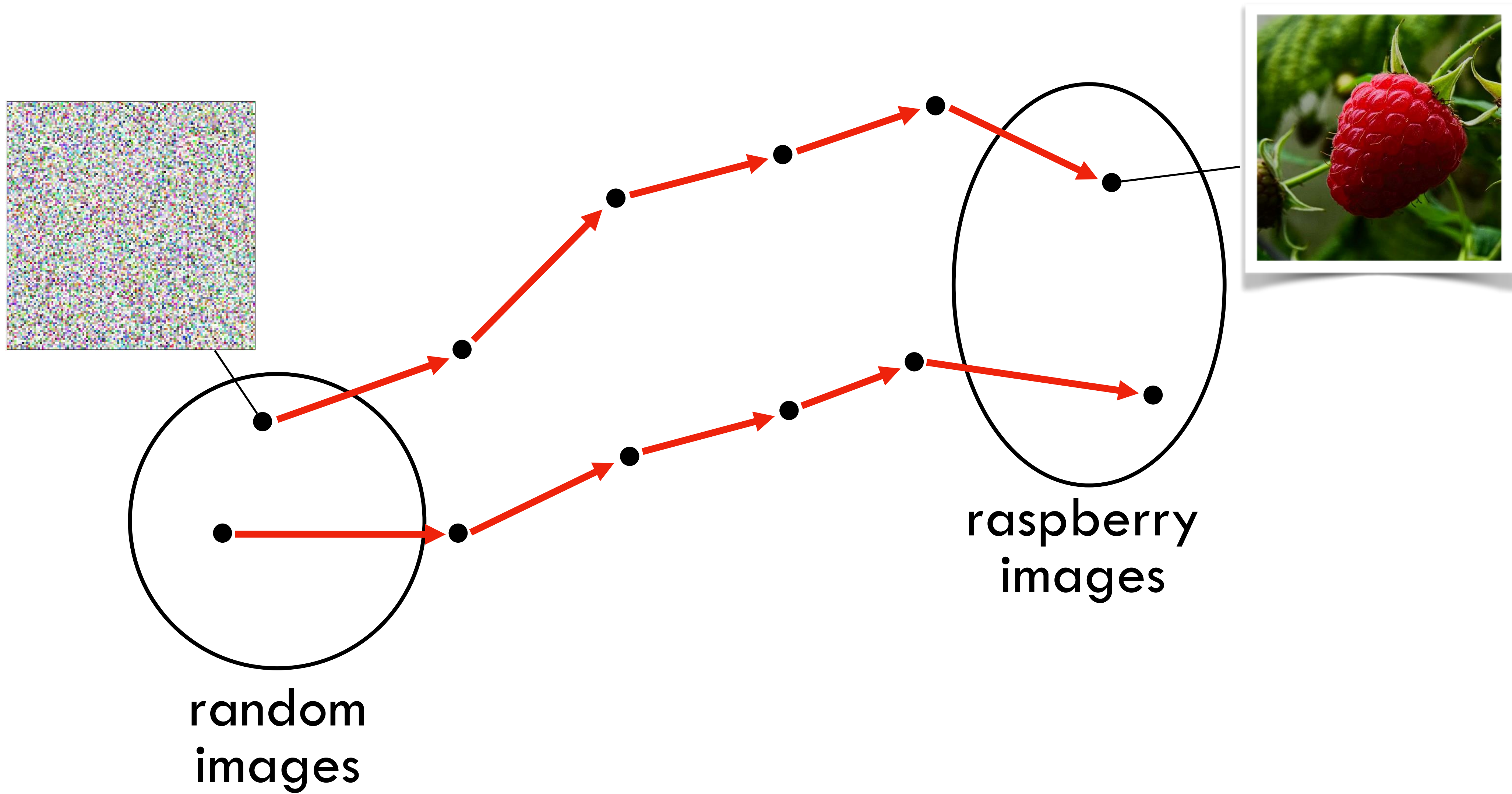


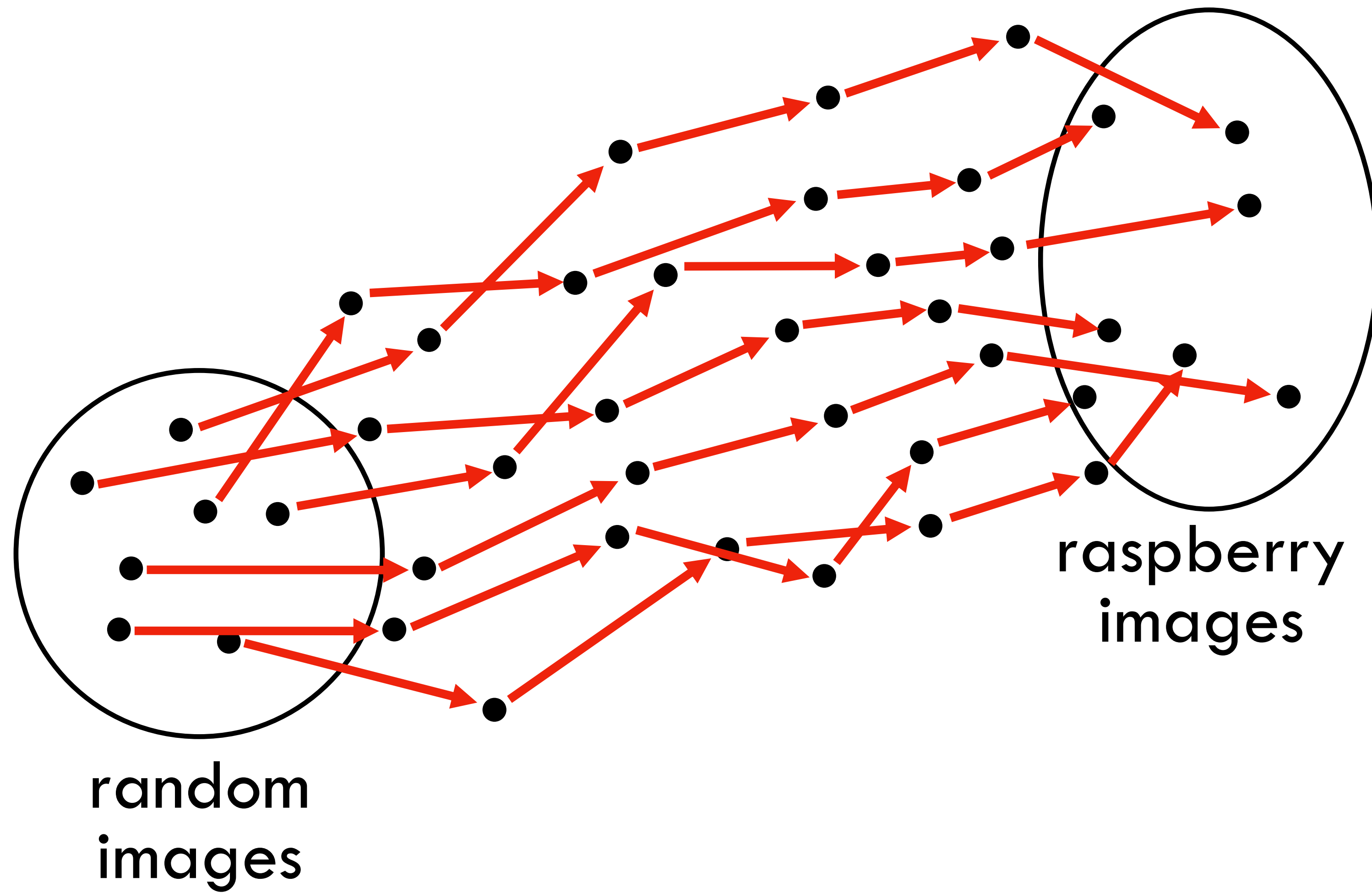


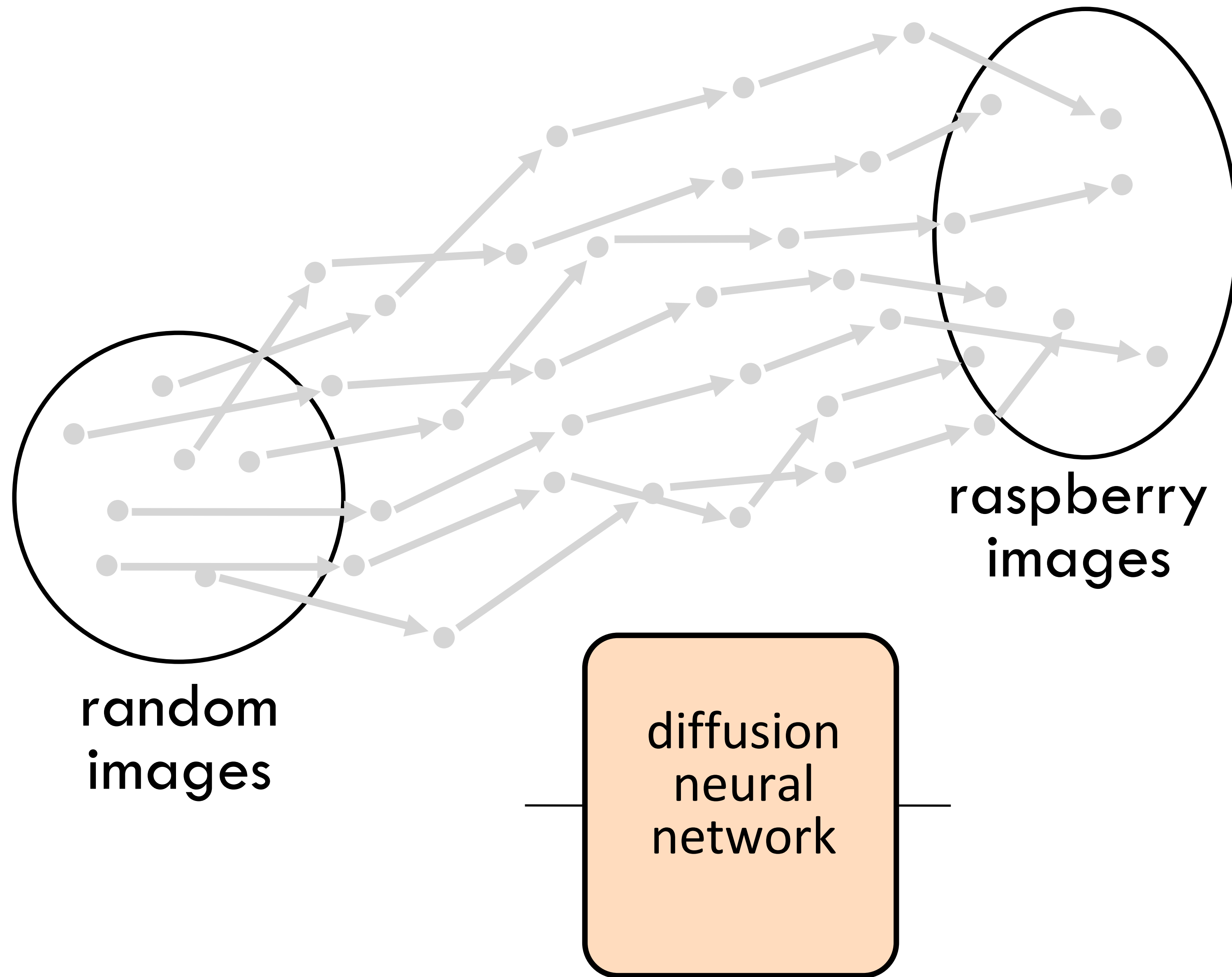


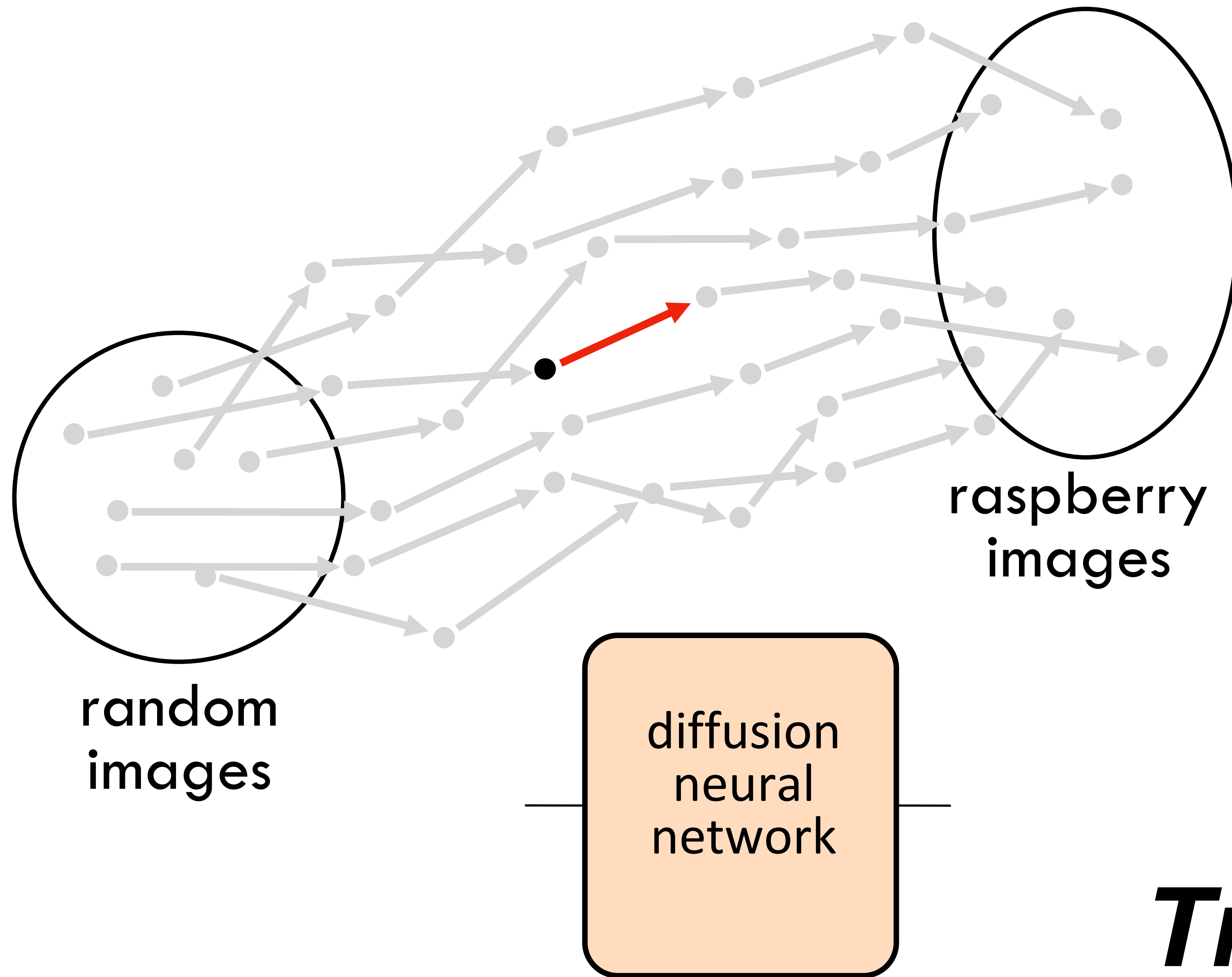






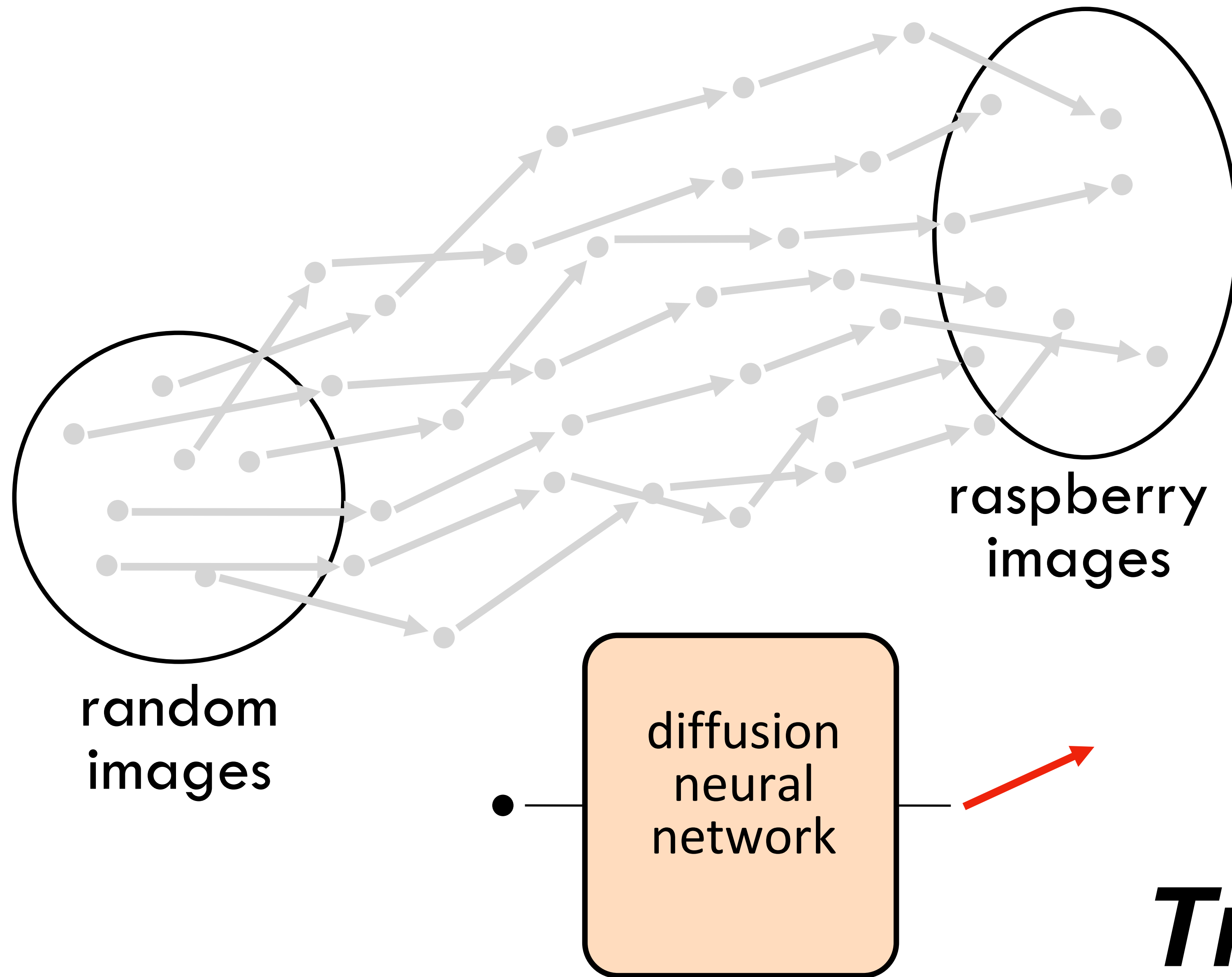






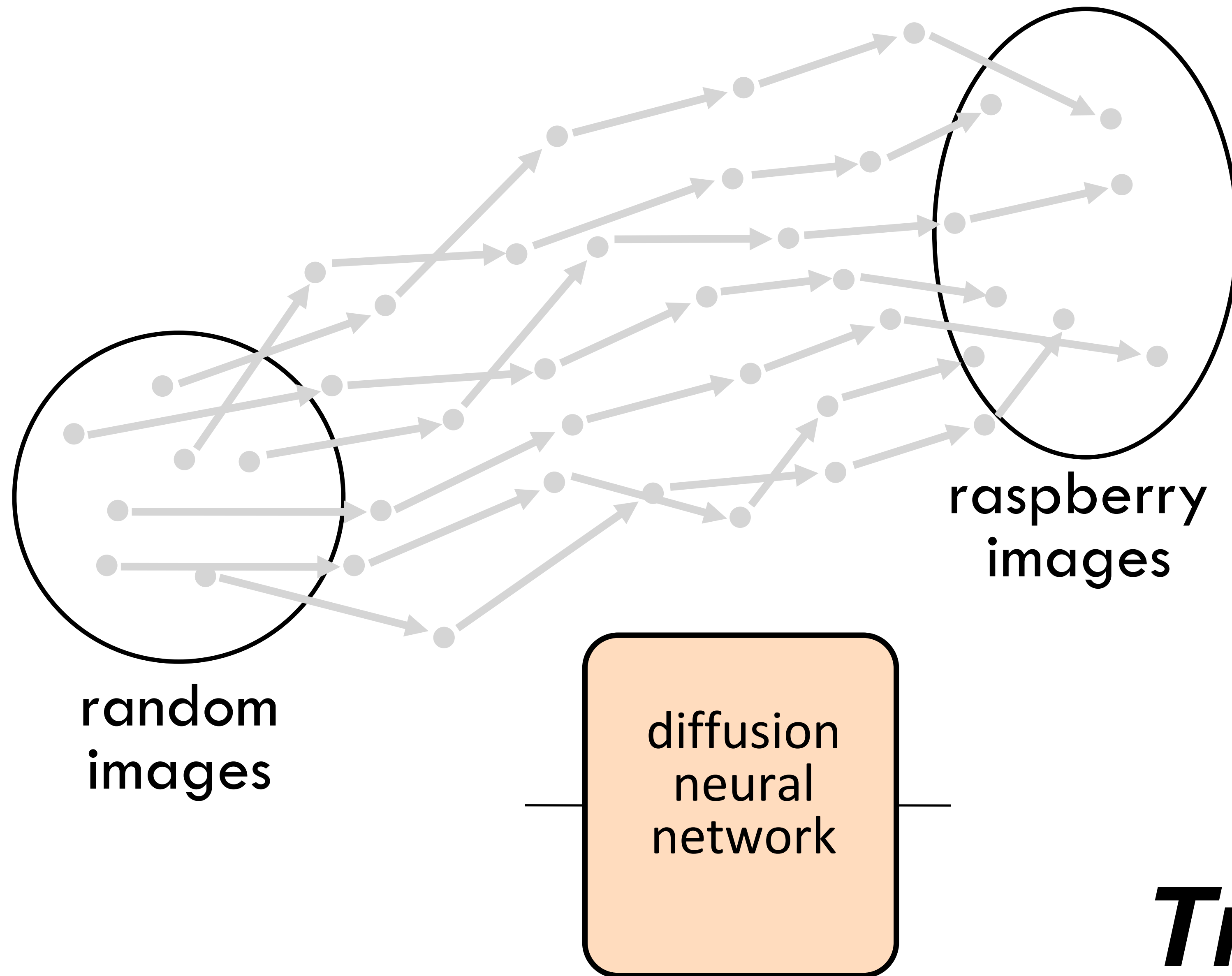
Training

slide from Steve Seitz's [video](#)



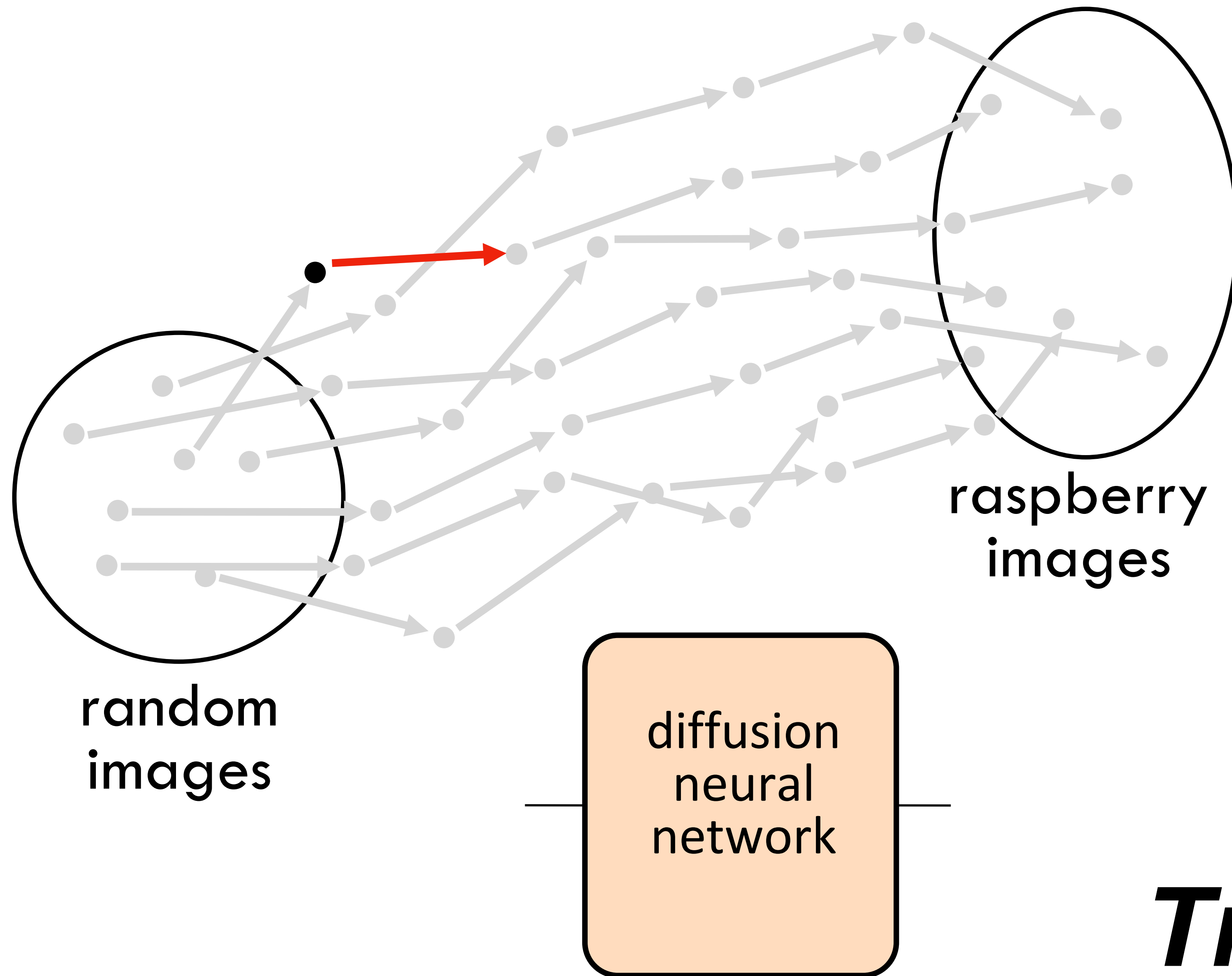
