

# Inteligencia Artificial

## Redes Neuronales Problemas Multilabel

Juan Malo

Jorge Márquez

Roberto Romero

Erika Villa

# Descripción del Dataset Original

**Cantidad de Imágenes:** 118287 imágenes.

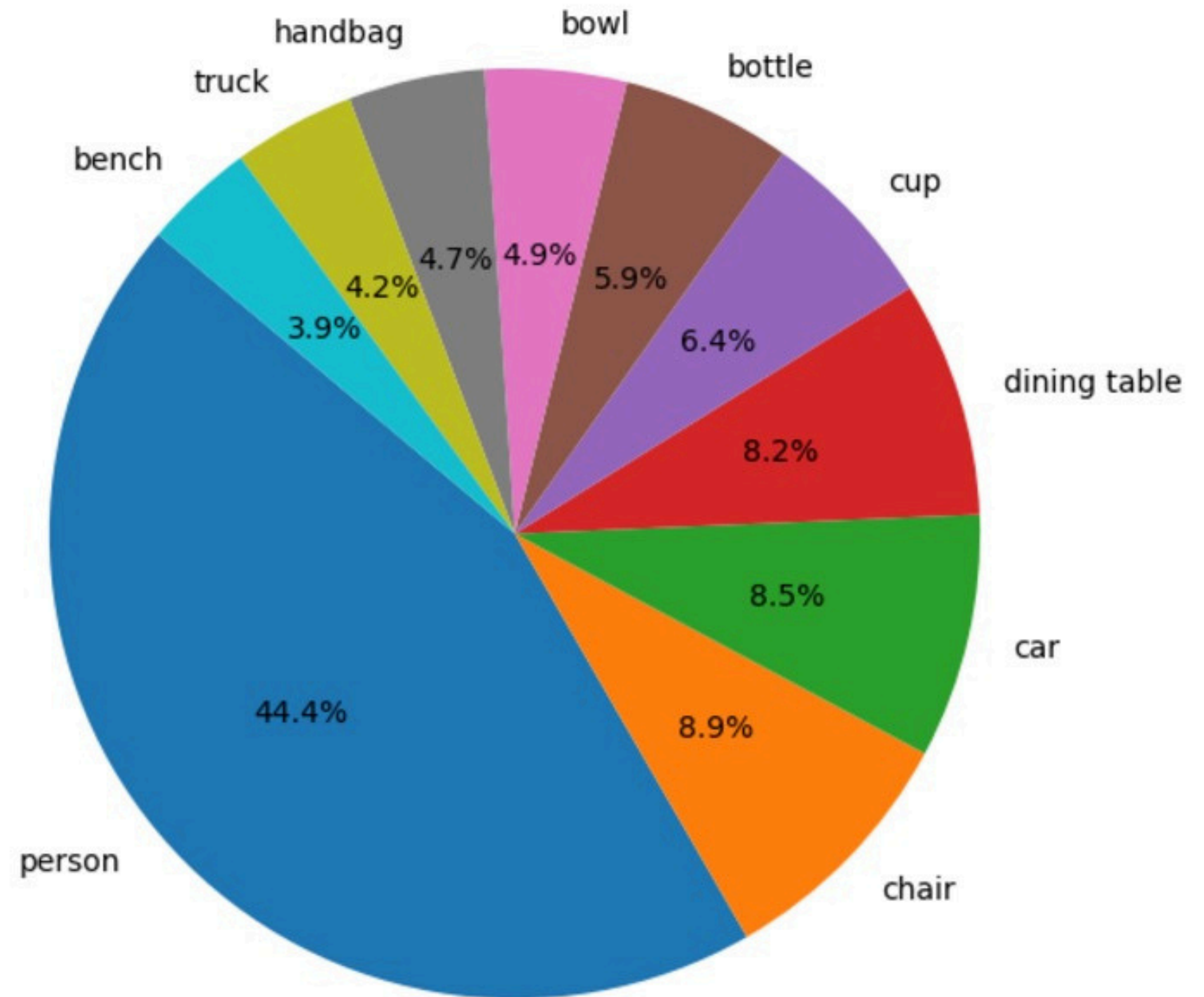
**Dimensiones de las Imágenes:** Tamaño variable

