



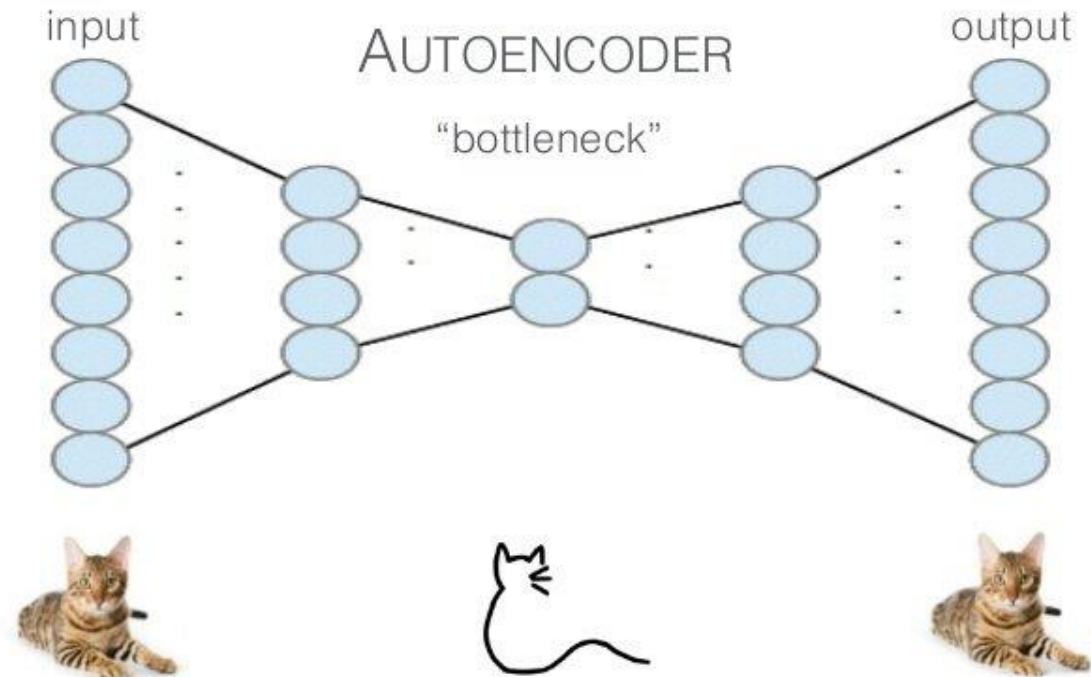
GENERATIVE MODELS

Santiago López

VARIATIONAL AUTOENCODERS

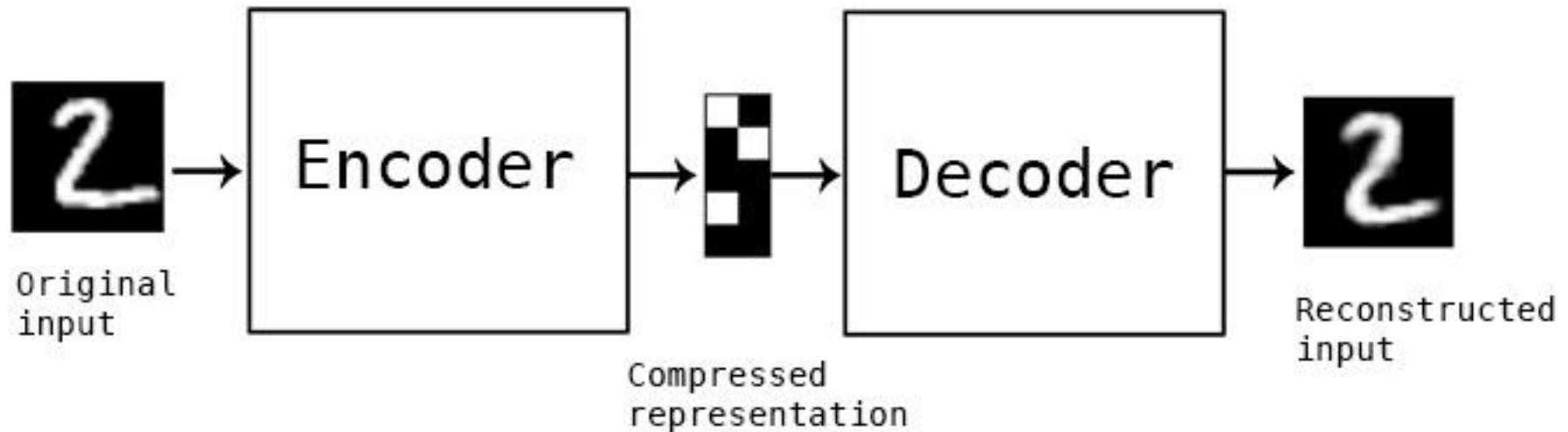
AUTOENCODER

- Modelo computacional por un encoder y un decoder.
- El encoder reduce la dimensionalidad de los datos. Típicamente es un CNN.
- El decoder reconstruye la imagen a partir de los datos codificados.



