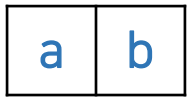
What is Latent Space?

Latent Space Explained by Gio Paik

Gio Paik
giopaik@naver.com

Probability Space

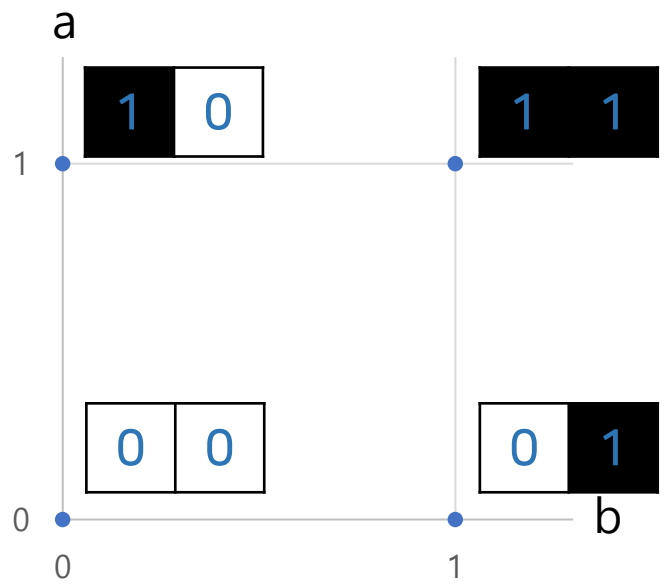
- Before we discuss about Latent Space, We should talk about Probability Space.
- Think about an image consists of 1×2 pixel with two possible pixel value (0, 1).



How many Images can there be?

Probability Space

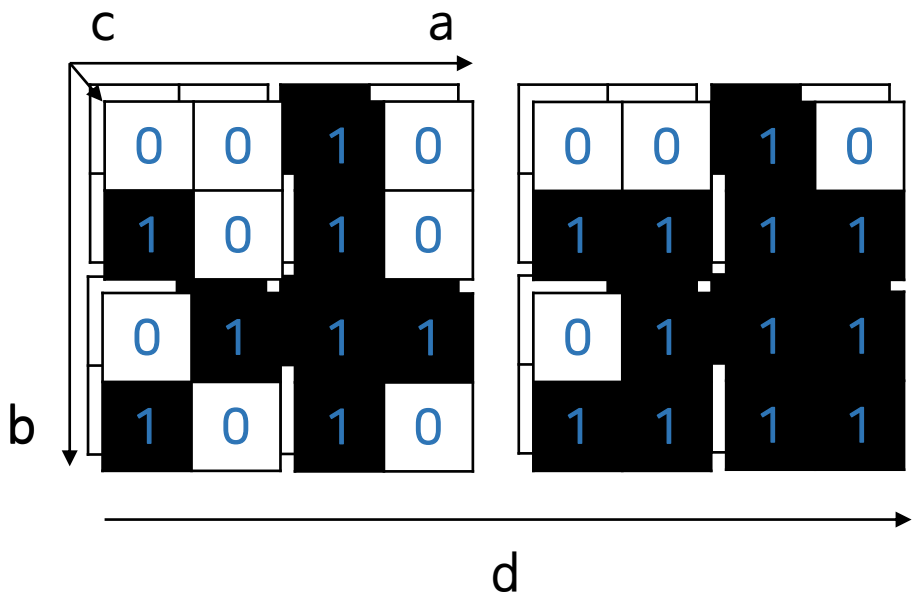
- We have $n = 1 \times 2$ pixels.
- Each pixels have two possible values, (0, 1).
- We have 2 axis (a, b) with 2 possible values (0,1).



Cite

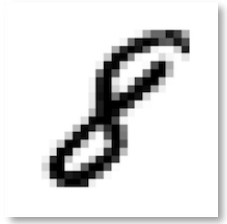
Probability Space

- What if we have $n = 2 \times 2$ pixels?
- Each pixels have two possible values, (0, 1).
- We have 4 axis (a, b, c, d) with 2 possible values.