**Variable de Estudio (Y):** 80 categorías (etiquetas).

# Análisis de frecuencia de categorías

Total de imágenes dataset

Distribución de las 10 categorías principales



# Descripción del Dataset Balanceado

**Cantidad de Imágenes:** 45432 imágenes para entrenamiento.

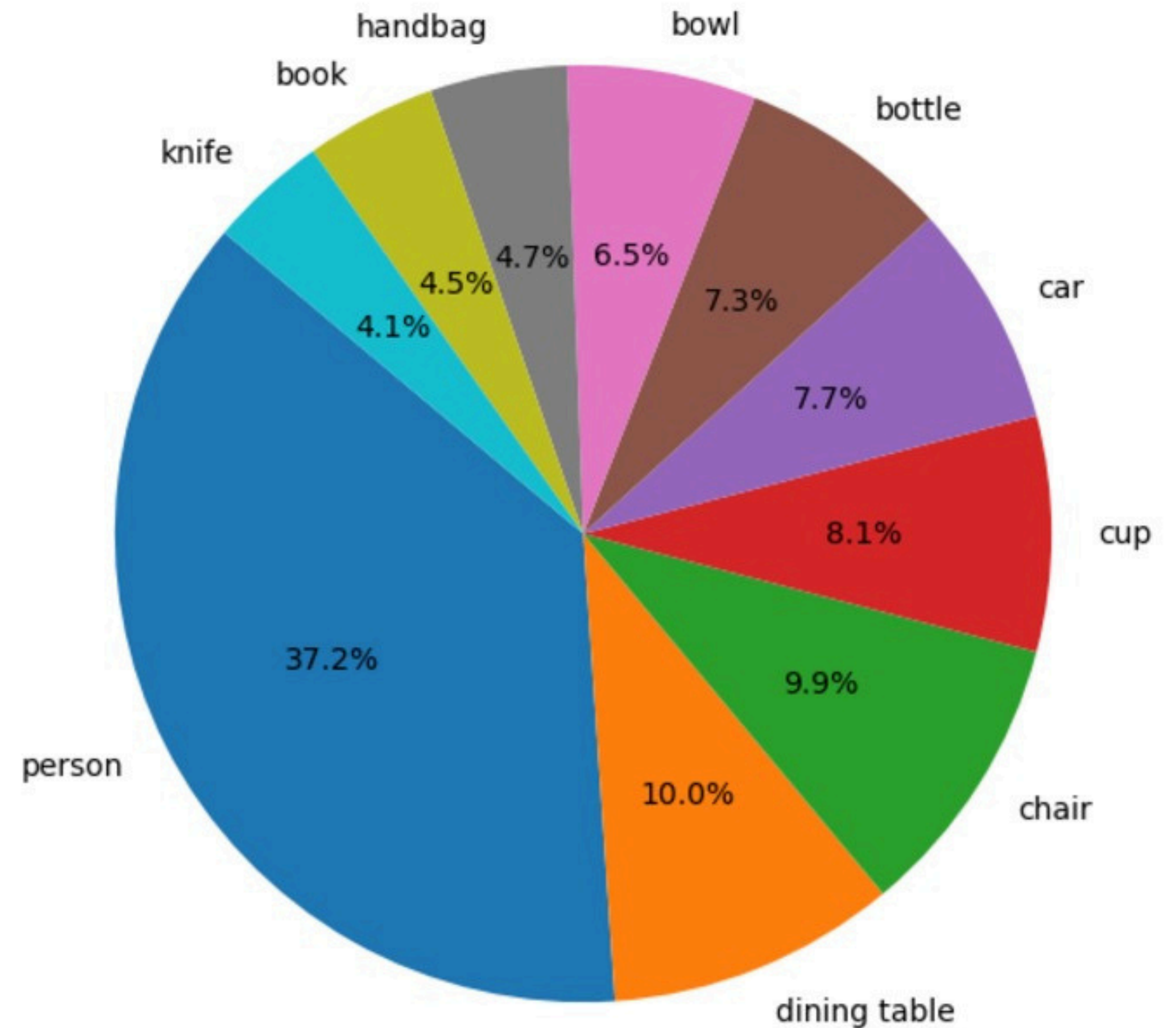
**Dimensiones de las Imágenes:** Tamaño variable