Training

slide from Steve Seitz's [video](#)



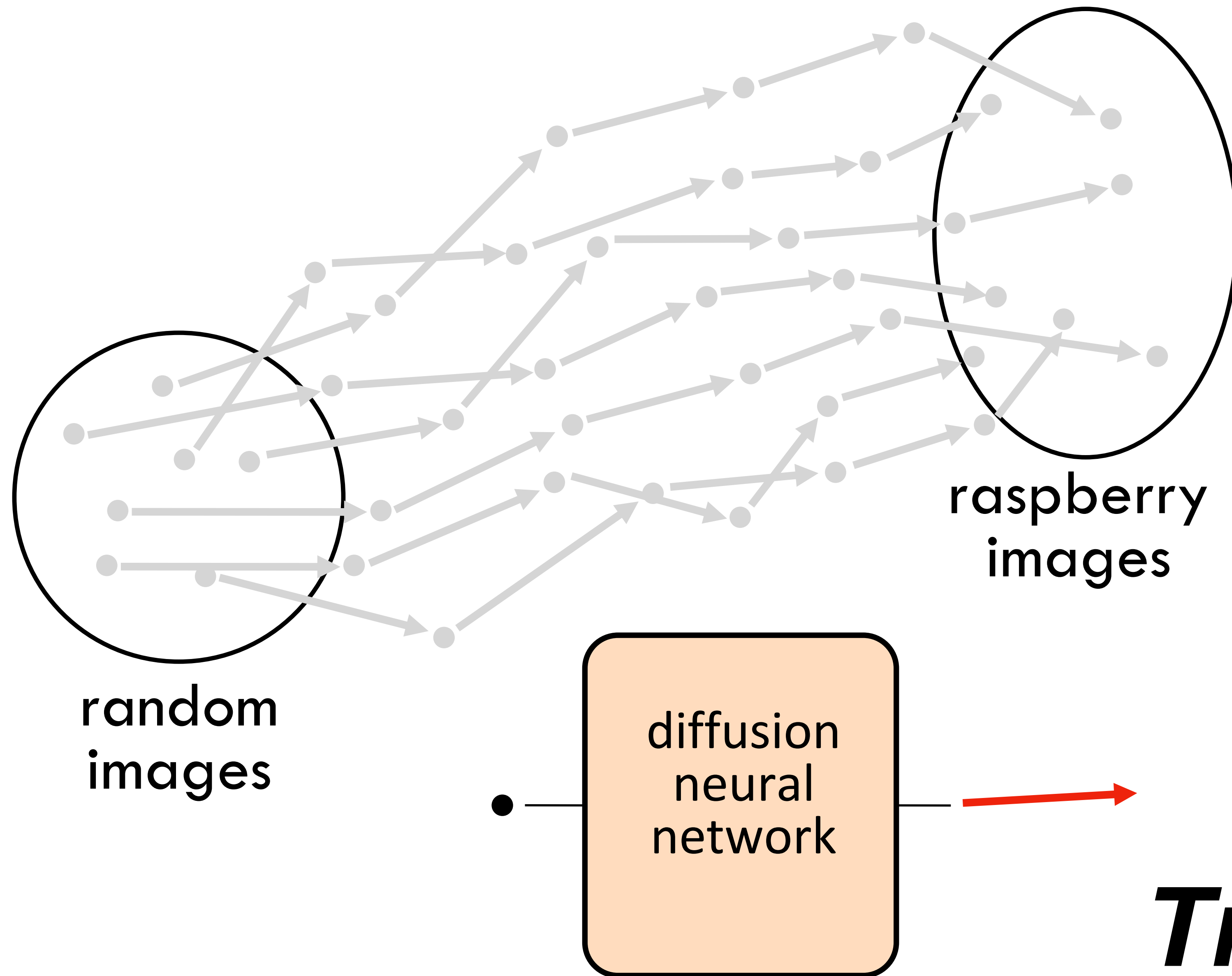
Training

slide from Steve Seitz's [video](#)



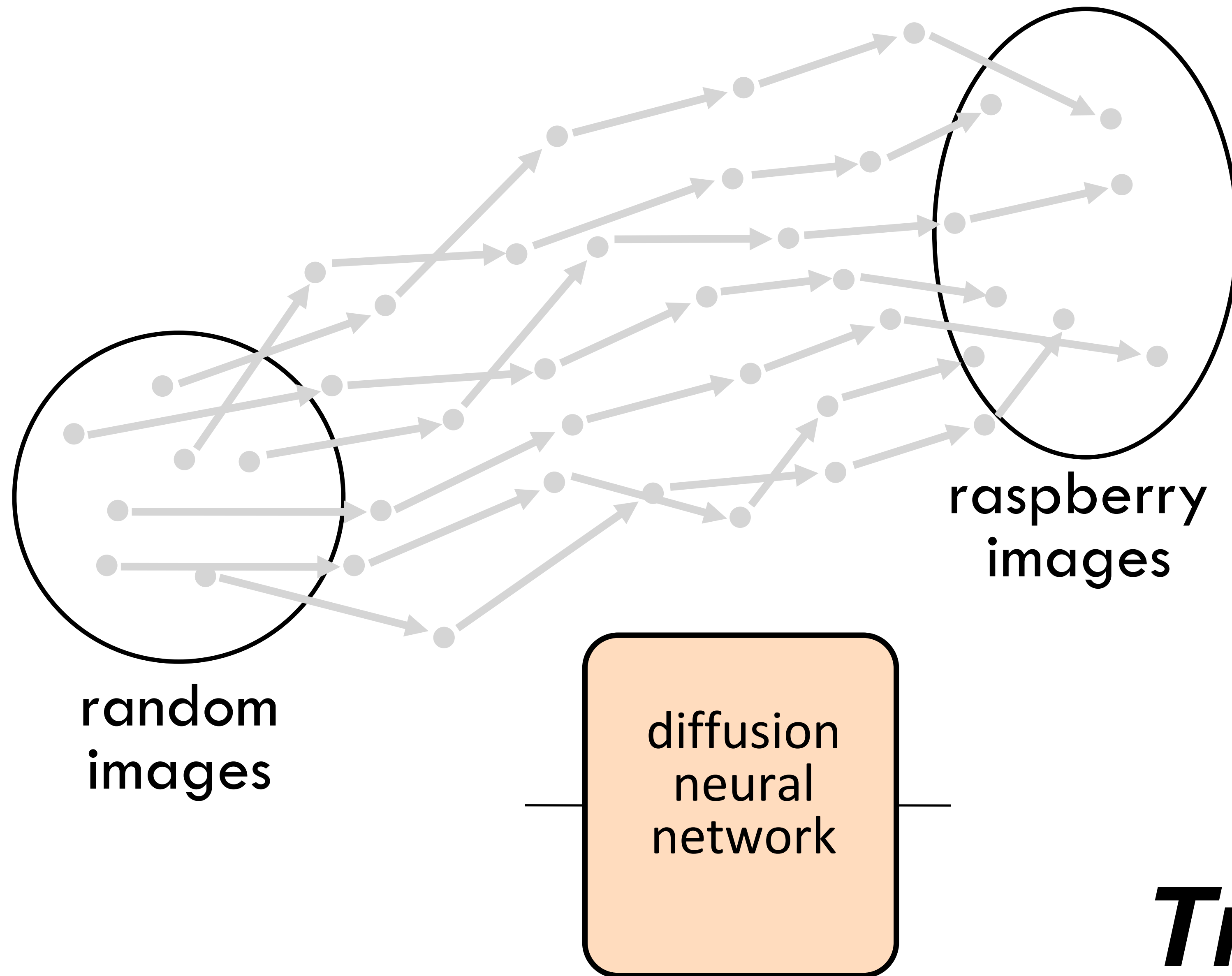
Training

slide from Steve Seitz's [video](#)



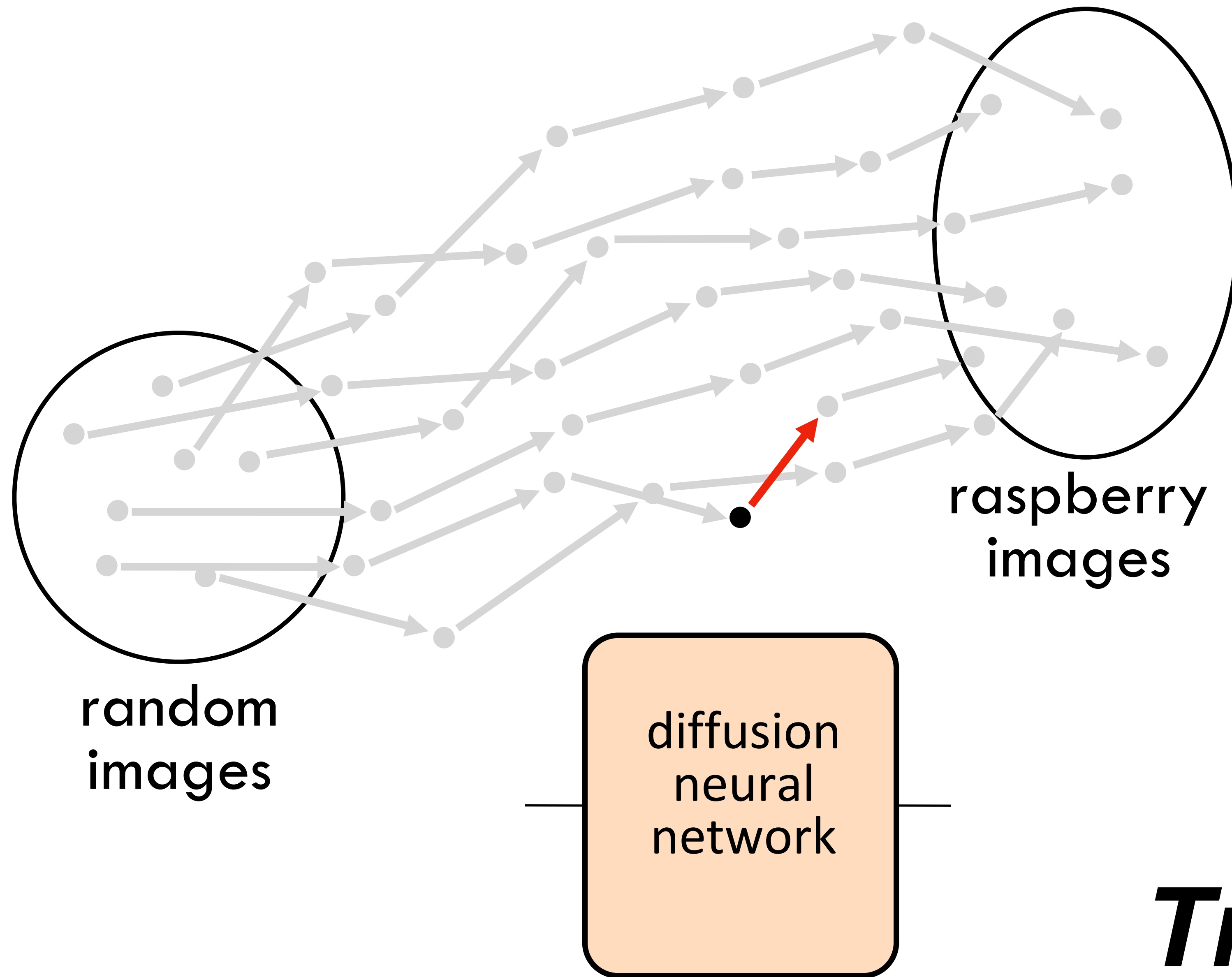
Training

slide from Steve Seitz's [video](#)



