

## Caso 2: Calcular m2 piscinas

La siguiente propuesta aborda el problema de identificar si en una dirección específica existe una casa y, de ser así, detectar si hay una piscina y calcular su superficie aproximada. Para ello, se propone un flujo de trabajo con dos modelos principales:

### 1. Detección de Casas

Para determinar si se encuentra una casa en la dirección especificada, se propone realizar un fine-tuning del modelo YOLOv8, entrenándolo con un conjunto de imágenes satelitales que incluyan casas en distintos escenarios y contextos. Este enfoque aprovecha el transfer learning, permitiendo un entrenamiento más rápido y eficiente, ya que YOLOv8 cuenta con pesos preentrenados en tareas similares de detección de objetos. Esto no solo reduce la cantidad de datos necesarios para alcanzar un buen desempeño, sino que también garantiza una alta precisión en la detección, incluso en escenarios variados.

El fine-tuning de YOLOv8-seg es una decisión adecuada para este problema debido a sus capacidades avanzadas de detección en tiempo real y su optimización para trabajar con bounding boxes. Su arquitectura moderna equilibra la velocidad y la precisión, lo que lo convierte en una solución ideal para analizar imágenes satelitales de alta resolución y detectar casas de forma precisa, incluso en contextos complejos con diferentes tipos de entornos (zonas rurales, urbanas o suburbanas). Además, YOLOv8 permite un pipeline de entrenamiento más simplificado en comparación con modelos más complejos, lo que reduce los costos computacionales.

El dataset utilizado para el entrenamiento estará compuesto por imágenes satelitales etiquetadas, con bounding boxes que marquen claramente las casas en cada imagen. Esto permitirá al modelo aprender a localizar y clasificar casas en una variedad de escenarios. Para garantizar la calidad del entrenamiento y la validez del modelo, se emplearán las siguientes métricas de evaluación:

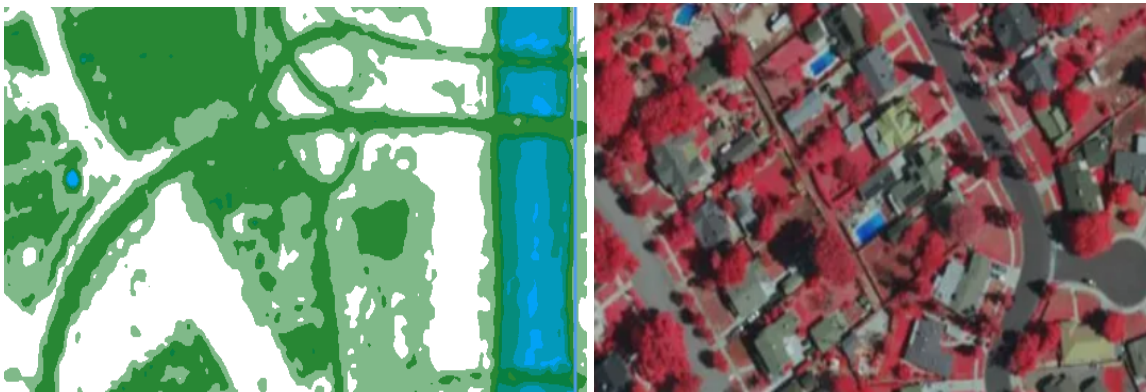
- **Precisión:** Mide qué porcentaje de las detecciones realizadas por el modelo son correctas (True Positives) en relación con todas las detecciones realizadas. Una alta precisión indica que el modelo comete pocos falsos positivos.
- **IoU (Intersection over Union):** Evalúa la superposición entre los bounding boxes predichos y los reales. Se calcula como:

### 2.Detección Piscinas

Para esta segunda fase, se propone una segmentación por instancias utilizando la arquitectura Mask R-CNN para identificar piscinas en imágenes satelitales de las direcciones proporcionadas. Este enfoque permite generar máscaras únicas para cada instancia detectada, recolectando todos los píxeles correspondientes a la superficie de una piscina. Esto es fundamental para calcular los m<sup>2</sup> reales de la piscina utilizando la resolución de la imagen satelital. La segmentación por instancias con Mask R-CNN se ha elegido porque otros enfoques, como los bounding boxes o la segmentación semántica, no cumplen

con los requisitos del proyecto. Los bounding boxes no ofrecen precisión píxel a píxel, ya que incluyen áreas innecesarias, mientras que la segmentación semántica no permite diferenciar piscinas individuales ni distinguir entre distintos tipos de cuerpos de agua. Mask R-CNN combina detección de objetos con segmentación precisa, lo que facilita la diferenciación y un análisis más detallado.

El dataset estará compuesto por imágenes satelitales etiquetadas con piscinas, lagos y ríos, para que el modelo aprenda a diferenciar entre estos tipos de cuerpos de agua. En lugar de utilizar las bandas RGB, se empleará falso color con las bandas NIR, Red y Green. Este esquema permite resaltar los cuerpos de agua frente a su entorno, manteniendo información relevante del contexto. Se opta por no utilizar índices como NDWI en esta etapa, ya que este podría simplificar demasiado la información y dificultar la diferenciación entre superficies de agua y otros elementos brillantes.



Para la evaluación del modelo, se emplearán dos métricas principales:

- **IoU (Intersection over Union):** Mide qué tan bien se superponen las máscaras predichas con las reales, garantizando precisión píxel a píxel.
- **mAP (mean Average Precision):** Calcula la media de la precisión para distintos niveles de solapamiento entre máscaras y para todas las clases detectadas, proporcionando una evaluación integral del desempeño del modelo.

### 3. Evaluación

Para evaluar el rendimiento del modelo en la detección de piscinas, se puede comparar la superficie predicha en metros cuadrados con la superficie real obtenida del catastro. Esto se haría calculando el error relativo medio (Mean Relative Error, MRE), que mide el promedio del porcentaje de error entre todas las predicciones realizadas en comparación a las superficies reales del catastro.

### 4. Justificación de éxito

El éxito del proyecto se basa en el uso de modelos avanzados y probados, como YOLOv8 y Mask R-CNN, que han demostrado un excelente rendimiento en tareas de detección de objetos y segmentación por instancias, respectivamente. Además, el enfoque propuesto incluye un preprocesamiento robusto con bandas espectrales NIR, Red y Green, lo que mejora la diferenciación entre piscinas y cuerpos de agua naturales, reduciendo los falsos positivos.

## **Fuentes**

Paper Mask R-CNN (Papers with Code): <https://arxiv.org/pdf/1703.06870v3>

Código Mask R-CNN (Papers with Code):

[https://github.com/facebookresearch/detectron2/blob/601d7666faaf7eb0ba64c9f9ce5811b13861fe12/detectron2/modeling/roi\\_heads/mask\\_head.py#L154](https://github.com/facebookresearch/detectron2/blob/601d7666faaf7eb0ba64c9f9ce5811b13861fe12/detectron2/modeling/roi_heads/mask_head.py#L154)

Documentación YOLOv8: <https://docs.ultralytics.com/es/models/yolov8/#usage-examples>

Artículo detección piscinas via imagenes satelitales:

<https://medium.com/geoai/swimming-pool-detection-and-classification-using-deep-learning-aaf4a3a5e652>

Foro: solucionar problema fuentes de agua naturales:

<https://forums.fast.ai/t/swimming-pool-detection/17922>