



Generative Adversarial Network

對抗式生成網路

李宏毅 & 教研處

「版權聲明頁」

本投影片已經獲得作者授權台灣人工智慧學校得以使用於教學用途，如需取得重製權以及公開傳輸權需要透過台灣人工智慧學校取得著作人同意；如果需要修改本投影片著作，則需要取得改作權；另外，如果有需要以光碟或紙本等實體的方式傳播，則需要取得人工智慧學校散佈權。

課程內容

1. Conditional GAN
 - Conditional GAN原理理論說明
 - Conditional GAN 程式實作(Pix2Pix)

各時段預計完成內容

PART I	Conditional GAN原理講解
PART II	Conditional GAN實作演練
PART III	GAN怎麼評估生成影像好壞與適用範圍

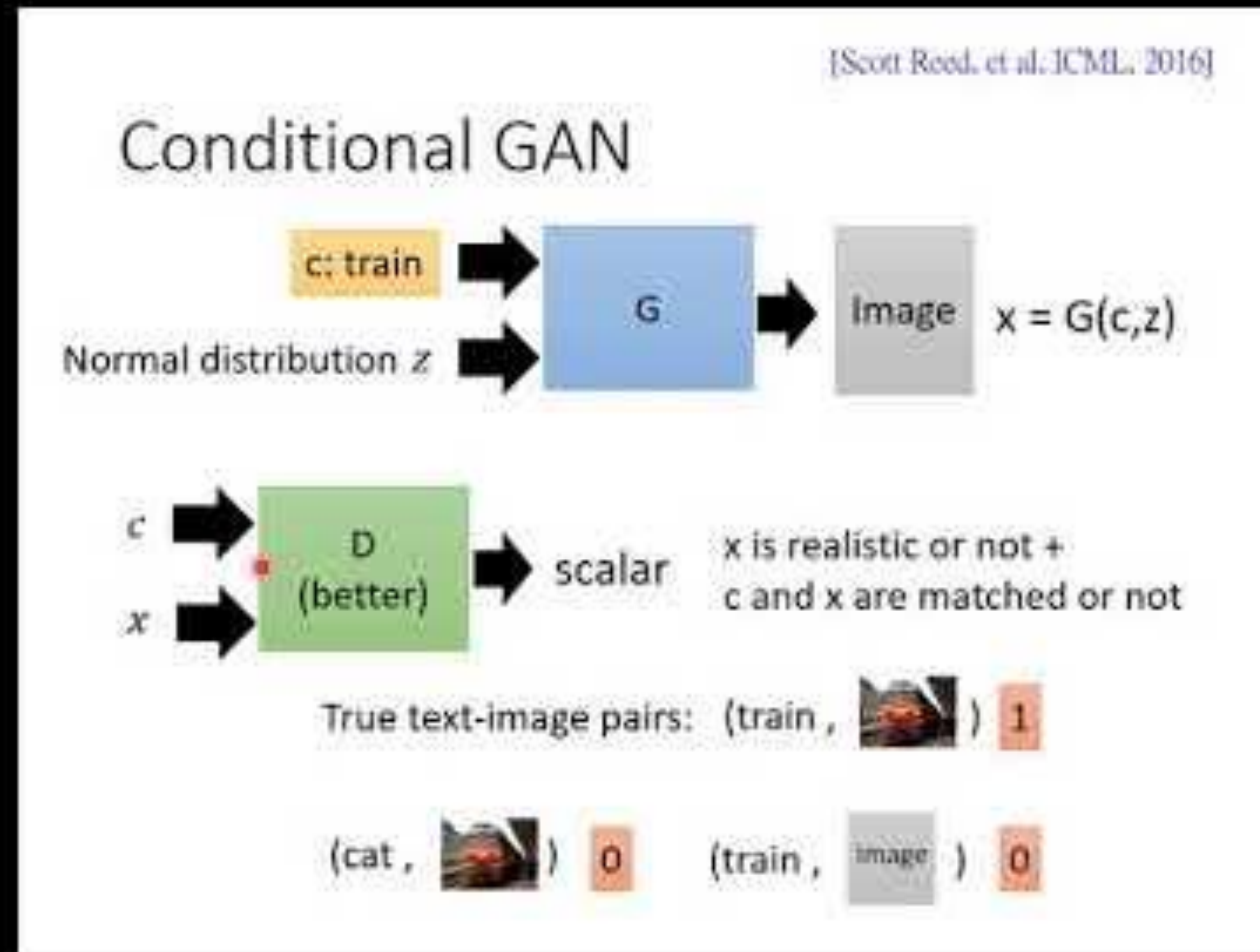
[講師投影片](#)
[影片播放列表](#)