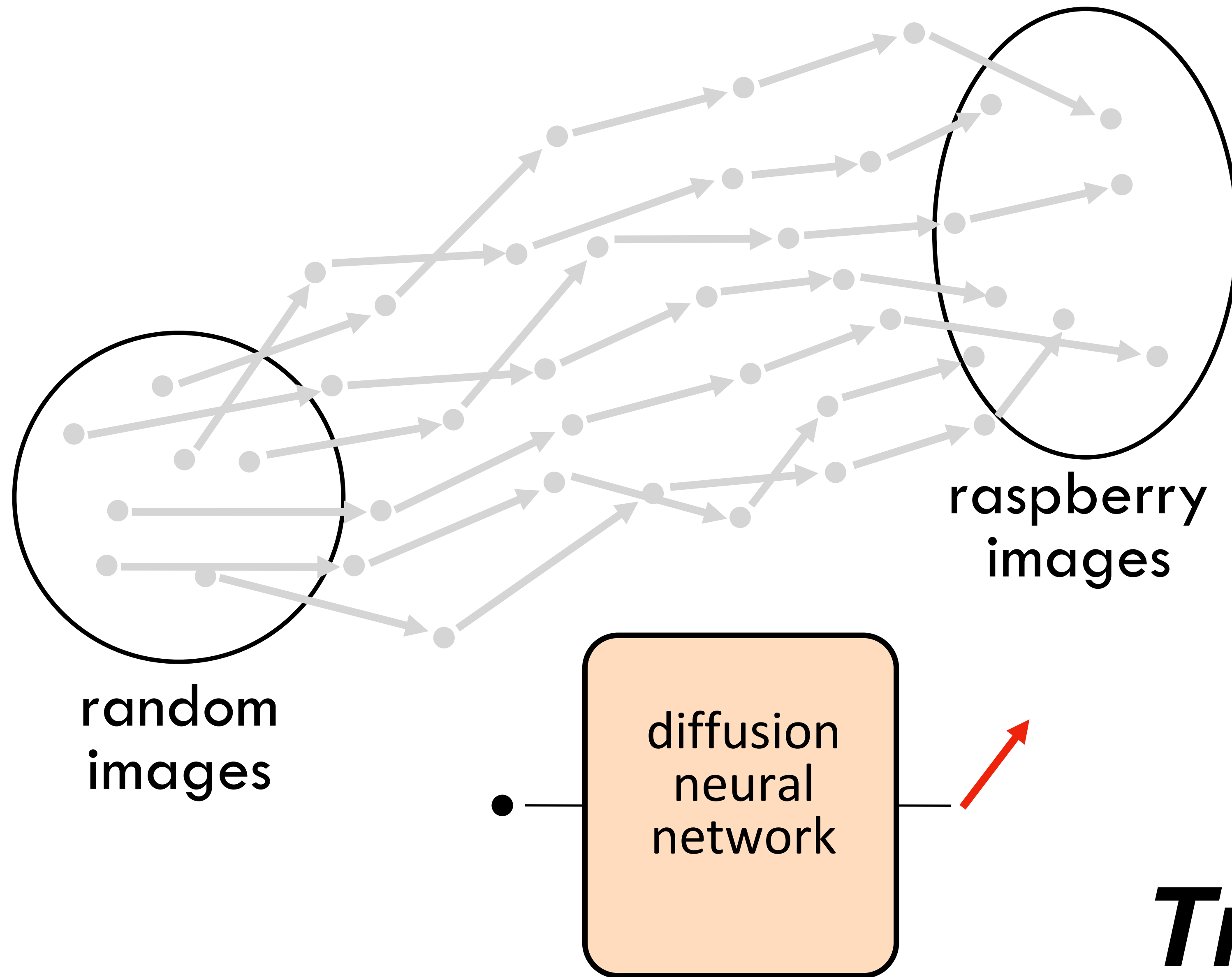
Training

slide from Steve Seitz's [video](#)



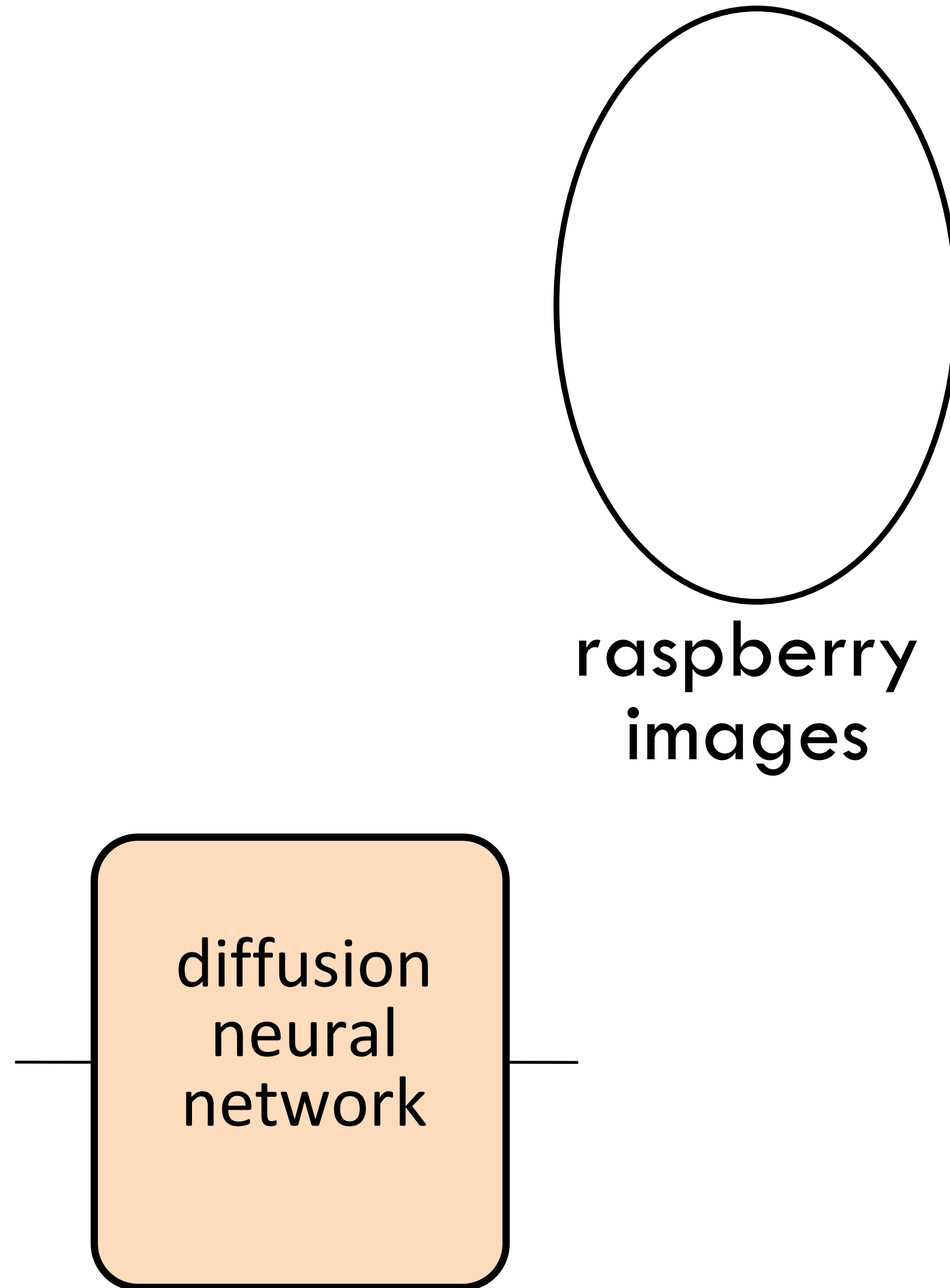
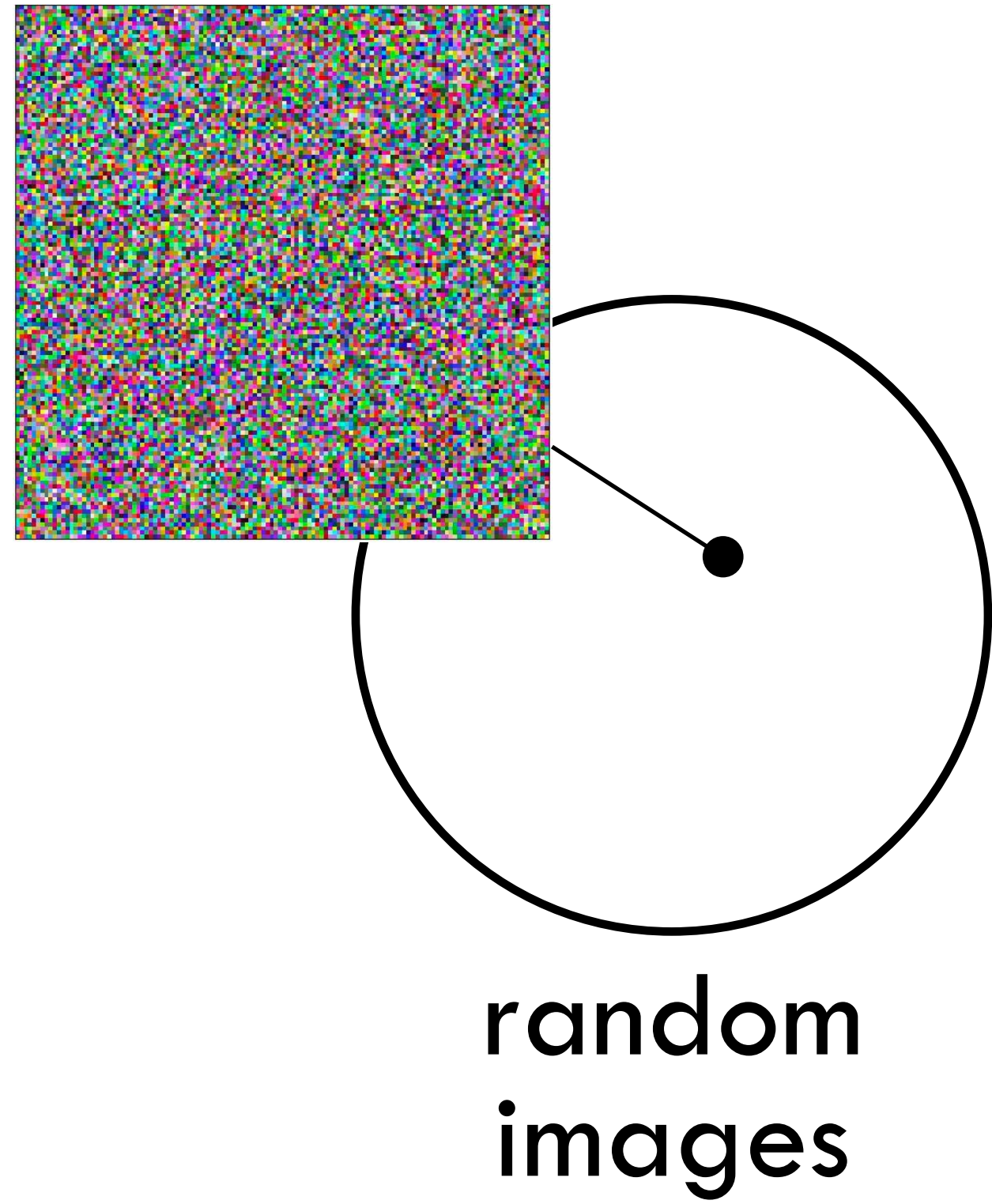
Training

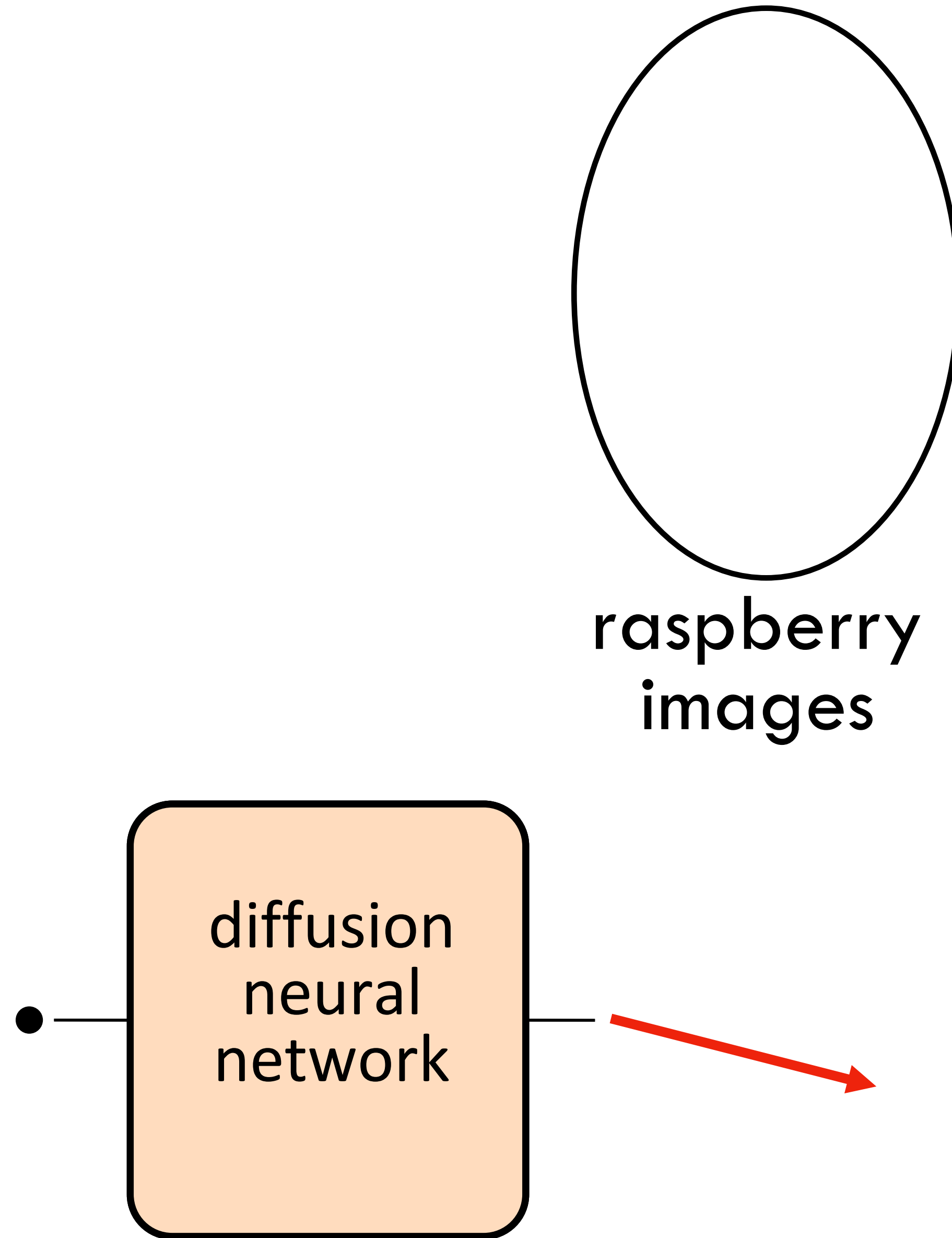
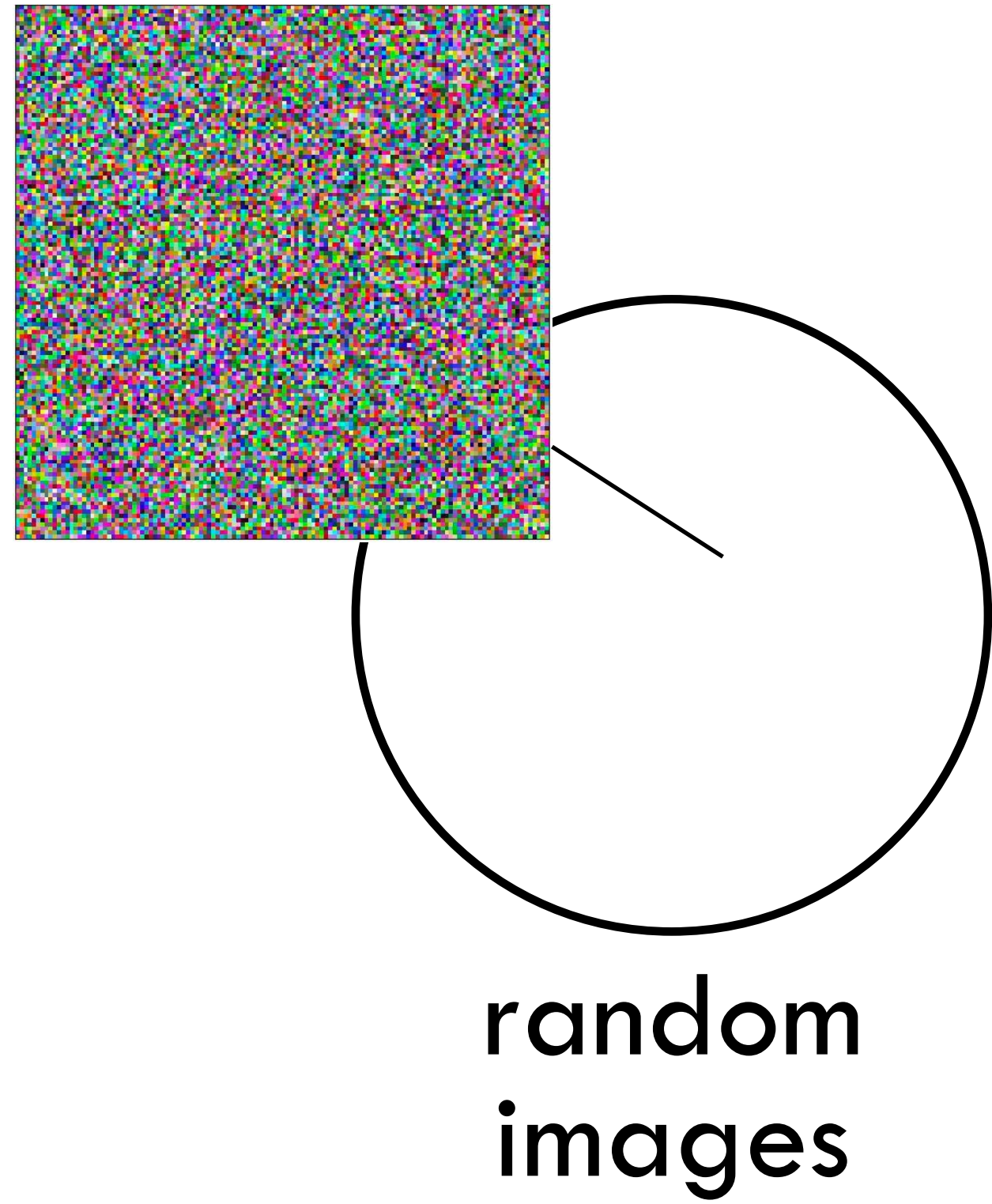
slide from Steve Seitz's [video](#)

