

Portada

Instituto Tecnológico de Estudios Superiores de Monterrey



Maestría en Inteligencia Artificial Aplicada (MNA) Curso: TC5035 Proyecto Integrador

Profesores: Dra. Grettel Barceló Alonso y Dr. Luis Eduardo Falcón

Proyecto: Reconocimiento y Clasificación de Objetos basado en Colores para Detección en Vivo en Cámaras de Seguridad

Avance 1. Análisis exploratorio de datos

Equipo 43

Ronald David Zuniga Sanchez - A01686240

1 - Introducción

En el presente avance del Proyecto Integrador **Reconocimiento y Clasificación de Objetos basado en Colores para Detección en Vivo en Cámaras de Seguridad**, se realiza un análisis exploratorio de los datos describiendo los mismos utilizando técnicas estadísticas y de visualización para describir los aspectos más relevantes el conjunto de datos para justificar y planificar operaciones de preprocesamiento en próximos avances.

Se desarrolla la justificación de la selección de un nuevo dataset y las características principales del mismo así como la descripción a mayor detalle de las técnicas de ML a aplicar.

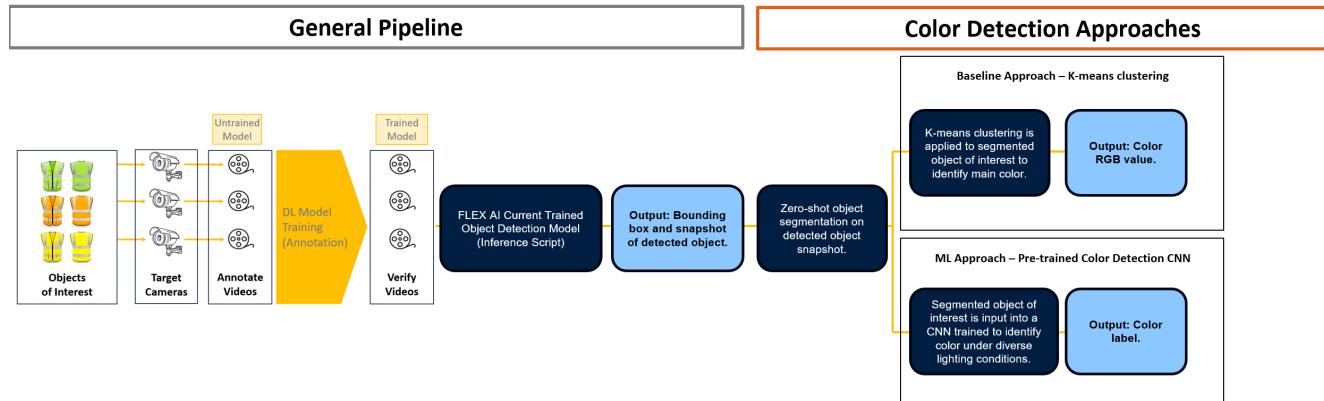
Contexto del Proyecto

Actualización de las técnicas de aprendizaje de máquina (ML) aplicables

Considerando una revisión complementaria de literatura [5][6][11], se describen a mayor detalle los 2 acercamientos propuestos para la investigación sobre las técnicas a aplicar para la resolución del problema propuesto donde se considera lo siguiente:

1. La implementación de una fase general de segmentación sobre el objeto de interés detectado por el modelo actual del producto. La segmentación se propone para aislar la información de color del objeto de interés del fondo o contexto de imagen inmediato dentro del bounding box region.
2. Enfoque 1 (Línea Base) - Técnicas de computer vision como agrupación de k-means para identificar el color principal.
3. Enfoque 2 (Utilizando ML) - Técnica de ML utilizando una red neuronal convolucional (CNN) pre entrenada para identificar el color en diversas condiciones de iluminación.

Proposed ML Approach



TC5035 – CAPSTONE PROJECT MNA – PROGRESS REPORT 1

Actualización de la selección del dataset

Se expuso en el avance anterior, que se utilizaría como conjunto de datos base para el proyecto el dataset CIFAR-10 considerando su conjunto de imágenes multiclas. Sin embargo, posterior a revisiones de documentación relevante asociada al tema de investigación, se encontrarán los siguientes retos al considerar el dataset de CIFAR-10:

1. Dataset directamente aplicable a tareas de clasificación de imagen. El dataset de CIFAR-10 no cuenta con información de contexto, es decir metadatos por imagen asociados a detección de objetos (bounding box, labels, object and instance segmentation).
2. Tamaño de dataset pequeño y de poca diversidad de objetos (10 clases, 60K imágenes) en comparación otros datasets estándar para uso en tareas de visión computacional.
3. Falta de imágenes de escenarios de mundo real, objetos en contexto.

Considerando estas limitantes, se determinó seleccionar un dataset alternativo para continuar con el proyecto.

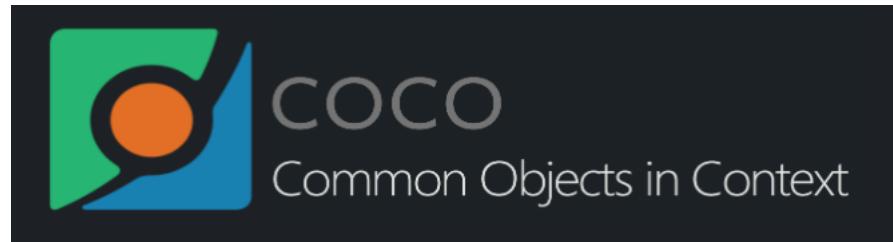
Selección de Dataset

Considerando el enfoque de las técnicas de aprendizaje de máquina (ML) aplicables propuestas y considerando el benchmark propuesto en literatura de referencia [1][2], se determina utilizar el dataset de Common Objects in Context (COCO) en su versión de 2017.

Table 1. Overview of the six most commonly used benchmark computer vision datasets (2004–2018) (based on Everingham et al., 2010; Fei-Fei et al., 2004; Griffin et al., 2007; Kuznetsova et al., 2018; Lin et al., 2014; Russakovsky et al., 2015). Citations from Google Scholar (4 December 2020).

Name	Categories	Tasks	Images	Instances	Average instances/ image	Provenance	Citations
Caltech 101	101	• image classification	9247 (30–400 image/ category)	n/a	n/a	Google Image Search	3825
Caltech 256	256	• image classification	30607 (80–800 image/category)	n/a	n/a	Google Image Search/ PicSearch	2161
PASCAL VOC	20	• image classification • object detection	11000	270000	2.3	Flickr	9626
ImageNet	• 21841 synsets • 1000 image classification • 200 object detection	• image classification • single object localization • object detection	14197122		3.0	Flickr	19745
MS COCO	91	• object detection • object segmentation • semantic scene labelling	328000	2500000	7.7	Flickr	12701
Open Images	• 19794 image classification • 600 object detection • 326 visual relationships	• image classification • object detection • visual relationships	9178275	15440132 (bounding boxes)	8.1 (object detection)	Flickr	376

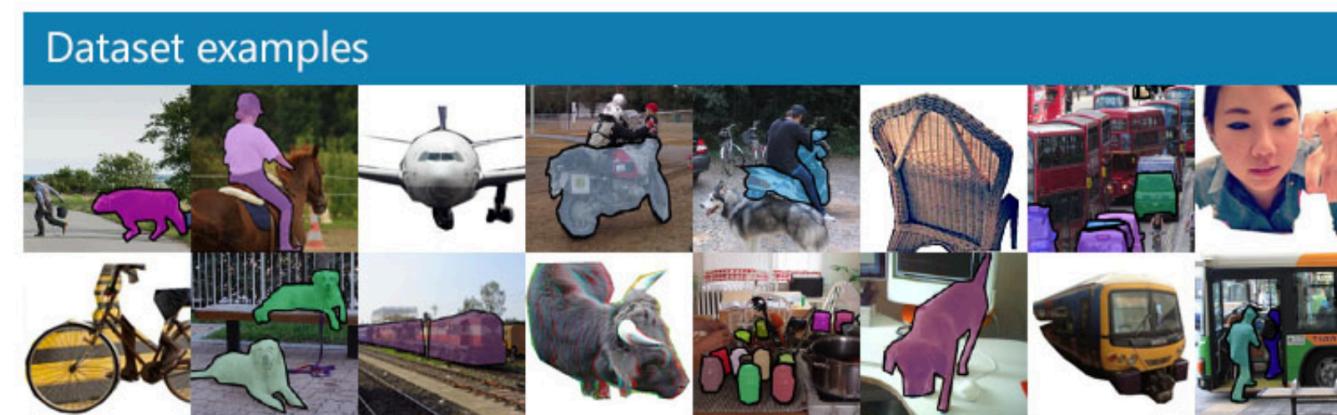
Descripción de Dataset Seleccionado - Common Objects in Context (COCO)



El dataset COCO (Common Objects in Context) es un conjunto de datos de imagen clave para la visión computacional. Desarrollada por Microsoft, contiene más de 200,000 imágenes etiquetadas (330,000 en total) en 80 categorías de objetos y 91 categorías de materiales. Estas imágenes están contextualizadas en entornos reales, lo que lo convierte en un estándar para proyectos de detección de objetos, segmentación de imágenes y generación de descripciones.

