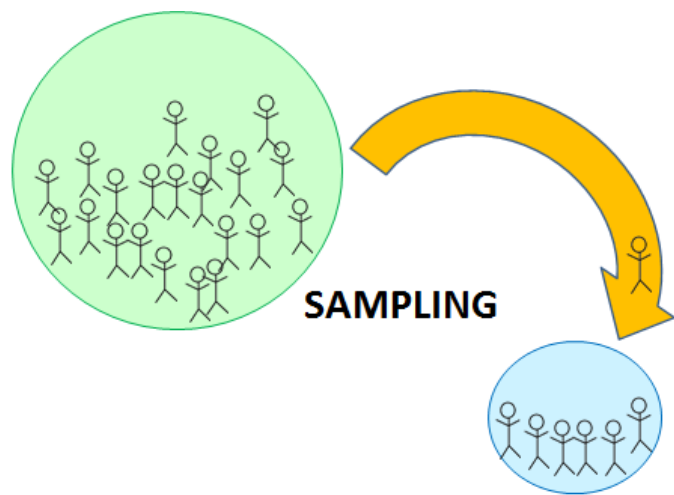


Generative Neural Network



Generative Model

- 生成性(generative)：模型、規則(產生新事物)
- 創造性(creative)：洞察力、想像力(產生新事物)
- 人工智能的生成模型：機率分佈、隨機取樣



創造力來源(心理學角度)：

1. 聯想主義(聯繫不同想法)
2. 分歧思維(多樣、非常規)
3. 流動狀態(立即回饋、創意迭代)

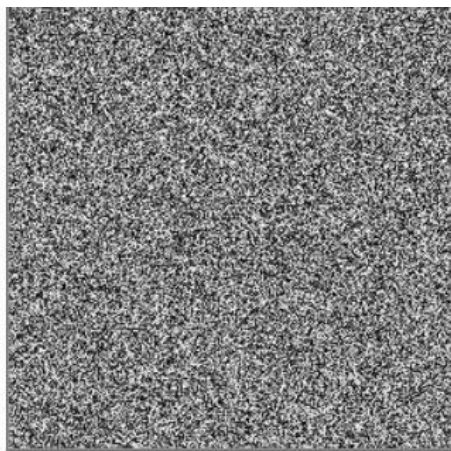
智力五要素：

1. 專注力(手機在視線外)
2. 觀察力(工具、好奇心)
3. 記憶力(藍莓、草莓)
4. 思考力(知識、邏輯)
5. 想像力(個人、抽象，無單一具體訓練方法)