PROBLEMAS DEL AUTOENCODER

- El espacio latente no es continuo, lo cual no permite interpolar para generar imágenes
- Es como si el modelo se aprendiera las reconstrucciones de memoria y no generaliza.



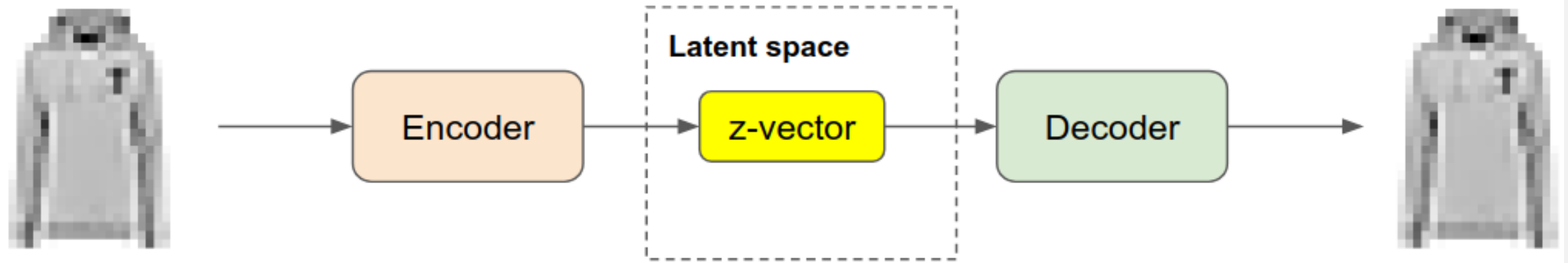
VARIATIONAL AUTOENCODER

- Representa el espacio latente como una distribución de probabilidad.
- Esto permite tomar muestras de la distribución para generar imágenes

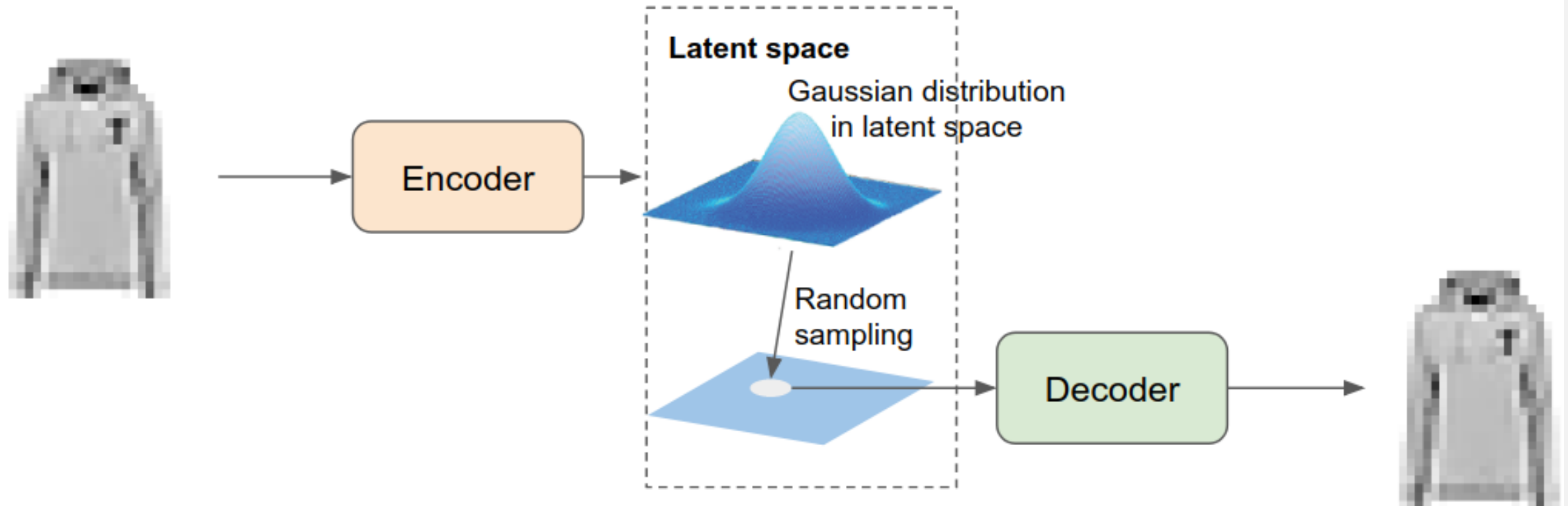
LOSS FUNCTION

- Tienes dos partes: una se encarga de la distribución del espacio latente, y otra de la reconstrucción de la imagen.
- KL divergence se usa para evaluar la función de probabilidad.
- Para evaluar la reconstrucción de la imagen se usan funciones típicas como el MSE (mean squared error) entre los píxeles

