

Práctica 2: Trabajo de investigación

3ºGCED

Procesamiento de Imagen, Vídeo y Audio

Andrés Lires Saborido

andres.lires@udc.es



UNIVERSIDADE DA CORUÑA

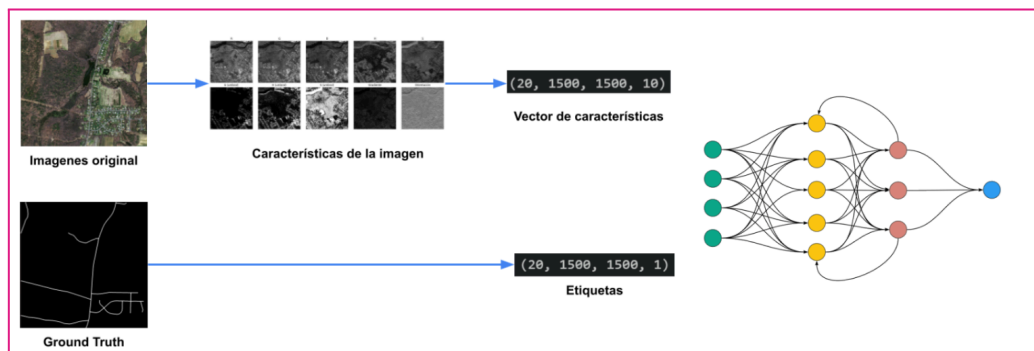
Introducción

Se desarrolló un *Jupyter Notebook* en *Visual Studio* en un entorno virtual (*Python 3.11.9*) usando las librerías *NumPy* (2.1.3), *SciPy* (1.15.2), *scikit-image* (0.25.2), *scikit-learn* (1.6.1) y *Tensorflow* (2.19.0) para resolver los dos ejercicios. El código se intercala con explicaciones o conclusiones en *Markdown*, documentando metodologías, implementación y resultados de forma más detallada. Este documento sirve como breve manual del trabajo realizado.

Ejercicio 1: Segmentación de carreteras en imagen aérea de alta resolución.

Esquema global del método.

- **Objetivo resumido:** Desarrollar un modelo capaz de segmentar carreteras mediante la clasificación de píxeles basándose en características extraídas con técnicas de procesamiento de imagen.
1. Se aplican una serie de procesamientos con los cuales se obtiene un **vector de valores que caracterizan cada píxel** de la imagen.
 2. Uso de imágenes ground truth para obtener un vector de **etiquetas binarias**.
 3. Entrenamiento de un **clasificador binario supervisado**.



Procedimiento.

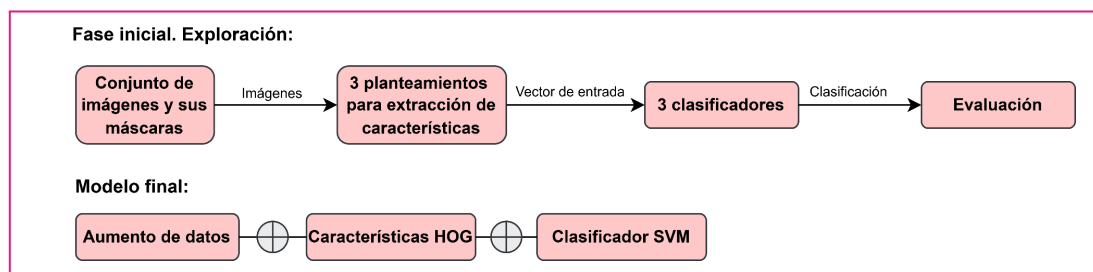
1. **Carga de datos.**
2. **Experimentación de técnicas de procesamiento de imagen.**
 - Exploración de **canales de color**: espacios RGB, HSV y Lab*.
 - **Umbralización** en canales H, S y b (expandido) con distintas técnicas.
 - Cálculo de **gradientes** (exploración de parámetros e imágenes).
 - Detección de bordes con **Canny** (escala de grises, H y S).
3. **Selección de características.**
 - Canales **R, G, B** (RGB), **H, S** (HSV), umbralización **manual** (canal **b**) y **Otsu** (en canales **H** y en **S**), magnitud y dirección del **gradiente** de la imagen **gray** ($\sigma = 1, 8 \text{ ángulos}$), siendo los que ofrecen mejor diferenciación.

4. División de datos en entrenamiento y test.
5. Modelado (CNN).
 - **Arquitectura:** 2 capas convolucionales + capa de salida.
 - **Métricas:** AUC (datos desbalanceados), precisión, sensibilidad, especificidad.
 - **Salida:** Probabilidad por píxel [0–1] de ser carretera.
6. Resultados.
 - *Precisión:* 0.86, **AUC: 0.89** en las imágenes de test.
 - En el *Notebook* se pueden visualizar las segmentaciones predichas.
7. Técnicas de postprocesado seleccionadas:
 - Umbralización (*probabilidad* > 0.6 → Carretera).
 - Eliminación de ruido mediante componentes conexas.

Ejercicio 2: Reconocimiento de objetos.

Esquema global del método.

- **Objetivo resumido:** Desarrollar un sistema que a partir de la extracción de características de las imágenes y su máscara construya un vector de entrada con el cual entrenar un clasificador que distinga entre imágenes de “*elefante*” y “*rinoceronte*”.
- Experimentación con **3 diferentes vectores de características** como entrada.
 - Entrenamiento de **3 clasificadores diferentes para cada vector de entrada**.
 - Optimización de un modelo **clasificador definitivo**.



Procedimiento.

1. Carga de datos.
2. Creación de un vector de etiquetas.
 - Cada imagen tendrá una etiqueta 0 (*elefante*), 1 (*rinoceronte*) o 2 (*otro*).
3. Entrenamiento de diferentes modelos de clasificación.
 - Diferentes planteamientos para el **vector de entrada**:
 - Imágenes recortadas y **aplanadas**.
 - **Características generales** de las imágenes (área, perímetro...).
 - Características **HOG** de las imágenes.
 - Para ambos casos se probarán los siguientes **clasificadores**:

- Modelo simple de **red de neuronas**.
- **SVM** con kernel lineal.
- **Random Forest** con 100 árboles.
- Presentación de resultados (precisión, matriz de confusión...). En general satisfactorios pero posiblemente mejorables (precisión entre 0.6 y 0.8).

4. Modelo final.

- **Aumento del conjunto de datos** mediante transformaciones sobre las originales.
- Elección de características **HOG** (exploración de parámetros) y clasificador **SVM** como modelo final.
- Buenos resultados obtenidos (*Precisión: 0.92* y un único error entre clases objetivo).

Para cualquier tipo de duda se puede consultar el *notebook*, en donde se justifica más detalladamente cada decisión tomada. Durante la defensa se aclararán posibles consultas.