Características principales:

- Las imágenes se dividen en 80 categorías de objetos diferentes, desde personas y animales hasta objetos y vehículos.
- Proporciona anotaciones precisas para cada objeto en las imágenes, incluyendo sus posiciones y contornos.
- Las imágenes están recopiladas de diferentes fuentes, lo que garantiza una amplia variedad de contextos y escenas.
- El conjunto de datos COCO se utiliza para una variedad de tareas en visión por computadora, como detección de objetos, segmentación semántica y generación de texto a partir de imágenes.
- Se utiliza como referencia en la evaluación de algoritmos y modelos de aprendizaje automático en visión por computadora.
- COCO se actualiza regularmente, lo que garantiza que sea un recurso valioso y actualizado para la comunidad de investigación y desarrollo en visión por computadora.



▼ 2 - Importando el Dataset

Para la descarga y uso avanzado del dataset seleccionado (COCO), se determina utilizar la herramienta open-source FIFTY ONE de Voxel 51 [7] [9]. Fifty One opera como un paquete de Python que facilita el acceso, visualización, y evaluación de modelos en el dataset COCO.

Fifty One permite descargar subconjuntos específicos de una serie de datasets, visualizar datos y etiquetas, y ejecutar análisis de componentes y exploración de datos con múltiples opciones. En el presente informe se utilizan las herramientas básicas de exploración de datos y se inicia el uso de agrupación por conjuntos de imágenes utilizando embeddings (Clustering Images with Embeddings) [10].



▼ Instalación de dependencias

```
# Instalar aplicación y librerías de Fifty One para la importación y análisis del dataset.
!pip install fiftyone
```

```
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas->fiftyone) (2024.1)
Requirement already satisfied: networkx>=2.2 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (3.3)
Requirement already satisfied: imageio>=2.4.1 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (2.31.6)
Requirement already satisfied: tifffile>=2019.7.26 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (2024.5.3)
Requirement already satisfied: PyWavelets>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (1.6.0)
Requirement already satisfied: joblib>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from scikit-learn->fiftyone) (1.4.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn->fiftyone) (3.5.0)
Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3.4.0->starlette>=0.24.0->fiftyone) (3.7)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3.4.0->starlette>=0.24.0->fiftyone) (1.3.1)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3.4.0->starlette>=0.24.0->fiftyone) (1.2.1)
Collecting hyperframe<7,>=6.0 (from h2>=3.1.0->hypercorn>=0.13.2->fiftyone)
  Downloading hyperframe-6.0.1-py3-none-any.whl (12 kB)
Collecting hpack<,>4.0 (from h2>=3.1.0->hypercorn>=0.13.2->fiftyone)
  Downloading hpack-4.0.0-py3-none-any.whl (32 kB)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-packages (from httpx>=0.10.0->universal-analytics-python3<2,>=1.0.1->fiftyone) (2024.2.2)
Collecting httpcore==1.* (from httpx>=0.10.0->universal-analytics-python3<2,>=1.0.1->fiftyone)
  Downloading httpcore-1.0.5-py3-none-any.whl (77 kB)
  77.9/77.9 kB 13.3 MB/s eta 0:00:00
Requirement already satisfied: attrs>=19.2.0 in /usr/local/lib/python3.10/dist-packages (from jsonlines->voxel51-eta<0.13,>=0.12.6->fiftyone) (23.2.0)
Collecting texttable (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading texttable-1.7.0-py2.py3-none-any.whl (10 kB)
Collecting pycryptodomex>=3.16.0 (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading pycryptodomex-3.20.0-cp35-abi3-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (2.1 MB)
  2.1/2.1 MB 99.9 MB/s eta 0:00:00
Collecting pyzstd>=0.15.9 (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading pyzstd-0.15.10-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (411 kB)
  411.2/411.2 kB 46.4 MB/s eta 0:00:00
Collecting pyppmd<1.2.0,>=1.1.0 (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading pyppmd-1.1.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (138 kB)
  138.9/138.9 kB 21.0 MB/s eta 0:00:00
Collecting pybcj<1.1.0,>=1.0.0 (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading pybcj-1.0.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (49 kB)
  49.7/49.7 kB 7.4 MB/s eta 0:00:00
Collecting multivolumefile>=0.2.3 (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading multivolumefile-0.2.3-py3-none-any.whl (17 kB)
Collecting inflate64k<1.1.0,>=1.0.0 (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading inflate64-1.0.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (93 kB)
  93.1/93.1 kB 13.7 MB/s eta 0:00:00
Collecting brotli>=1.1.0 (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone)
  Downloading Brotli-1.1.0-cp310-cp310-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux_2_12_x86_64.manylinux2010_x86_64.whl (3.0 MB)
  3.0/3.0 kB 103.0 MB/s eta 0:00:00
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->voxel51-eta<0.13,>=0.12.6->fiftyone) (3.3.2)
Building wheels for collected packages: fiftyone-db
  Building wheel for fiftyone-db (pyproject.toml) ... done
    Created wheel for fiftyone-db: filename=fiftyone_db-1.1.2-py3-none-manylinux1_x86_64.whl size=42156164 sha256=9edd01eb4d6fb3ed150b3728ebb1fb8875505a1fb8d13bc65f17a4686372cdb6
    Stored in directory: /root/.cache/pip/wheels/29/52/b9/14b9d344410c63c0447ba16bf47c0a064c55d735fa14ecf95c
Successfully built fiftyone-db
Installing collected packages: texttable, sseclient-py, pprintpp, kaleido, brotli, xmltodict, taskgroup, retrying, rarfile, pyzstd, pyppmd, pycryptodomex, pybcj, priority, multivol
Successfully installed Deprecated-1.2.14 aiofiles-23.2.1 argcomplete-3.3.0 boto3-1.34.103 botocore-1.34.103 brotli-1.1.0 dacite-1.7.0 dill-0.3.8 dnspython-2.6.1 fiftyone-0.23.8 fif
```

