

# Creación de modelo YOLO a partir de dataset personalizado

## Resumen

Este informe presenta el proceso completo para crear un modelo de detección de objetos utilizando YOLO (You Only Look Once) a partir de un dataset personalizado. El proceso incluye la recolección de imágenes, el etiquetado utilizando la herramienta Roboflow, y el entrenamiento del modelo en un entorno de Google Colab. Además, se detallan las razones detrás de cada paso, consejos y mejoras para la obtención de un modelo robusto y que cumpla con las necesidades esperadas

## 1. Recolección de Imágenes

El modelo YOLO es uno de los algoritmos más populares para la detección de objetos en imágenes. Su capacidad para realizar detección en tiempo real lo hace ideal para muchas aplicaciones prácticas. Sin embargo, para entrenar un modelo YOLO, se necesita un conjunto de datos bien preparado. Este tutorial describe cómo construir un modelo YOLO a partir de un conjunto de datos personalizado, comenzando desde la recolección de imágenes.

El conjunto de datos debe ser relevante para el problema de detección de objetos que estamos tratando de resolver. Se deben capturar o recopilar imágenes que representen las clases de objetos que queremos detectar. El dataset debe incluir imágenes con diferentes perspectivas, iluminación y fondos para garantizar que el modelo aprenda a generalizar correctamente.



Figura 1: Ejemplo de distintas perspectivas en caras

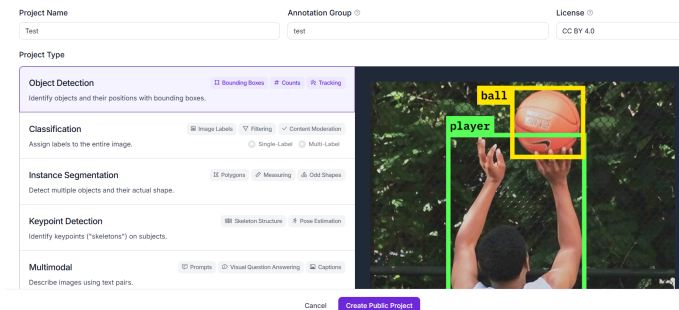
Algunas recomendaciones para tener una buena muestra de imágenes son las siguientes:

1. **Representación equilibrada de categorías:** Es fundamental asegurarse de que cada clase o categoría esté representada de manera equilibrada en el conjunto de datos. Esto evitará que el modelo tenga un sesgo hacia las clases con más imágenes, lo que podría reducir la capacidad del modelo para generalizar correctamente a nuevas instancias de clases menos representadas.
2. **Diversidad en las perspectivas:** El conjunto de datos debe incluir imágenes tomadas desde diferentes ángulos y perspectivas de los objetos. Esto permite que el modelo aprenda a reconocer los objetos independientemente de su orientación en el espacio, mejorando la robustez del modelo a diferentes situaciones.
3. **Variedad en condiciones de iluminación:** Incluir imágenes tomadas en diversas condiciones de iluminación (luz natural, luz artificial, sombras, etc.) es crucial para garantizar que el modelo pueda reconocer los objetos bajo diferentes condiciones de visibilidad. Esta variabilidad ayuda a mejorar la capacidad del modelo para trabajar en entornos del mundo real.
4. **Escalabilidad y variación en el tamaño de los objetos:** Se recomienda que el conjunto de datos incluya imágenes con los objetos de diferentes tamaños. Esto puede lograrse al capturar objetos tanto cerca como lejos de la cámara. Esta variación ayudará al modelo a identificar objetos correctamente a diferentes distancias y tamaños.
5. **Diversidad de fondos y entornos:** Para evitar que el modelo aprenda a identificar un objeto en un contexto específico, es importante incluir imágenes con diversos fondos y entornos. Esto incluye tanto fondos simples como complejos, urbanos, naturales, etc., lo que permite al modelo generalizar a diferentes escenarios.
6. **Incluir imágenes con oclusiones:** Incorporar imágenes en las que los objetos estén parcialmente ocultos por otros objetos o elementos del entorno es esencial para que el modelo aprenda a manejar oclusiones y a realizar predicciones incluso cuando no se pueda ver el objeto en su totalidad.

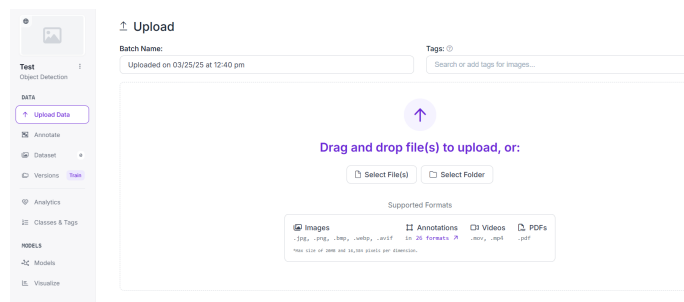
## 2. Etiquetado con Roboflow

Una vez que se han recolectado las imágenes, el siguiente paso es etiquetar cada objeto en las imágenes con sus respectivas clases. Para esto, se recomienda utilizar *Roboflow*, una herramienta en línea que permite etiquetar imágenes de manera eficiente, además de entregar el dataset en el formato compatible para su entrenamiento posterior.

