**Selected model** – Grounding DINO and SAM

Justification - Because they offer a zero-shot approach. This means they can identify and segment objects in images without having seen examples of these objects during training.

Others reason –

Performance – better than Mask-RCNN

Versatility – can adopt new unseen objects, making them highly flexible.

tools, libraries –

- Google Colab with GPU connection
- Torch
- Torchvision – used for mainly image transformations.
- Supervision – for visualizing

References –

https://python.plainenglish.io/building-zero-shot-object-detection-with-grounding-dino-eea6dc4adf36

https://pytorch.org/docs/stable/index.html

https://supervision.roboflow.com/latest/#quickstart

Output –