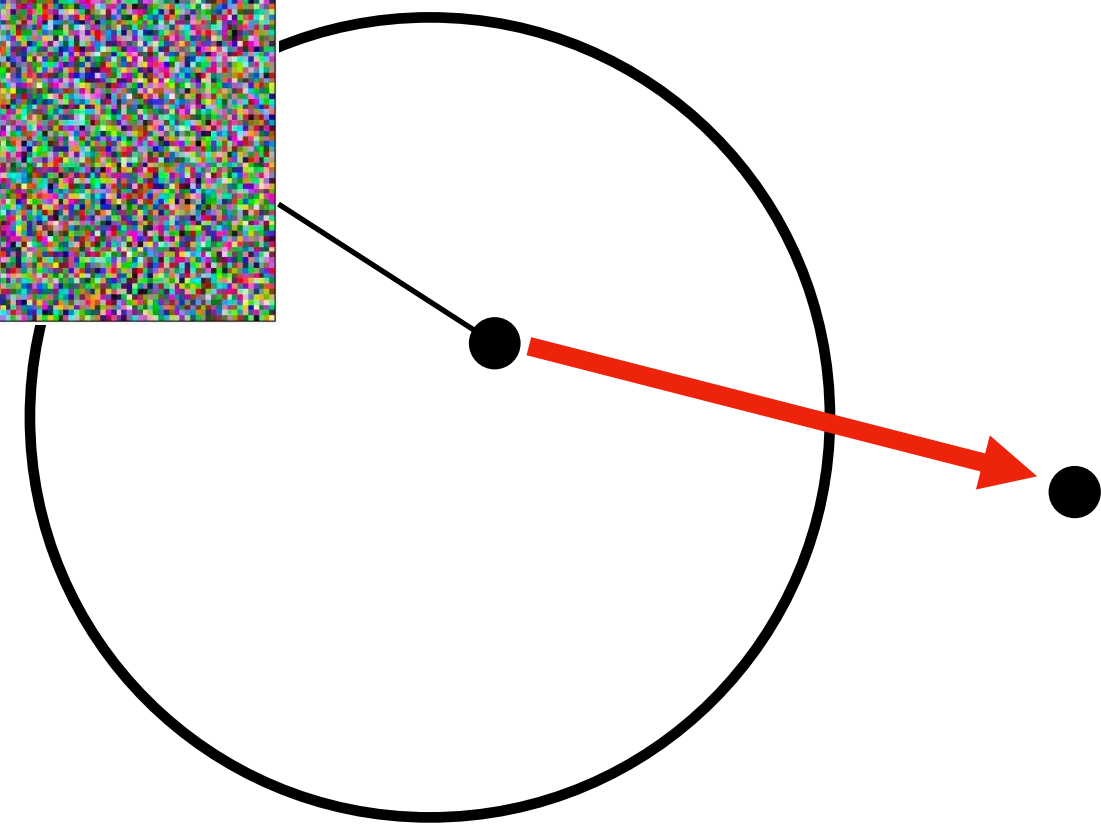
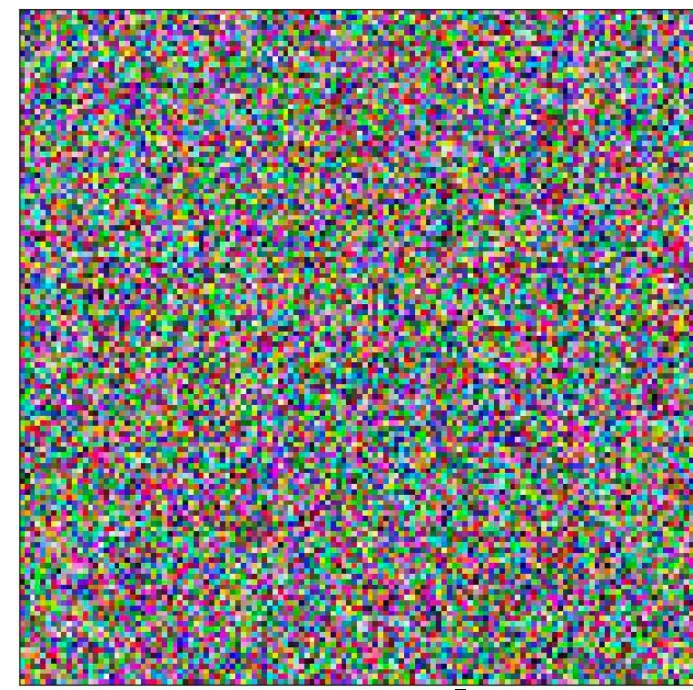


Training

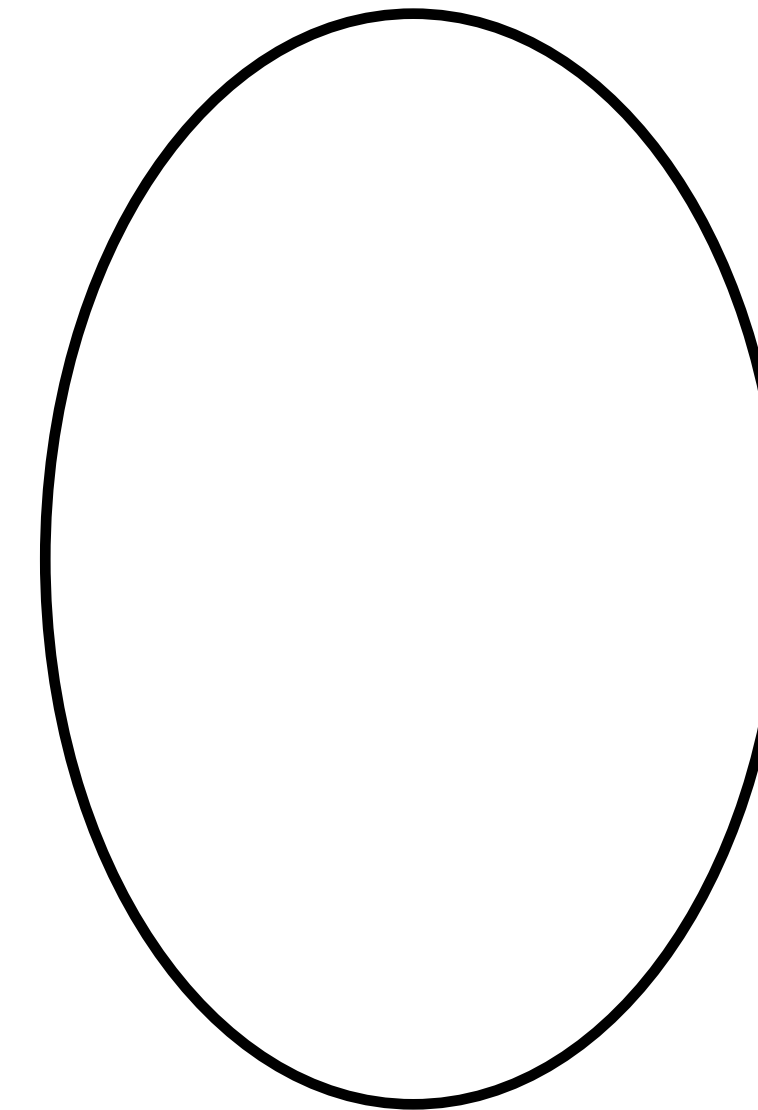
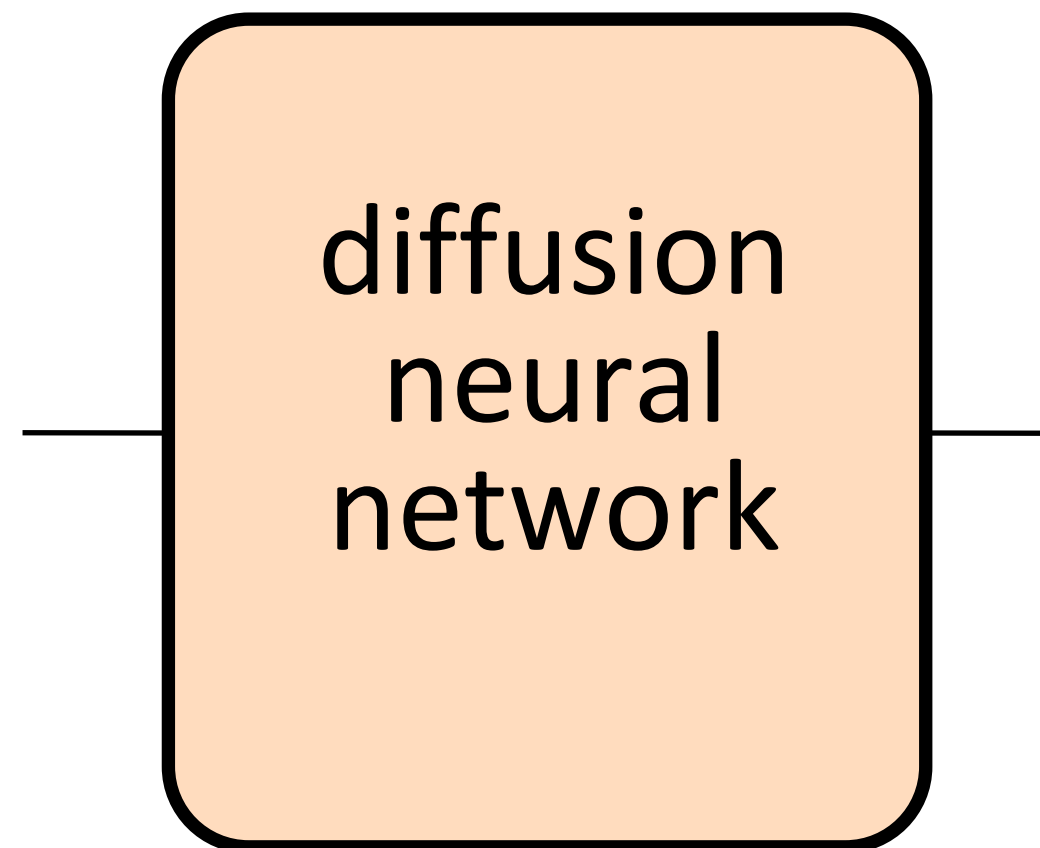
slide from Steve Seitz's [video](#)



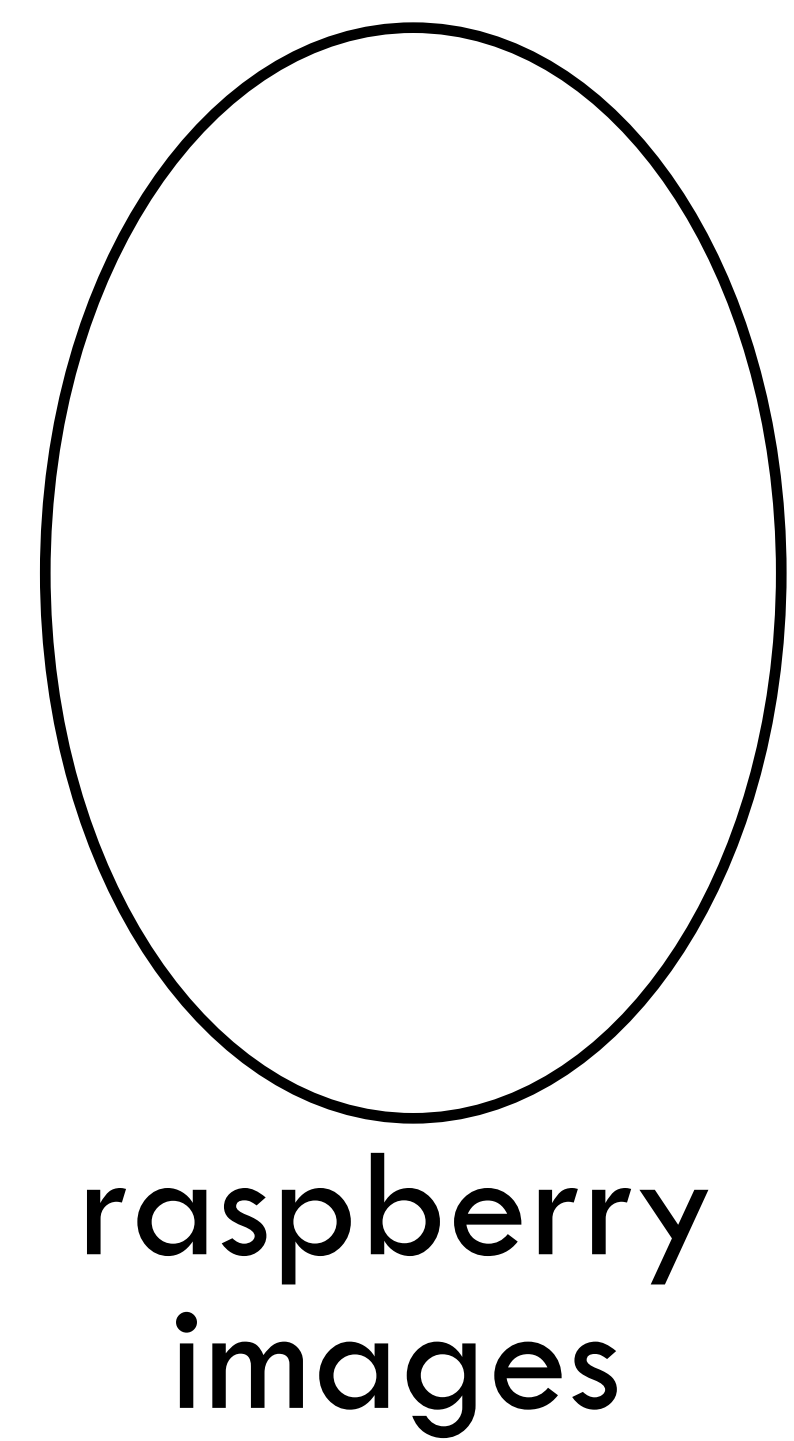
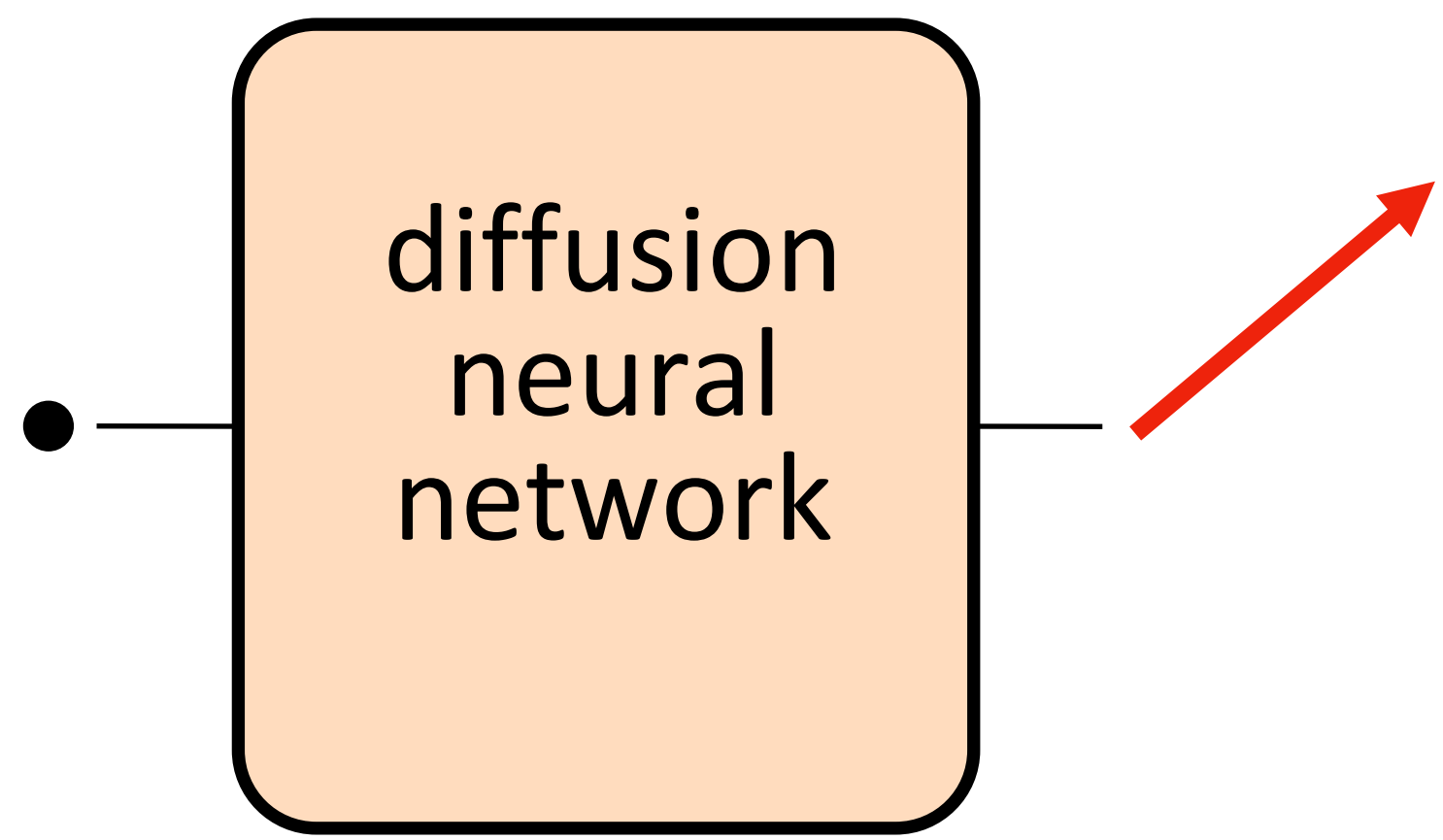
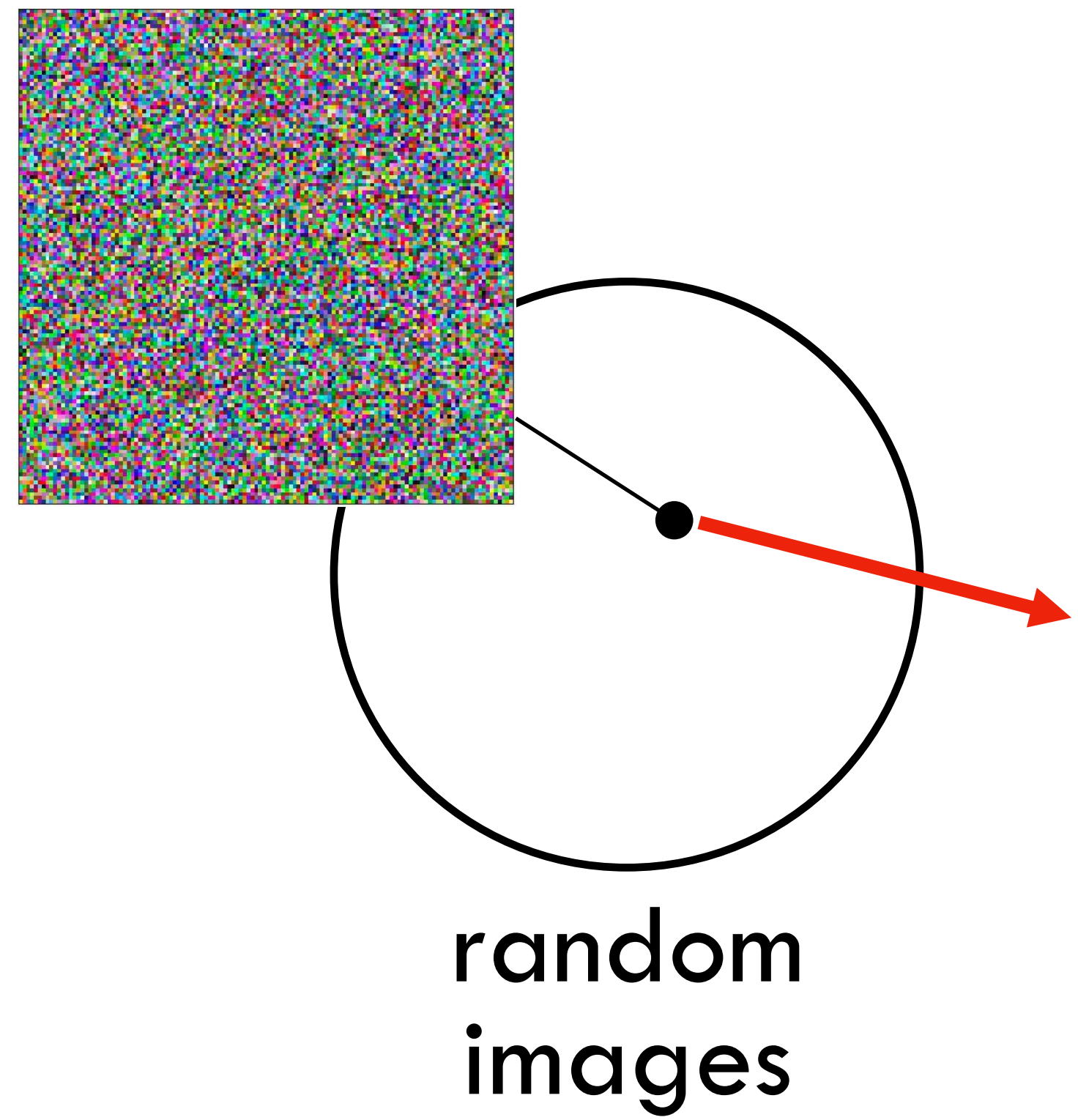