Generative Model

舊約聖經創世紀1章1-2節

Q：生成式模型的輸入來源是什麼？



雜訊？



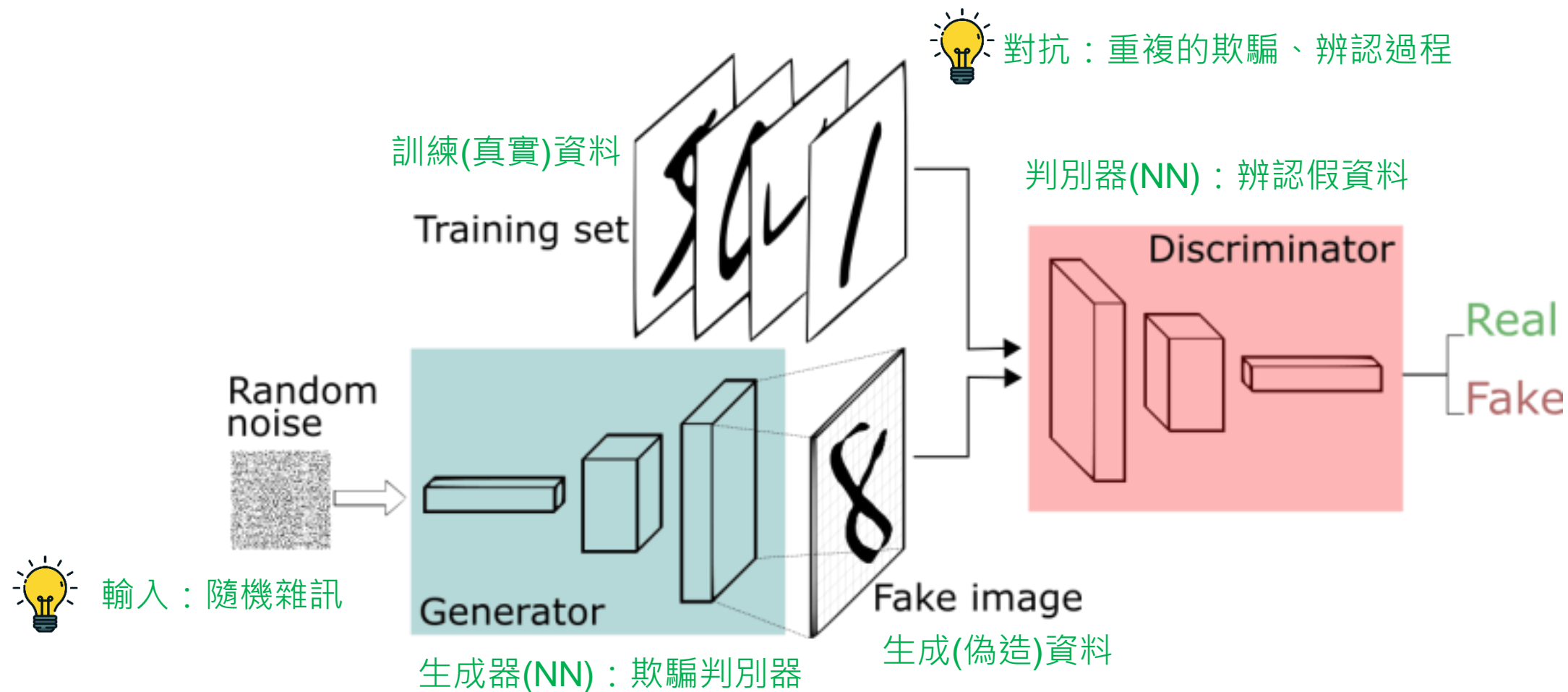
¹起初，神創造天地。
²地是空虛混沌，淵面黑暗；神的靈運行在水面上。

人工智能目前的想像力與人類的想像力是有很大差異的。人類的想像力源自於複雜的情感、經驗、文化背景和無限的聯想能力，而AI則基於數據、算法和預先訓練的模式來生成內容。