**Variable de Estudio (Y):** 78 categorías (etiquetas), dos excluidas (toaster, hair drier).

# Análisis de frecuencia de categorías

Dataset balanceado

Distribución de las 10 categorías principales





# Exploración de Imágenes

Imagen: 000000531474.jpg

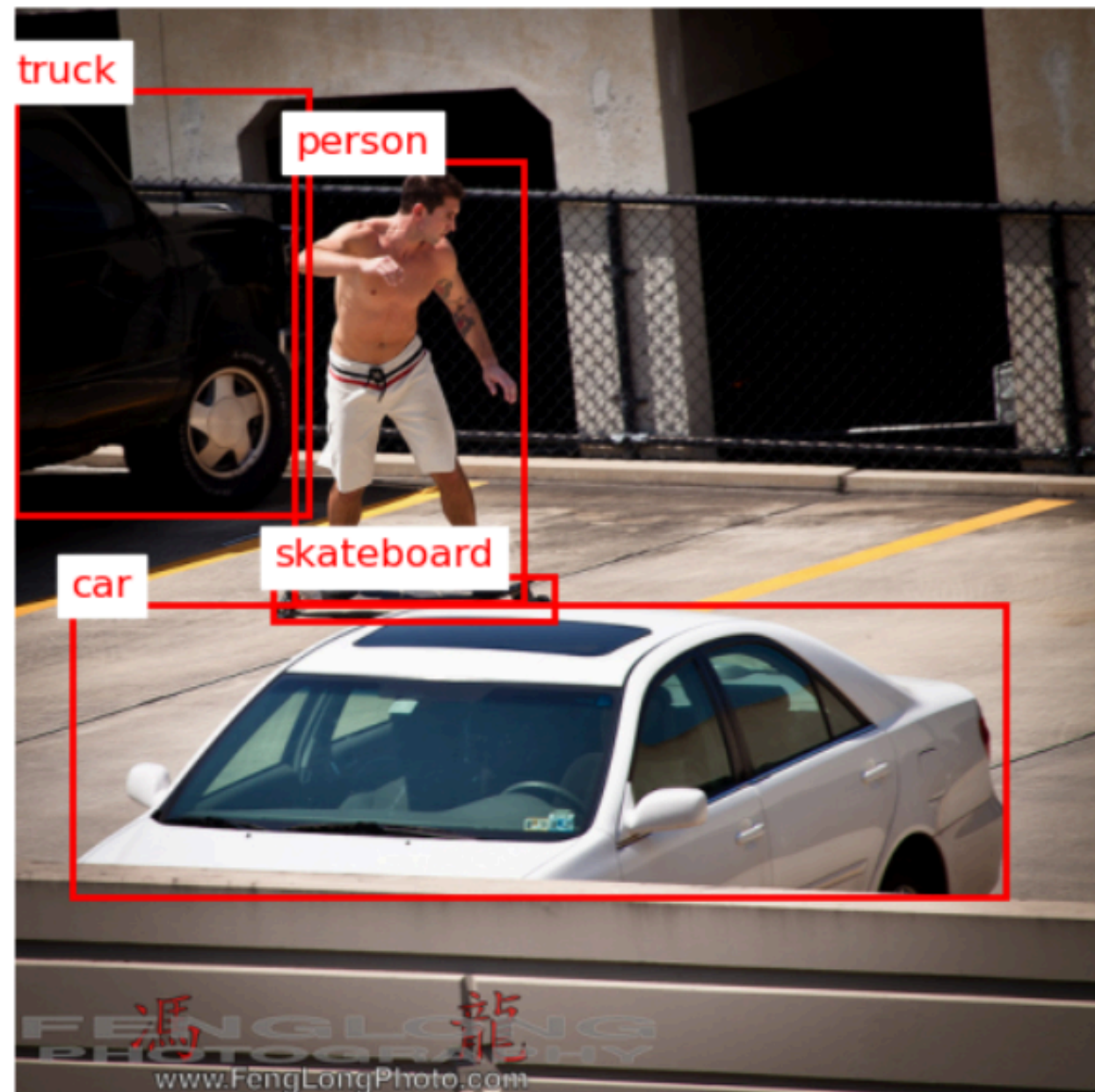
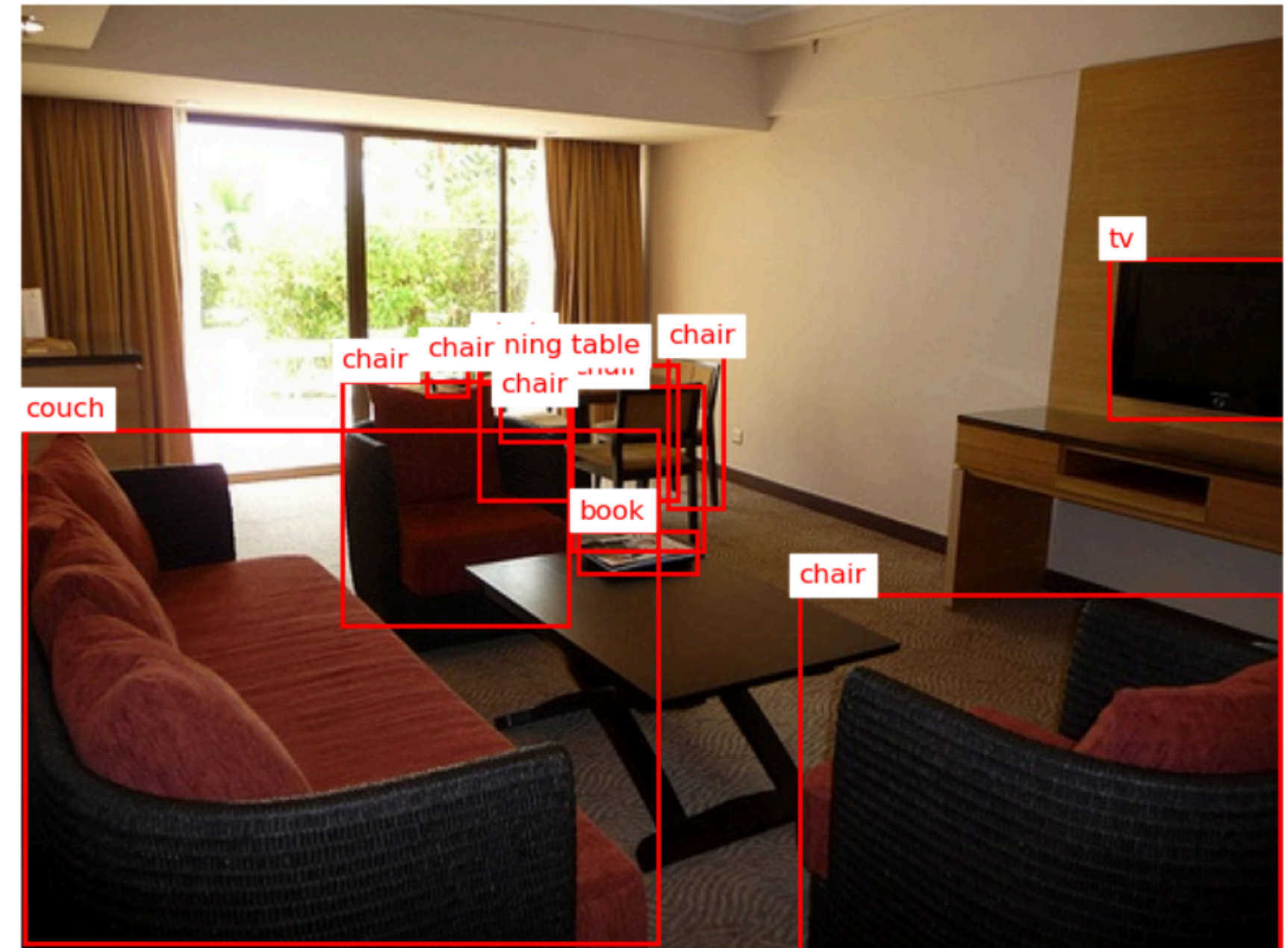