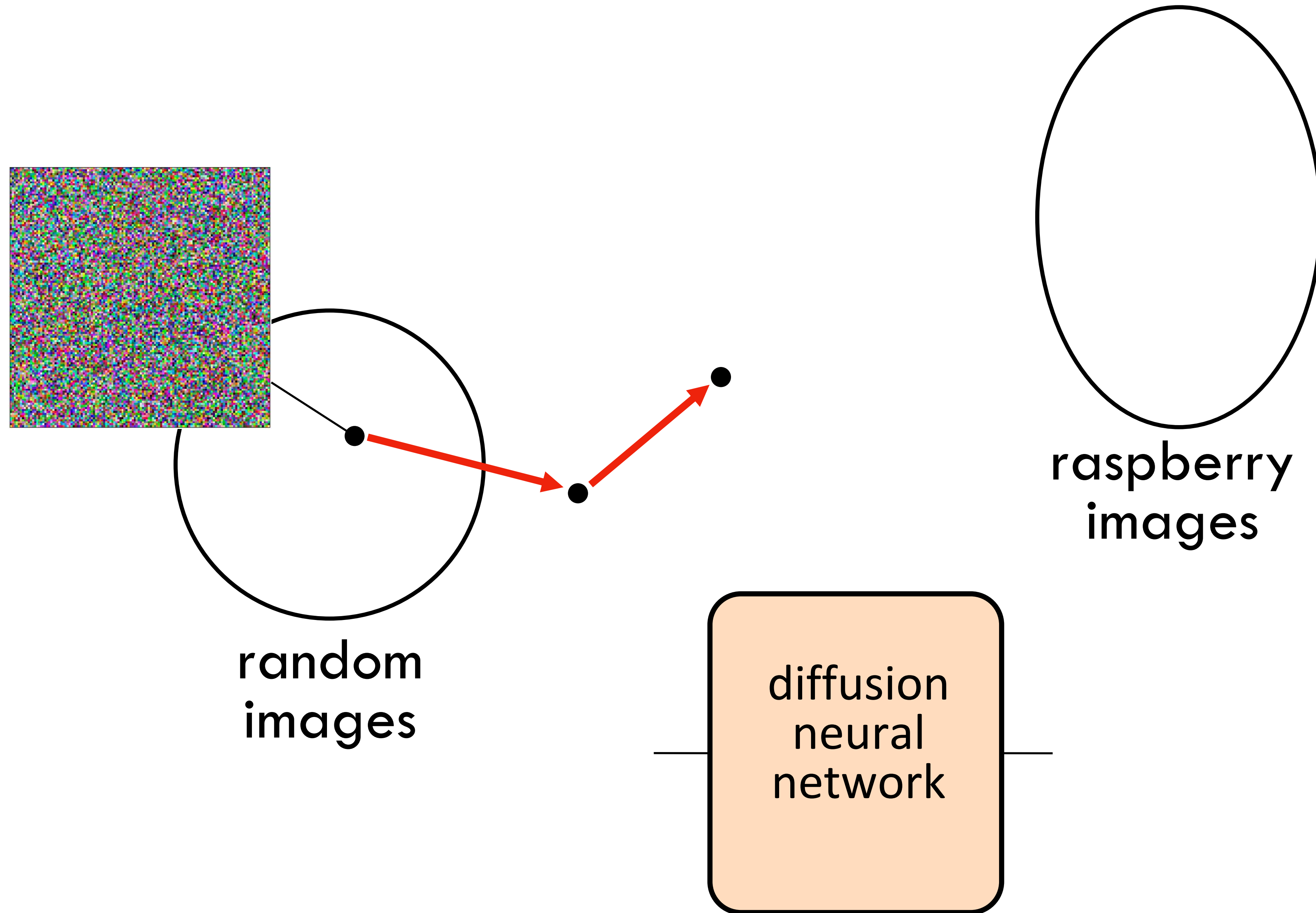


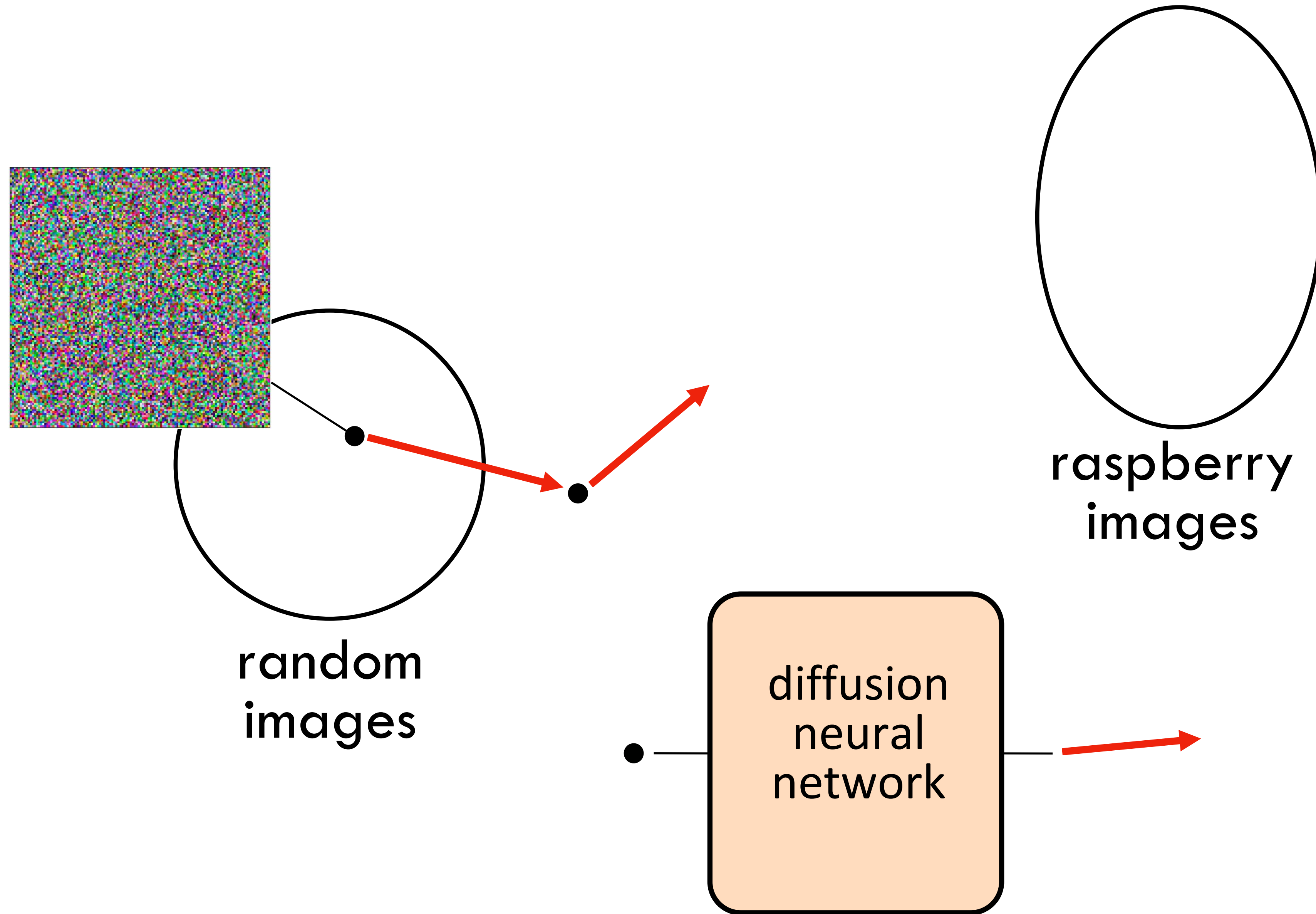
random
images

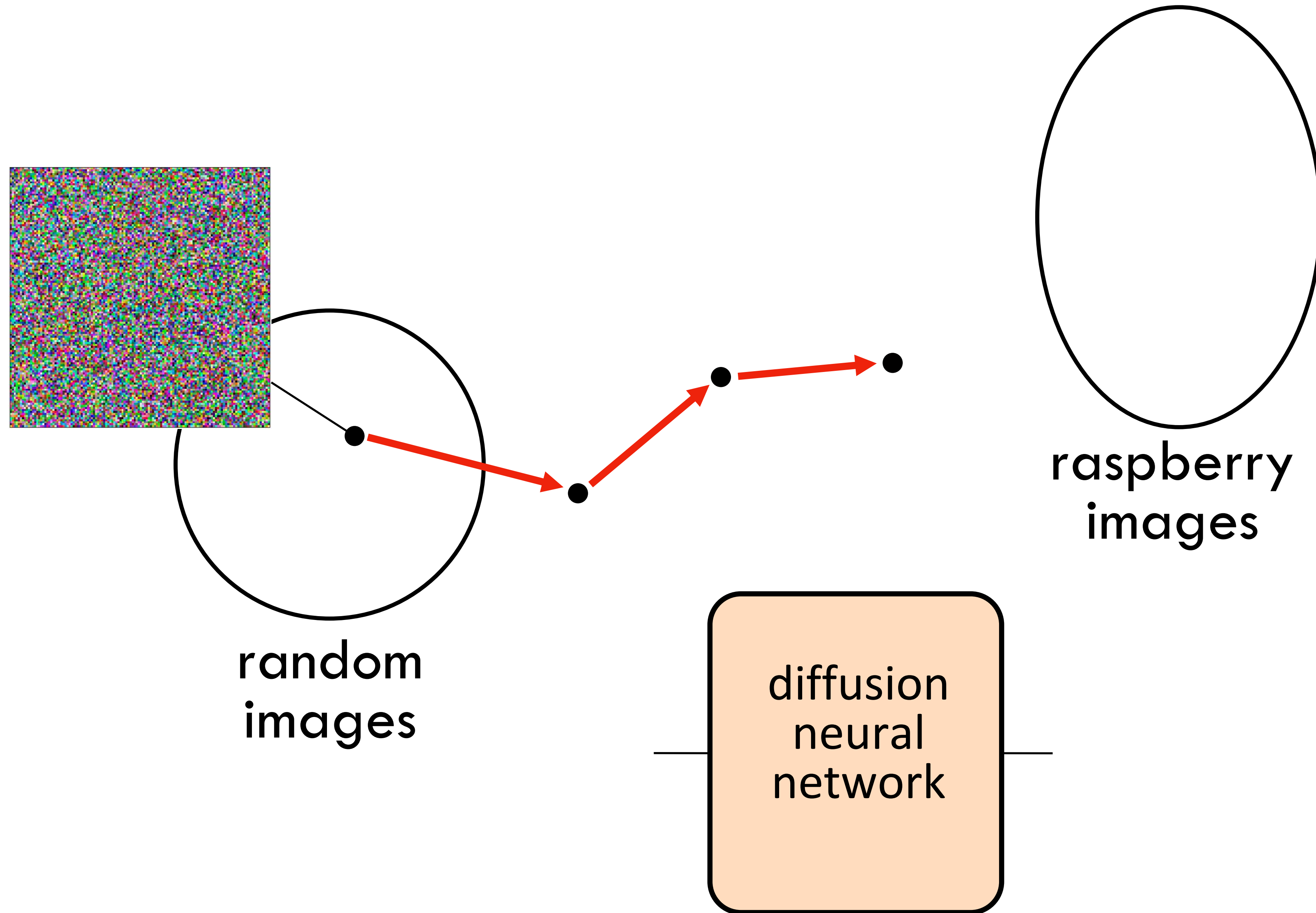


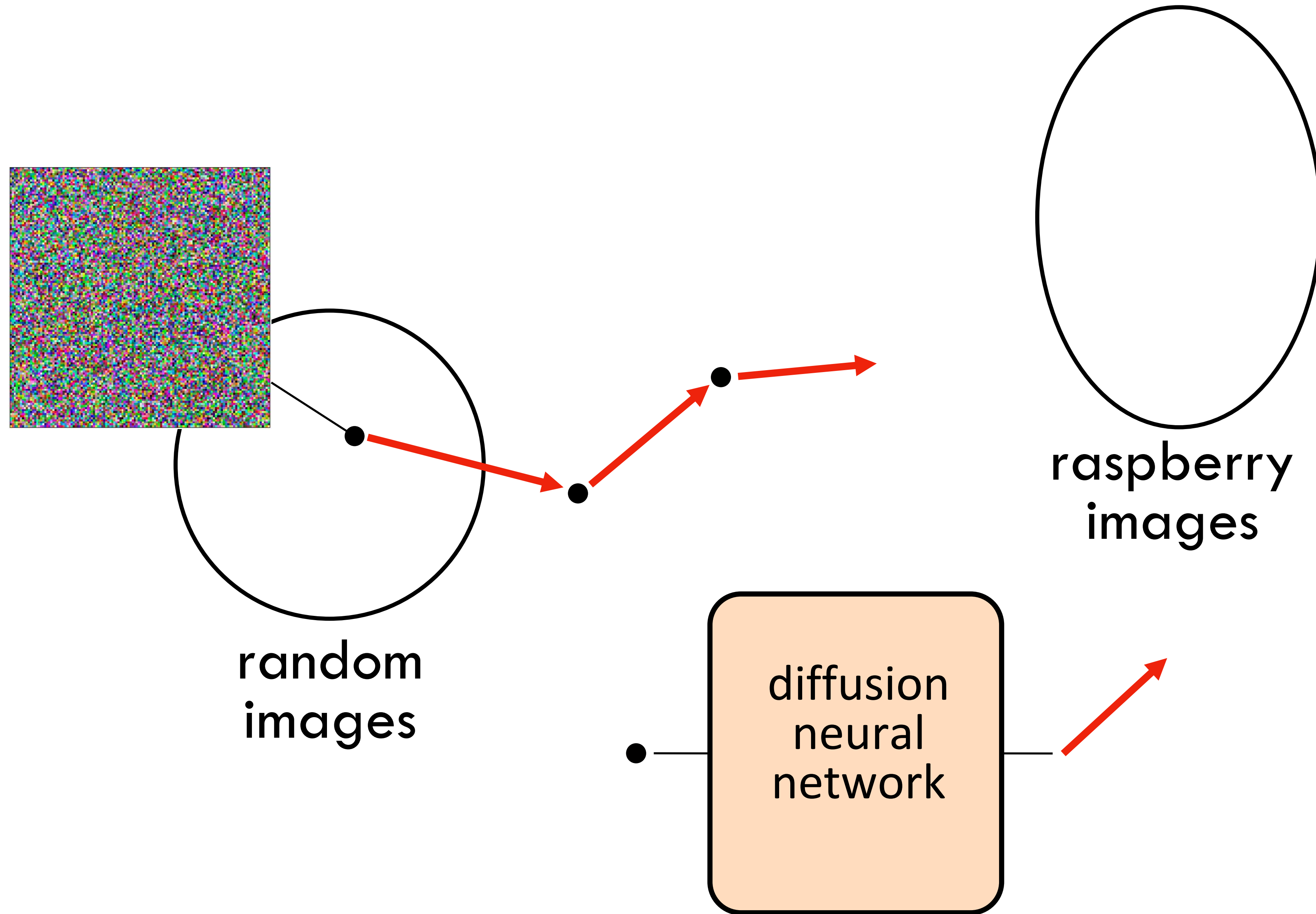
raspberry
images







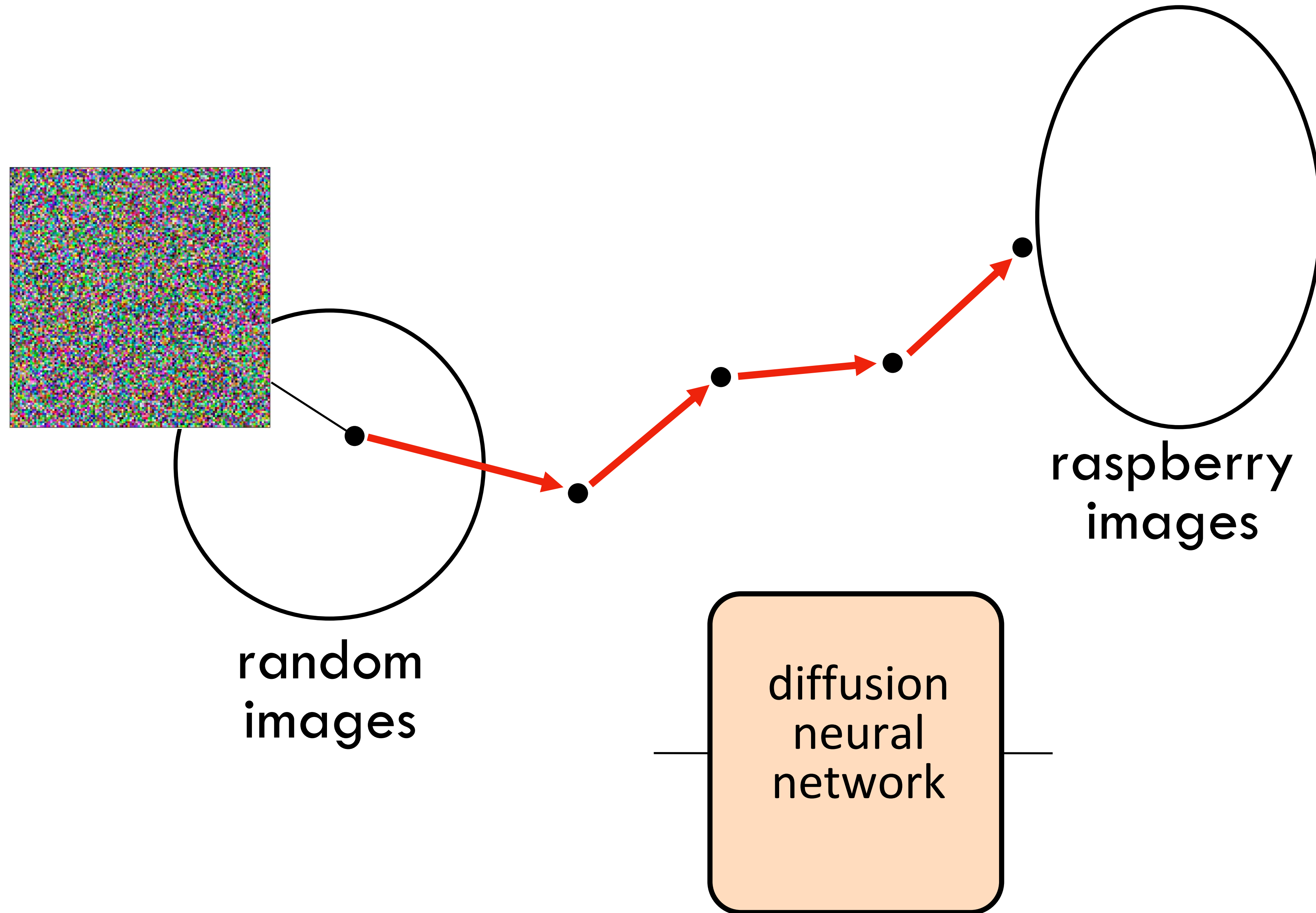


