

Tarea 05: Generación de imágenes con autoencoder: Pytorch Lightning - Hydra - WandB

Curso de Inteligencia Artificial
Escuela de Ingeniería en Computación
Instituto Tecnológico de Costa Rica

I. OBJETIVO

Implementar un modelo de inteligencia artificial de generación de imágenes basado en la experimentación de dos arquitectura un autoencoder clásico y un autoencoder de Unet (skip connections) para la reconstrucción de imágenes, utilizando herramientas de desarrollo como Pytorch Lightning para la personalización/estructuración de entrenamiento de modelos, Hydra para la ejecución de múltiples configuraciones y WandB para la comparación de resultados.

II. PASOS A SEGUIR / INSTRUCCIONES

A. Dataset

Para el desarrollo de esta tarea se propone el uso del dataset propuesto en *MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection*. Un dataset de escenarios industriales reales con diferentes tipos de anomalías en la forma de detección de defectos en objetos o texturas.

Para el efecto de esta tarea y con el objetivo de reducir el tiempo de entrenamiento solamente se seleccionaron 4 clases de objetos (**cable**, **capsule**, **screw** y **transistor**). Además utilice las imágenes con tamaños de 128x128 con 3 canales.

B. Configuraciones y entrenamientos

El estudiante deberá estructurar el proyecto utilizando la librería **Hydra** para la gestión modular de configuraciones, asegurando la correcta separación de hiper parámetros entre el modelo, el entrenamiento y los registros experimentales.

La estructura mínima recomendada del proyecto es la siguiente:

```
conf/
  - config.yaml
  - model/
    - vae.yaml
  - trainer/
    - default.yaml
  - logger/
    - wandb.yaml
```

Cada módulo de configuración deberá permitir la ejecución de experimentos con distintos parámetros del modelo, tales como:

- Dimensión del espacio latente (z).
- Cantidad de épocas, tamaño de batch, o cualquier hiperparametro que requiera.

Para cada modelo de autoencoder realice una experimentación con las siguientes funciones de pérdida:

- L1
- L2
- SSIM
- SSIM + L1

Los experimentos deberán ser registrados y monitoreados en **Weights & Biases (WandB)**, con la finalidad de comparar métricas entre configuraciones, tales como:

- Pérdida de entrenamiento y validación.
- Ploteo de comparación de reconstrucción de imágenes originales y reconstruidas por el modelo del set de validación (16)
- Ploteo de comportamiento del vector latente con el algoritmo (t-SNE) por el modelo del set de validación.
- Ploteo de reconstrucción de imágenes buenas y con anomalías del set de prueba (16)

C. Evaluación del modelo

La evaluación del modelo deberá incluir dos aspectos principales:

- 1) **Reconstrucción:** Se deberá visualizar un conjunto de imágenes originales y sus correspondientes reconstrucciones. Se espera que las reconstrucciones sean más precisas para imágenes sin defectos y presenten errores notorios en aquellas con anomalías.

Asimismo, se deberá generar un informe final con las siguientes visualizaciones y análisis:

- Histogramas del error de reconstrucción entre clases normales y defectuosas para cada subclase de defecto del set de pruebas.
- Comparación de resultados entre configuraciones (funciones de pérdida). Incluyendo análisis de reconstrucción de imágenes correctas y aquellas que son defectuosas.

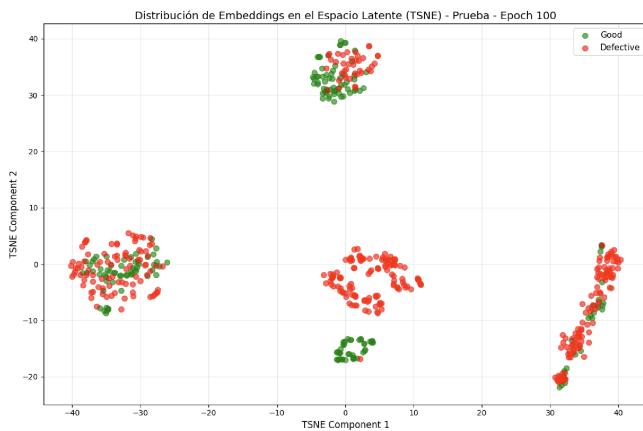
D. Entrega

Para esta tarea no se solicita informe escrito, pero si debe venir el paso a paso realizado en el notebook con anotaciones textuales, esto en el fin de facilitar la entrega y enfocarse en el entendimiento del modelo.

Todas las imágenes deben de ser producidas a través de WandB y deben de estar reinsertadas en el notebook para su evaluación.

TABLE I
TAREA DE INTELIGENCIA ARTIFICIAL GENERATIVAS EN IMÁGENES

Criterio	Descripción	Puntos
1. Modelo y entrenamiento para reconstrucción	Construcción del modelos utilizando Pytorch Lightning y Hydra.	20 pts
2. Entrenamiento autoencoder y u-net	Ejecutar al menos 4 entrenamientos con mismos hiperárametros pero cambiando la función de pérdida para cada modelo controlados todos con archivos de configuración con Hydra donde se compare utilizando WandB los mejores modelos de reconstrucción mostrando comparación de gráficas de la función de pérdida (train/val) entre las 4 ejecuciones.	50 pts
4. Presentación, organización y análisis crítico	El notebook está claramente estructurado, con anotaciones explicativas de cada paso y resultados bien presentados. Incluye análisis visual (reconstrucciones) y una interpretación crítica de los resultados obtenidos. Se valorará el orden, claridad y coherencia del flujo de trabajo.	30 pts
Total		100 pts



RÚBRICA

Si el trabajo no se encuentra debidamente ordenado y presentado siguiendo una adecuada estructura, puede ser considerado como incompleto y cualquiera de las rúbricas se puede ver afectada.