Probability of “Real Images”

- Now, Let’s think about probability space of “Real Images”.
- Probability Space has n dimension with m possible values.



mnist dataset

$28 \times 28 \times 1 = 784$ axis with 256 values

$28 \times 28 \times 256 = 200{,}704$ possible images

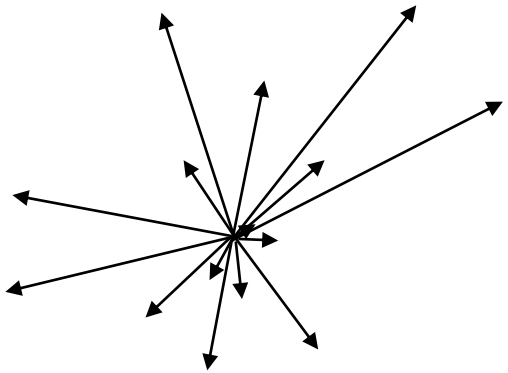
Probability of “Real Images”

- In these days, our image has color (means x3 bigger images), and higher resolution.

Image of me (Mask on!)
$960 \times 1280 \times 3 = 3.68$ million axis with 256 values
1 of 943,718,400 possible images

3.68 Million Axis????

- Each pixel could have three R,G,B values with 256 possible values.
- Can you imagine that huge possibility space?



Forget about pixel space.

- Possibility of pixel data is too big and not useful.
- Let's find something more simple, **efficient way to describe** an image.

Described by pixel data →

"An image with 960 width and 1280 heights,
pixel data is (128, 214, 25), (127, 213, 26)…"

Described by Natural Language →

"Image of a handsome guy with black hair.
He is wearing a mask, a black shirt and a white cardigan."

Latent Space

- We should focus on the nature of the whole image, not on each pixel.
- Our model learns how to describe an attribute of image.

This is an “Black Hair”. Not (245, 240, 232), (244, 241, 233) … kind data.

Latent Space

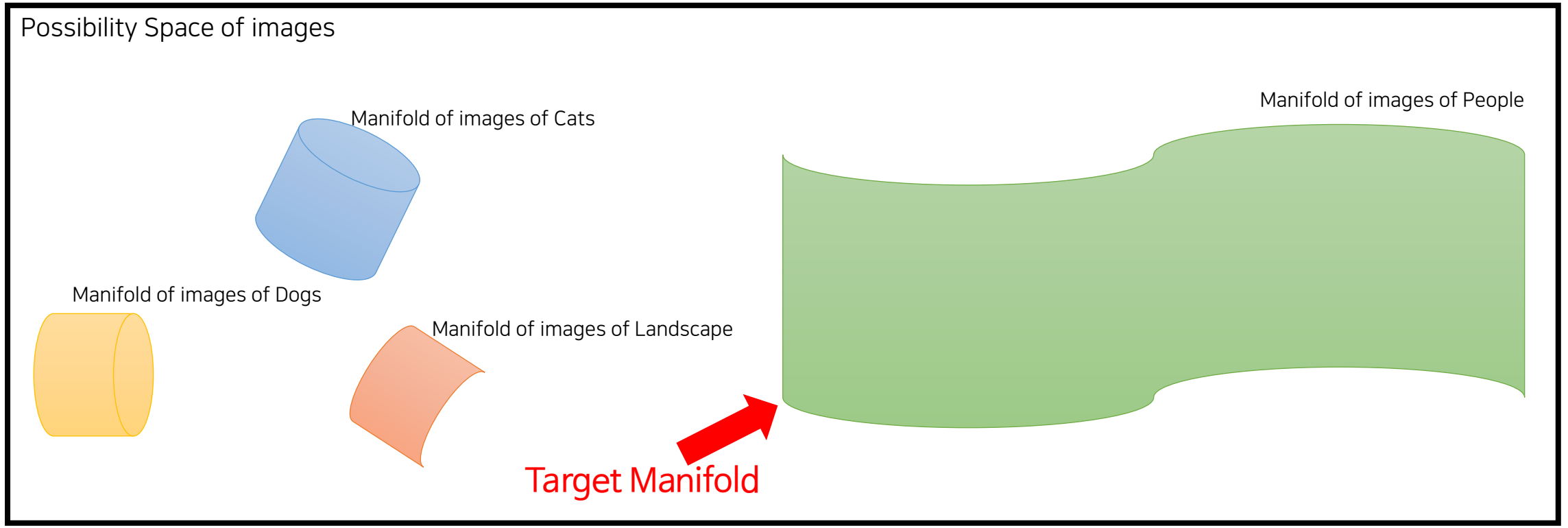
- Now we can describe an image on latent space of people image.
- Each axis stands for an attribute of people like hair color, eye color, hair length, nose, mood, etc …
- Images can be described with much smaller axis than pixel space.