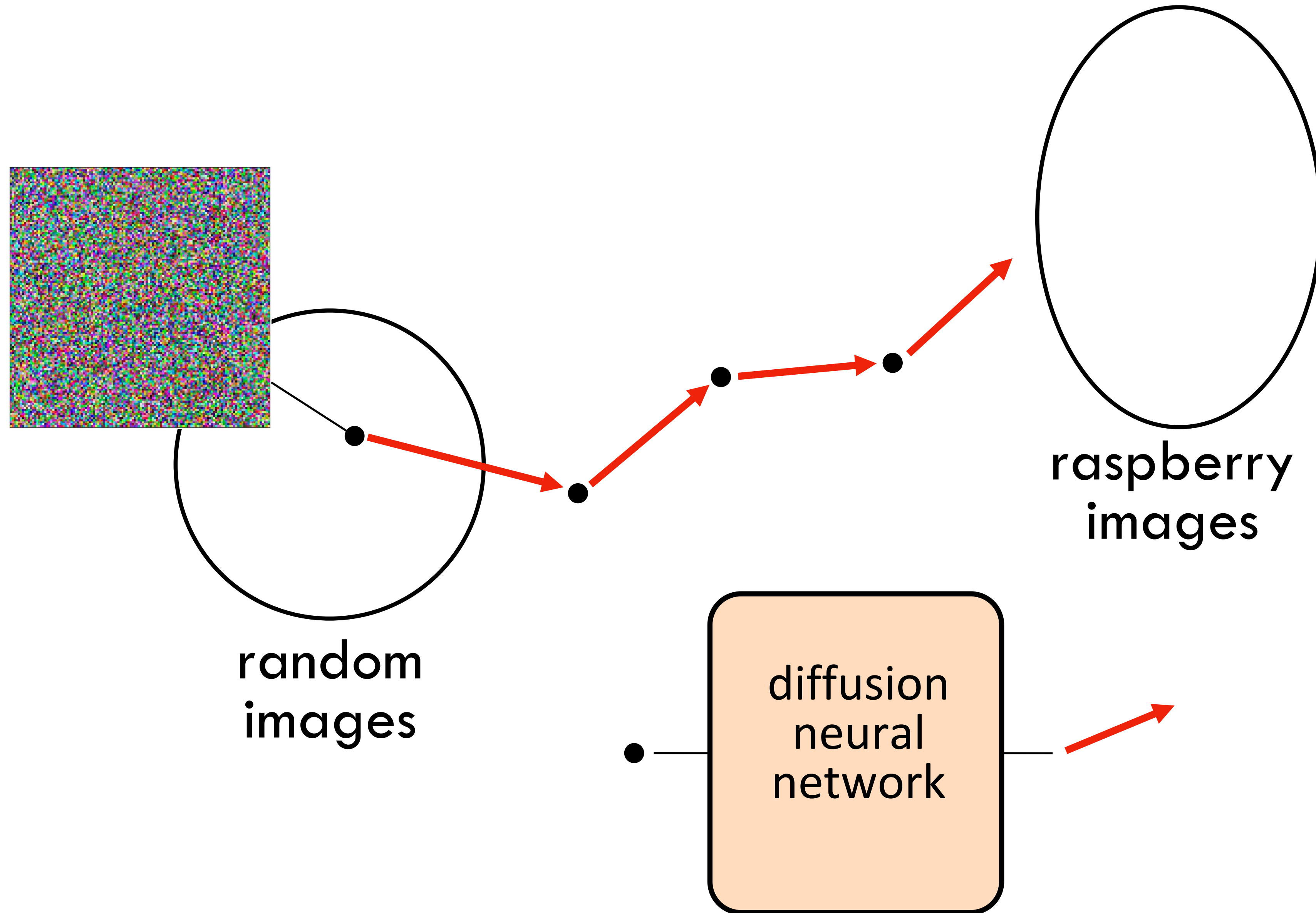


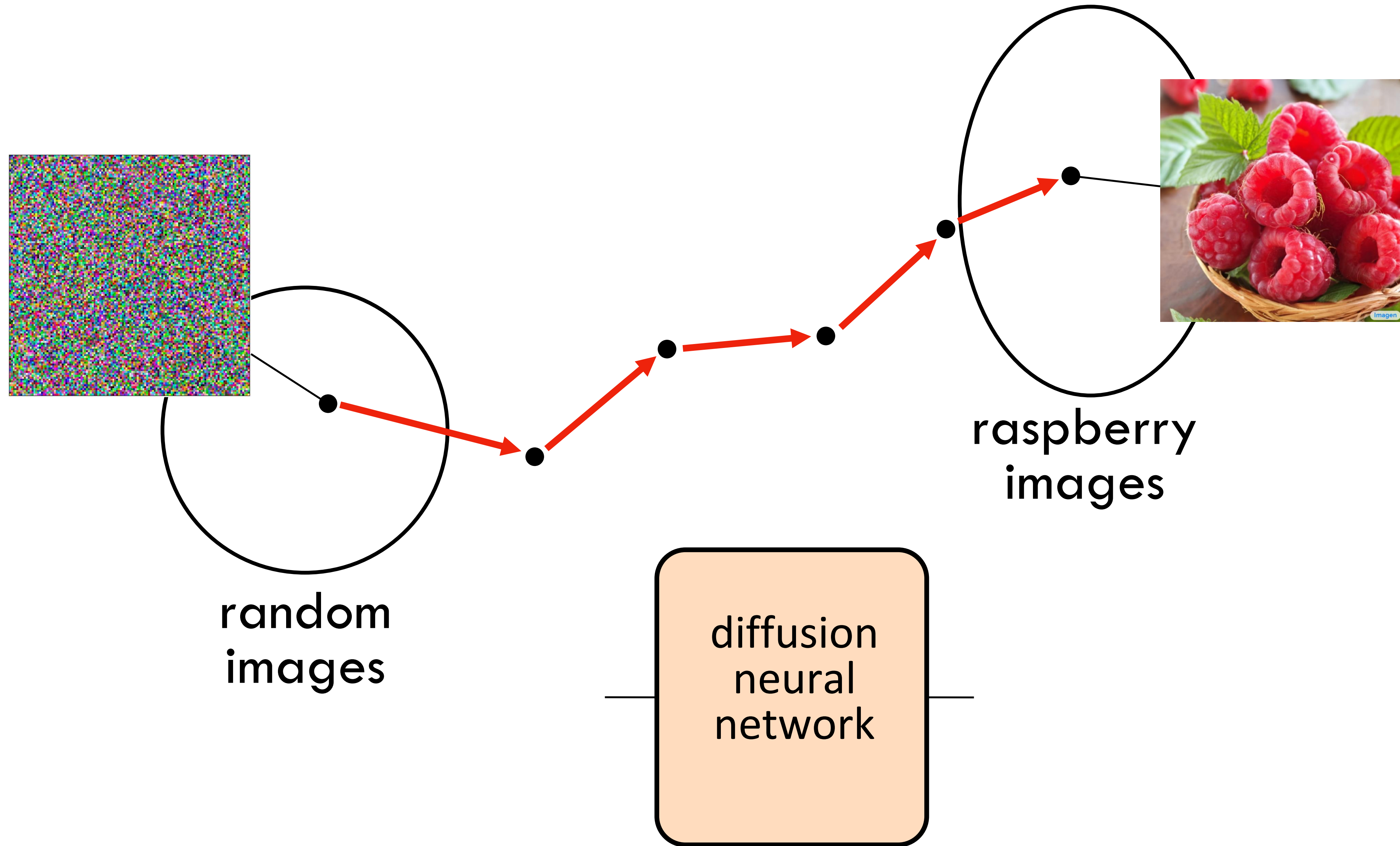
random
images

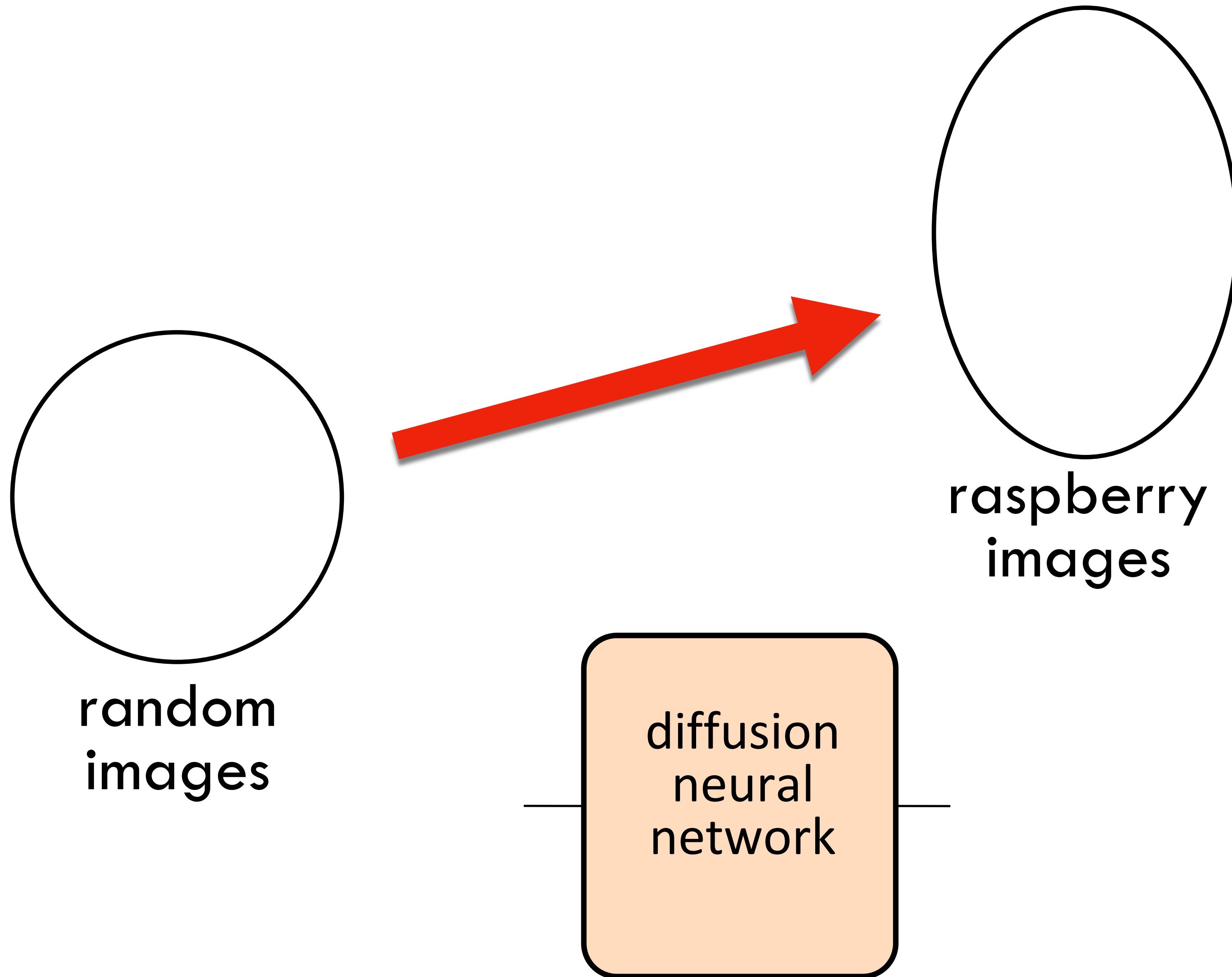
raspberry
images

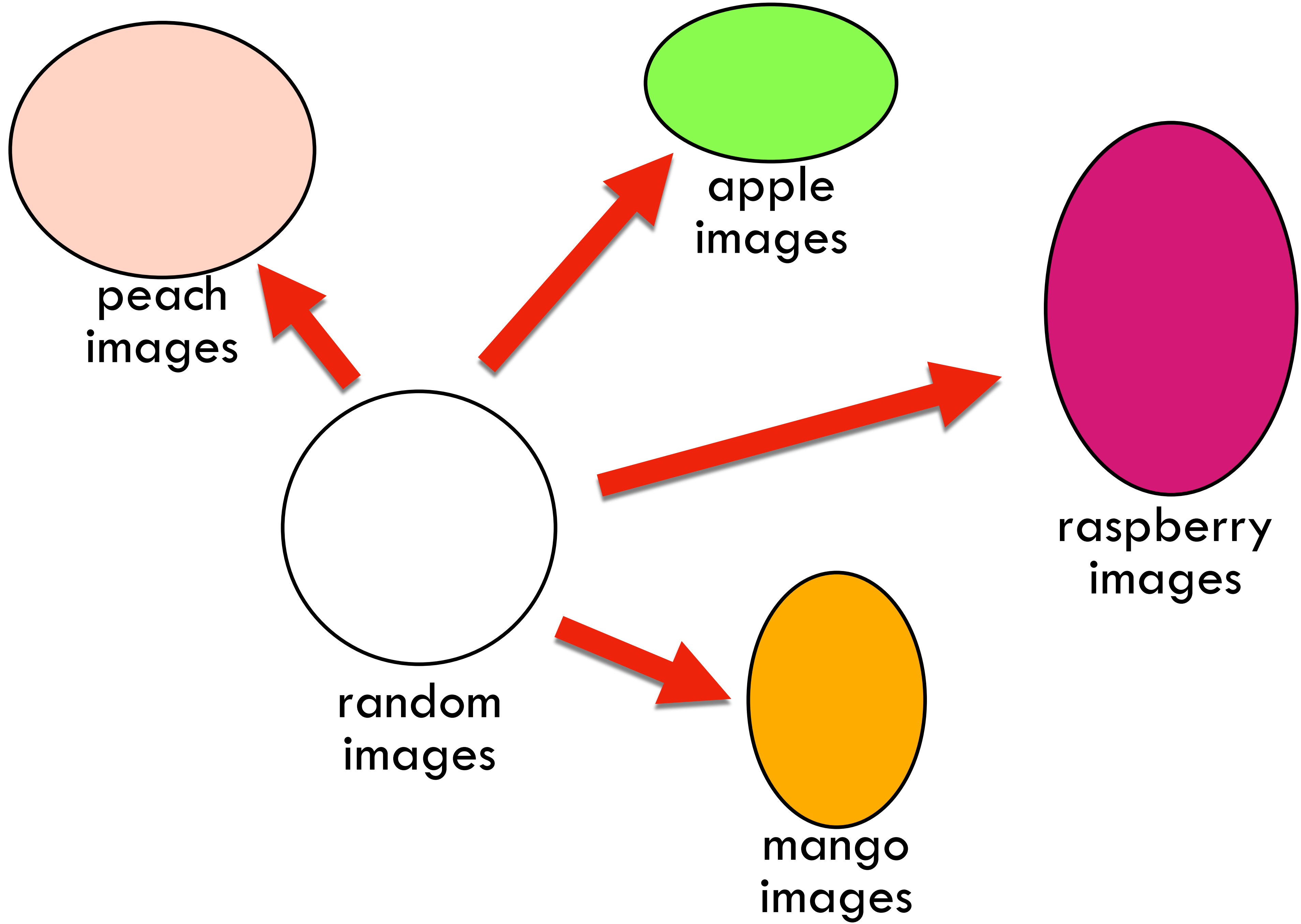
diffusion