理論講授

Conditional GAN原理講解

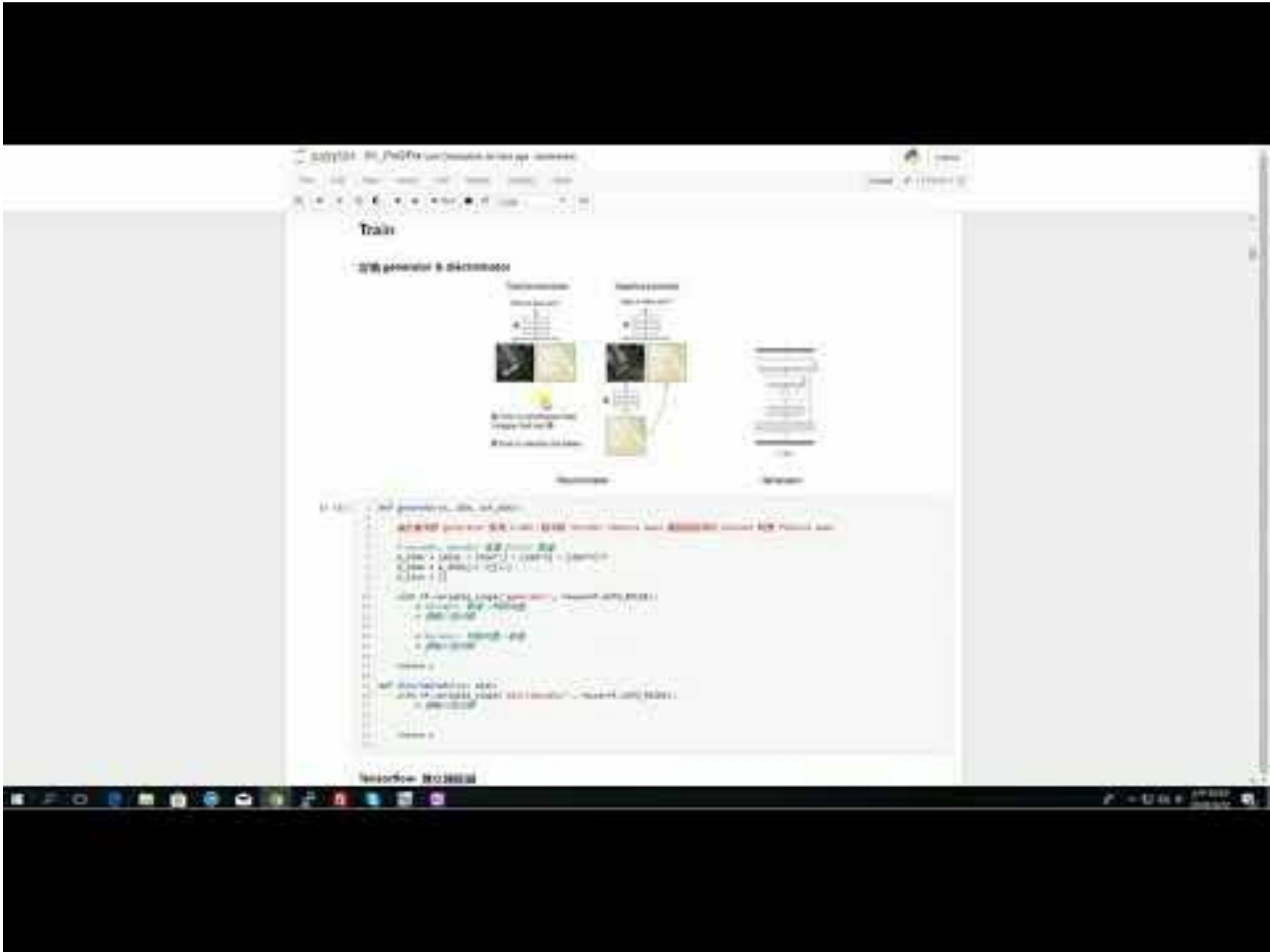
Conditional GAN原理說明

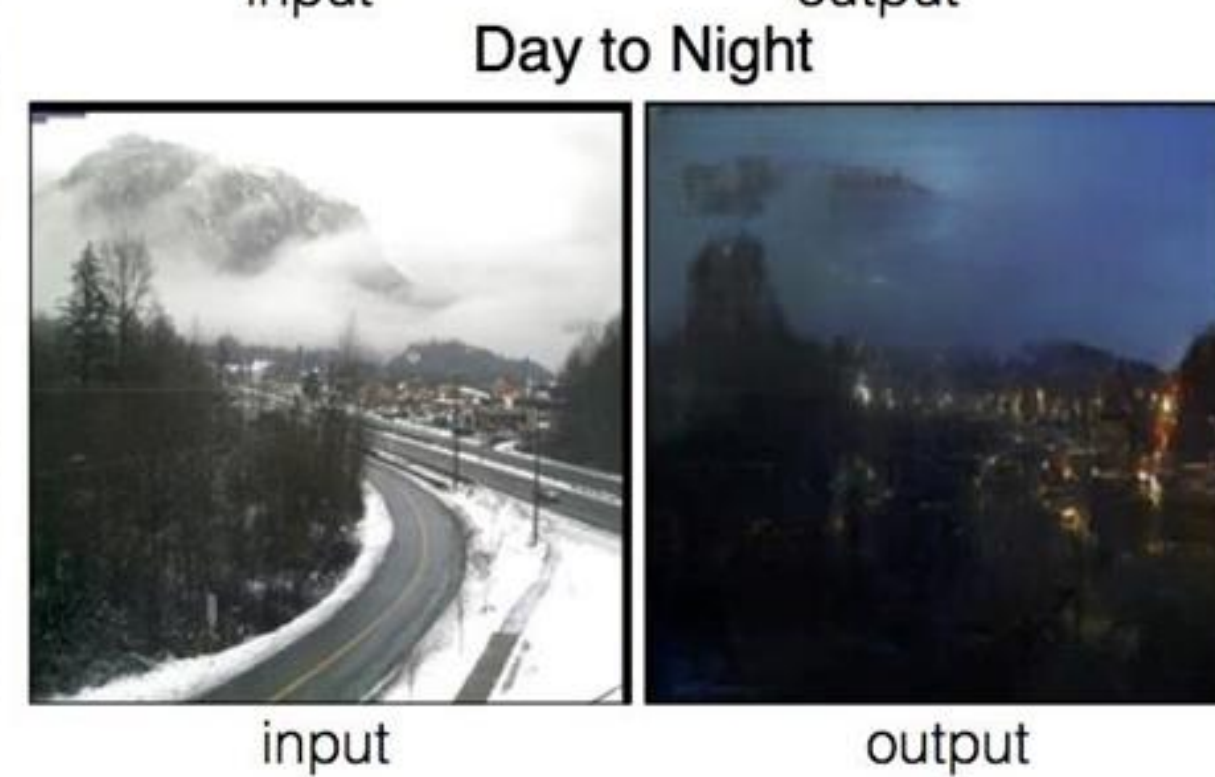