Instalación de dependencias para Clustering Analysis

```
!pip install -U scikit-learn fiftyone
```

```
Requirement already satisfied: taskgroup in /usr/local/lib/python3.10/dist-packages (from hypercorn>=0.13.2->fiftyone) (0.0.0a4)
Requirement already satisfied: tomli in /usr/local/lib/python3.10/dist-packages (from hypercorn>=0.13.2->fiftyone) (2.0.1)
Requirement already satisfied: wsproto>=0.14.0 in /usr/local/lib/python3.10/dist-packages (from hypercorn>=0.13.2->fiftyone) (1.2.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2>=3->fiftyone) (2.1.5)
Requirement already satisfied: tenacity>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from plotly>=4.14->fiftyone) (8.3.0)
Requirement already satisfied: pymongo<3.0.0,>=1.16.0 in /usr/local/lib/python3.10/dist-packages (from pymongo>=3.12->fiftyone) (2.6.1)
Requirement already satisfied: anyio<5,>=3.4.0 in /usr/local/lib/python3.10/dist-packages (from starlette>=0.24.0->fiftyone) (3.7.1)
Requirement already satisfied: httpx>=0.10.0 in /usr/local/lib/python3.10/dist-packages (from universal-analytics-python3<2,>=1.0.1->fiftyone) (0.27.0)
Requirement already satisfied: dill in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (0.3.8)
Requirement already satisfied: future in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (0.18.3)
Requirement already satisfied: glob2 in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (0.7)
Requirement already satisfied: jsonlines in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (4.0.0)
Requirement already satisfied: py7zr in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (0.21.0)
Requirement already satisfied: rarfile in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (4.2)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (2.31.0)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (1.16.0)
Requirement already satisfied: sortedcontainers in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (2.4.0)
Requirement already satisfied: tzlocal in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (5.2)
Requirement already satisfied: urllib3 in /usr/local/lib/python3.10/dist-packages (from voxel51-eta<0.13,>=0.12.6->fiftyone) (2.0.7)
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.10/dist-packages (from beautifulsoup4->fiftyone) (2.5)
Requirement already satisfied: botocore<1.35.0,>=1.34.103 in /usr/local/lib/python3.10/dist-packages (from boto3->fiftyone) (1.34.103)
Requirement already satisfied: jmespath<2.0.0,>=0.7.1 in /usr/local/lib/python3.10/dist-packages (from boto3->fiftyone) (0.1.0)
Requirement already satisfied: s3transfer<0.11.0,>=0.10.0 in /usr/local/lib/python3.10/dist-packages (from boto3->fiftyone) (0.10.1)
Requirement already satisfied: wrapt<2,>=0.1.10 in /usr/local/lib/python3.10/dist-packages (from Deprecation->fiftyone) (1.14.1)
Requirement already satisfied: wcwidth<0.3.0,>=0.2.12 in /usr/local/lib/python3.10/dist-packages (from ftfy->fiftyone) (0.2.13)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->fiftyone) (1.2.1)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib->fiftyone) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->fiftyone) (4.51.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->fiftyone) (1.4.5)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->fiftyone) (3.1.2)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas->fiftyone) (2024.1)
Requirement already satisfied: networkx>=2.2 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (3.3)
Requirement already satisfied: imageio>=2.4.1 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (2.31.6)
Requirement already satisfied: tifffile>=2019.7.26 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (2024.5.3)
Requirement already satisfied: PyWavelets>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from scikit-image->fiftyone) (1.6.0)
Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3.4.0->starlette>=0.24.0->fiftyone) (3.7)
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3.4.0->starlette>=0.24.0->fiftyone) (1.3.1)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio<5,>=3.4.0->starlette>=0.24.0->fiftyone) (1.2.1)
Requirement already satisfied: hyperframe<7,>=6.0 in /usr/local/lib/python3.10/dist-packages (from h2>=3.1.0->hypercorn>=0.13.2->fiftyone) (6.0.1)
Requirement already satisfied: hpack<,>4.0 in /usr/local/lib/python3.10/dist-packages (from h2>=3.1.0->hypercorn>=0.13.2->fiftyone) (4.0.0)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-packages (from httpx>=0.10.0->universal-analytics-python3<2,>=1.0.1->fiftyone) (1.0.5)
Requirement already satisfied: attrs>=19.2.0 in /usr/local/lib/python3.10/dist-packages (from jsonlines->voxel51-eta<0.13,>=0.12.6->fiftyone) (23.2.0)
Requirement already satisfied: texttable in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (1.7.0)
Requirement already satisfied: pycryptodomex>=3.16.0 in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (3.20.0)
Requirement already satisfied: pyzstd>=0.15.9 in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (0.15.10)
Requirement already satisfied: pyppmd<1.2.0,>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (1.1.0)
Requirement already satisfied: pybcj<1.1.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (1.0.2)
Requirement already satisfied: multivolumefile>=0.2.3 in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (0.2.3)
Requirement already satisfied: inflate64k<1.1.0,>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (1.0.0)
Requirement already satisfied: brotli>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from py7zr->voxel51-eta<0.13,>=0.12.6->fiftyone) (1.1.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->voxel51-eta<0.13,>=0.12.6->fiftyone) (3.3.2)
Installing collected packages: scikit-learn
  Attempting uninstall: scikit-learn
    Found existing installation: scikit-learn 1.2.2
    Uninstalling scikit-learn-1.2.2:
      Successfully uninstalled scikit-learn-1.2.2
  Successfully installed scikit-learn-1.4.2
```

```
fiftyone plugins download https://github.com/jacobmarks/clustering-plugin
```

```
Migrating database to v0.23.8
Downloading jacobmarks/clustering-plugin...
Copying plugin '@jacobmarks/clustering' to '/root/fiftyone/_plugins_/@jacobmarks/clustering'
```

```
!pip install umap-learn git+https://github.com/openai/CLIP.git
```

```
Collecting git+https://github.com/openai/CLIP.git
  Cloning https://github.com/openai/CLIP.git to /tmp/pip-req-build-uxzvaj85
  Running command git clone --filter=blob:none --quiet https://github.com/openai/CLIP.git /tmp/pip-req-build-uxzvaj85
  Resolved https://github.com/openai/CLIP.git to commit a1d071733d7111c9c014f024669f959182114e33
  Preparing metadata (setup.py) ... done
Collecting umap-learn
  Downloading umap_learn-0.5.6-py3-none-any.whl (85 kB)
```

```
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.10/dist-packages (from umap-learn) (1.25.2)
Requirement already satisfied: scipy>=1.3.1 in /usr/local/lib/python3.10/dist-packages (from umap-learn) (1.11.4)
Requirement already satisfied: scikit-learn>0.22 in /usr/local/lib/python3.10/dist-packages (from umap-learn) (1.2.2)
Requirement already satisfied: numba>=0.51.2 in /usr/local/lib/python3.10/dist-packages (from umap-learn) (0.58.1)
Collecting pynndescent>=0.5 (from umap-learn)
  Downloading pynndescent-0.5.12-py3-none-any.whl (56 kB)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from umap-learn) (4.66.4)
Collecting ftfy (from clip==1.0)
  Downloading ftfy-6.2.0-py3-none-any.whl (54 kB)
Requirement already satisfied: regex in /usr/local/lib/python3.10/dist-packages (from clip==1.0) (2023.12.25)
Requirement already satisfied: torch in /usr/local/lib/python3.10/dist-packages (from clip==1.0) (2.2.1+cu121)
Requirement already satisfied: torchvision in /usr/local/lib/python3.10/dist-packages (from clip==1.0) (0.17.1+cu121)
Requirement already satisfied: llvmlite<0.42,>=0.41.0dev0 in /usr/local/lib/python3.10/dist-packages (from numba>=0.51.2->umap-learn) (0.41.1)
Requirement already satisfied: joblib>=0.11 in /usr/local/lib/python3.10/dist-packages (from pynndescent>=0.5->umap-learn) (1.4.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn>=0.22->umap-learn) (3.5.0)
Requirement already satisfied: wcwidth<0.3.0,>=0.2.12 in /usr/local/lib/python3.10/dist-packages (from ftfy->clip==1.0) (0.2.13)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from torch->clip==1.0) (3.14.0)
Requirement already satisfied: typing-extensions>=4.8.0 in /usr/local/lib/python3.10/dist-packages (from torch->clip==1.0) (4.11.0)
Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-packages (from torch->clip==1.0) (1.12)
Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch->clip==1.0) (3.3)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from torch->clip==1.0) (3.1.4)
Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages (from torch->clip==1.0) (2023.6.0)
Collecting nvidia-cuda-nvrtc-cu12==12.1.105 (from torch->clip==1.0)
  Using cached nvidia_cuda_nvrtc_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (23.7 MB)
Collecting nvidia-cuda-runtime-cu12==12.1.105 (from torch->clip==1.0)
  Using cached nvidia_cuda_runtime_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (823 kB)
Collecting nvidia-cuda-cupti-cu12==12.1.105 (from torch->clip==1.0)
  Using cached nvidia_cuda_cupti_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (14.1 MB)
Collecting nvidia-cudnn-cu12==8.9.2.26 (from torch->clip==1.0)
  Using cached nvidia_cudnn_cu12-8.9.2.26-py3-none-manylinux1_x86_64.whl (731.7 MB)
Collecting nvidia-cublas-cu12==12.1.3.1 (from torch->clip==1.0)
  Using cached nvidia_cublas_cu12-12.1.3.1-py3-none-manylinux1_x86_64.whl (410.6 MB)
Collecting nvidia-cufft-cu12==11.0.2.54 (from torch->clip==1.0)
  Using cached nvidia_cufft_cu12-11.0.2.54-py3-none-manylinux1_x86_64.whl (121.6 MB)
Collecting nvidia-curand-cu12==10.3.2.106 (from torch->clip==1.0)
  Using cached nvidia_curand_cu12-10.3.2.106-py3-none-manylinux1_x86_64.whl (56.5 MB)
Collecting nvidia-cusolver-cu12==11.4.5.107 (from torch->clip==1.0)
  Using cached nvidia_cusolver_cu12-11.4.5.107-py3-none-manylinux1_x86_64.whl (124.2 MB)
Collecting nvidia-cusparse-cu12==12.1.0.106 (from torch->clip==1.0)
  Using cached nvidia_cusparse_cu12-12.1.0.106-py3-none-manylinux1_x86_64.whl (196.0 MB)
Collecting nvidia-nccl-cu12==2.19.3 (from torch->clip==1.0)
  Using cached nvidia_nccl_cu12-2.19.3-py3-none-manylinux1_x86_64.whl (166.0 MB)
Collecting nvidia-nvtx-cu12==12.1.105 (from torch->clip==1.0)
  Using cached nvidia_nvtx_cu12-12.1.105-py3-none-manylinux1_x86_64.whl (99 kB)
Requirement already satisfied: triton==2.2.0 in /usr/local/lib/python3.10/dist-packages (from torch->clip==1.0) (2.2.0)
Collecting nvidia-nvjitlink-cu12 (from nvidia-cusolver-cu12==11.4.5.107->torch->clip==1.0)
```

```
# Carga de aplicación y herramienta de importación de datasets de Fifty One.
import fiftyone as fo
import fiftyone.brain as fob
import fiftyone.zoo as foz
from fiftyone import ViewField as F
```