Imagen: 000000113879.jpg



# Resultados de calidad de train

## Test 1:

<b>Accuracy</b>	<b>0.7980</b>
<b>Precision</b>	<b>0.8548</b>
<b>Recall</b>	<b>0.8426</b>
<b>F1 Score:</b>	<b>0.8476</b>

## Test 2:

<b>Accuracy</b>	<b>0.4476</b>
<b>Precision</b>	<b>0.7419</b>
<b>Recall</b>	<b>0.6372</b>
<b>F1 Score:</b>	<b>0.6796</b>



# Resultados

Predicciones: bird



Predicciones: bowl, apple, sandwich, keyboard



Predicciones: person, toilet, tv, laptop, refrigerator, book, vase





## Conclusiones

- Se presentaron **dificultades** para **equilibrar** el **dataset** debido a una notable discrepancia en la distribución de las **etiquetas**. Por ejemplo, la etiqueta "persona" contenía 64,115 instancias, mientras que otras etiquetas contaban con menos de 189. Este **desajuste** generó problemas al momento de realizar el **balanceo**.
- Adicionalmente, se encontraron **limitaciones** significativas de **hardware** que impidieron llevar a cabo entrenamientos más extensos o utilizar modelos de mayor capacidad, lo cual **afectó el rendimiento general**.
- Otro desafío importante fue la presencia de **malas etiquetaciones en algunas imágenes**. En varios casos, se agrupaban múltiples personas bajo una única etiqueta, lo que dificultaba la correcta interpretación de los datos. Además, en ocasiones las etiquetas **no representaban con precisión los objetos** a los que hacían referencia, lo que comprometió la calidad de las anotaciones y, en consecuencia, el entrenamiento del modelo.

## Conclusiones

- Se presentó una discrepancia durante la validación del entrenamiento debido a la distribución desequilibrada de las imágenes en el dataset, lo que dificultó obtener resultados consistentes.
- Por otro lado, la librería **PyTorch** resultó ser más **fácil** de utilizar en comparación con **TensorFlow**. Con esta última, experimentamos varios inconvenientes al intentar configurar el entrenamiento para hacer uso de la tarjeta gráfica, lo que generó retrasos y complicaciones en el proceso.

## Bibliografía

- PyTorch. (2023). PyTorch: Un marco flexible para el aprendizaje profundo. PyTorch Foundation. Recuperado de <https://pytorch.org/>
- Torchvision. (2023). Torchvision: Biblioteca de visión por computadora para PyTorch. PyTorch Foundation. Recuperado de <https://pytorch.org/vision/stable/>
- YouTube. (s/f). Youtu.Be. Recuperado el 28 de noviembre de 2024, de [https://youtu.be/V\\_xro1bcAuA?si=JQyp5lyzFP1Ng-BE](https://youtu.be/V_xro1bcAuA?si=JQyp5lyzFP1Ng-BE)
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. arXiv preprint arXiv:1405.0312. <https://arxiv.org/abs/1405.0312>
- Deep Learning: clasificando imágenes con redes neuronales. (2020, mayo 25). LIS Data Solutions. <https://www.lisdatasolutions.com/es/blog/deep-learning-clasificando-imagenes-con-redes-neuronales/>



**GRACIAS !**