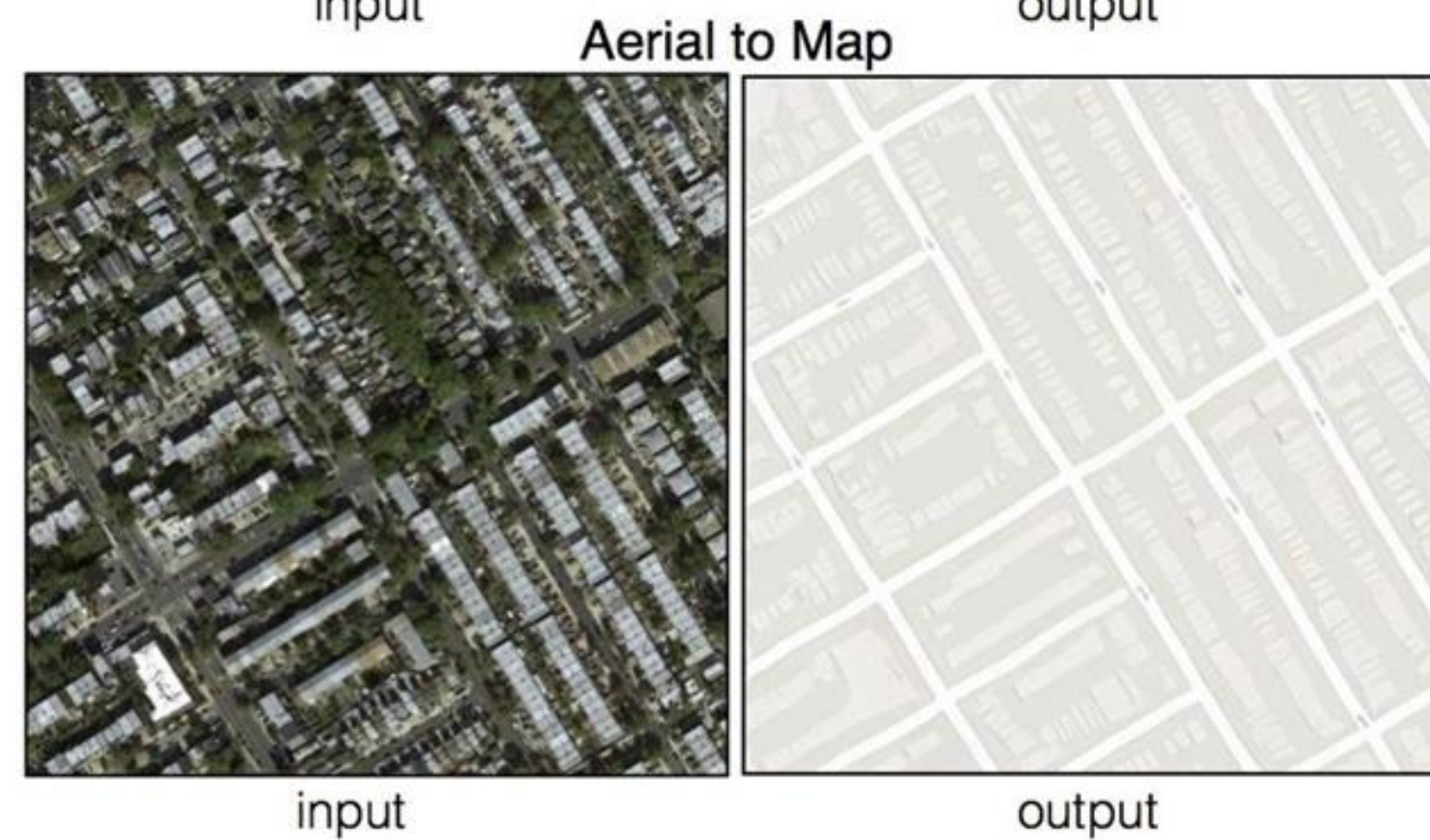
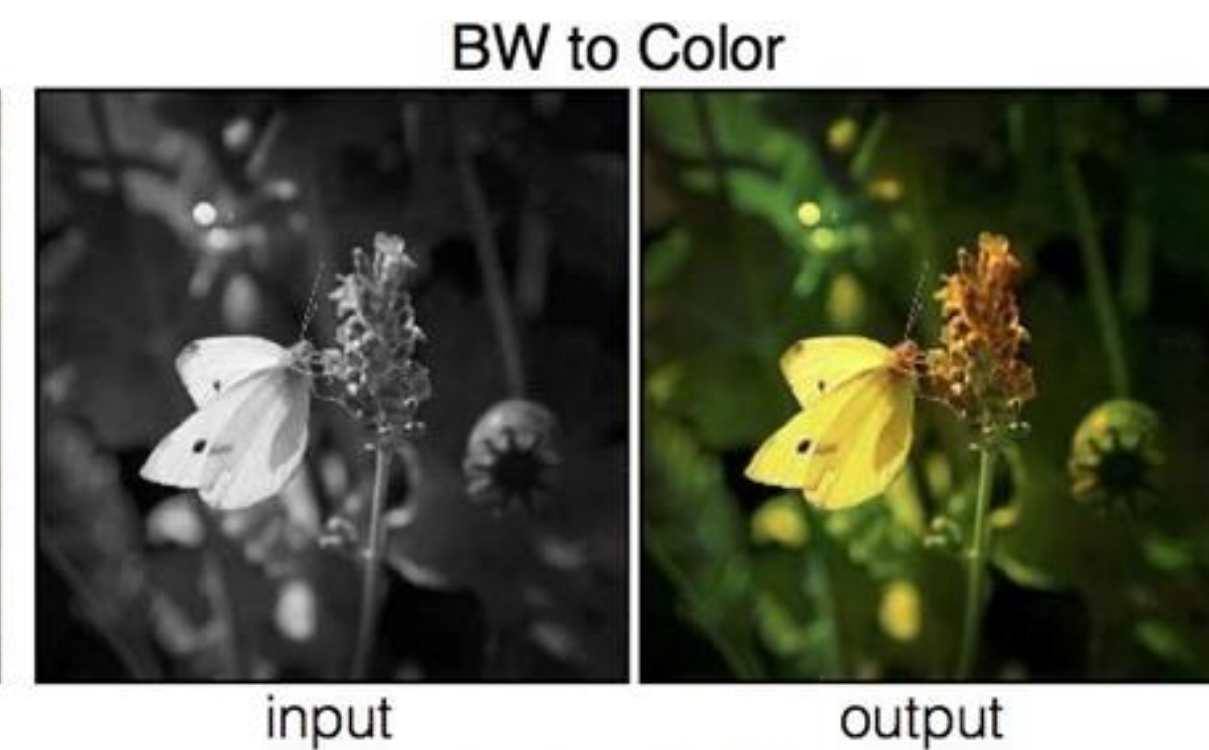
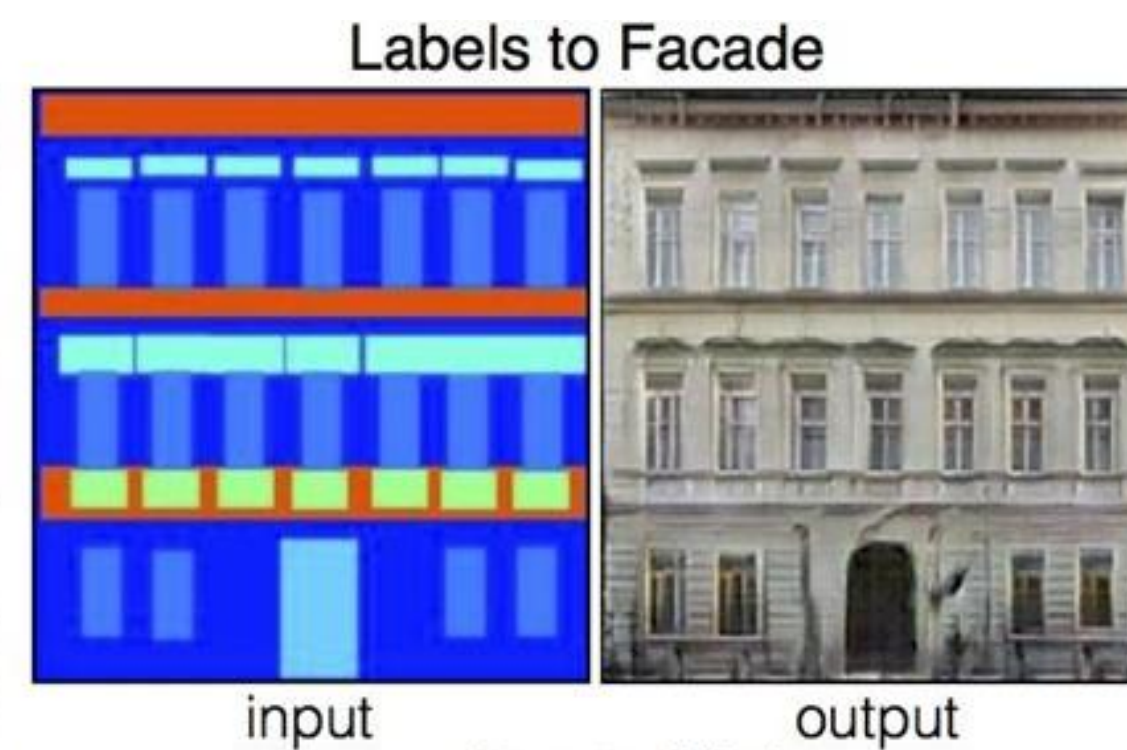