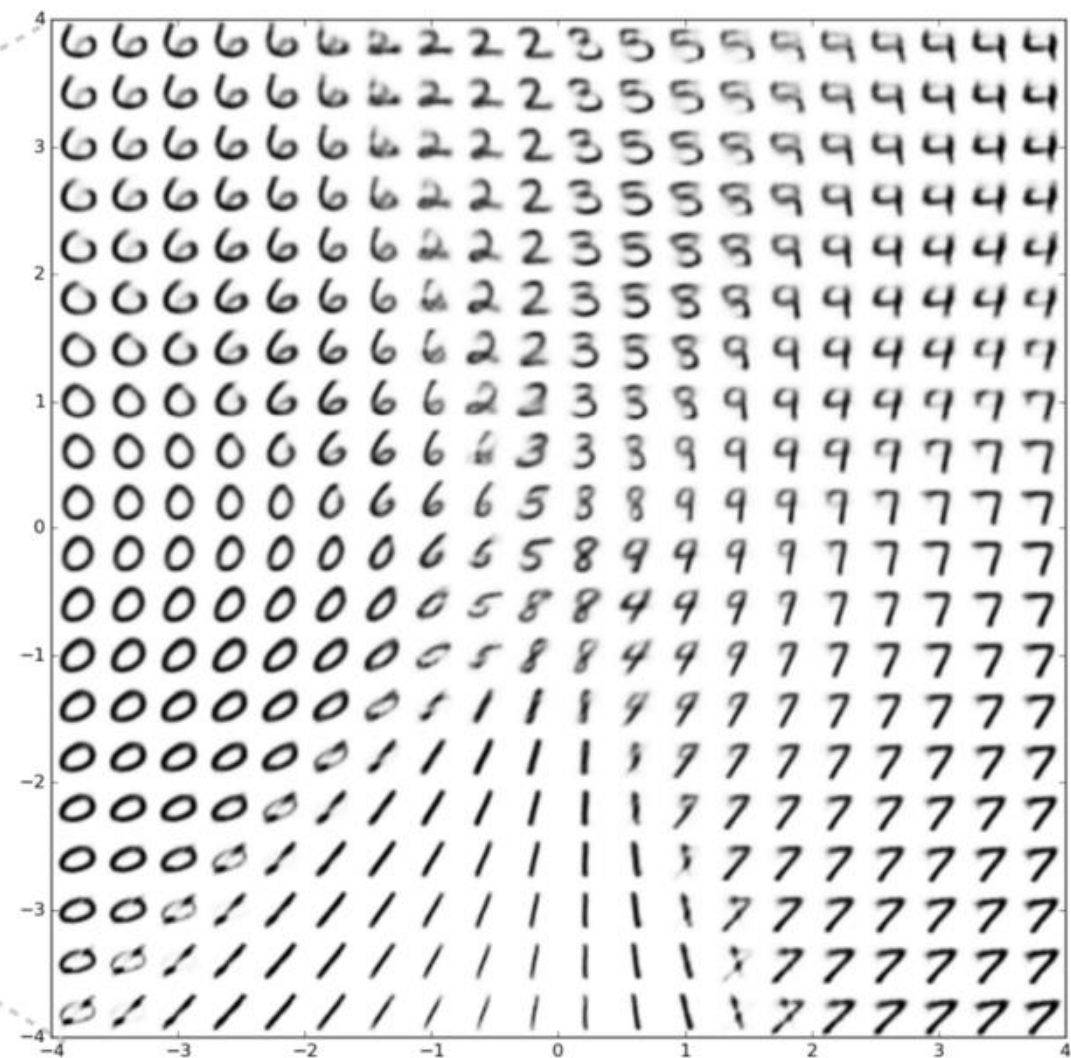