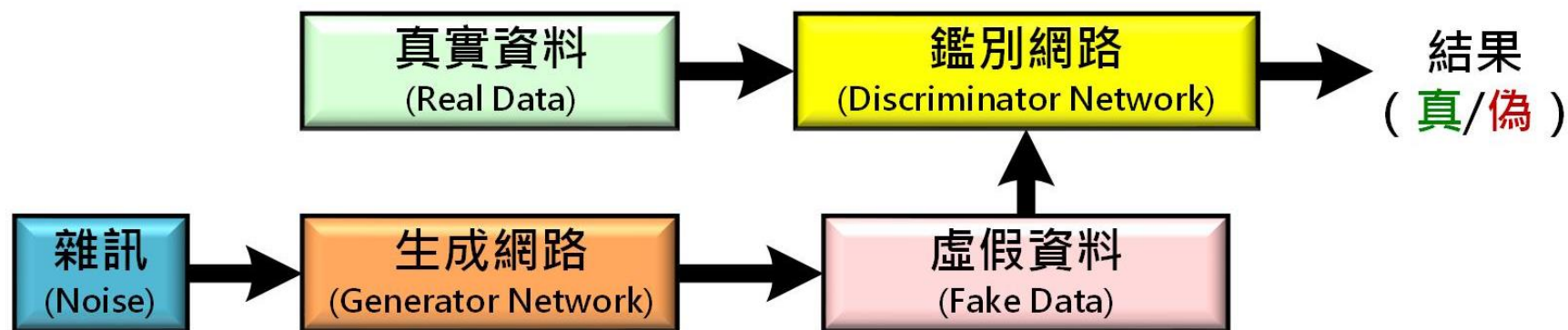
Generative Neural Network

- 生成式神經網路是一種機器學習模型，旨在生成新數據樣本，某種程度上模擬訓練數據的特徵
- 生成對抗網路(Generative Adversarial Network, GAN)：
 - 由生成器和判別器兩個部分組成，生成器負責生成假數據，而判別器則判斷數據是真實還是生成的。兩者通過對抗過程來提升彼此的性能。
- 自編碼器(Autoencoder)：
 - 編碼器(Encoder)：將輸入數據映射到潛在空間(latent space)
 - 解碼器(Decoder)：將潛在向量轉換回原始數據的近似值
- 變分自編碼器(Variational Autoencoder, VAE)：
 - 編碼器(Encoder)：將輸入數據映射到潛在空間(latent space)
 - 重參數化(Parameterization)：將隨機性引入潛在變數
 - 解碼器(Decoder)：將潛在向量轉換回原始數據的近似值
- 擴散模型(Diffusion Model)
 - 正向過程(Forward Process)：隨機添加雜訊，逐漸「破壞」數據結構
 - 逆向過程(Reverse Process)：逐步去除噪聲以生成與訓練數據相似的樣本

Generative Adversarial Network (GAN)



Generative Adversarial Network (GAN)



成年→兒童

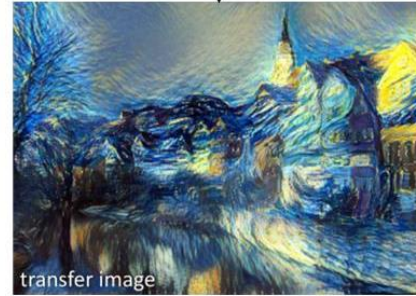
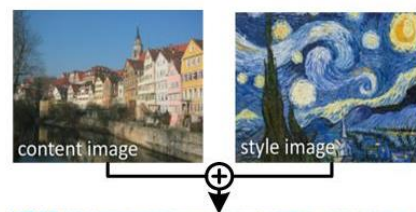
女生→男生



成年→老年

真實→卡通

人臉生成(變臉APP)