Manifold

- In deep learning, Our model learns the manifold of possible images.
- Manifold is subset of all possible images.
  (set of images of people, set of images of landscapes …)
- Learning Manifold means model learns how to describe the images.

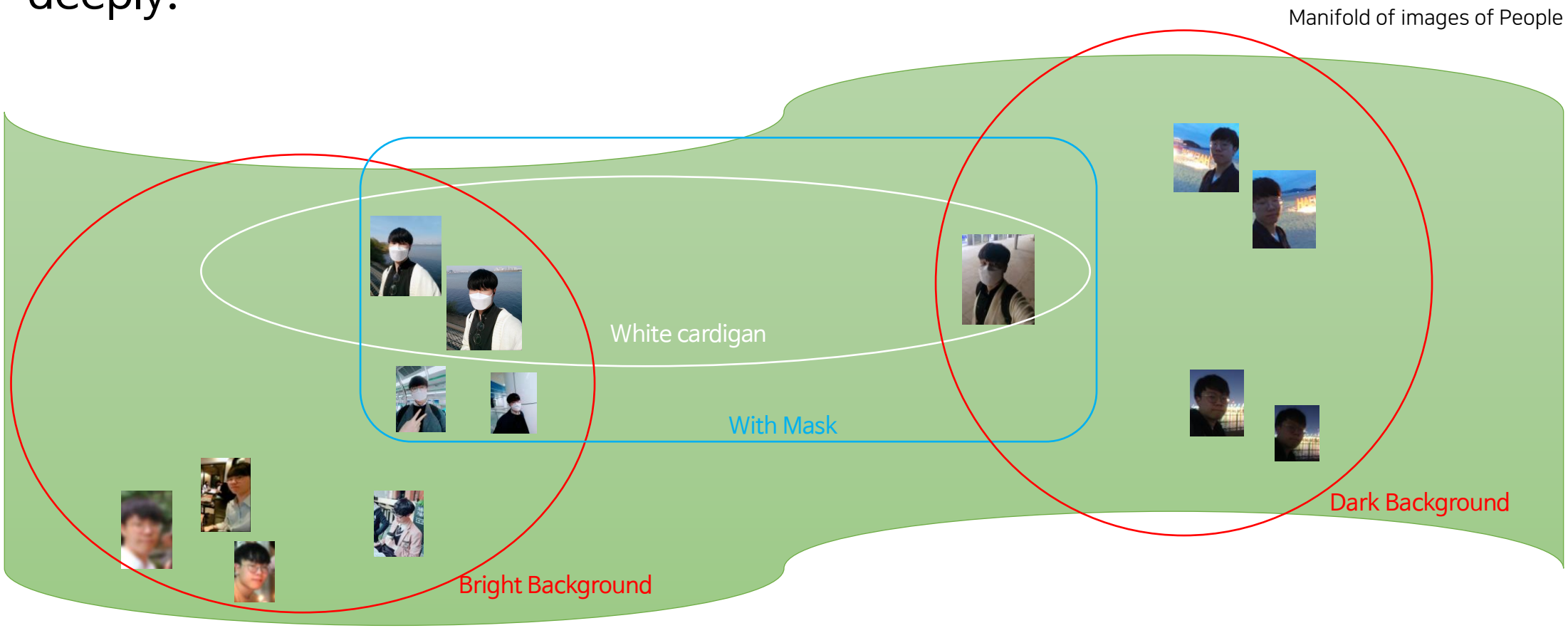
Visualization of Learning Manifold

- There are about 1,592 million full-hd color images.
- First, We find a set of images of people in that possibility space.



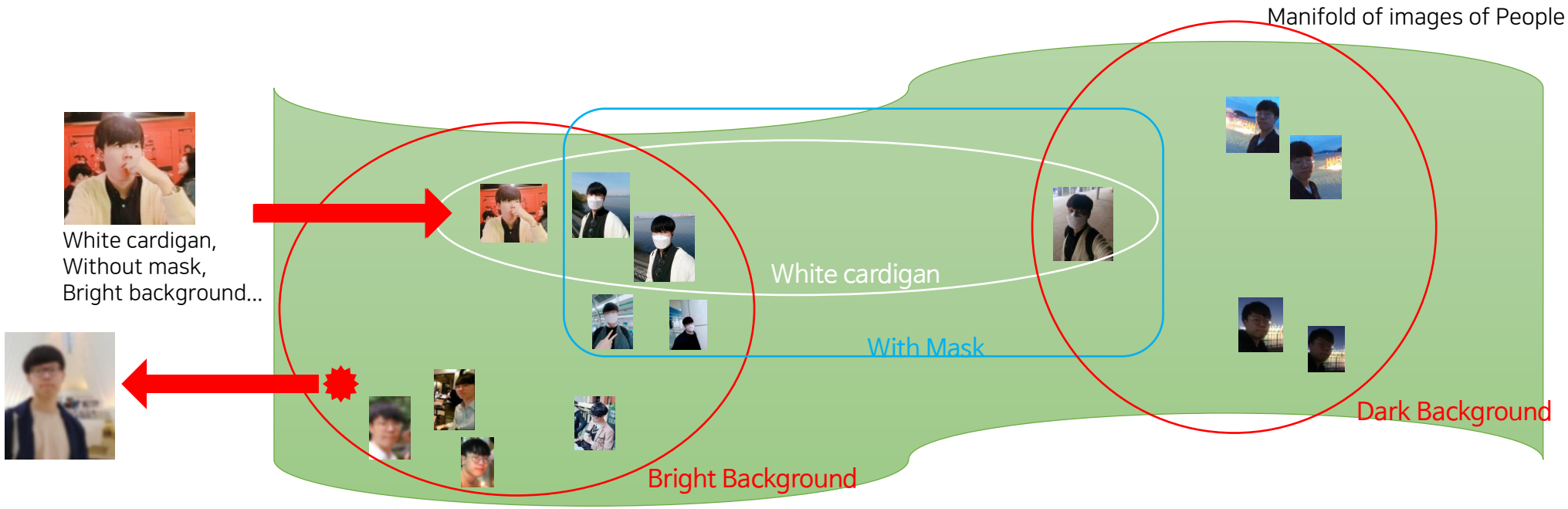
Visualization of Learning Manifold

- After we find our target manifold, model understand manifold more deeply.



Generative Model

- Now we can map image into manifold!
- Similarly, We can decode latent vector to image!



What we can do with latent space

- Training Latent Space representation to generative model let model create more natural image with higher efficiency.

Image Interpolation using
pixel space representation

Image Interpolation using
latent space representation

Figure from https://atcold.github.io/pytorch-Deep-Learning/en/week07/07-3/

What is Latent Space?

Thank you for watching!

Gio Paik
giopaik@naver.com