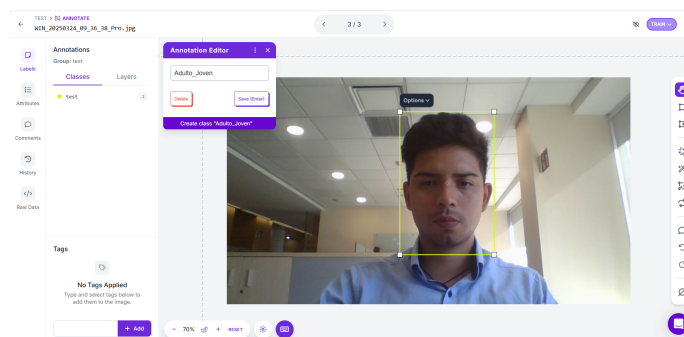
1. Crear una cuenta en Roboflow y crear un proyecto nuevo.



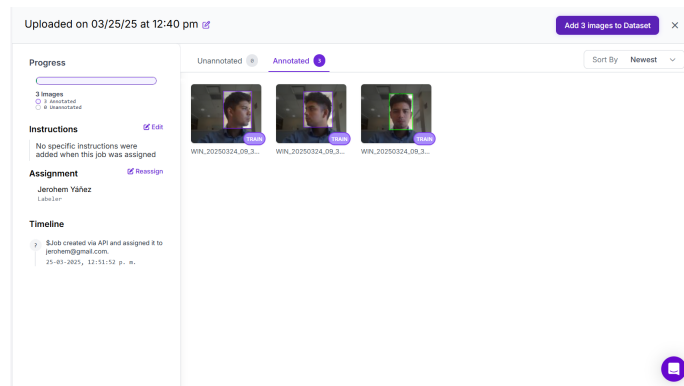
2. Subir las imágenes



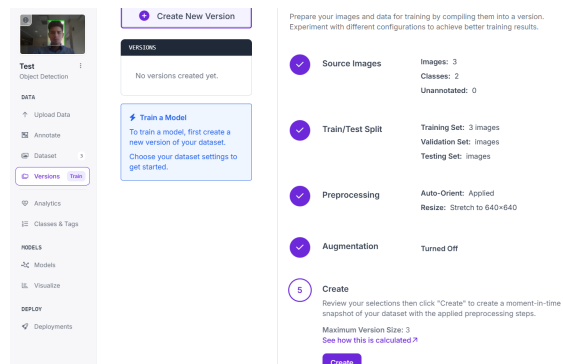
3. Escoger el etiquetado manual para que las personas a cargo de esto puedan realizarlo personalmente
4. Utilizar la interfaz de etiquetado para dibujar cuadros delimitadores alrededor de los objetos en las imágenes y asignar una clase a cada cuadro delimitador.



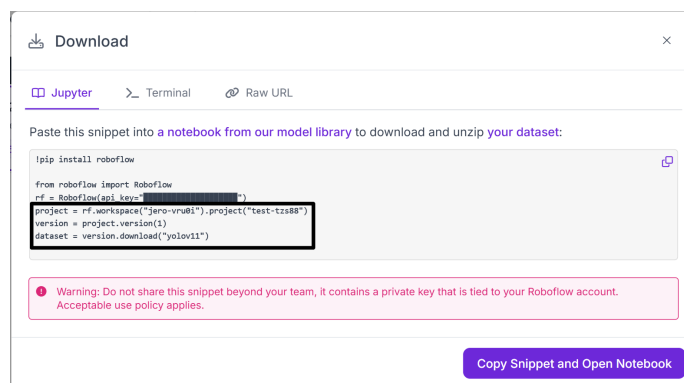
5. Añadir las imágenes ya etiquetadas al dataset



6. Crear una Versión del dataset, en donde aplicaremos técnicas de preprocesamiento y de data augmentation según se requiera



7. Descargar el dataset en formato YOLOv11. Las opciones de descarga dependen del dónde entrenaremos el modelo, esta vez mostraremos como se haría en un notebook de google colab, por lo tanto utilizaremos la opción "Show download code"
8. Con el código de descarga, solamente guardaremos las variables "project", "version", "dataset".



### 3. Entrenamiento del Modelo en Google Colab

Una vez que las imágenes están etiquetadas, el siguiente paso es entrenar el modelo. Para ello, se utilizará Google Colab, un entorno en la nube que proporciona acceso gratuito a GPUs, lo cual es esencial para entrenar modelos de aprendizaje profundo.

Aquí puedes encontrar el notebook con los comandos necesarios y explicaciones detalladas de los pasos a seguir para entrenar y obtener el modelo final.

Acceder [Notebook Google Colab](#)