neural
network

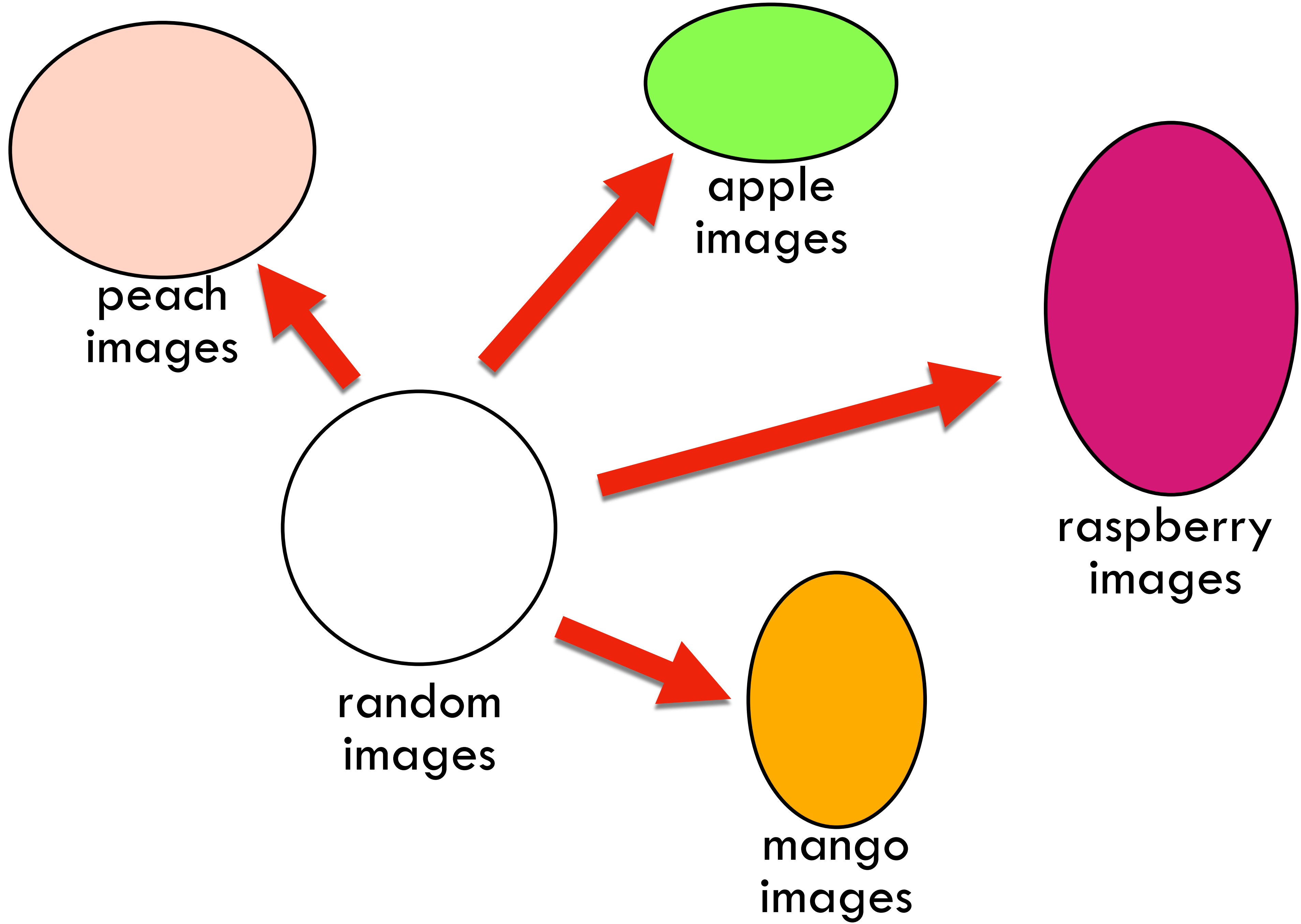


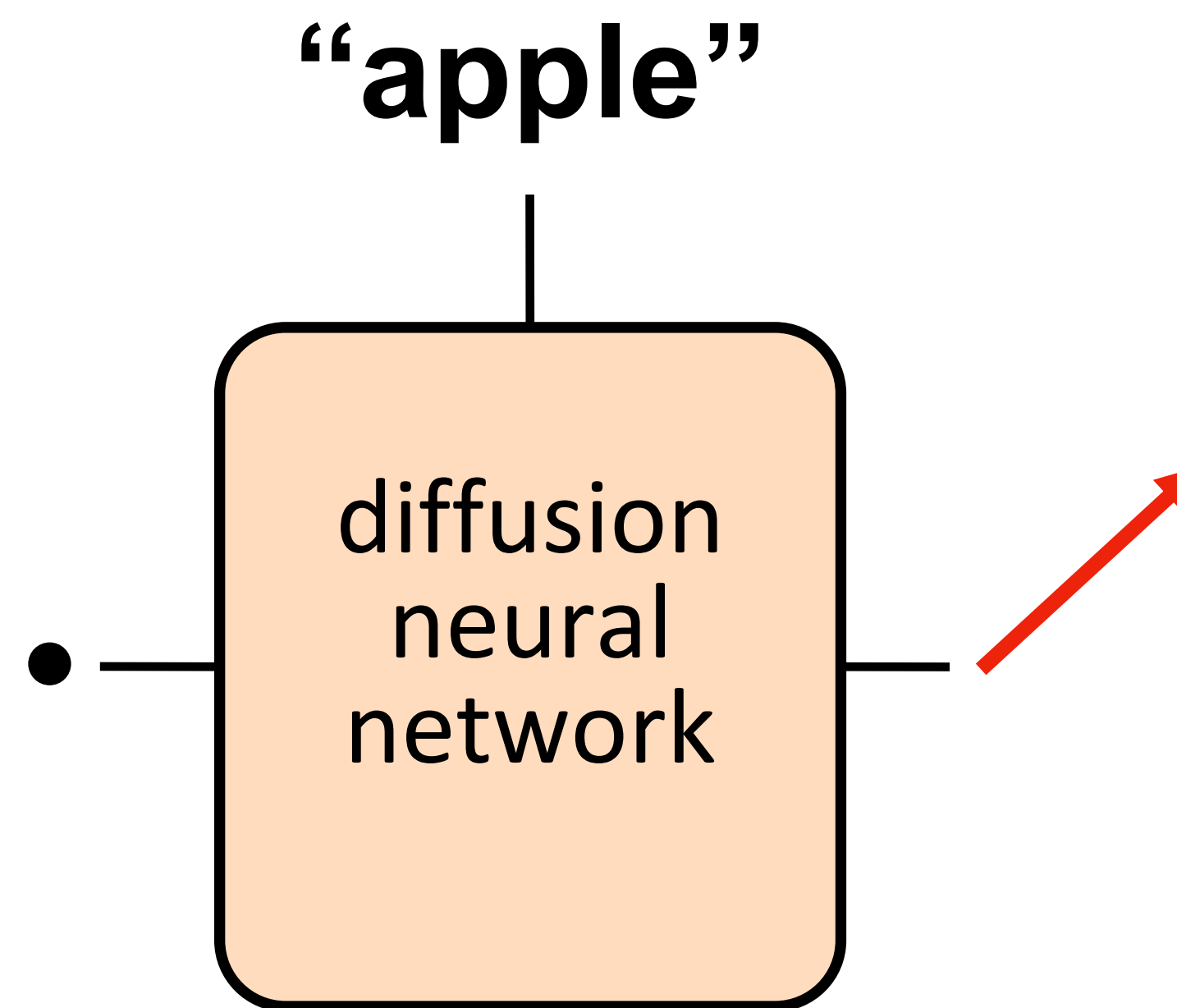
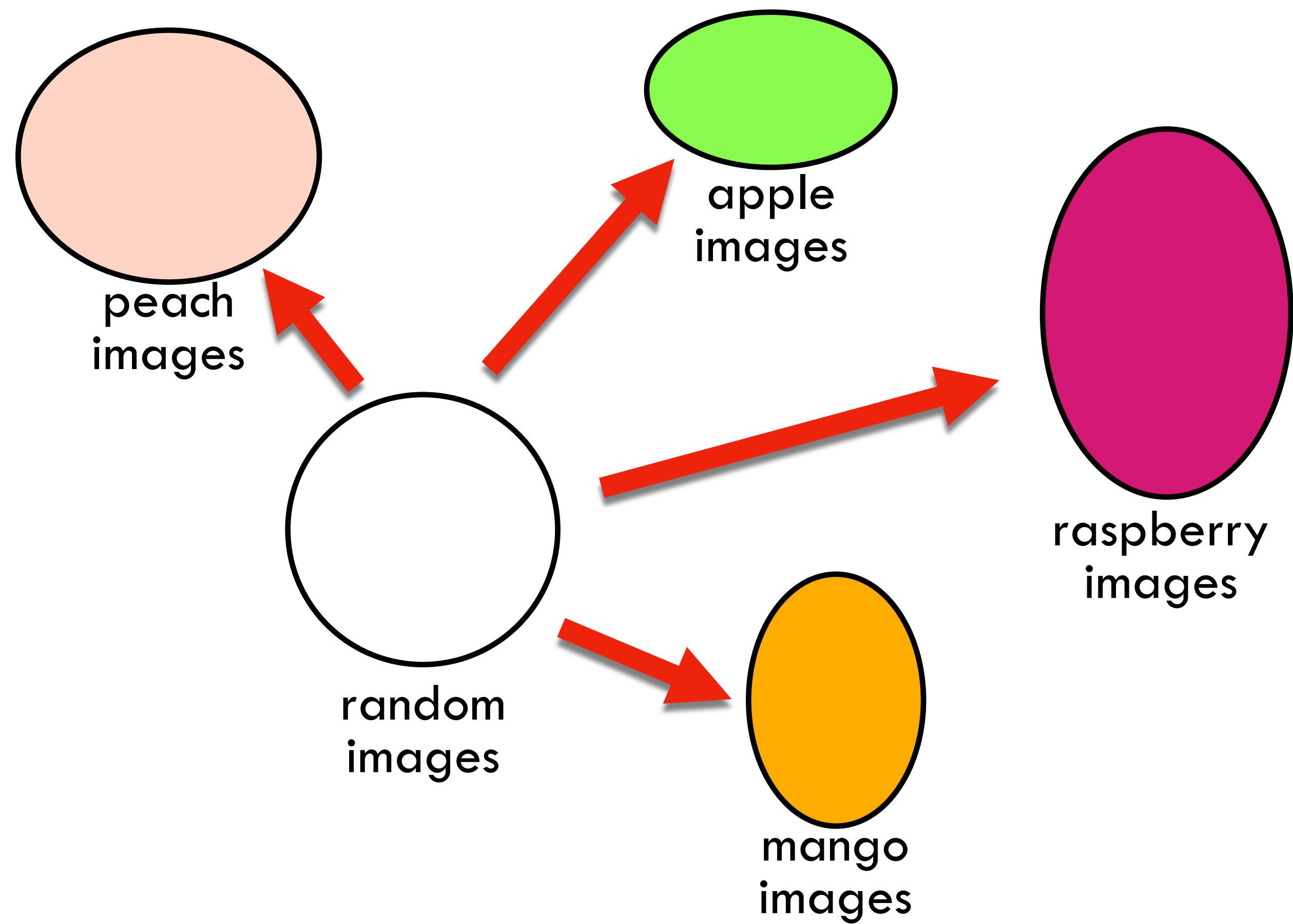


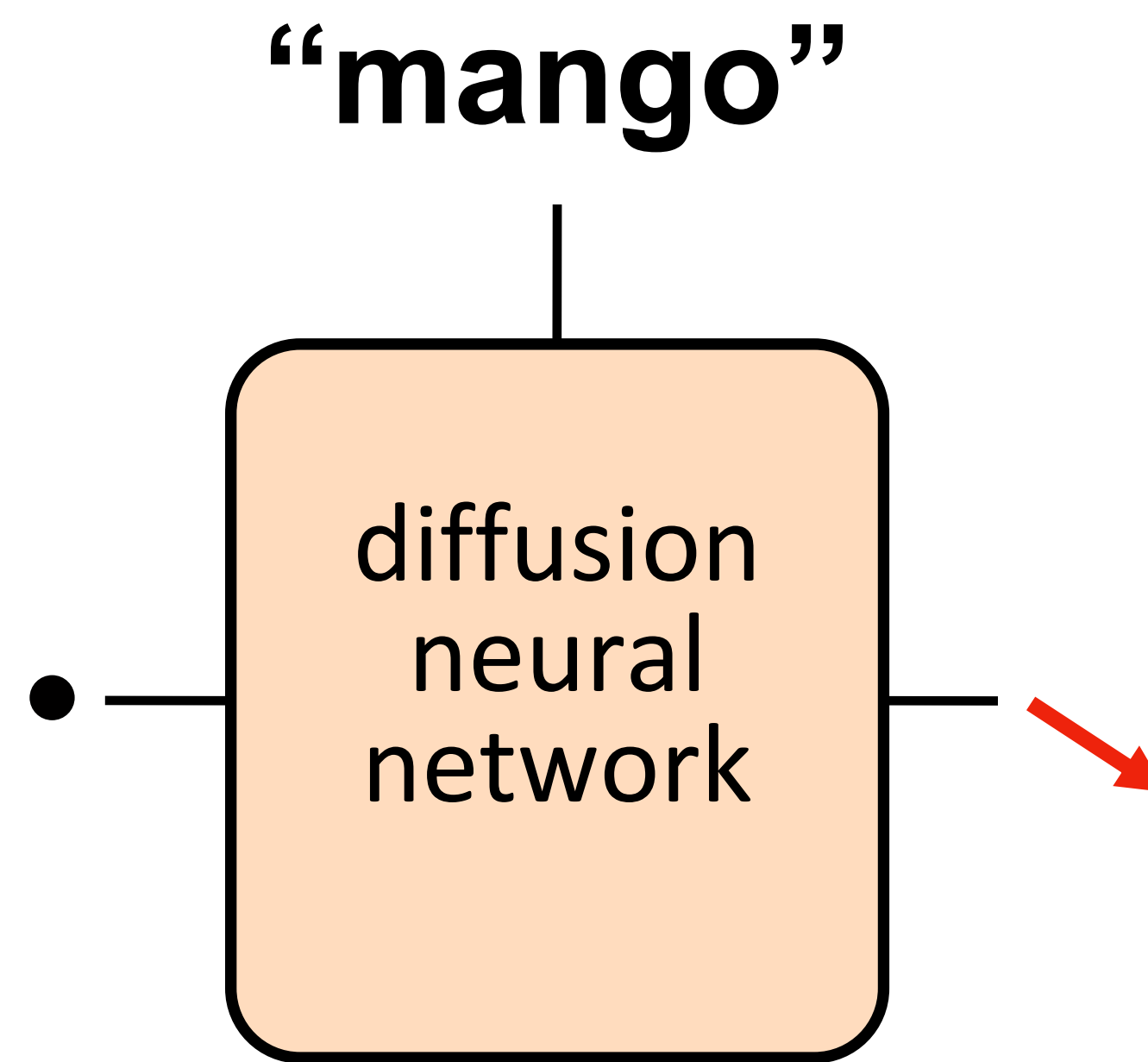
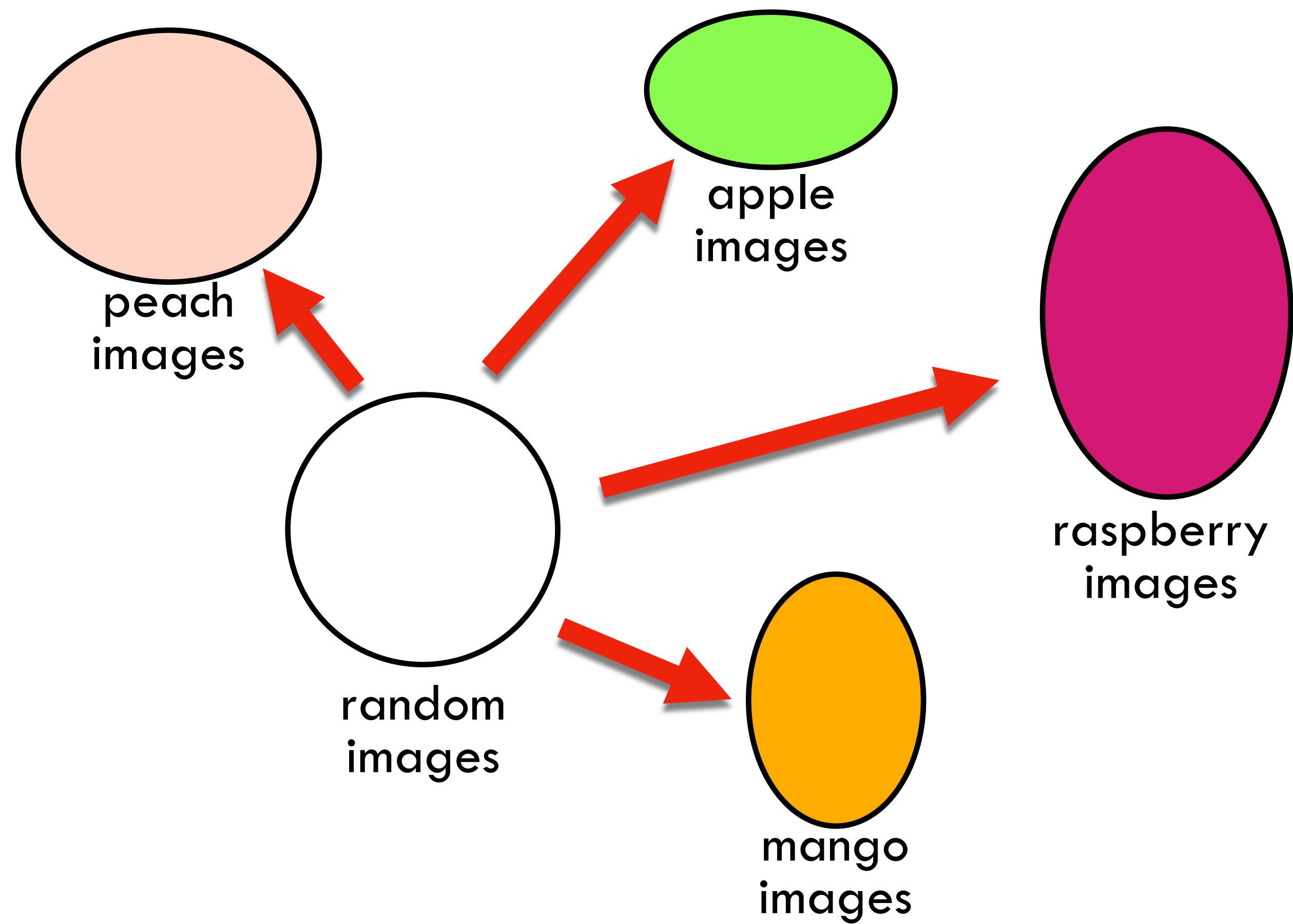