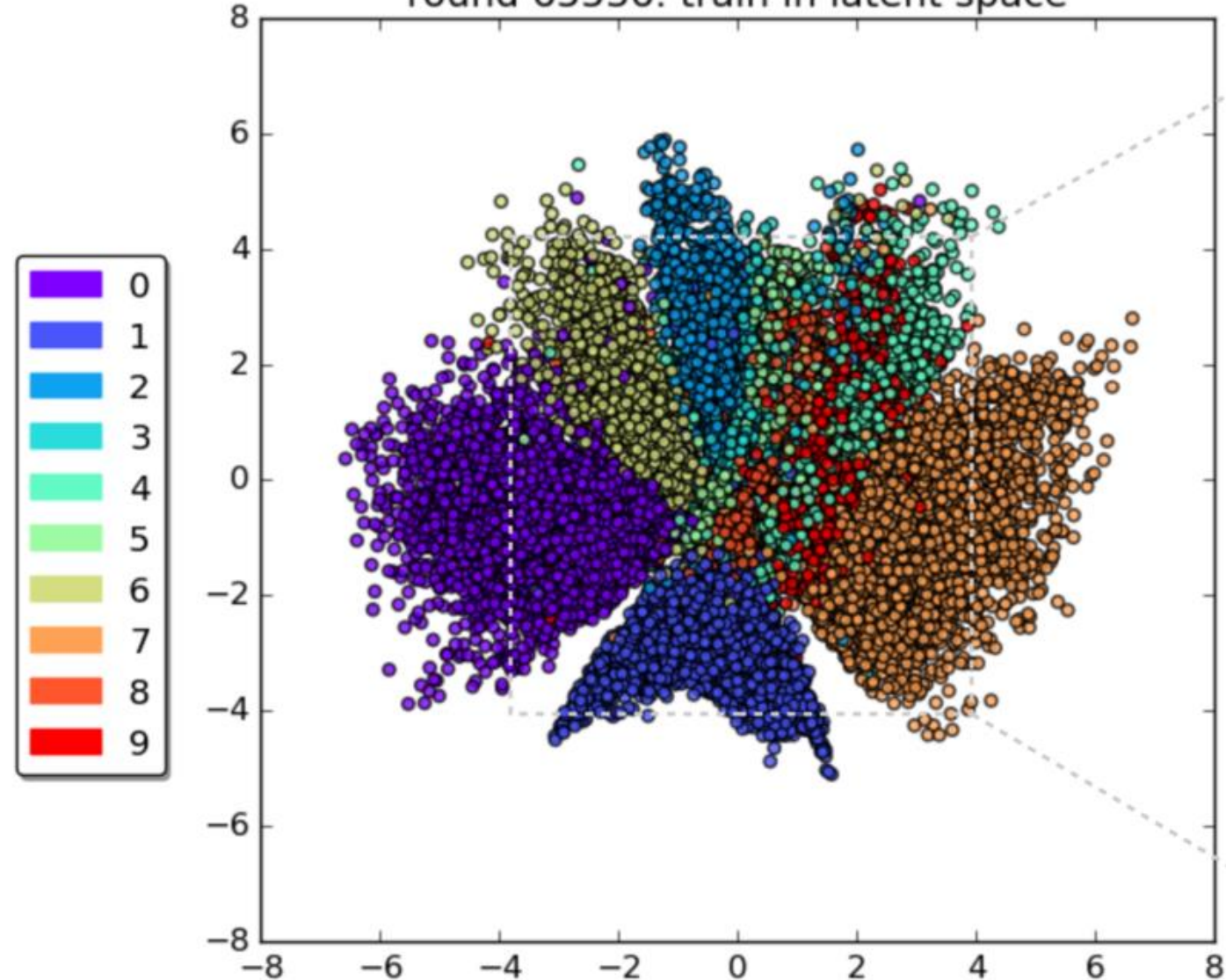
A. How a classical autoencoder works



