

L U M B R E



Modelos de difusión

Grado: 8

Materia: Inteligencia Artificial Generativa · Inteligencia Artificial

11/10/2026

Antes de empezar

Los modelos de difusión ganaron el trono de la generación de imagen con un idea sencilla: destruir poco a poco una muestra hasta volverla ruido, y entrenar una red para deshacer ese camino paso a paso. Generar es desruido desde el caos.

QUÉ APRENDERÁS

- Explicar el principio de los modelos de difusión: corromper datos con ruido y aprender a invertir el proceso para generar muestras nuevas.

Modelos de difusión

Esculpir quitando ruido

Una difusión aprende a esculpir a la inversa: parte de ruido puro y, paso a paso, va quitando granulado hasta revelar una imagen coherente. Es generación por refinamiento progresivo, no por un único salto.

VIDEO

Stable Diffusion explicado

Vídeo en español que desmonta paso a paso cómo un modelo de difusión convierte ruido en imágenes.



Escanea para verlo en Lumbre

https://www.youtube.com/watch?v=TIrb1p_Slz4

La intuición del ir y volver

Se define un proceso directo que añade ruido gaussiano a un dato hasta destruirlo. El modelo aprende el proceso inverso: dado un punto ruidoso, dar un pequeño paso hacia una muestra limpia.

Proceso directo (corrupción)

- Se añade ruido gaussiano en muchos pasos crecientes.
- Cada paso corrompe un poco más la muestra.
- Al final, el dato es indistinguible de ruido puro.

Proceso inverso (generación)

Comenzando de ruido puro, la red predice y resta ruido iteradamente. Cada iteración afina la imagen hasta obtener una muestra verosímil del dominio aprendido.

Entrenar prediciendo ruido

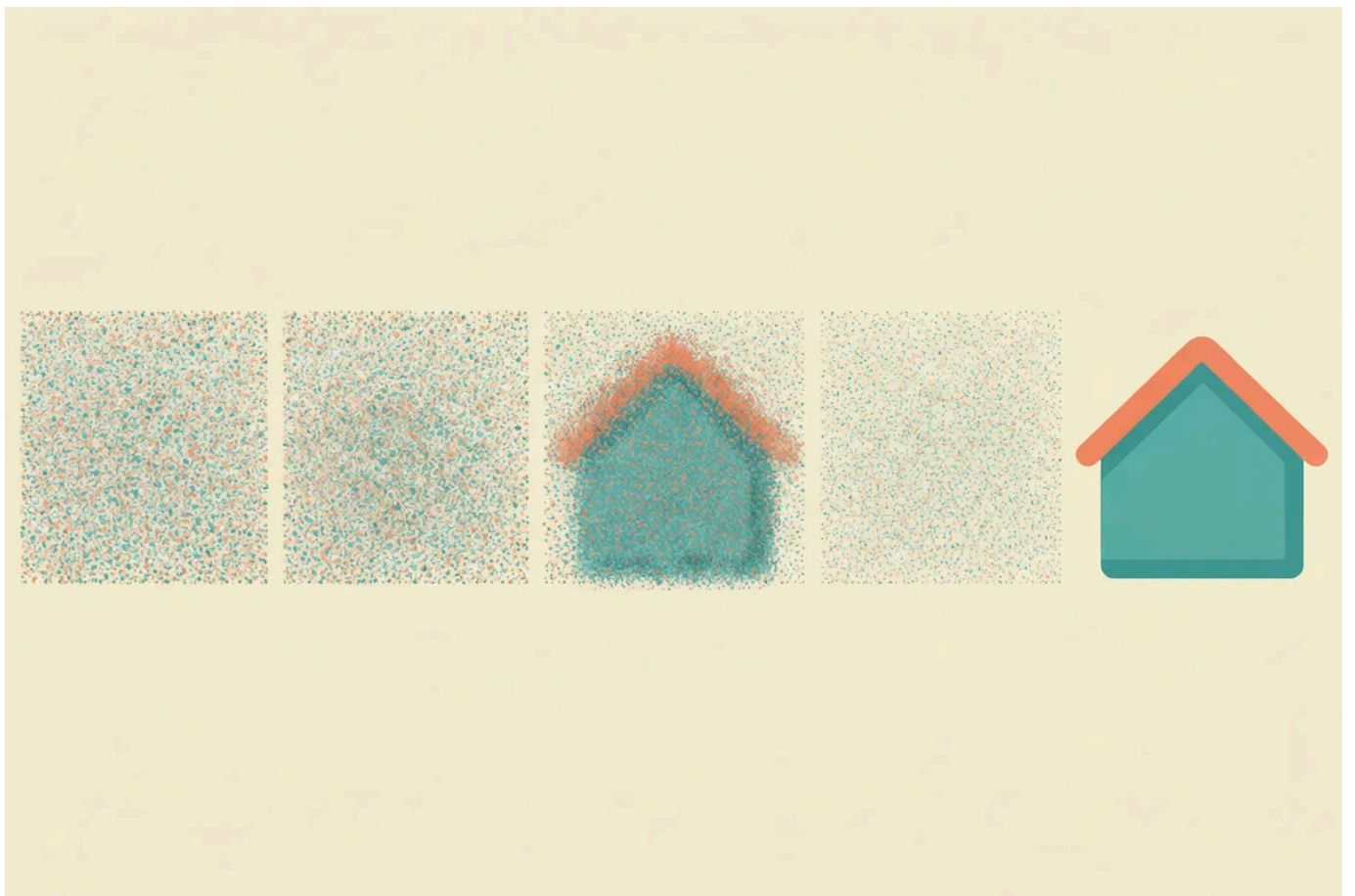
En vez de modelar la imagen limpia, la red se entrena para predecir el ruido añadido en un paso escogido al azar. Es una tarea de regresión simple y estable.

La pérdida simple

La función de coste mide cuánto se equivoca la red al estimar el ruido. Sorprendentemente, esta pérdida tan llana bastó para superar a las GAN en calidad.

Pasos de muestreo

La calidad sube con más pasos del proceso inverso, pero encarece la generación. Se usan solucionadores y aceleraciones para reducir el número de iteraciones.



Generar por difusión es recorrer el camino inverso al de la corrupción: quitar granulado hasta que aparezca la forma.

Latente: difundir en espacio comprimido

Difundir en píxeles es caro. Los modelos latentes difunden en un espacio comprimido por un codificador, abaratando enormemente el proceso.

Condiciones: texto a imagen

- Se guía la generación con una condición, como un prompt de texto.
- La condición modula los pasos de desruido hacia el contenido pedido.
- Así nacen los generadores de imagen desde descripción.

Ventajas frente a otras familias

Entrenamiento más estable que las GAN, sin el pulso adversario, y una variedad de muestras amplia. A cambio, generan despacio por sus múltiples pasos.

Despejar, no alzar

Si una GAN intenta producir la imagen entera de un golpe y falla en el equilibrio, la difusión pule poco a poco: cada paso es fácil, y la suma es asombrosa.

Errores frecuentes

- Confundir el proceso directo con el inverso: el primero solo añade ruido y no aprende nada; el modelo entrena el segundo.
- Esperar que la imagen salga de un solo salto: la calidad llega tras decenas de iteraciones de desruido.
- Difundir directamente en píxeles de alta resolución sin espacio latente, haciendo la generación prohibitivamente cara.

Ejemplo resuelto: generar un logo con el texto «zorro geométrico»

Un generador de imágenes por prompt es una difusión condicionada: el texto guía cada paso de desruido hacia el contenido pedido.

1. Codificas el prompt con un modelo de lenguaje y obtienes su vector de condición.
2. Partes de ruido gaussiano puro en el espacio latente comprimido.
3. En cada paso, la red predice el ruido de ese estado condicionada por el texto.
4. Restas el ruido estimado y avanzas al siguiente paso del proceso inverso.
5. Al terminar las iteraciones, un decodificador transforma el latente limpio en la imagen final.

¿Para qué sirve en la realidad?

Stable Diffusion, DALL-E o Midjourney generan ilustración publicitaria, concept art y prototipos de diseño en segundos; la medicina usa difusión para sintetizar imágenes médicas escasas, y el cine explora su versión para vídeo.

1. En un modelo de difusión, la red se entrena típicamente para:

- ☐ producir la imagen limpia en un solo paso
- ☐ predecir el ruido añadido en un paso escogido
- ☐ clasificar la imagen como real o falsa
- ☐ comprimir la imagen en un vector

2. La generación con difusión requiere múltiples pasos iterativos de desruido, lo que la hace más lenta por muestra que una única pasada de una GAN.

- ☐ Verdadero
- ☐ Falso

ACTIVIDAD INTERACTIVA DISPONIBLE EN LUMBRE

3. Difusión**4. Ordena el ciclo de una imagen generada por difusión:****Elementos:**

- partir de ruido gaussiano puro
- avanzar un paso del proceso inverso
- estimar el ruido a quitar en ese paso
- restar ruido y pasar al siguiente paso

Ruido	. . .
Paso	. . .
Predecir	. . .
Refinar	. . .

Modelo de difusión https://en.wikipedia.org/wiki/Diffusion_model Familia de modelos generativos por desruido iterativo.

5. Une cada pieza de la difusión con su papel.

5.	añade ruido gaussiano en pasos hasta destruir el dato	-----
5.	camino aprendido que quita ruido paso a paso	-----
5.	tarea de entrenamiento: estimar el ruido añadido	-----
5.	espacio comprimido donde difundir es barato	-----

Opciones: proceso directo · proceso inverso · predecir ruido · difusión latente

Esculpir quitando ruido

Modelos de difusión

Directo

ruido gaussiano

muchos pasos

dato destruido

Inverso

aprendido

predice y resta ruido

iterativo

Entrenar

regresión simple

pérdida estable

Acelerar

espacio latente

menos pasos

condición de texto

El mapa mental interactivo es interactivo en la versión online.



Stable Diffusion https://en.wikipedia.org/wiki/Stable_Diffusion El modelo latente de difusión que popularizó la generación de imagen desde texto.

6. ¿Por qué la difusión evita el problema de inestabilidad del entrenamiento adversario de una GAN? Razona sobre la simplicidad de predecir ruido.

Referencia curricular

Grado:	8
Materia:	Inteligencia Artificial Generativa · Inteligencia Artificial

Versión interactiva completa (con actividades, videos y autoevaluación):



<https://univ.lumbre.lat/object/cdia-geni-001-modelos-de-difusion>

Material educativo generado desde Lumbre · lumbre.lat