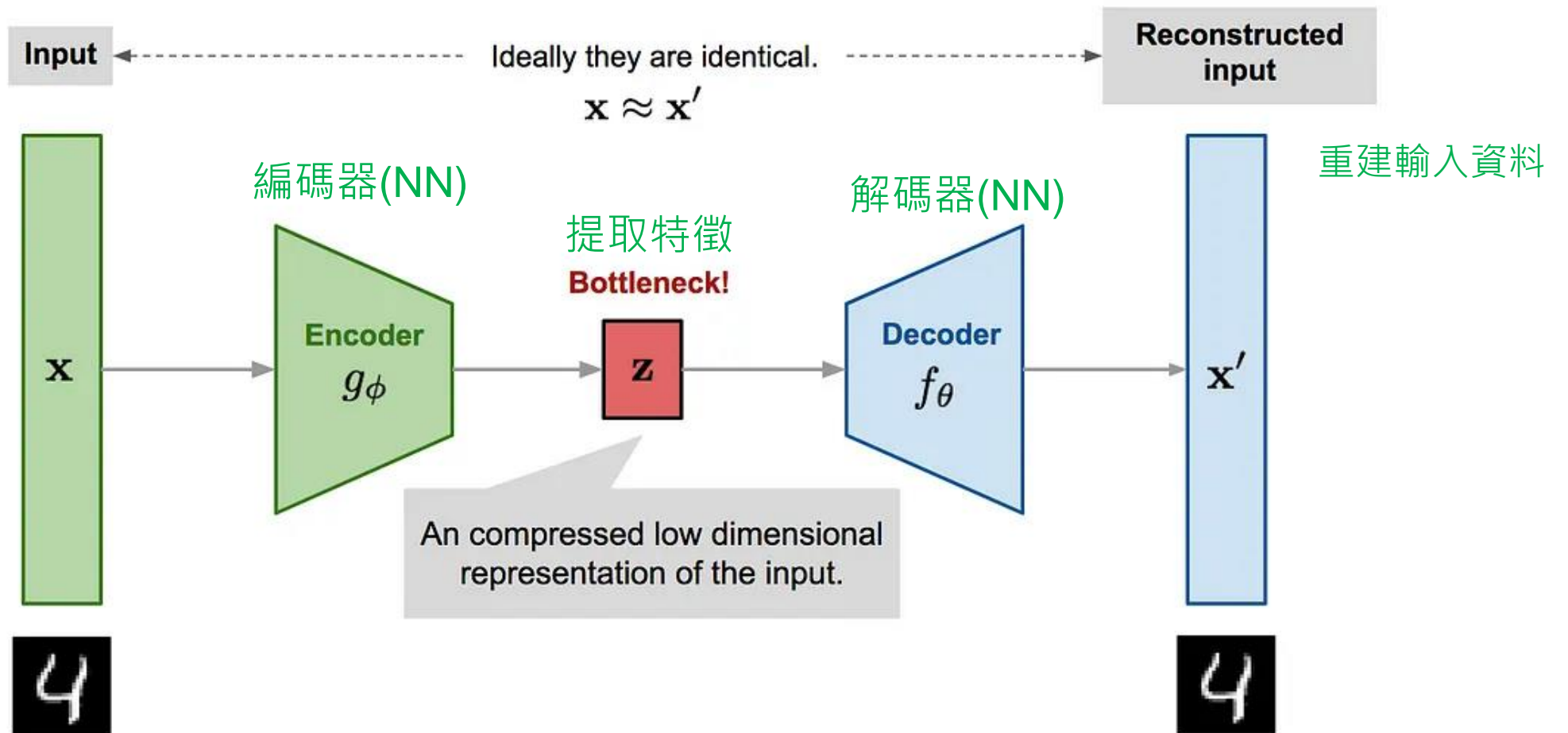


風格轉換



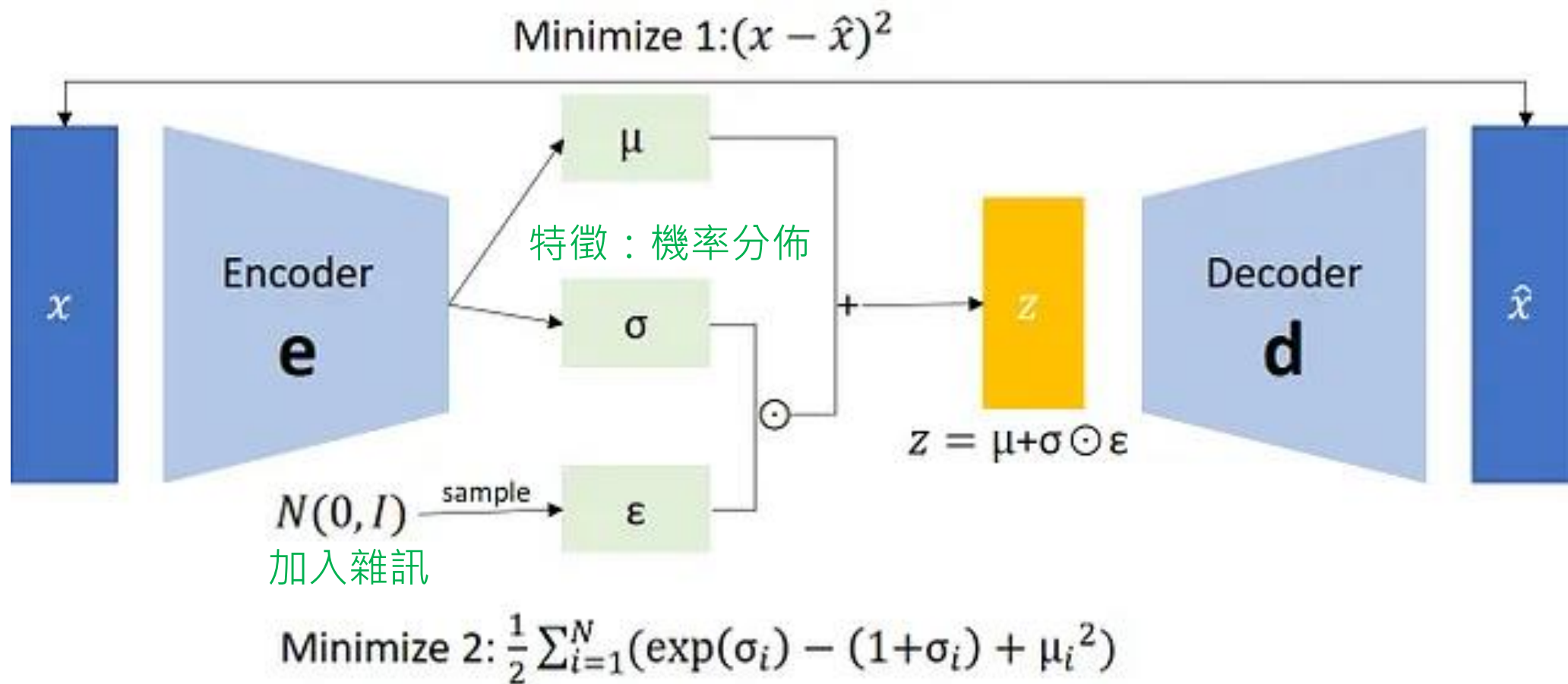
圖像修復(填補/上色/去糊)