B. How a variational autoencoder works



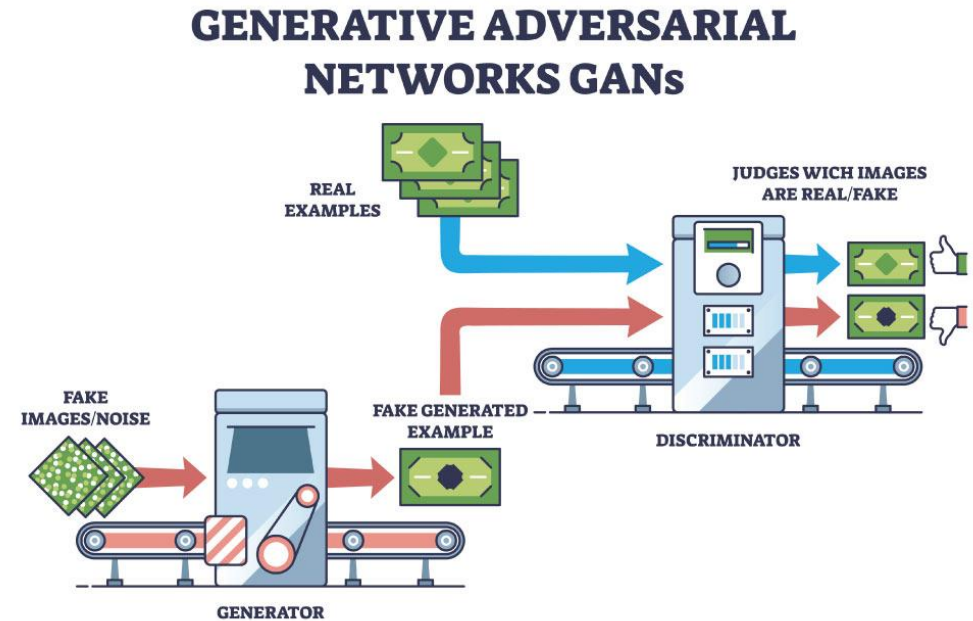
round 65536: train in latent space



GAN (GENERATIVE ADVERSARIAL NETWORK)

¿QUÉ ES?

- Los GANs (Generative Adversarial Networks) son un tipo de modelo generativo que utiliza dos redes neuronales: un Generador y un Discriminador, que compiten entre sí.

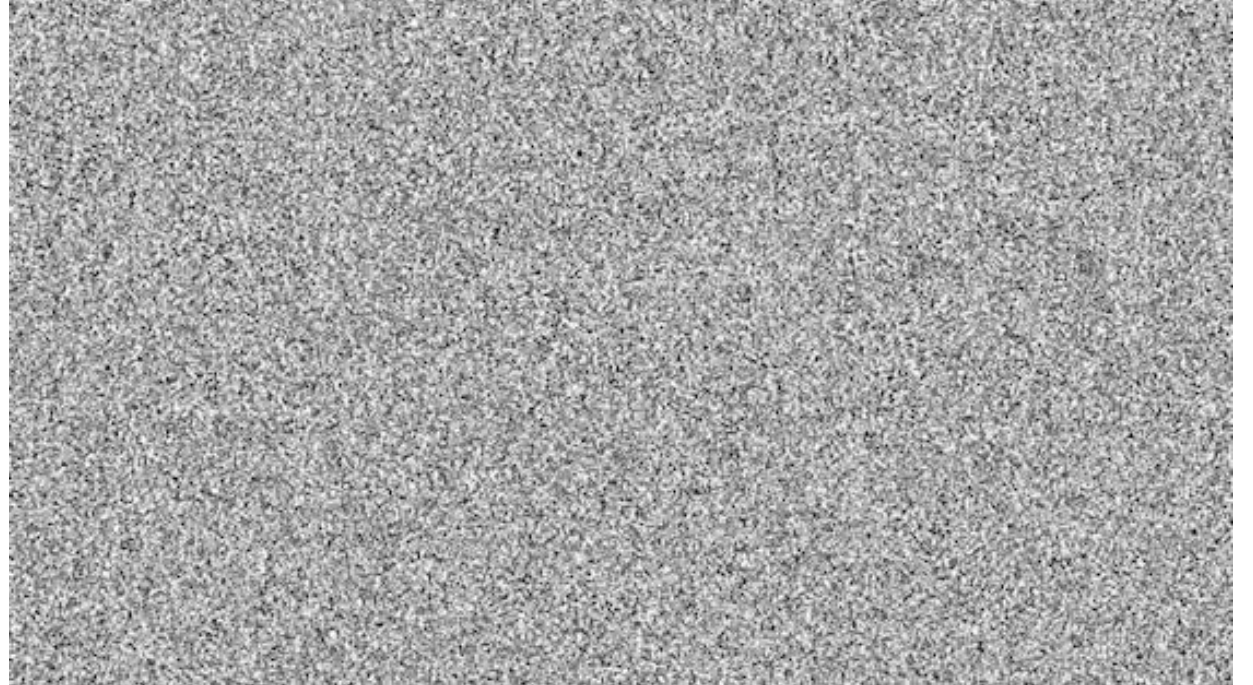


COMPONENTES PRINCIPALES

- Un GAN consta de dos componentes principales:
 1. Generador: Crea datos sintéticos a partir de ruido aleatorio.
 2. Discriminador: Intenta distinguir entre datos reales y datos falsos generados por el generador.

GENERADOR

- El generador toma ruido aleatorio (por ejemplo, de una distribución normal) como entrada.
- Produce datos sintéticos (como imágenes falsas).
- Su objetivo es engañar al discriminador para que clasifique sus salidas como reales.



DISCRIMINADOR

- El discriminador es un clasificador binario.
- Recibe datos reales y falsos (generados por el generador).
- Su objetivo es diferenciar correctamente entre datos reales y falsos.

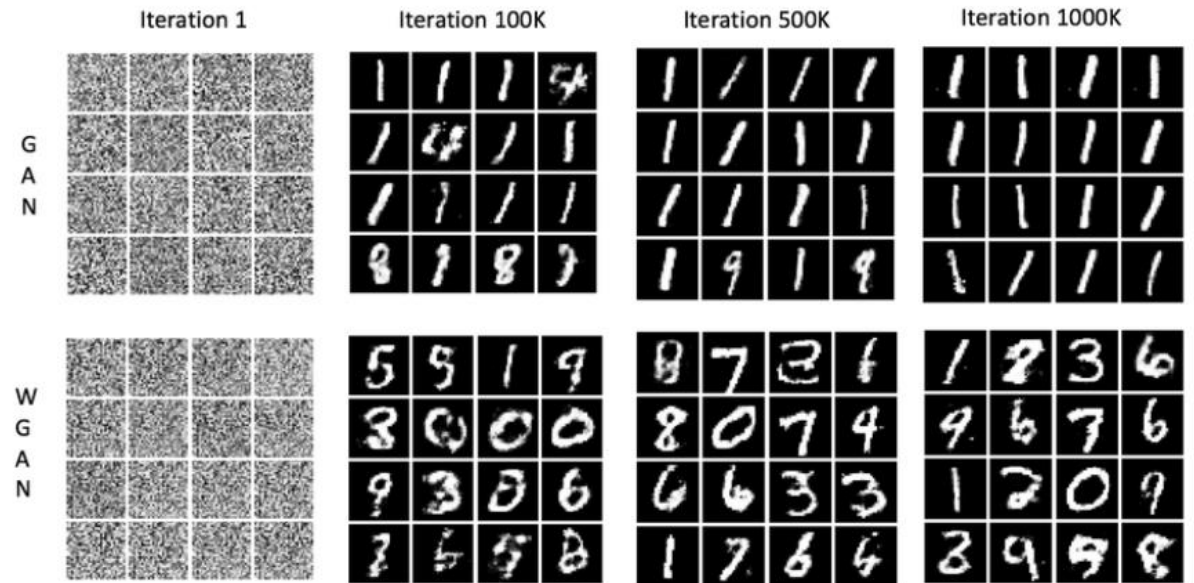


PROCESO DE ENTRENAMIENTO

1. El generador crea datos falsos.
2. El discriminador evalúa datos reales y falsos.
3. El discriminador se entrena para mejorar su precisión.
4. El generador se entrena para engañar mejor al discriminador.
5. Este proceso se repite hasta que el generador produce datos realistas.

DIFICULTADES AL ENTRENAR

- Inestabilidad: El equilibrio entre el generador y el discriminador es difícil de mantener.
- Colapso de moda: El generador produce poca variedad de salidas.



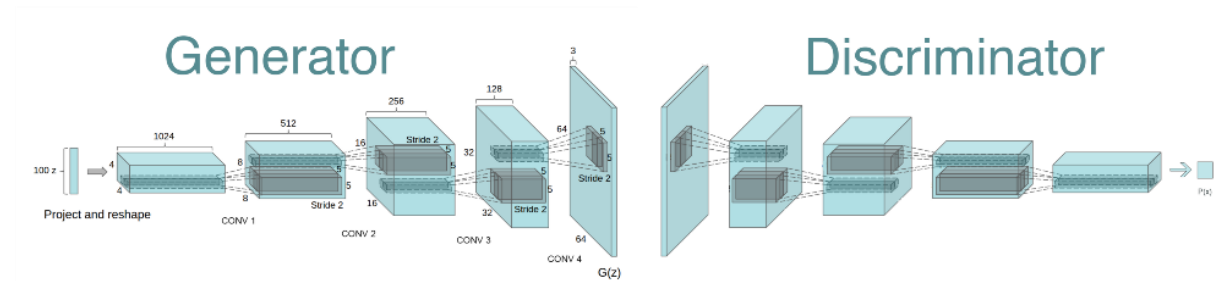
APLICACIONES

- Generación de imágenes: Crear imágenes realistas (por ejemplo, rostros).
- Traducción de imágenes: Convertir una imagen en otro estilo (por ejemplo, foto a pintura).
- Aumentación de datos: Generar datos sintéticos para entrenar otros modelos.
- Super-resolución: Mejorar la calidad de imágenes de baja resolución.
- <https://thispersondoesnotexist.com>

VARIACIONES

DCGAN (DEEP CONVOLUTIONAL GAN)

- Utiliza capas convolucionales en el generador y el discriminador.
- Diseñado específicamente para la generación de imágenes de alta calidad.



WGAN (WASSERSTEINGAN)

- Introduce la distancia de Wasserstein para medir la diferencia entre distribuciones.
- Mejora la estabilidad del entrenamiento y reduce problemas como el colapso de moda.

CYCLEGAN

- Permite la traducción de imágenes sin necesidad de pares de datos.
- Por ejemplo, convertir fotos de caballos en cebras y viceversa.



CONDITIONAL GAN

- Permite controlar la salida del generador mediante información adicional (por ejemplo, etiquetas de clase).
- Útil para generar datos específicos, como imágenes de un tipo particular.

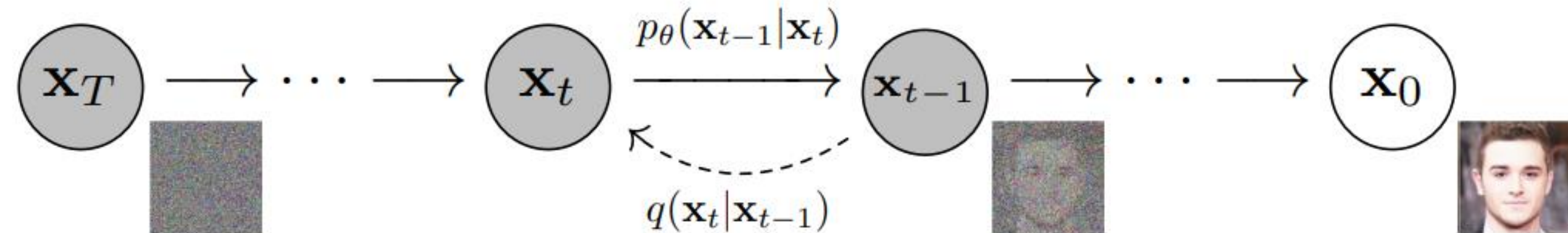
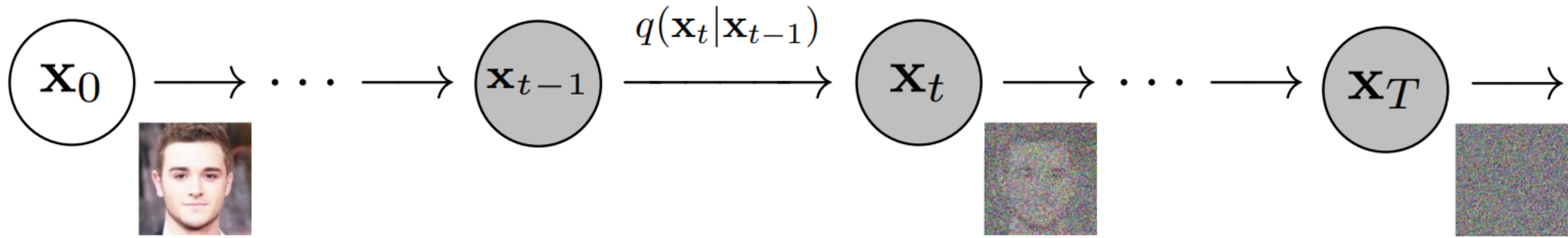
STYLEGAN

- Enfocado en generar imágenes de alta calidad y con control detallado sobre el estilo.
- Introduce características como el mapeo de estilos y ruido en diferentes capas.
- <https://thispersondoesnotexist.com>



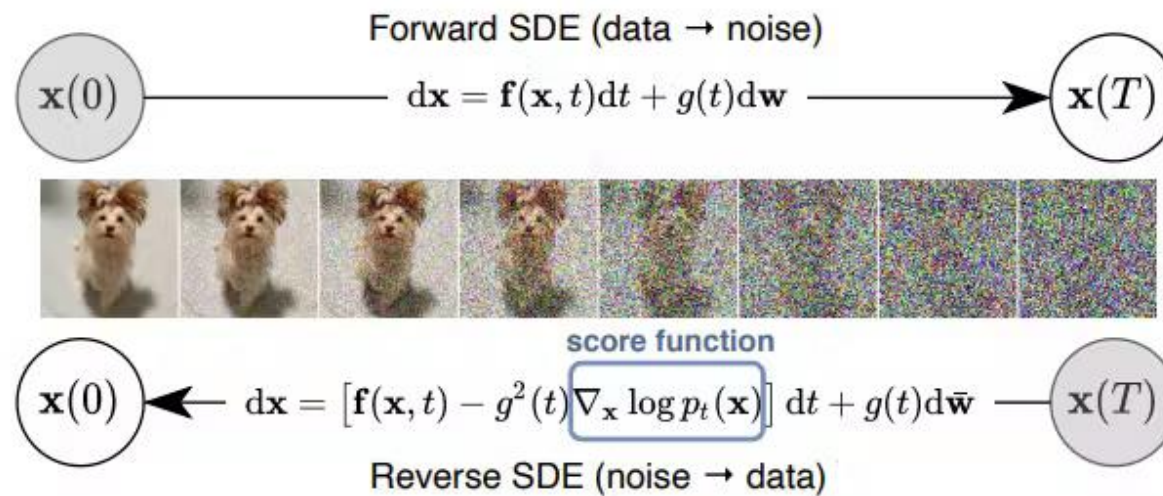
MODELOS DE DIFUSIÓN

¿EN QUÉ CONSISTEN?

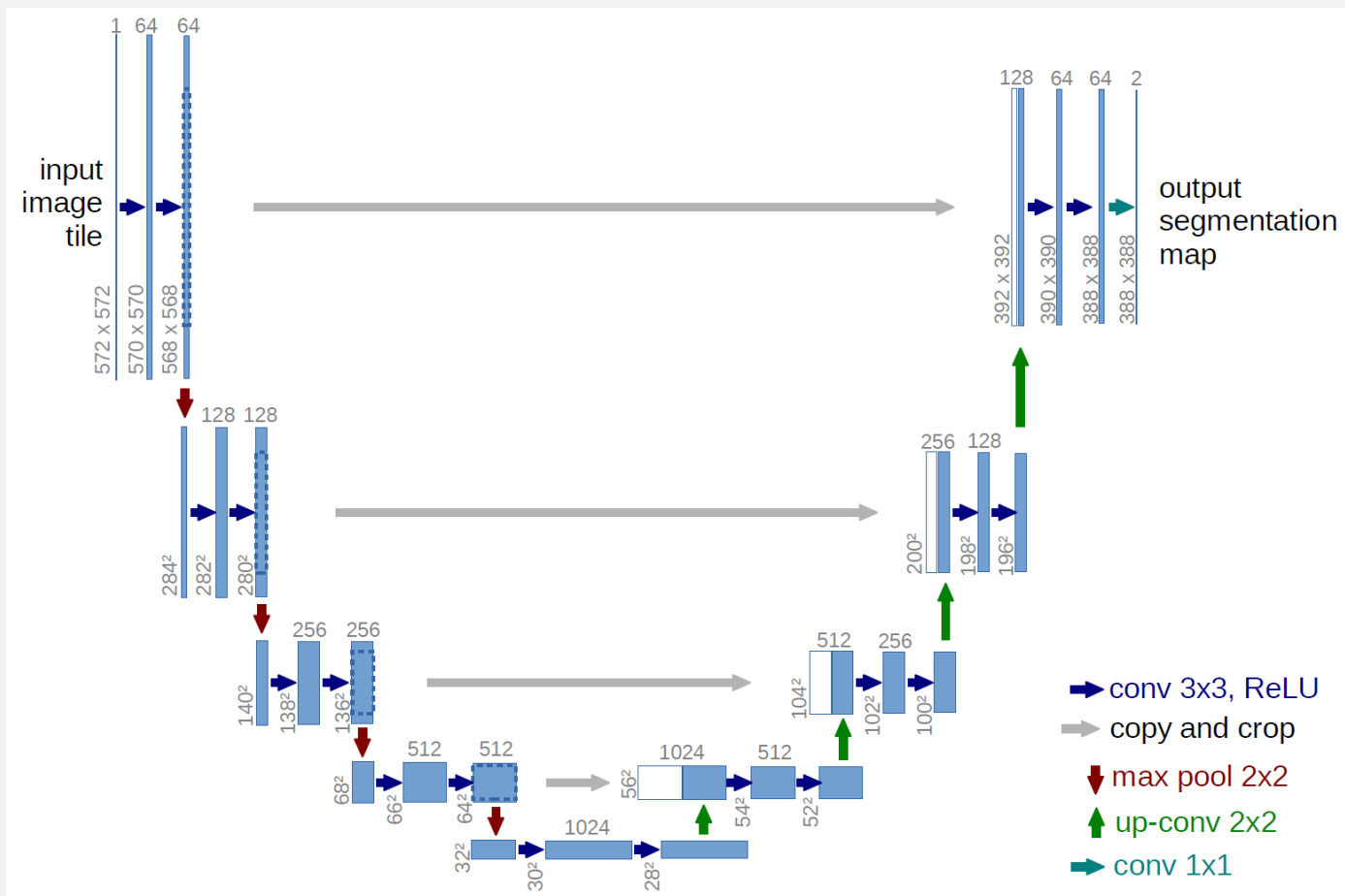


ENTRENAMIENTO

- De manera simplificada, el modelo se entrena para predecir el ruido que tiene la imagen en el paso actual.



ARQUITECTURA



VENTAJAS Y DESVENTAJAS

- Ventaja: Genera imágenes de muy alta calidad
- Desventaja: Es muy lento para generar. Tiene que predecir el ruido de cada paso, y durante el entrenamiento se hacen muchos pasos de ruido

MODELOS AUTOREGRESIVOS

x_1						x_n
			x_i			
						x_{n^2}

OpenAI



Image Generation