實作演練

Conditional GAN – Pix2Pix

Conditional GAN 實作演練





Pix2Pix

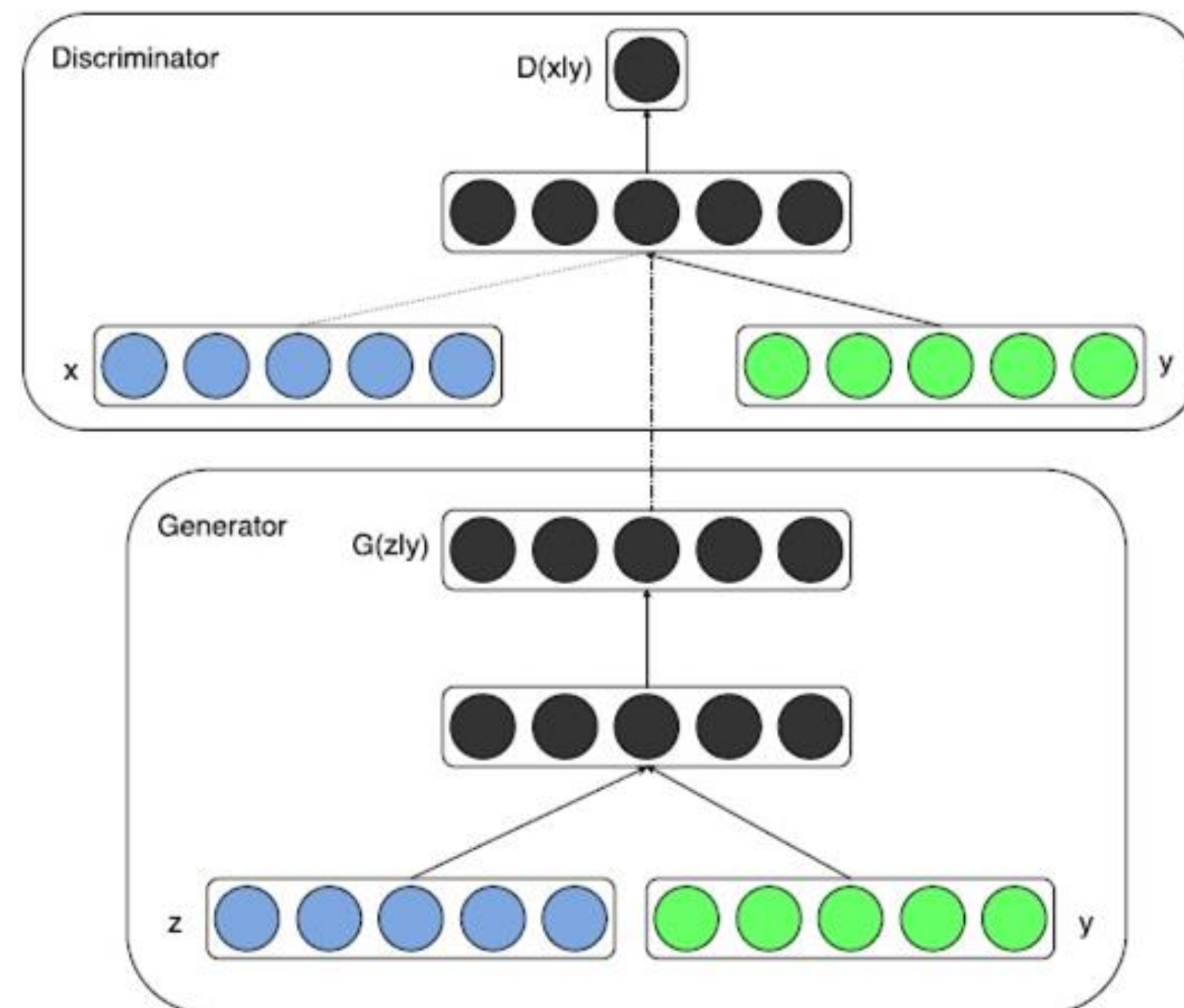
Conditional GAN 的一種應用，可用在風格轉換



Pix2Pix

Conditional GAN

- 在生成器和判別器中增加條件限制
- Pix2Pix 是 conditional GAN 的一個應用，可以用來做影像風格轉換

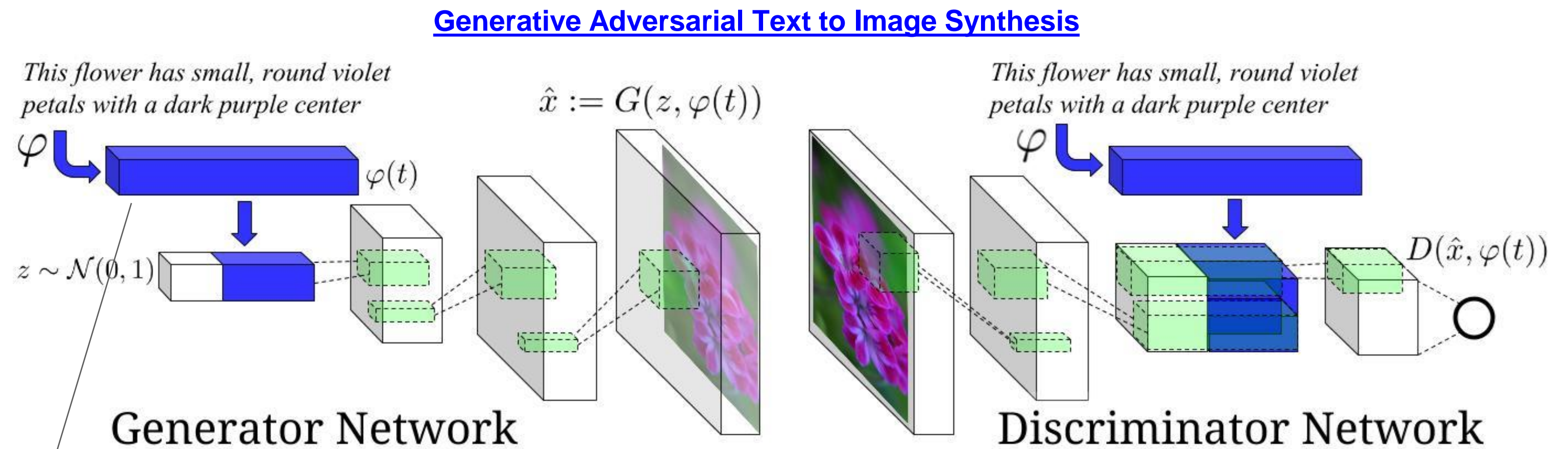


Pix2Pix

Conditional GAN 應用

- Text to Image

- condition: 文字



- Pix2Pix

- condition: 影像
 - input-output pair (input 、condition image)

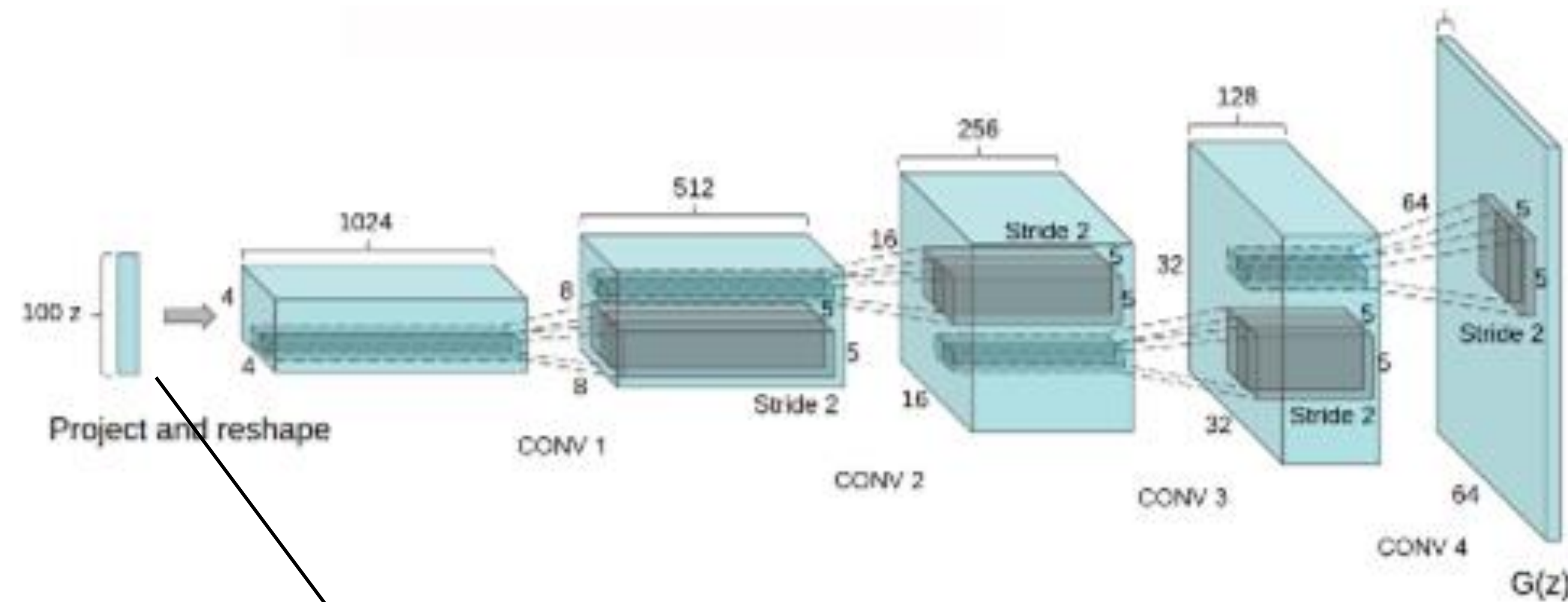
Image-to-Image Translation with Conditional Adversarial Networks