Autoencoder



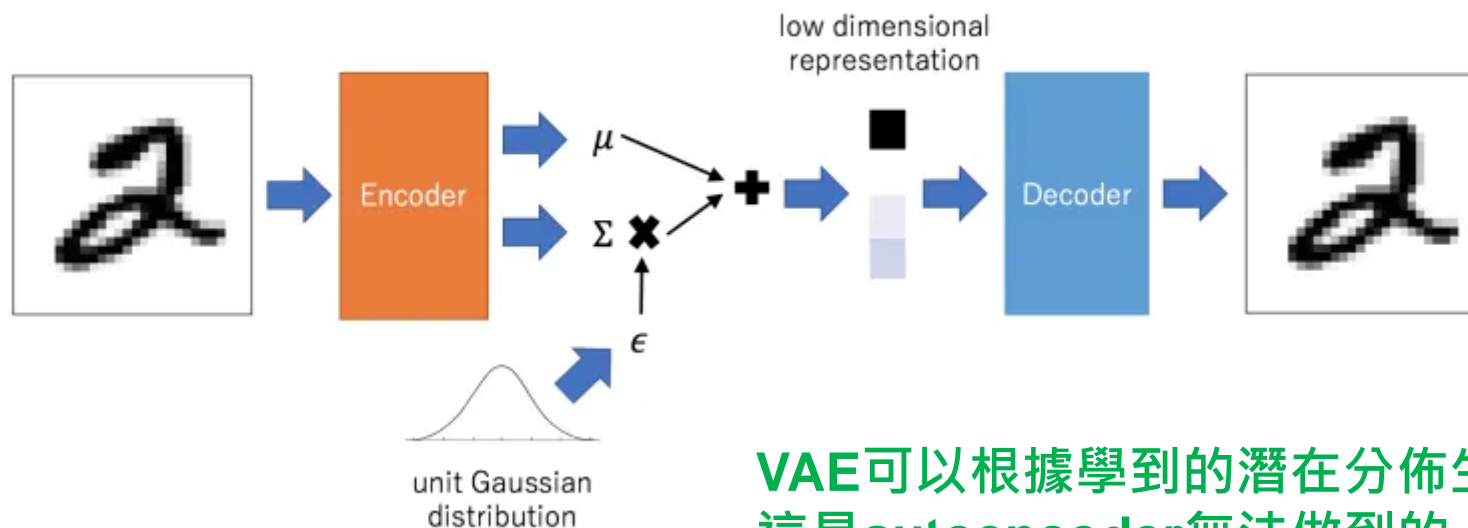
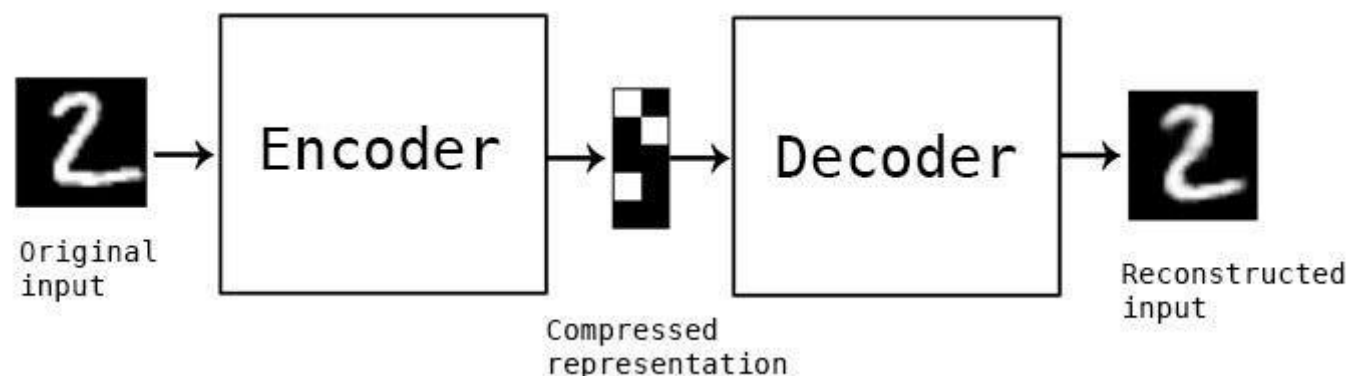
直接從編碼器輸出重建圖像：生成圖像與輸入圖像非常相似

Variational Autoencoder (VAE)



從潛在空間的機率分佈中採樣，然後通過解碼器生成圖像：生成圖像與輸入圖像類似，但不完全相同

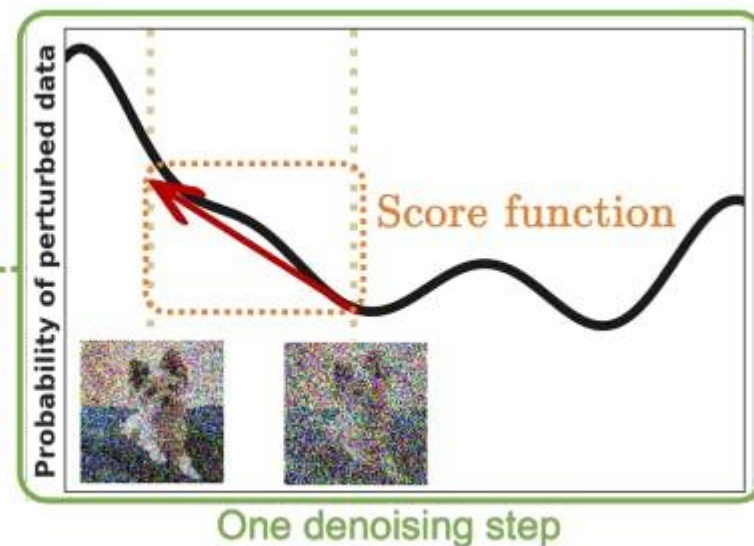
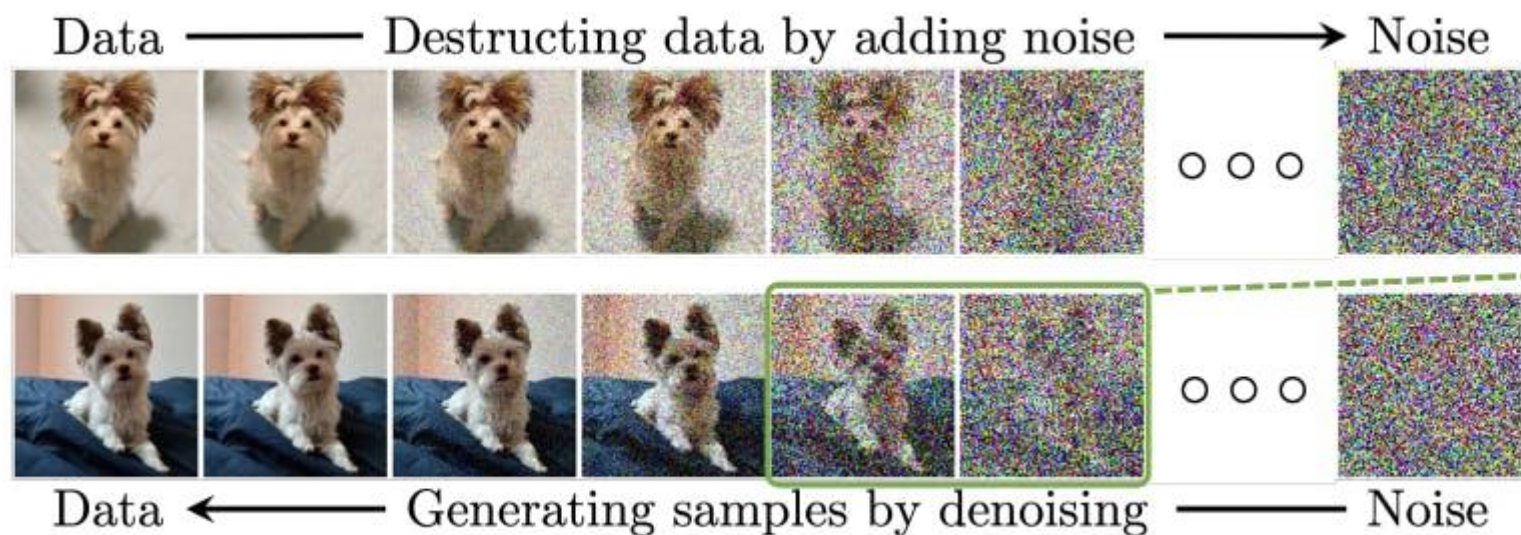
Autoencoder and VAE



VAE可以根據學到的潛在分佈生成全新的圖像，這是autoencoder無法做到的

Diffusion Model

正向過程：加雜訊



逆向過程：去雜訊

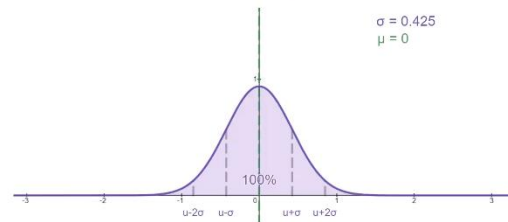
生成圖像並不一定會與輸入圖像完全相同

Probability Distribution



27:39

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]}$$



大白话啦

也就是随机变量所有取值概率的总和为100%



生成模型是一個機率分佈，經由隨機採樣產生輸出結果

生成模型	生成對抗網路	自編碼器	變分自編碼器	擴散模型
模型架構	由生成器和判別器組成 生成器學習數據分佈， 判別器學習區分真假樣本	編碼器-解碼器架構，編碼器壓縮數據，解碼器重構數據	編碼器-解碼器架構，並引入概率模型	單一模型，基於馬爾可夫鏈多步去除雜訊，生成高質量數據
生成過程	一步生成(生成器直接輸出)	一步解碼(從潛在空間生成數據)	一步生成(從潛在變量生成數據)	多步生成過程，從純雜訊開始逐步去除雜訊，經過多個步驟生成樣本
訓練方法	對抗訓練，生成器和判別器交替訓練，試圖互相改進	重構損失最小化	通過最大化潛在變量的證據下界來訓練	學習逆向擴散過程，通過多步去噪生成數據樣本，使用變分下界優化
潛在空間	無明確潛在空間結構	有固定的潛在空間	學習潛在空間的分佈，假設潛在空間為正態分佈	無需顯式潛在空間，生成樣本是從噪聲分佈中逐步生成
應用場景	常用於圖像生成、圖像翻譯、文本生成、視頻生成	數據壓縮、特徵提取、降維等應用	數據壓縮、圖像生成、生成式數據增強等任務	圖像生成、數據補全、圖像修復、風格轉換等應用 特別適合高質量圖像生成的任務
損失函數	生成器和判別器的對抗損失	重構損失	重構損失+KL散度	基於變分下界的多步去噪損失

生成模型比較

- **生成過程**：GAN 是透過對抗學習生成圖像；VAE 是透過概率取樣生成；擴散模型則是透過逆向噪聲過程逐步生成。
- **圖像質量**：GAN 和擴散模型生成的圖像質量更高，細節更豐富。VAE 生成的圖像相對模糊。
- **潛在空間特性**：VAE 和擴散模型學習的潛在空間有明確的概率意義，GAN 的潛在空間則沒有這種連續性或概率性。
- **穩定性**：VAE 和擴散模型的訓練過程較穩定，而 GAN 的訓練過程可能會出現不穩定情況，如模式崩潰。
- **生成速度**：GAN 的生成速度最快，擴散模型生成速度較慢。