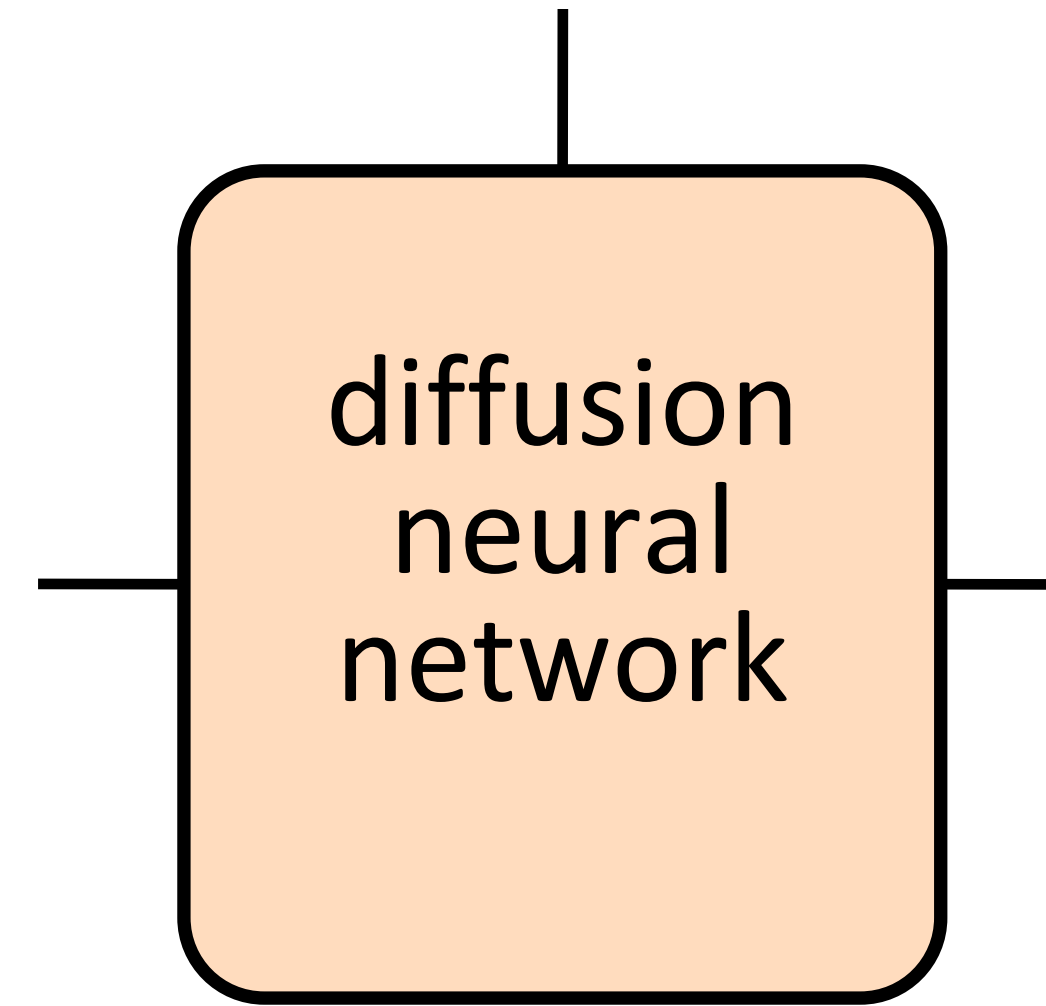


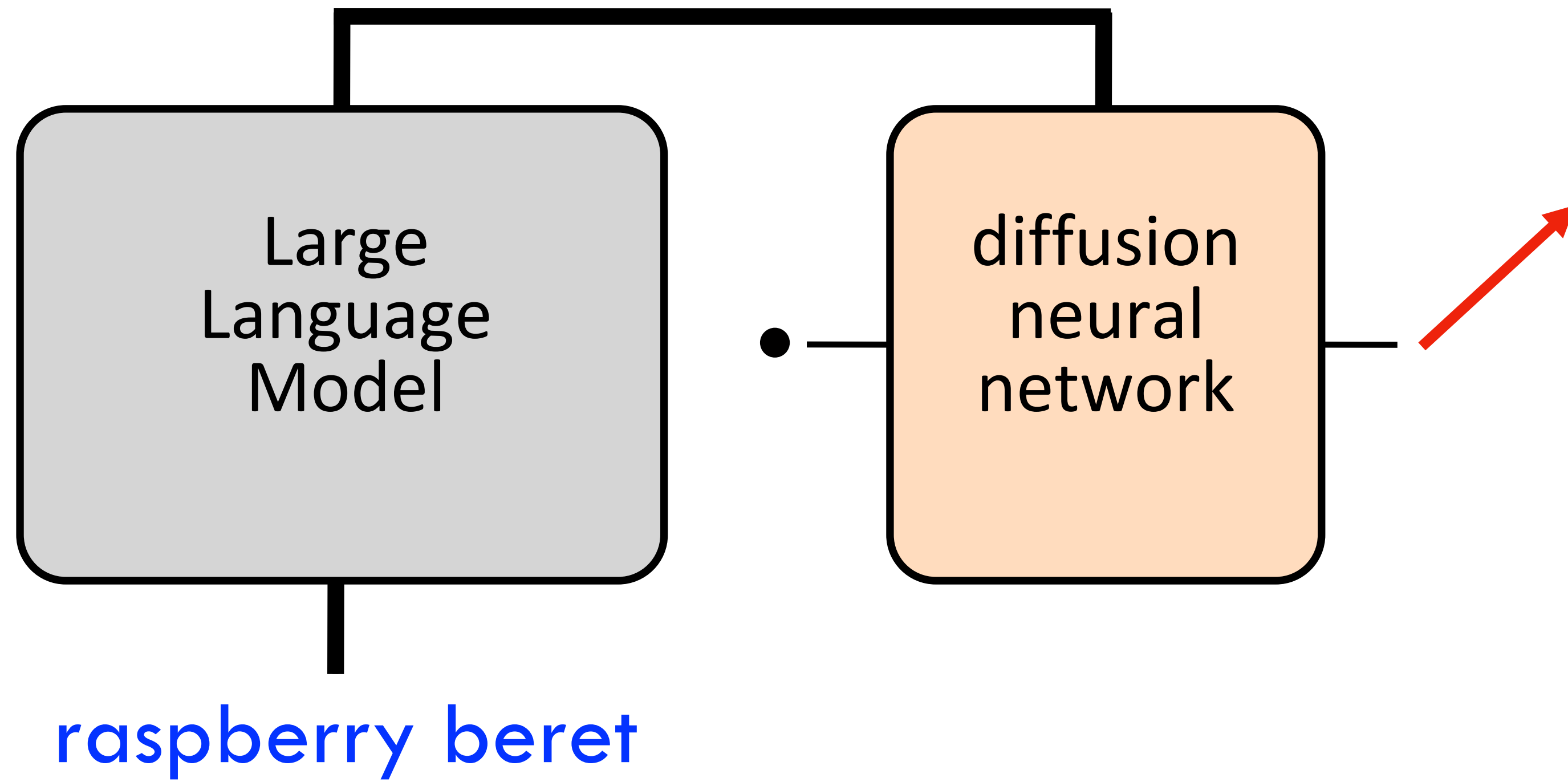






raspberry beret







raspberry beret



raspberry beret

slide from Steve Seitz's [video](#)

beret of raspberries



beret of raspberries

slide from Steve Seitz's [video](#)



beret of raspberries

slide from Steve Seitz's [video](#)



chocolate guacamole pancakes

slide from Steve Seitz's [video](#)

squirrel inside
a nutshell



Imagen

slide from Steve Seitz's [video](#)



An astronaut riding a horse in a photorealistic style (Dall-E 2)

slide from Steve Seitz's [video](#)

**Language
Generator**

(+ pixels)

Parti

Diffusion
(+ language)

Imagen

**Language
Generator**

+

Diffusion

Dall-E 2

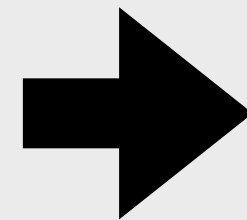
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 6 6 6 1 1 6 6 4 4 4 9 9 9 9 9 9
1 1 1 1 1 7 7 1 1 1 1 1 1 1 1 1 6 6 6 6 6 6 6 6 4 1 4 9 9 9 9 9 9
1 1 7 1 1 1 7 1 1 1 1 1 1 1 1 6 6 6 6 6 6 6 6 4 1 4 9 9 9 9 9 9
1 1 7 1 1 1 1 1 1 1 1 1 1 1 2 6 2 6 6 2 2 6 2 9 9 9 9 9 9 9 9
1 1 1 7 7 1 1 1 1 1 1 1 1 1 6 6 6 6 2 5 2 2 4 9 9 9 9 9 9 9 9
1 1 1 7 1 1 1 1 1 1 1 1 1 2 6 6 6 2 5 2 2 0 9 9 9 9 9 9 9 9
1 1 1 1 1 1 2 2 2 2 2 1 2 6 6 6 2 6 2 6 2 0 9 9 9 9 9 9 9 9
2 1 1 1 1 2 2 1 2 2 2 1 2 6 6 6 6 6 2 6 4 9 9 9 9 9 9 8 8 8
2 2 1 1 1 1 1 1 1 1 1 1 1 6 6 2 6 6 6 4 4 9 9 9 9 9 9 8 8
1 1 1 1 1 1 1 1 1 1 1 1 1 1 6 6 0 0 0 0 0 4 9 9 9 9 9 9 8
1 1 1 1 1 1 1 1 1 7 1 5 2 2 2 0 0 0 0 0 4 9 9 9 9 9 9 8
1 1 1 1 1 1 1 1 1 1 5 2 2 2 2 0 0 0 0 4 4 6 9 9 9 9 9 9
1 1 1 1 1 1 1 1 1 2 2 2 2 2 5 2 0 5 4 6 9 9 9 9 9 9 9 9
1 1 1 1 7 1 1 1 1 2 2 6 6 5 5 2 9 9 9 9 9 9 9 9 9 9 9 9
1 1 1 1 7 2 2 2 1 1 2 2 6 5 5 0 2 9 9 9 9 9 9 9 9 9 9 9
1 1 1 7 2 2 2 2 1 1 2 6 6 5 5 0 0 4 4 4 9 9 0 0 0 0 9 9 9 9
1 1 0 1 2 1 2 1 1 1 1 1 1 0 0 0 0 4 0 0 0 4 0 0 0 0 9 9 9 8
1 1 1 1 1 1 1 1 1 1 7 1 1 0 0 0 0 0 0 0 0 0 0 0 0 9 9 9 8
1 1 1 1 1 1 1 1 1 7 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 9 9 8 8
1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 4 4 9 9 8 9
1 1 1 1 1 7 7 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 4 4 9 9 5 5 9
1 1 1 1 1 7 7 7 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 4 4 9 9 5 4 3
1 1 1 1 1 7 7 3 3 3 3 3 3 3 4 0 4 0 0 4 0 0 0 4 4 9 9 5 4 4
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 5 0 0 0 0 0 0 4 9 9 9 4 3 3
3 3 3 3 3 3 3 3 3 3 3 4 4 4 3 3 3 5 3 0 4 4 4 4 4 4 9 9 4 3 8
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 8 3 4 4 4 4 4 9 3 9 4 8 8
3 3 8 3 3 3 3 3 3 3 3 3 3 3 8 3 8 8 3 3 8 3 8 4 4 8 4 8 3 8 3 3 3
8 8 3 8 8 3 8 3 3 3 8 3 8 3 3 3 3 3 8 3 8 4 8 8 8 3 3 3 3 3 3
3 3 8 3 3 3 3 3 3 8 8 8 8 3 3 8 8 3 3 8 8 8 8 8 8 3 8 3 8
3 3 8 8 3 3 3 8 3 3 3 3 3 3 3 8 8 3 3 8 3 3 3 8 3 8 8 8 3 8
3 3 8 8 3 3 3 8 8 3 8 3 3 3 3 3 3 3 3 3 3 3 3 8 3 8 8 8 8 8
3 3 8 3 3 3 3 8 3 8 8 3 3 3 3 3 3 3 3 3 3 3 3 8 8 8 8 8 8

Parti

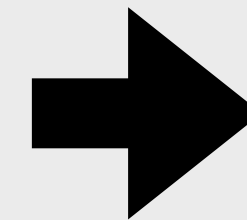
734
37
520
323
1023
14
592
842
164
325
296
820
893
621
924
724
422
342
290
76
•
•
•
891

Dall-E 2

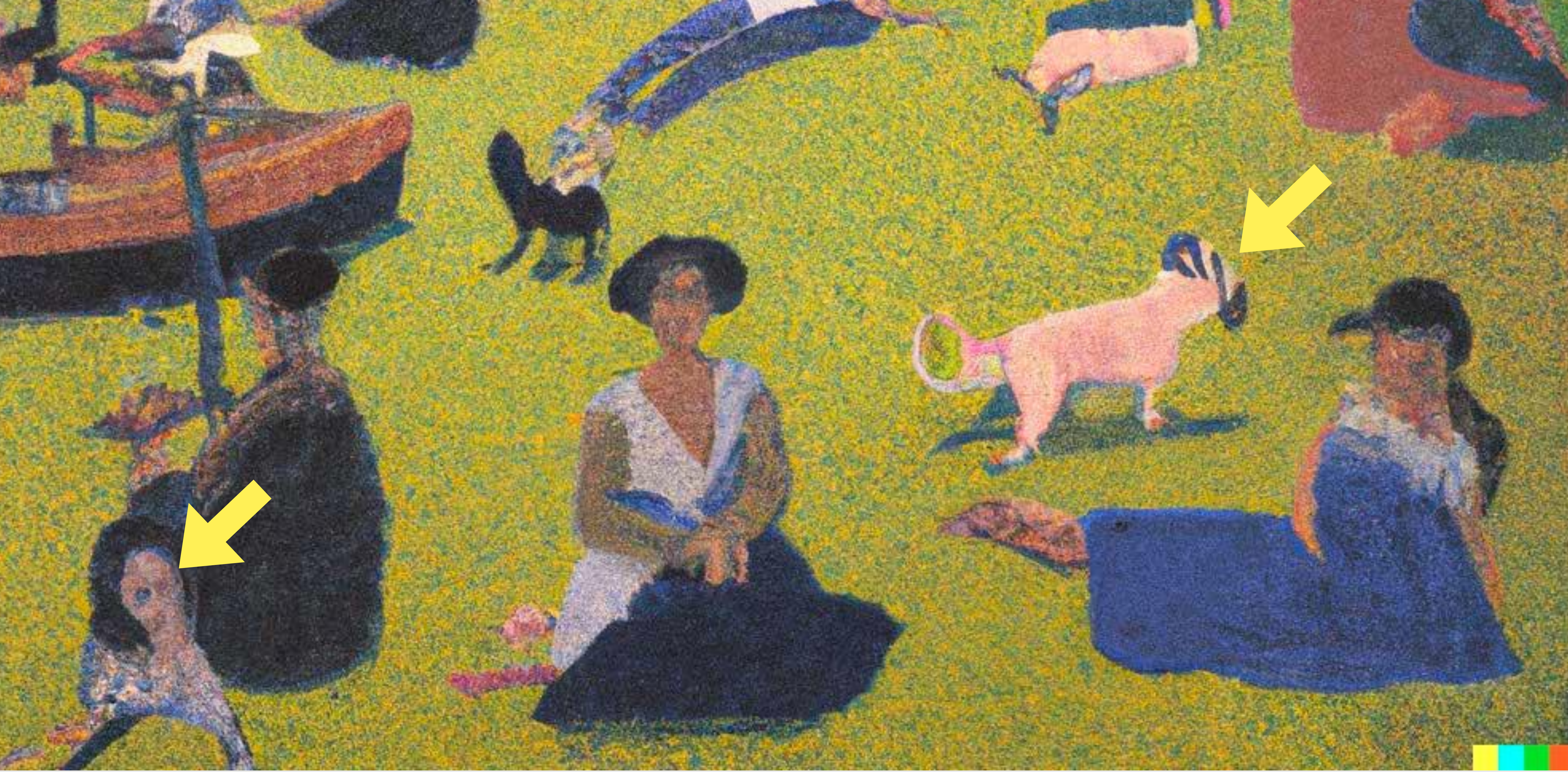
slide from Steve Seitz's [video](#)



- 734
- 37
- 520
- 323
- 1023
- 14
- 592
- 842
- 164
- 325
- 296
- 820
- 893
- 621
- 924
- 724
- 422
- 342
- 290
- 76
- 584
-
-
-
- 891









Squirrel reaching for a nut. Latte art

slide from Steve Seitz's [video](#)



A teddy bear making chocolate guacamole pancakes

slide from Steve Seitz's [video](#)

A dog looks curious in the mirror,
seeing a cat



A bowl of soup
that looks like a monster
knitted out of wool



Dall-E 2

Impressive compositionality:

