

Apuntes Semana 13 - 20/05/2025

Jose Pablo Barquero Díaz - 2022119938 ¹

Abstract—Este resumen presenta los autoencoders, sus componentes clave, variantes más comunes y aplicaciones prácticas, destacando su versatilidad en el aprendizaje profundo. También se menciona la arquitectura U-Net como una extensión especializada en segmentación de imágenes biomédicas.

I. RESPUESTAS DEL QUIZ

1) Describa qué es una capa de convolución y una capa de pooling.

- La capa de convolución es una capa que extrae características de una imagen aplicando filtros. Cada filtro recorre la imagen realizando multiplicaciones y sumas para detectar patrones como bordes, texturas o formas.
- La capa de pooling es una capa que reduce el ancho y alto del mapa de activación, ayudando a disminuir el número de parámetros y el tiempo de cómputo sin afectar los canales/profundidad.

2) Mencione dos ventajas por la cual se prefiere aplicar pequeñas convoluciones anidadas en lugar de una convolución grande, si su salida fuera equivalente

- La cantidad de parámetros es menor.
- Mejor capacidad de analizar la imagen, debido a que se emplean funciones no lineales.

3) Describa parte por parte un módulo de Inception

- Convolución 1x1
- Convolución 3x3
- Convolución 5x5
- Max pooling 3x3 + convolución 1x1

4) Describa cada variable de $\left\lceil \frac{m-k+2p}{s} \right\rceil + 1$

- m: Tamaño de los píxeles.
- k: Tamaño del Kernel
- p: Tamaño del Padding
- s: Tamaño del Stride

II. AUTOENCODERS

Un autoencoder es un tipo de red neuronal no supervisada diseñada para aprender una representación comprimida de los datos y luego reconstruir la entrada original a partir de esa codificación.

- Objetivo: Reconstruir la entrada con la menor pérdida posible.
- No requiere etiquetas (unsupervised learning).
- Utiliza comparación entre la entrada y la salida para ajustar sus parámetros mediante una función de pérdida.

A. Estructura general del autoencoder

Un autoencoder está compuesto por las siguientes etapas: entrada → encoder → cuello de botella → decoder → salida reconstruida, como se ilustra en la Figura 1

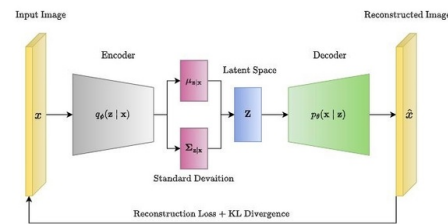


Fig. 1. Diagrama de un autoencoder

1) Encoder

- Conjunto de bloques convolucionales seguidos de módulos de pooling.
- Reduce la dimensión de los datos, extrayendo las características más importantes.
- Resultado: Vector latente.

2) Cuello de botella (Bottleneck)

- Es la representación en el espacio latente: una versión compacta y abstracta de la entrada.
- Es la parte más importante, ya que contiene toda la información necesaria para la reconstrucción.

3) Decoder

- Conjunto de convoluciones y operaciones de up-sampling.
- Toma el vector latente y reconstruye la entrada original a partir de él.

B. Tareas comunes que resuelve un autoencoder

1) Reducción de dimensionalidad

- Similar al PCA, pero más poderoso por ser no lineal. Se puede observar la diferencia en la Figura 2
- Aprende representaciones más eficientes que pueden usarse en otras tareas de aprendizaje automático.

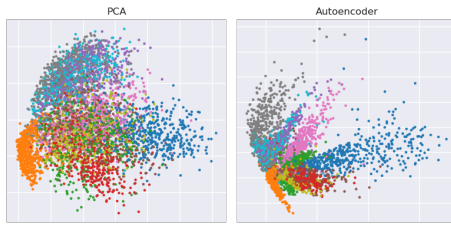


Fig. 2. PCA vs Autoencoder

2) Detección de anomalías

- Se entrena solo con datos normales (por ejemplo, transferencias bancarias legítimas).
- Si una entrada anómala se intenta reconstruir, el vector latente será deficiente y el error de reconstrucción será alto.
- Se define un umbral de error para detectar cuándo algo es potencialmente fraudulento.

3) Denoising

- Se entrena alimentando una entrada con ruido y esperando como salida la versión limpia como en la Figura 3.
- Aprende a reconstruir versiones más limpias de imágenes.



Fig. 3. Ejemplo de Denoising

4) Superresolución

- A partir de una imagen de baja resolución como la de la Figura 4, el autoencoder genera una imagen de mayor resolución (upsampling), usando la información latente.

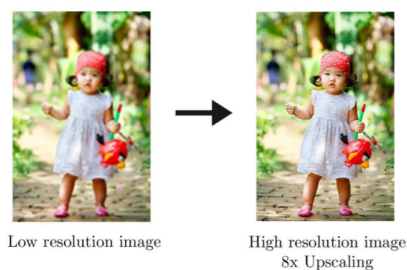


Fig. 4. Ejemplo de Superresolución

C. Tipos de autoencoders

- **Undercomplete Autoencoder:** El espacio latente tiene menor dimensión que la entrada, forzando al modelo a

aprender solo la información más relevante.

- **Sparse Autoencoder:** Introduce regularización para que solo un subconjunto de neuronas se active, incluso sin reducir la dimensión como se muestra en la Figura 5. Esto ayuda a obtener representaciones más interpretables y selectivas.

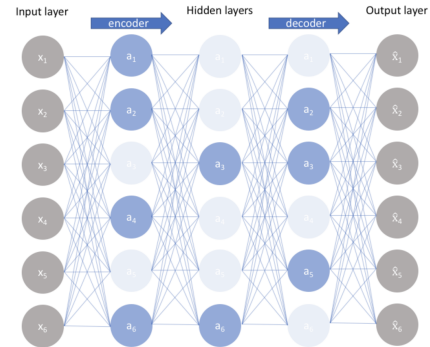


Fig. 5. Diagrama de Sparse Autoencoder

- **Denoising Autoencoder:** Se entrena con entradas alteradas con ruido, pero su objetivo es reconstruir la versión limpia. Ideal para tareas de restauración de datos como el ejemplo de la Figura 6.

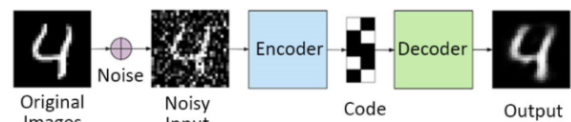


Fig. 6. Ejemplo de Denoising Autoencoder

- **Variational Autoencoder (VAE):** A diferencia de los autoencoders clásicos, el VAE aprende una distribución probabilística continua sobre el espacio latente. Esto permite realizar interpolaciones como la de la Figura 7 y generar nuevas muestras de forma coherente.

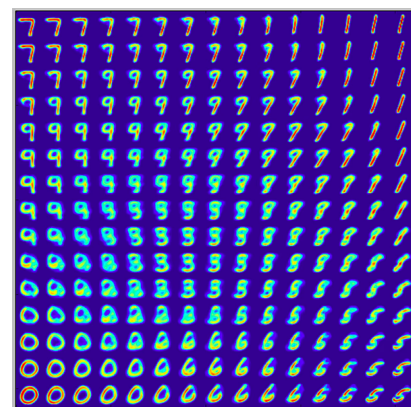


Fig. 7. Interpolación en el espacio latente de un autoencoder entrenado con el dataset MNIST

D. Hiperparámetros relevantes

- **Número de capas:** Afecta la capacidad del modelo. Más capas pueden mejorar el aprendizaje de patrones complejos, pero también incrementan el tiempo de entrenamiento.
- **Tamaño del vector latente:** Controla el grado de compresión. Una codificación demasiado pequeña puede perder información crítica, mientras que una demasiado grande podría sobreajustar.
- **Función de pérdida:** Define qué tan buena es la reconstrucción. Algunas funciones comunes son el error cuadrático medio (MSE) o el cross-entropy-loss.

E. U-Net

U-Net es una arquitectura basada en autoencoders, diseñada específicamente para segmentación de imágenes biomédicas. Introduce skip connections entre el encoder y el decoder, permitiendo conservar detalles espaciales importantes durante la reconstrucción. Esto la hace especialmente efectiva para segmentar tejidos u órganos en imágenes médicas.

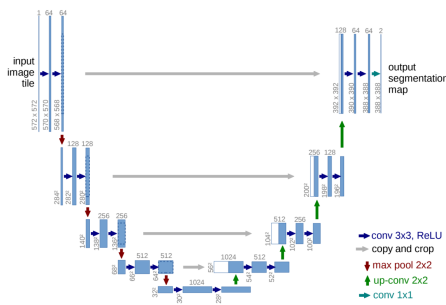


Fig. 8. Diagrama típico de una U-Net

La arquitectura U-Net recibe su nombre debido a la forma en “U” que se observa en su diagrama estructural de la Figura 8. Esta forma refleja claramente cómo está organizada la red