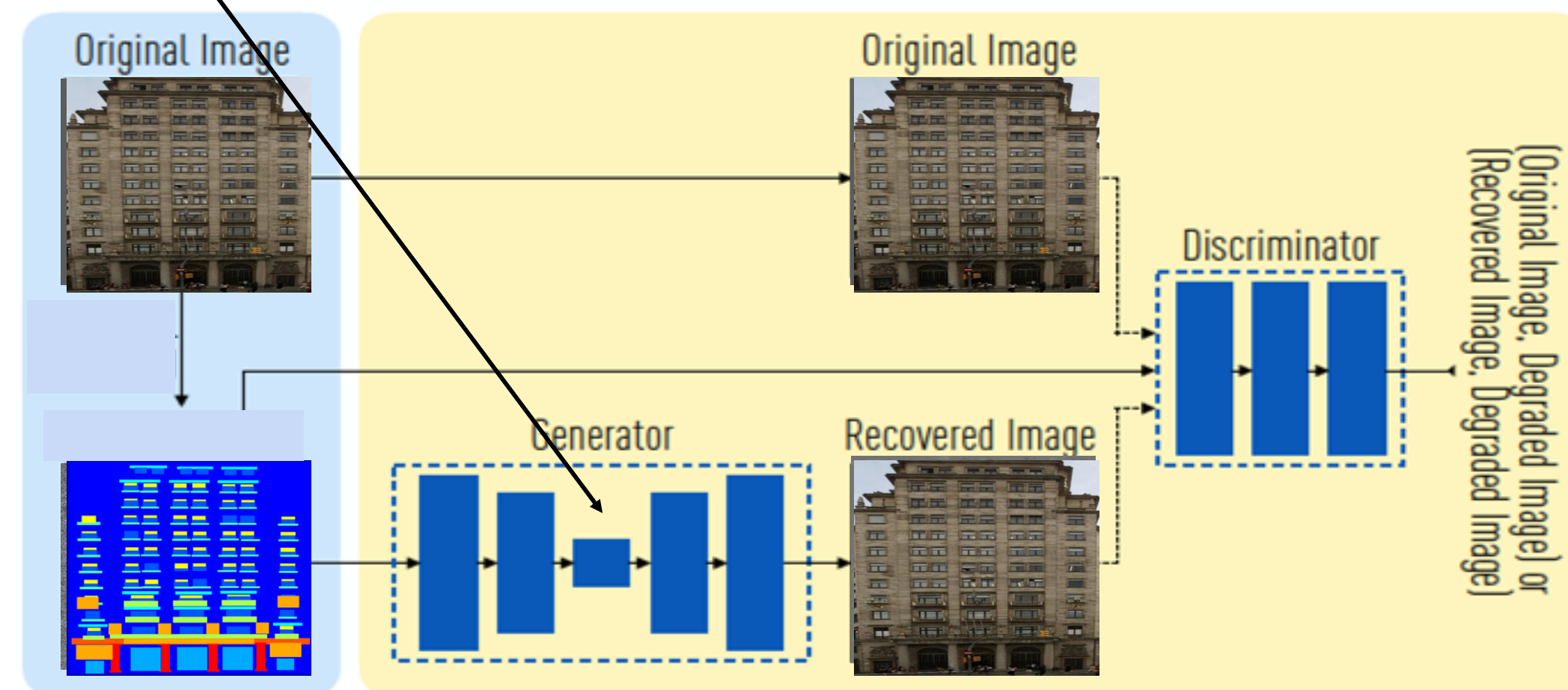
Pix2Pix

Pix2Pix 先抽取影像特徵，再放大還原成另一個風格影像

DCGAN
generator



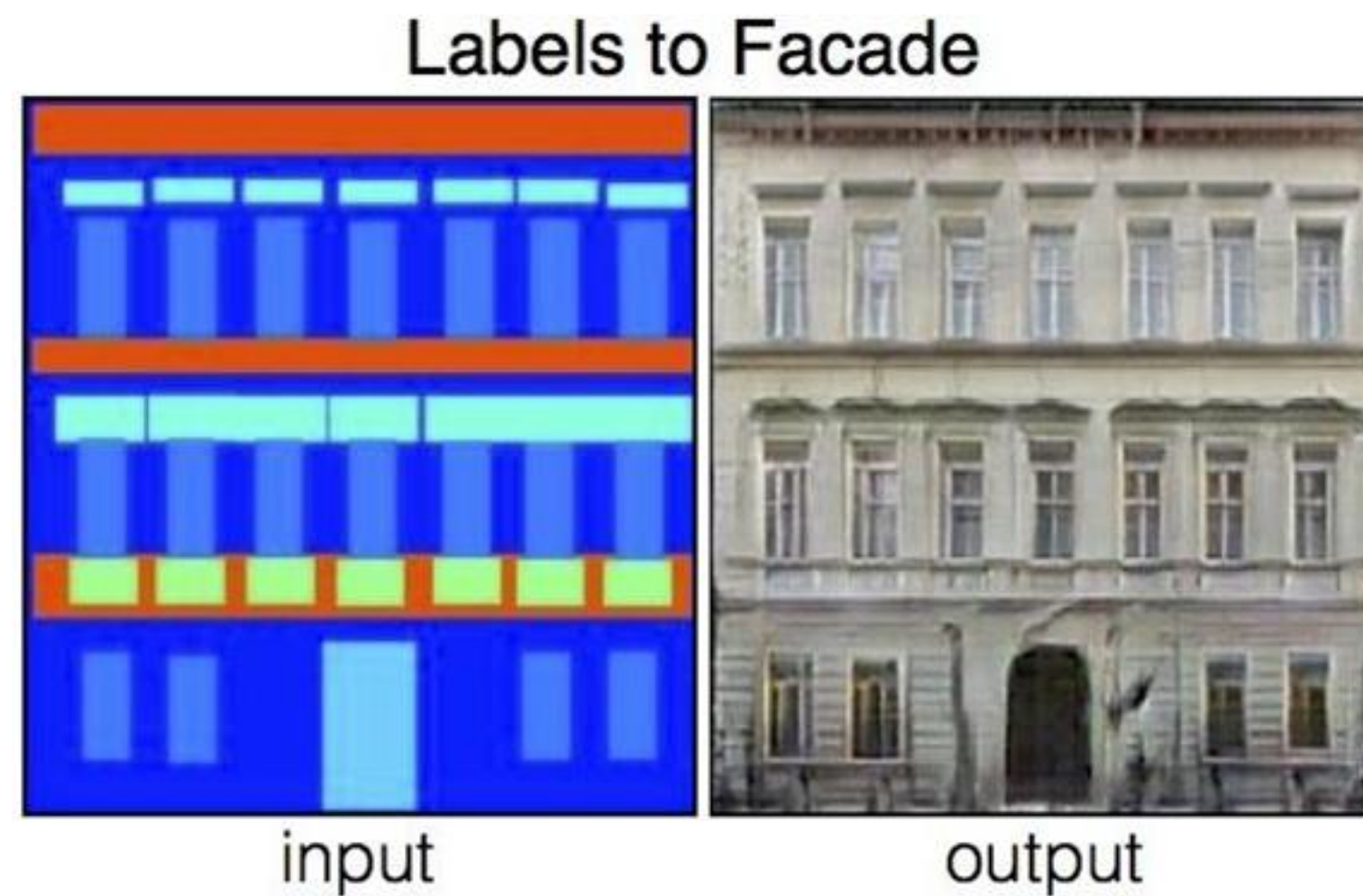
Pix2Pix



Pix2Pix

03_Condition_GAN/01_Pix2Pix.ipynb

- 手把手帶大家刻 Pix2Pix
- 資料集: Labels to Facade (256x256, RGB)