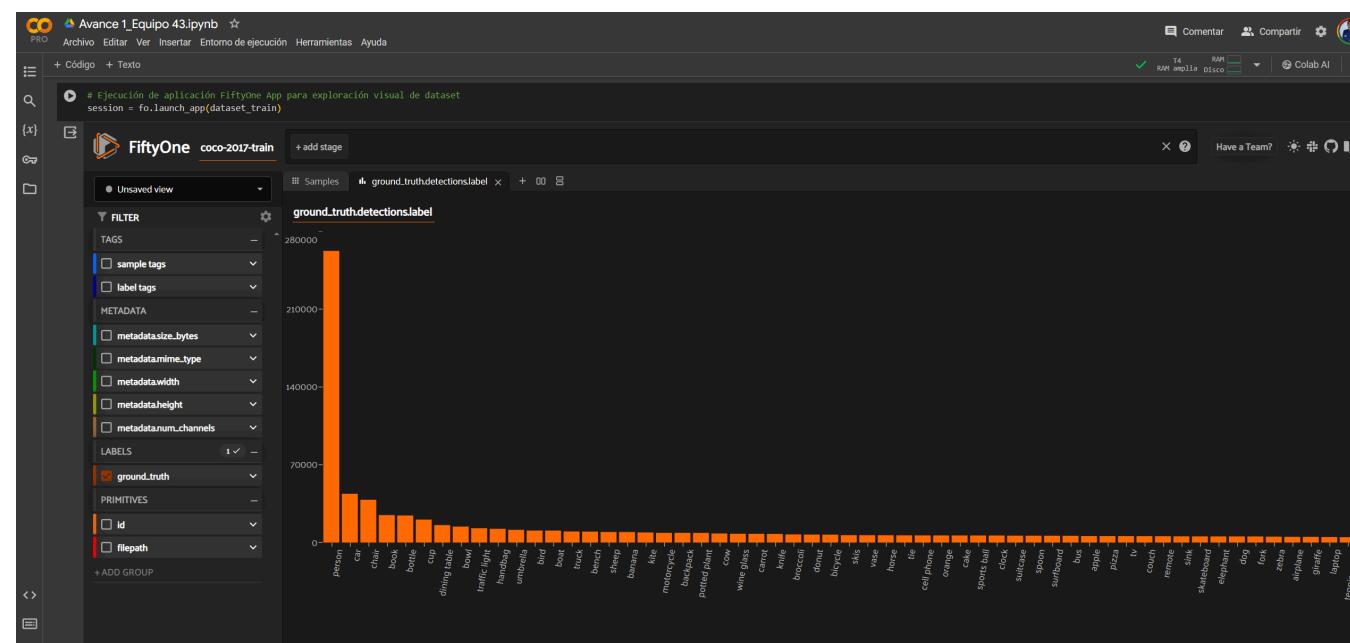
▼ COCO Dataset - Train Split

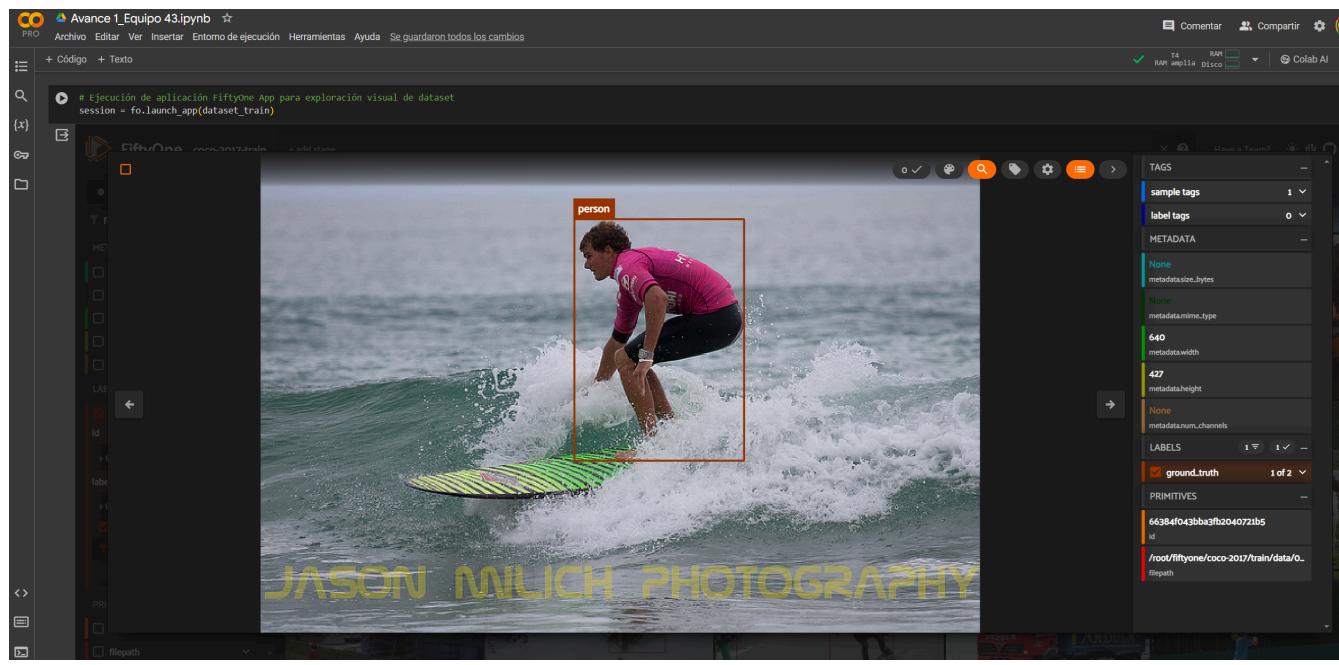
```
# Download and load the train split of COCO-2017
dataset_train = foz.load_zoo_dataset("coco-2017", split="train", label_types=["segmentations"])

# Resumen general del dataset split creado
print(dataset_train)

# Ejecución de aplicación FiftyOne App para exploración visual de dataset
session = fo.launch_app(dataset_train)
```

Capturas de Exploración inicial



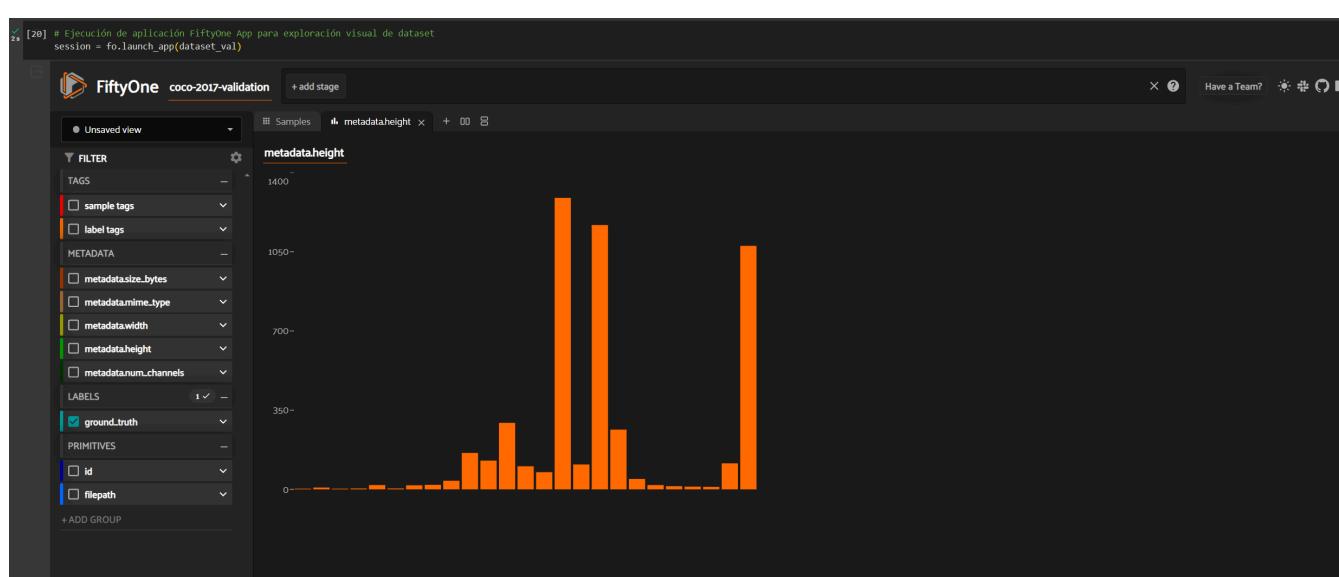
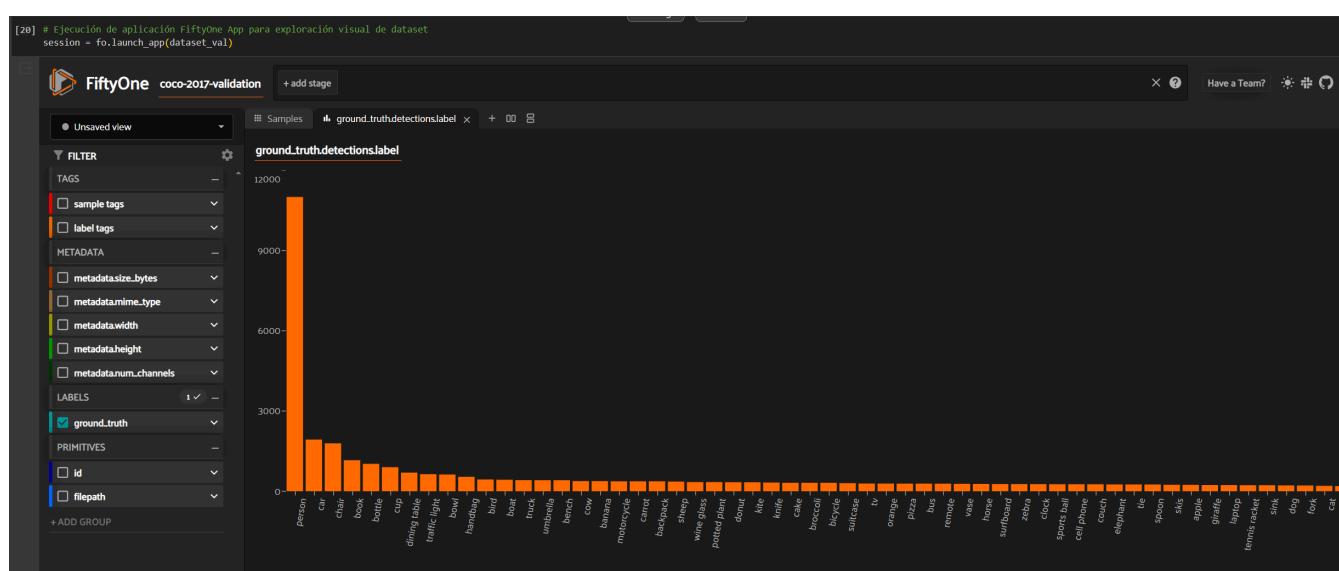
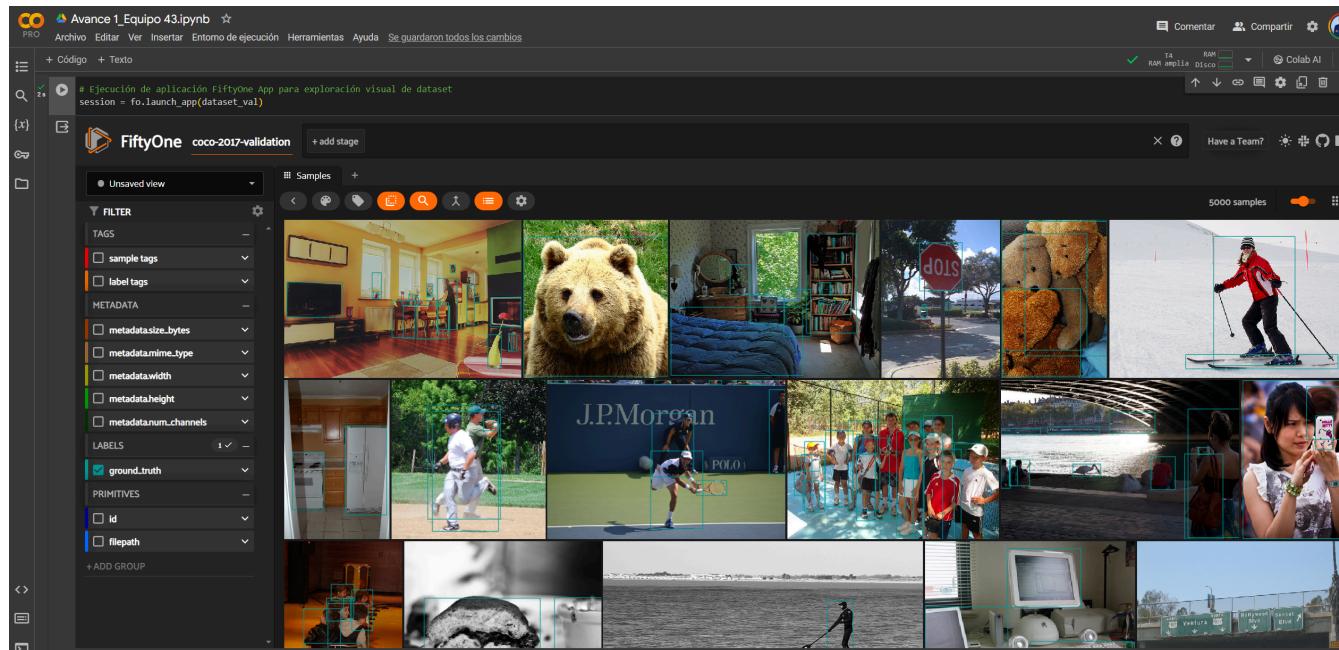


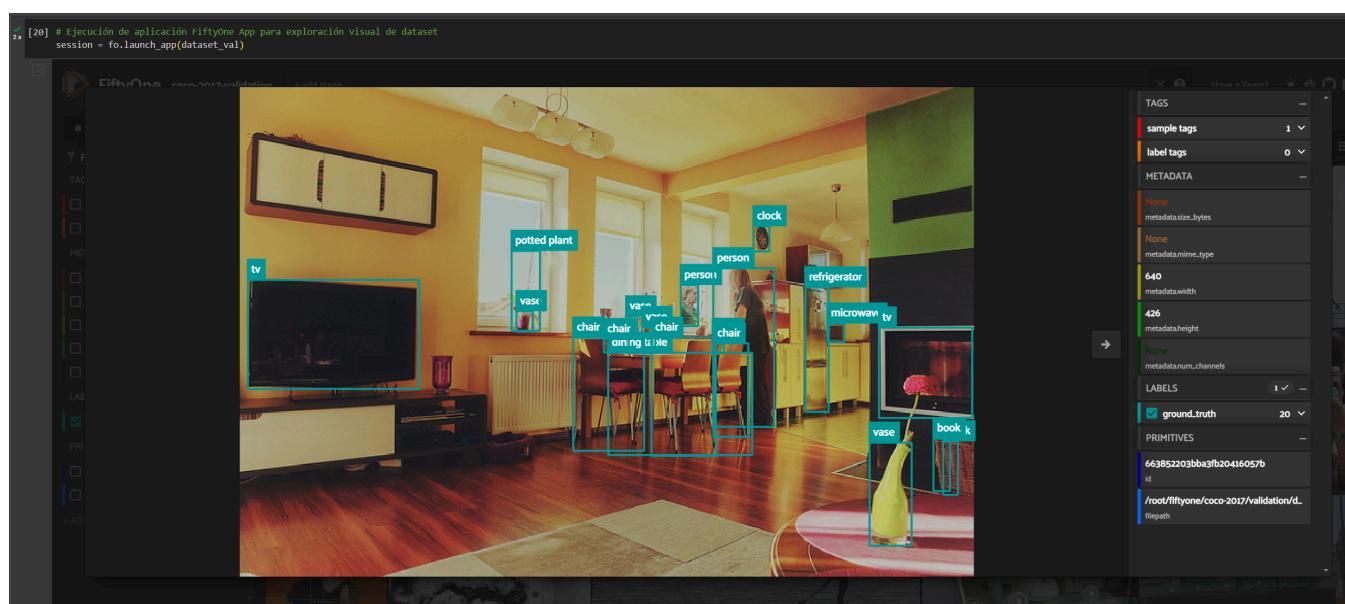
▼ COCO Dataset - Validation Split

```
# Download and load the validation split of COCO-2017
dataset_val = foz.load_zoo_dataset("coco-2017", split="validation", label_types=["segmentations"])

# Resumen general del dataset split creado
print(dataset_val)

# Ejecución de aplicación FiftyOne App para exploración visual de dataset
session = fo.launch_app(dataset_val)
```



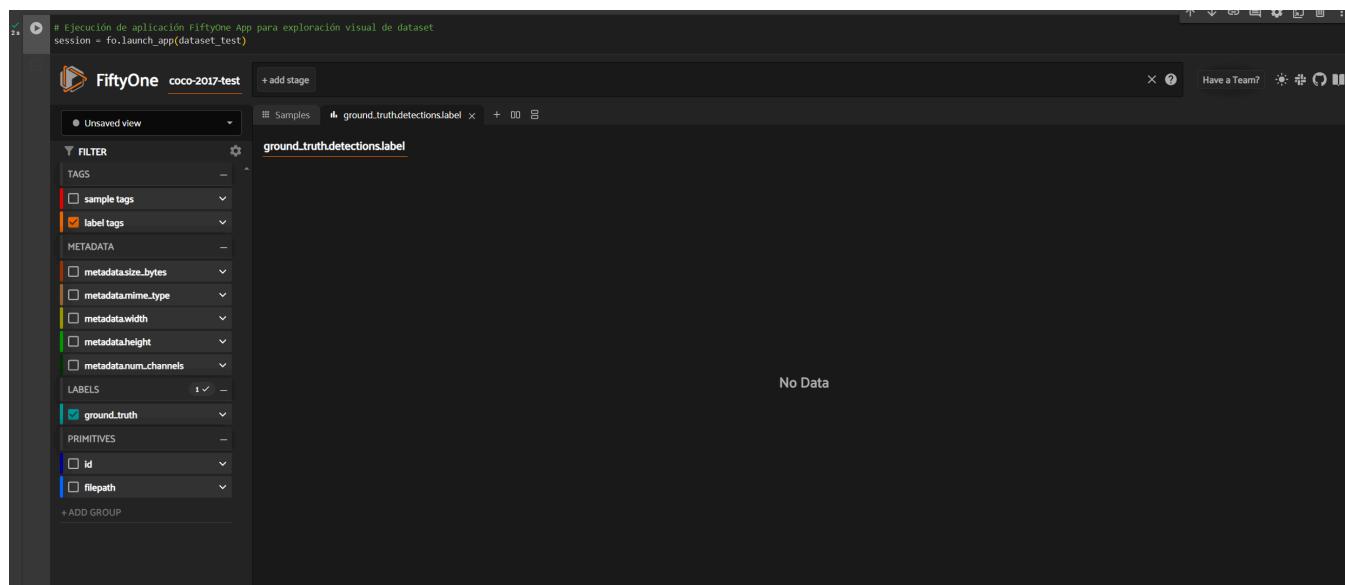
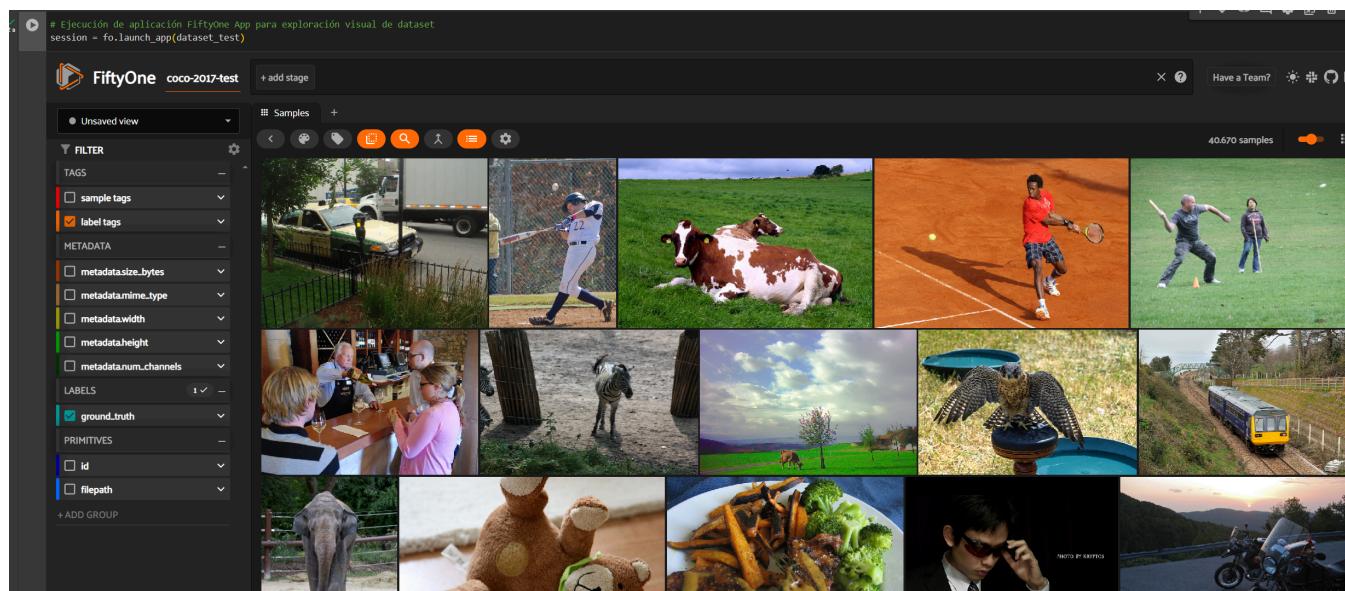


▼ COCO Dataset - Test Split

```
# Download and load the test split of COCO-2017
dataset_test = foz.load_zoo_dataset("coco-2017", split="test", label_types=["segmentations"])

# Resumen general del dataset split creado
print(dataset_test)

# Ejecución de aplicación FiftyOne App para exploración visual de dataset
session = fo.launch_app(dataset_test)
```



▼ 3 - Análisis Exploratorio de Datos

▼ Clustering Analysis

El clustering es una técnica esencial de aprendizaje no supervisado que permite descubrir patrones ocultos en los datos. El clustering, o análisis de grupos, es un conjunto de técnicas para agrupar puntos de datos. Los algoritmos de clustering toman una serie de objetos y devuelven asignaciones para cada objeto.

A diferencia de la clasificación, el clustering no comienza con una lista de clases para categorizar los objetos, obligando a los objetos a caer en grupos predefinidos. Más bien, el clustering intenta descubrir los grupos dados los datos. En otras palabras, el clustering trata de descubrir la estructura en los datos, no predecir etiquetas en una estructura preexistente [10].

Análisis sobre validation dataset

Para realizar el análisis por clustering se utilizan las siguientes librerías:

1. **scikit-learn** para realizar el clustering de datos una vez generados los features
2. **fiftyone** para ejecutar la función consolidadora "compute_visualization()"
3. **OpenAI's CLIP GitHub repo** para generar las características (features) utilizando el modelo CLIP ("clip-vit-base32-torch")
4. **umap-learn** para la reducción de dimensionalidad con el objetivo de permitir la visualización de características en un espacio de menor dimensión (2D en este caso)

Con estas herramientas se generarán las características utilizando CLIP y aplicamos una reducción de dimensionalidad UMAP para visualizar las características en 2D.

```
# delete labels to simulate starting with unlabeled data
dataset_cluster_val = dataset_val
dataset_cluster_val.select_fields().keep_fields()
dataset_cluster_val.name = "clustering-demo"
dataset_cluster_val.persistent = True
```

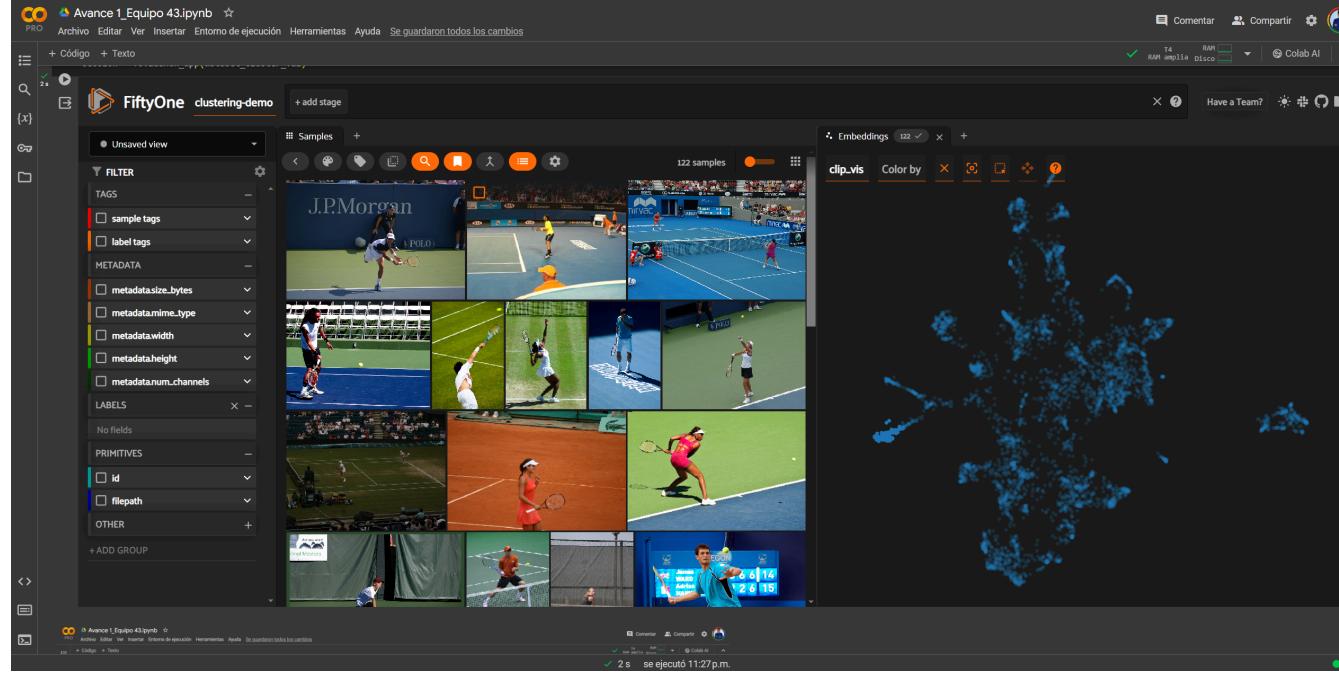
Con los vectores de características en mano, utilizamos el Plugin de Clustering de FiftyOne para estructurar nuestros datos. En la aplicación FiftyOne, ejecutamos el comando compute_clusters y seleccionamos "kmeans" como método de clustering y "clip_umap" como nuestros vectores de características. Establecimos el número de clusters en 20 y ejecutamos el algoritmo de clustering. Una vez completado, observamos en nuestras muestras un nuevo campo que contiene los números de cluster asignados a cada muestra.

```
# Creacion de variable local para uso en cálculo de embeddings
dataset = dataset_cluster_val

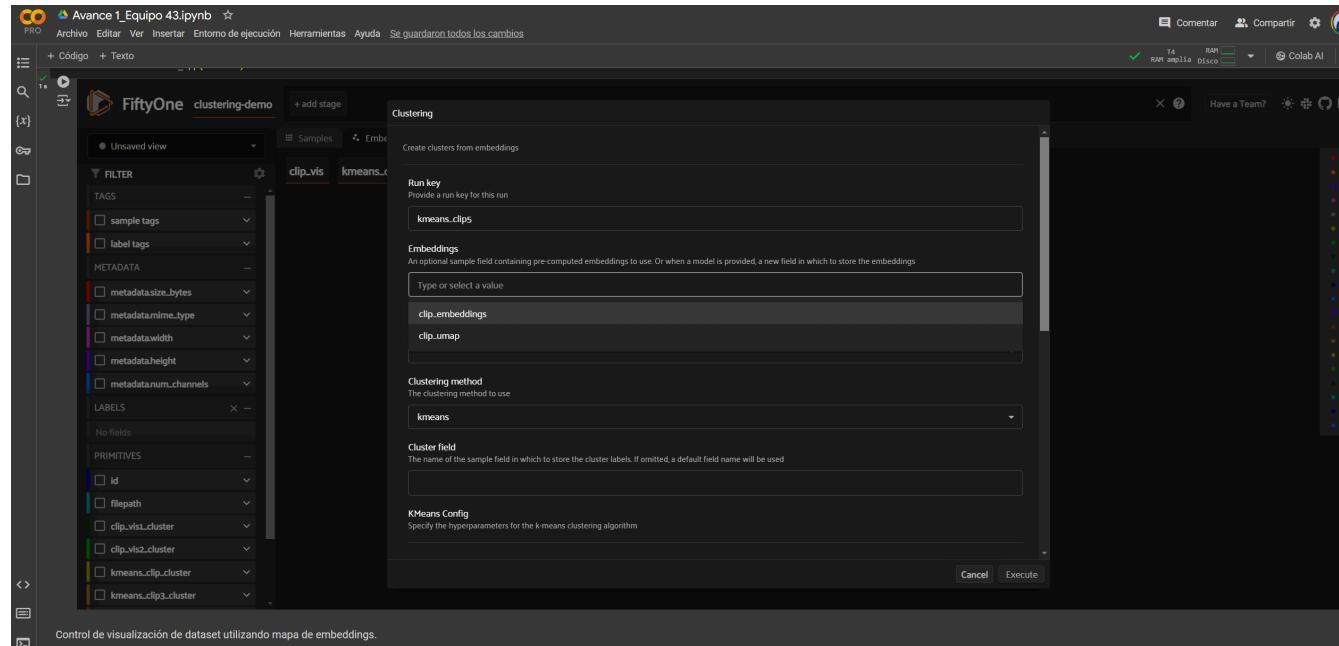
# Definición de reducción de dimensión por cálculo de embeddings utilizando CLIP
res = fob.compute_visualization(
    dataset,
    model="clip-vit-base32-torch",
    embeddings="clip_embeddings",
    method="umap",
    brain_key="clip_vis",
    batch_size=10
)
dataset.set_values("clip_umap", res.current_points)

# launch the app to visualize the dataset and embeddings
session = fo.launch_app(dataset)
```