隨機雜訊

- 生成模型(如生成對抗網絡和擴散模型)通常是透過從隨機分佈(通常是高斯雜訊)開始，然後逐步將這些噪聲轉換成特定的目標數據(例如圖像或文本)。以下是一些使用噪聲的主要原因：
- **隨機性與多樣性：**
 - 雜訊提供隨機起點，允許模型生成不同的樣本，這樣模型每次運行時都會產生不同的結果，從而提升生成數據的**多樣性**。
- **生成數據空間的探索：**
 - 雜訊是一種**低維的隱變量**，生成模型學習如何從這些低維雜訊中提取出**高維的複雜**數據(例如圖像)。這使得模型能夠有效探索數據空間，生成符合目標分佈的樣本

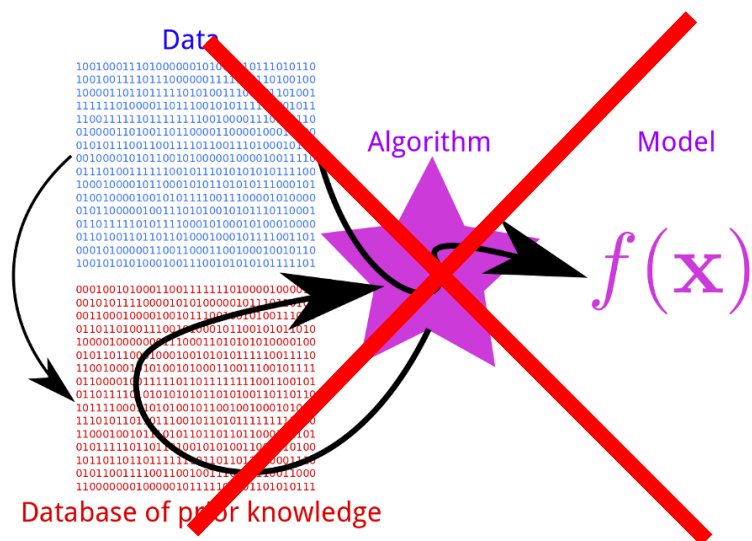


渾沌、太虛：一種原始、未經組織的狀態，一種混沌之中孕育出秩序的意象

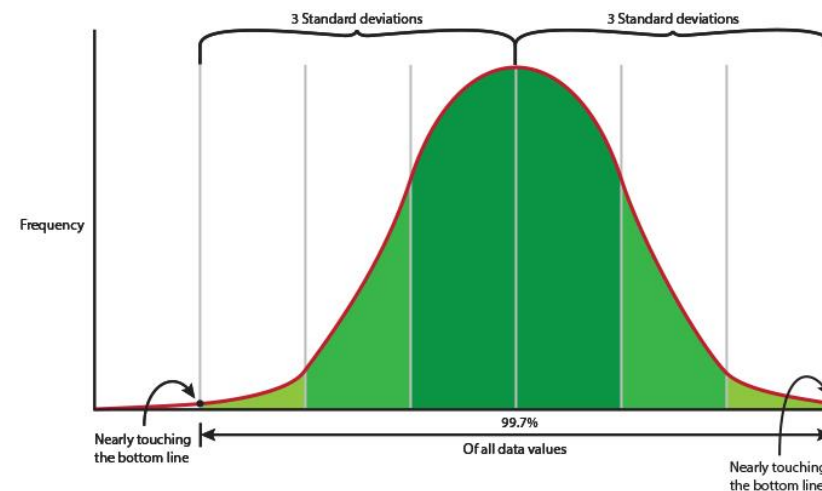
隨機雜訊生成圖片過程



生成模型



函數模型：映射 $z = f(x)$



μ 平均值

σ 標準差

機率模型：取樣 $z \sim p_{\theta}(z)$

AI Art



AI Art



Generative Adversarial
Network