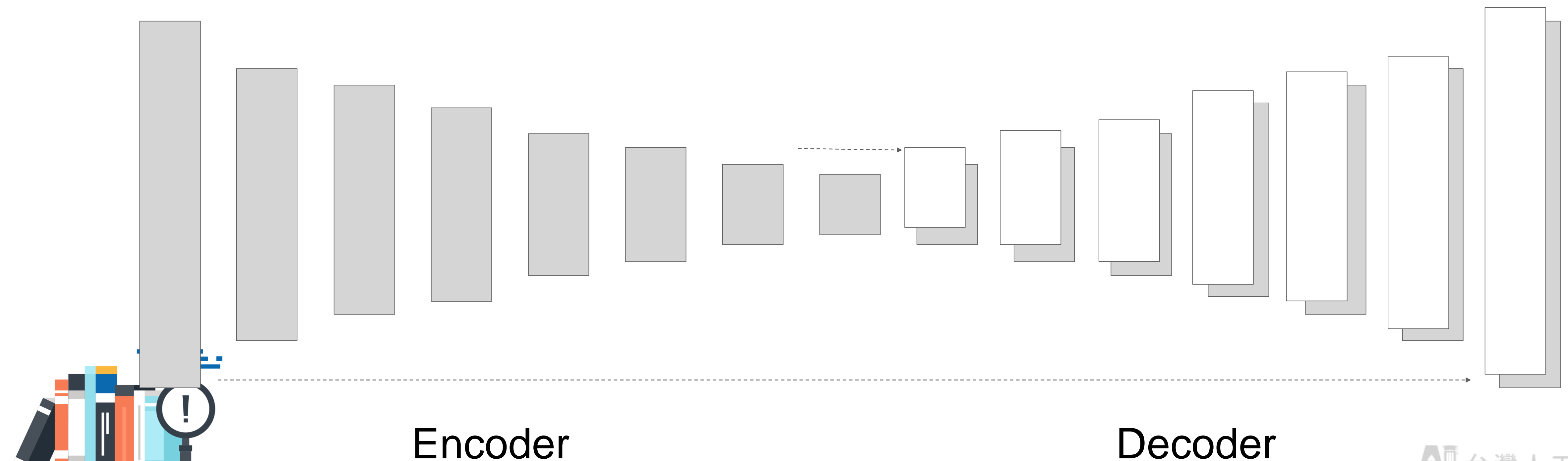


Pix2Pix

03_Condition_GAN/01_Pix2Pix.ipynb

Generator- Unet

- 前半段 Encoder feature maps 連結到後半段 Decoder 對應 feature maps



Pix2Pix

Generator

```
1 def generator(x, dim, out_dim):
2     '''
3     論文當中的 generator 採用 U-Net，前半段 Encoder feature maps 複製到後半段 Decoder 對應 feature maps。
4     ...
5     # encoder, decoder 每層 filter 數量
6     e_dims = [dim] + [dim*2] + [dim*4] + [dim*8]*5
7     d_dims = e_dims[::-1][1:]
8     e_list = []
9
10
11     with tf.variable_scope('generator', reuse=tf.AUTO_REUSE):
12         # Encoder: 影像->特徵向量
13         for i in range(8):
14             if i == 0:
15                 x = tf.layers.conv2d(x, e_dims[i], 4, strides=2, padding='SAME')
16             elif i == 7:
17                 x = tf.nn.leaky_relu(x)
18                 x = tf.layers.conv2d(x, e_dims[i], 4, strides=2, padding='SAME')
19             else:
20                 x = tf.nn.leaky_relu(x)
21                 x = tf.layers.conv2d(x, e_dims[i], 4, strides=2, padding='SAME')
22                 x = tf.layers.batch_normalization(x, training=True)
23
24             e_list.append(x)
25
26         e_list = e_list[::-1][1:]
27
28         # Decoder: 特徵向量->影像
29         for i in range(7):
30
31             x = tf.nn.relu(x)
32             x = tf.layers.conv2d_transpose(x, d_dims[i], 4, strides=2, padding='SAME')
33             x = tf.layers.batch_normalization(x, training=True)
34
35             if i < 3:
36                 x = tf.nn.dropout(x, 0.5)
37
38             x = tf.concat([x, e_list[i]], 3)
39
40             x = tf.nn.relu(x)
41             x = tf.layers.conv2d_transpose(x, out_dim, 4, strides=2, padding='SAME')
42
43             x = tf.nn.tanh(x)
44         return x
```



Pix2Pix

03_Condition_GAN/01_Pix2Pix.ipynb

Discriminator

```
46 def discriminator(x, dim):
47     with tf.variable_scope('discriminator', reuse=tf.AUTO_REUSE):
48
49         x = tf.layers.conv2d(x, dim, 4, strides=2, padding='SAME')
50         x = tf.nn.leaky_relu(x)
51
52         for i in range(3):
53             if i == 2:
54                 stride = 1
55             else:
56                 stride = 2
57
58             x = tf.layers.conv2d(x, dim*(2**(i+1)), 4, strides=stride, padding='SAME')
59             x = tf.layers.batch_normalization(x, training=True)
60             x = tf.nn.leaky_relu(x)
61
62         x = tf.layers.flatten(x)
63         x = tf.layers.dense(x, 1)
64
65     return x
66
```



生成影像好壞評估



台灣人工智慧學校_對抗式生成網路

File Edit View Insert Format Slide Arrange Tools Add-ons Help Last edit was made 19 minutes ago by 黃凱翔

Background... Layout... Theme... Transition...

不適合使用 GAN (個人經驗)

生成分布與特定分布相似性要求極高的領域

- 例子: 蛋白質影像去噪



Click to add speaker notes



GAN 怎麼衡量生成影像好壞？

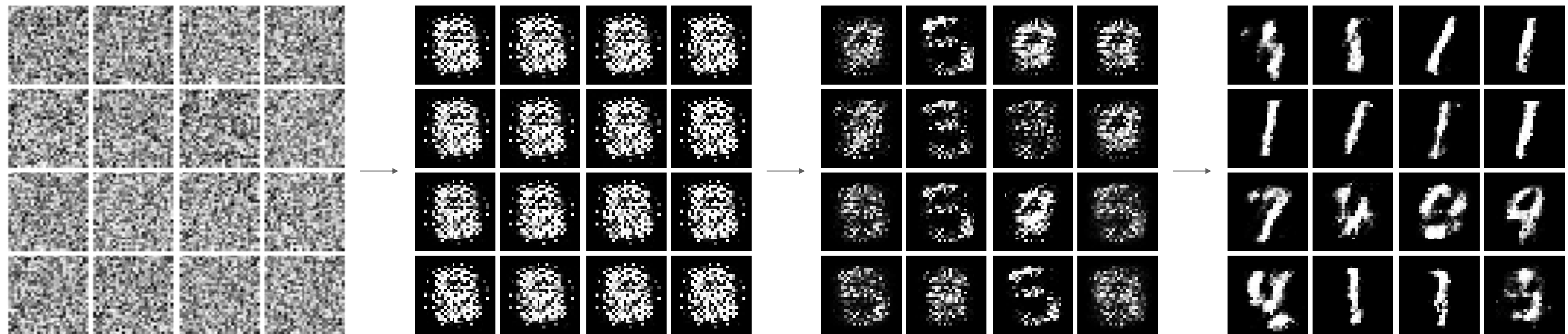


答案是...

Visual quality (視覺質量)
簡單來說就是用人眼看！



Visual quality



Visual quality

缺點

- 每個人主觀標準不同
- 好壞無法量化



客觀指標

1. Inception Score (IS)
2. Frechet Inception Distance (FID)



客觀指標 - Inception Score

- **單一樣本的獨特性**
 - 一張生成影像能夠被明確歸類到某個類別
- **多樣本之間的差異性**
 - 把多張生成影像擺在一起評估的時候，他們不應該長得一樣。應該要被分類成不同的類別