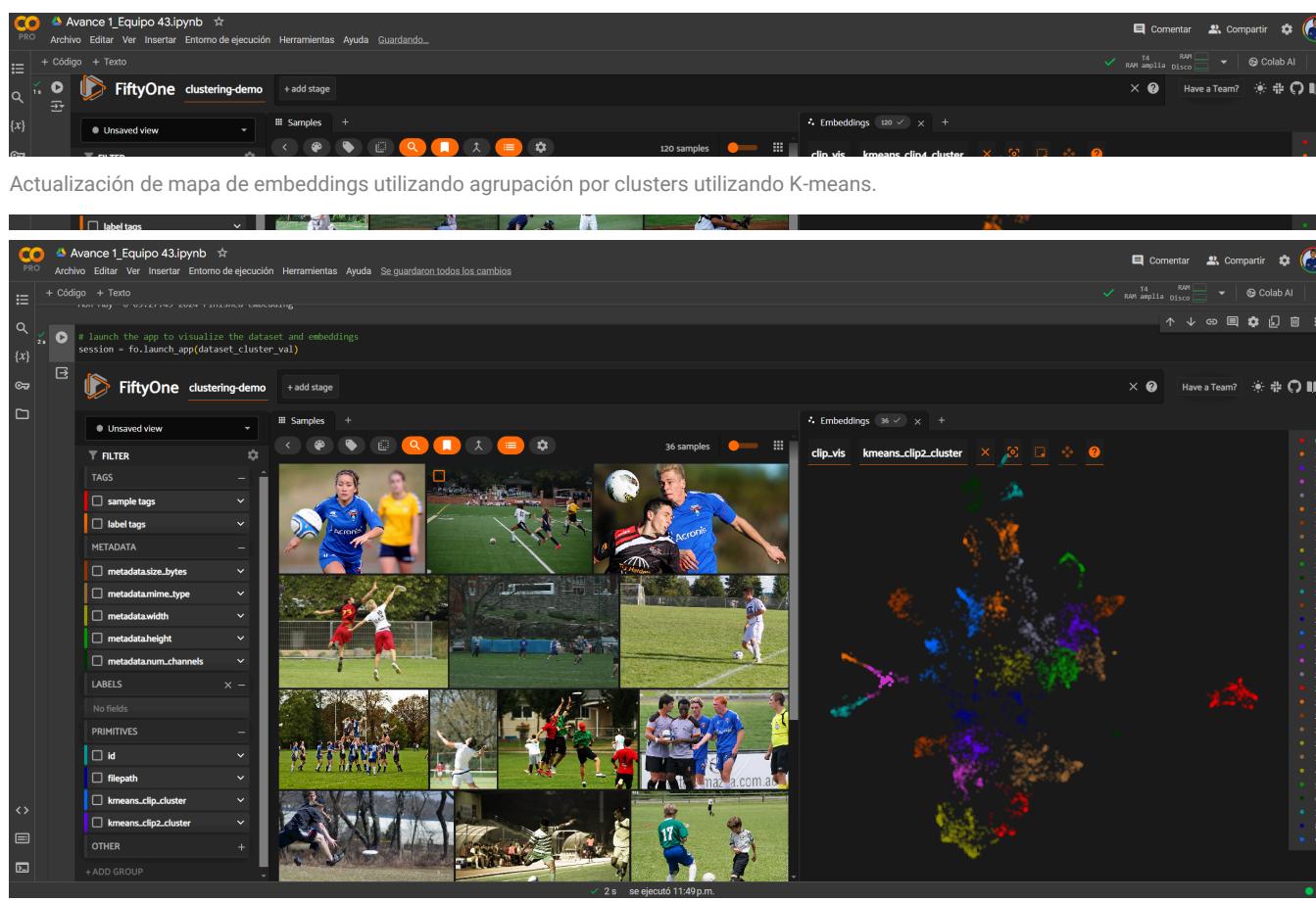
Control de visualización de dataset utilizando mapa de embeddings.



Segmentación de visualización en clusters utilizando CLIP Embeddings



Al visualizar los clusters, confirmamos que el clustering separó claramente las diferentes clases de imágenes. En el ejemplo se muestra la separación en 20 clusters donde se muestran las diferentes clases entre imágenes de deporte.



▼ Preguntas de Análisis

3.1 - ¿Cuáles son las estadísticas resumidas del conjunto de datos?

El dataset de COCO contiene más de 200,000 imágenes y más de 2 millones de instancias de objetos segmentado en 3 splits (train, validation, test). Se divide en 80 categorías de objetos y 91 categorías de "stuff" o materiales.

3.2 - ¿Cuál es la distribución de clases en el conjunto de datos?

La distribución de clases en COCO es desigual, con algunas categorías que contienen muchas más instancias que otras. La clase predominante es la de personas.

3.3 - ¿Se deberían normalizar las imágenes para visualizarlas mejor?

Se determina explorar técnicas de augmentación de datos para generar variantes de las imágenes del dataset que simulen diferentes condiciones de iluminación con el objetivo de preparar el modelo a entrenar para alta precisión.

3.4 - ¿Cuál es la variabilidad en términos de tamaño y resolución de las imágenes en el dataset?

as imágenes en COCO varían en tamaño y resolución, con un promedio de 640x480 píxeles, pero con imágenes tanto más grandes como más pequeñas.

3.5 - ¿Qué tipo de anotaciones están disponibles en el dataset?

El dataset proporciona anotaciones para detección de objetos, segmentación de imágenes, keypoints, y captions.

3.6 - ¿Qué estrategias de preprocesamiento podrían ser útiles para mejorar la detección del color en el dataset?

Para mejorar la detección del color, se considerará aplicar normalización por canal, ajuste de contraste y brillo, así como redimensionamiento para garantizar una entrada consistente al modelo.

▼ 4 - Referencias

1. Smits, Thomas & Wevers, Melvin. (2021). The agency of computer vision models as optical instruments. *Visual Communication*. 21. 147035722199209. 10.1177/147035722199209.
2. Top 10 Datasets for Computer Vision · Syntic AI. (2022). Retrieved from <https://www.syntic.ai/post/unlocking-computer-vision-the-10-best-datasets>
3. Common Objects in Context. (n.d.). Retrieved from <https://cocodataset.org/#home>
4. Meel, V. (2024). What is the COCO Dataset? What you need to know in 2024. Retrieved from <https://viso.ai/computer-vision/coco-dataset/>
5. Maitlo, N., Noonari, N., Ghanghro, S. A., Duraisamy, S., & Ahmed, F. (2024). Color Recognition in Challenging Lighting Environments: CNN Approach. ArXiv. [/abs/2402.04762](https://arxiv.org/abs/2402.04762)
6. De, K., & Pedersen, M. (2021). Impact of Colour on Robustness of Deep Neural Networks. 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 21–30. <https://doi.org/10.1109/ICCVW54120.2021.00009>
7. FiftyOne by Voxel 51. (n.d.). Retrieved from <https://docs.voxel51.com/>
8. COCO Dataset Integration in Fifty One. (n.d.). Retrieved from <https://docs.voxel51.com/integrations/coco.html?highlight=coco>