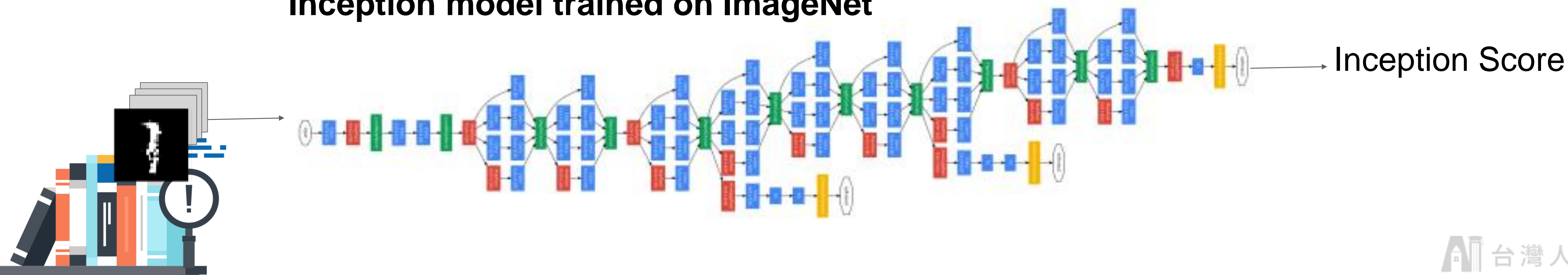


客觀指標 - Inception Score

實際上如何計算 IS ？

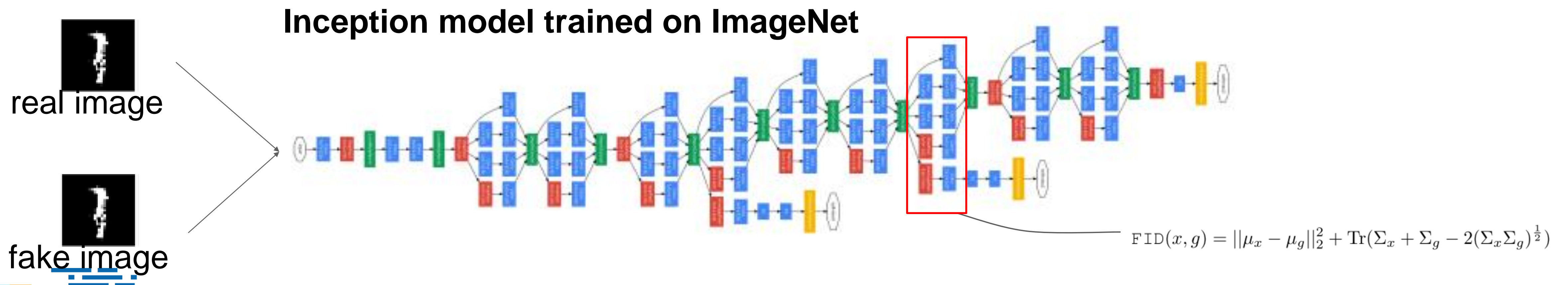
- 使用在 ImageNet 上面訓練好的 Inception model 當作評估模型
- 此評估模型能夠辨識大部分物體類別
- 將一批影像輸入模型，取 softmax 層
- 考慮單一樣本分類機率分布算 **entropy** (越小代表能夠生成明確類別的影像)
- 考慮所有樣本的平均分類機率分布算 **entropy** (越大代表這批影像出現很多類別)
- 詳細請參考 [paper](#), [github](#)

Inception model trained on ImageNet



客觀指標 - Frechet Inception Distance

- 將 real image、fake image 分別輸入 Inception model
- 在特定層取出 feature maps
- 計算兩者分布的差異 (平均及共變異數，假設多變量高斯分布)
- 詳細請參考 [paper](#)



GAN 怎麼衡量生成影像好壞？

小結

- 雖然有量化的指標，但其在很多情況下無法有效衡量生成影像的表現
- 相關研究還在發展中
- 目前主要還是以視覺質量評斷生成質量的高低



GAN 適用範圍

真的什麼都可以用 GAN 來生成嗎？



適合使用 GAN (個人經驗)

生成分布與特定分布相似性要求不高的領域

- 藝術
- 音樂
- 詩詞



送娃寄窗天，悠爐凋陽薇
來醒咎好地，已慣遇吾堪

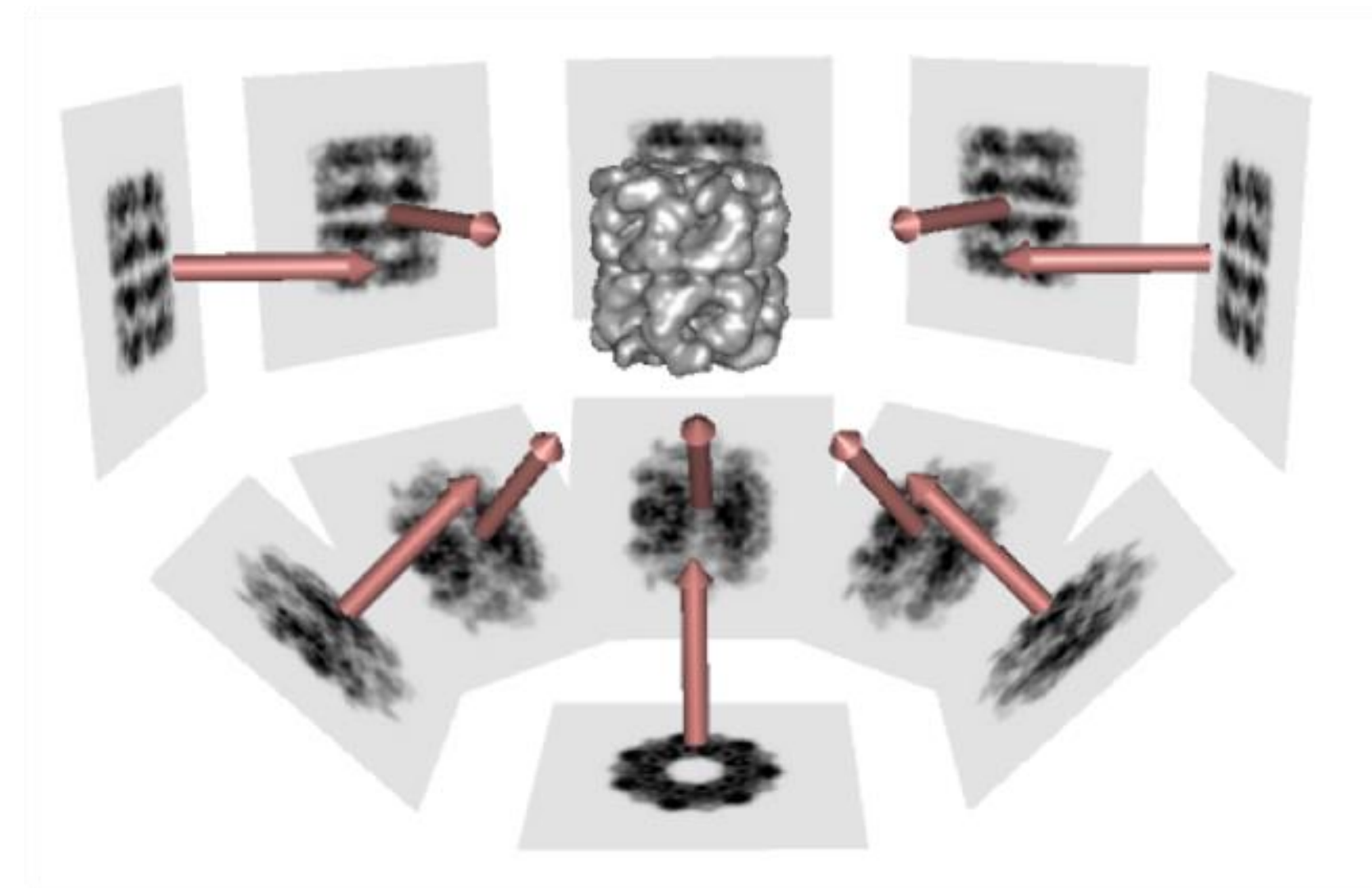
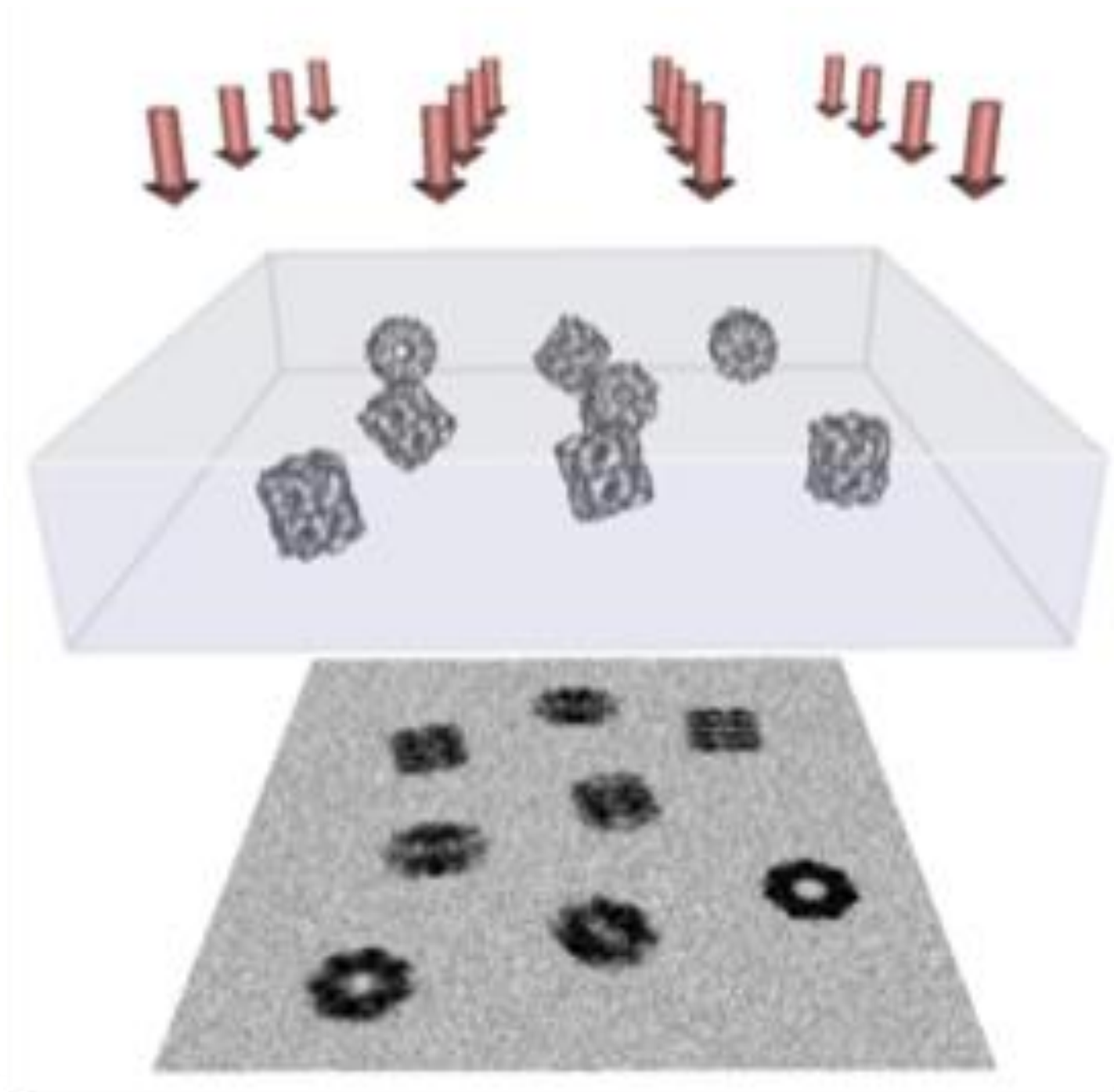
○ 可接受



不適合使用 GAN (個人經驗)

生成分布與特定分布相似性要求極高的領域

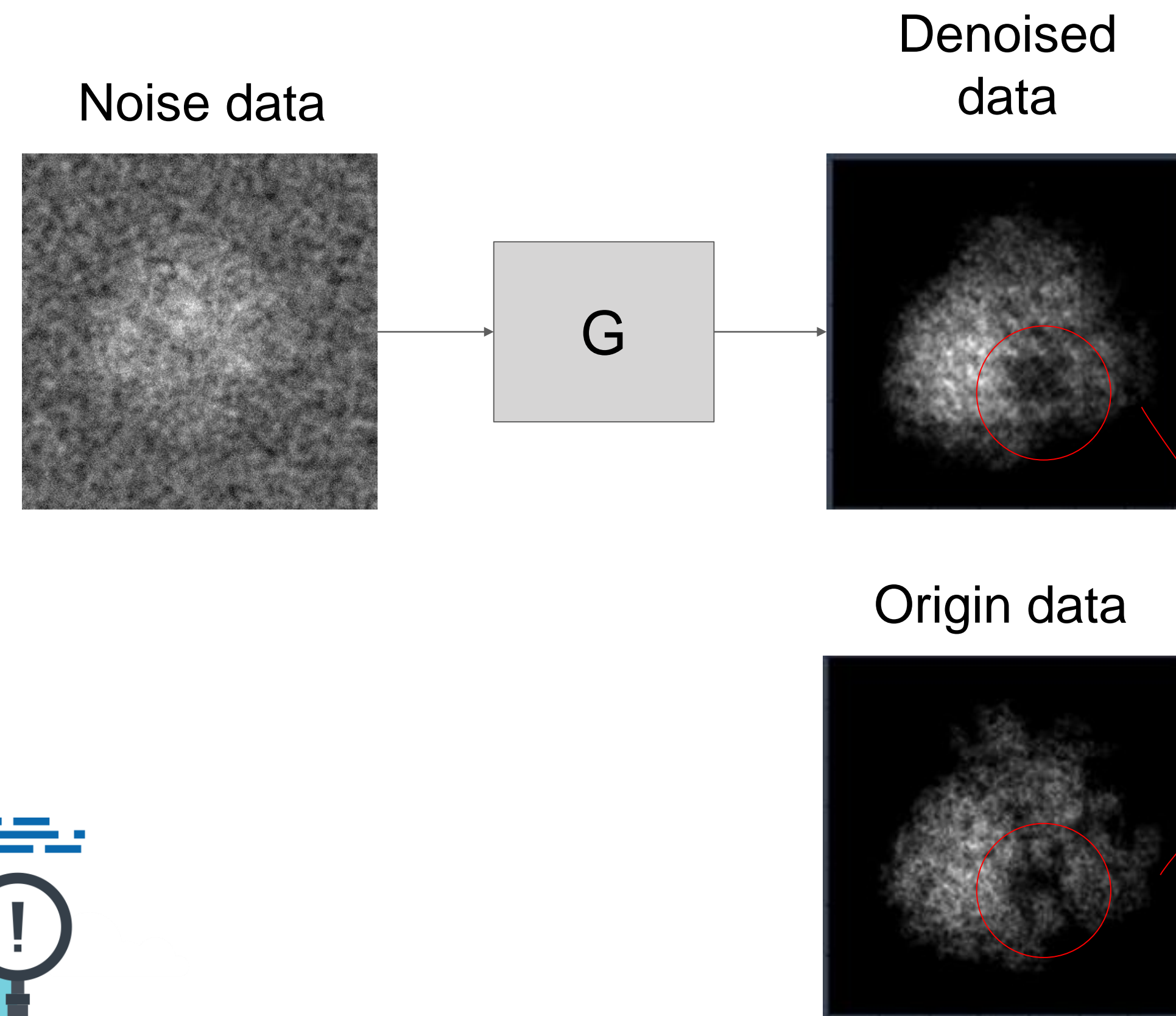
- 例子: 蛋白質影像去噪



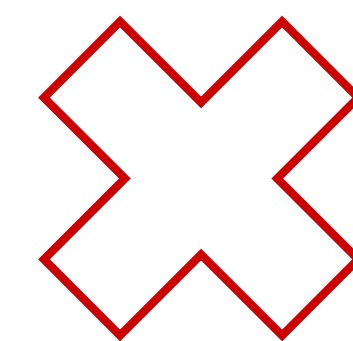
不適合使用 GAN (個人經驗)

生成分布與特定分布相似性要求極高的領域

- 例子: 蛋白質影像去噪



生成的影像多了原本沒有的部分



不可接受

