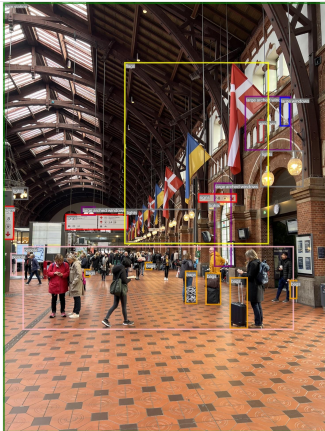


Visual Description Fine-tuned Pipeline Examples						
Name	Image	Notes	Segmentation - Grounding DINO (base) (example)	Phi3-mini (original)	Phi3-mini-VD (finetuned)	LLaVA-v1.6-vicuna-7B-VD (finetuned)
dog		from COCO validation	 <p>In the image, there is a dirt path that runs horizontally across the bottom of the scene. Along this path, there are two dogs. The first dog is positioned slightly to the right of the center of the path, with its body stretched out along the path. The second dog is located further to the right, with its body also stretched out along the path. Both dogs are wearing collars, with the first dog's collar positioned slightly above its body and the second dog's collar positioned slightly below its body. The dogs are connected by a leash, which runs from the first dog's collar to the second dog's collar. The leash is positioned slightly above the first dog's body and slightly below the second dog's body. The scene also features a river that runs horizontally across the top of the image, with the river's width being slightly less than the width of the image. The sky is present above the river, covering the top portion of the image.</p>	<p>In this serene scene, a river flows gently from left to right, taking up most of the space. Along the riverbank, there are patches of brown weeds and green grass, with some dried grass towards the right edge. Above the river, a blue sky stretches out. In the foreground, two dogs are standing near each other, one slightly in front of the other. The dog closer to the viewer is holding a leash, which extends from its body to the other dog. Reflected in the calm waters of the river is a portion of the scene, including the dogs and the surrounding environment.</p>	<p>Here is a descriptive paragraph of the dog walk scene:</p> <p>In this serene setting, two dogs are enjoying a walk along a winding path that follows the edge of a calm pond. The path is lined with lush green grass and weeds, which stretch from the left side to the right side of the image. The dogs are positioned on either side of the path, one closer to the pond and the other further away. The pond's surface reflects the surrounding environment, creating a sense of tranquility.</p>	
sheep		from COCO validation	 <p>In the scene, there is a road stretching from the left to the right side of the image. The sky is present above the road, with clouds scattered across it. Below the sky, there are mountains that stretch from the left to the right side of the image. The road runs horizontally across the middle of the image, with a suv parked on the left side and another suv parked on the right side. There are three cars parked on the road, with one car parked on the left side and two cars parked on the right side. In the center of the image, there are sheep grazing on the grassy area. The sheep are located near the middle of the image, with the grassy area stretching from the left to the right side. The sheep are positioned towards the center of the road, with some space between them. The road is not crowded with vehicles, allowing the sheep to graze comfortably. The trees are positioned on both sides of the road, providing shade and shelter for the sheep. The sky is clear, with no clouds present.</p>	<p>In the image, there is a highway scene with various objects present. The sky is located at the top of the image, with clouds scattered across it. Below the sky, there are mountains that stretch from the left to the right side of the image. The road runs horizontally across the middle of the image, with a suv parked on the left side and another suv parked on the right side. There are three cars parked on the road, with one car parked on the left side and two cars parked on the right side. In the center of the image, there are sheep grazing on the grassy area. The sheep are located near the middle of the image, with the grassy area stretching from the left to the right side. The sheep are positioned towards the center of the road, with some space between them. The road is not crowded with vehicles, allowing the sheep to graze comfortably. The trees are positioned on both sides of the road, providing shade and shelter for the sheep. The sky is clear, with no clouds present.</p>	<p>Here is a descriptive paragraph of the scene:</p> <p>In this scene, there are two cars parked on the side of the road, one near the left edge and another closer to the right edge. The road stretches across the entire width of the image, with a clear path for vehicles to travel. Along the roadside, there are several trees scattered about, some of which are quite tall. In the distance, you can see a mountain range rising up from the horizon. Closer to the foreground, there is a sheep grazing peacefully on the grassy verge. The sky above is visible, taking up the top portion of the image.</p>	

central station



In this indoor train station scene, there are several objects present. The scene is dominated by large arched windows, which are located on the left side of the image. These windows are positioned at the top, middle, and bottom sections of the image. There are also smaller windows on the right side of the image, located at the top and bottom sections.

In the center of the image, there are lights that illuminate the scene. These lights are positioned at the top, middle, and bottom sections of the image. There are also signs present, which are located on the left side of the image. These signs are positioned at the top, middle, and bottom sections of the image.

In the middle of the image, there are people present. These people are positioned at the top and bottom sections of the image. There are also bags present, which are scattered throughout the image. These bags are positioned at various locations, with some near the top, middle, and bottom sections of the image.

At the top of the image, there are flags present. These flags are positioned at the top section of the image. The entire scene is contained within a train station, which occupies the entire image.

You are standing in a bustling train station with a large crowd of people moving around you. Above, there are ceiling beams stretching across the room. On the walls, you'll find several lighting fixtures at regular intervals.

To your left, near the edge of the room, there's a flag hanging from the ceiling. On the floor, you'll notice a series of floor tiles that cover the entire area.

As you look around, you'll see multiple backpacks scattered throughout the crowd, some placed on the floor and others held by people. There are also several handbags and luggage items, with a few people carrying them.

On the wall to your right, there's a clock mounted at a comfortable height. The atmosphere is lively, with people moving in all directions, but you can still hear announcements over the loudspeakers.

You are standing in a large train station with a high ceiling. The ceiling is supported by multiple beams that stretch across the room. Above you, there are several flags waving gently.

To your left, you'll find a row of benches where people can sit and rest. There are also several poles scattered throughout the area, likely used for support or as part of the station's infrastructure.

As you look around, you'll notice many people moving about, some standing still while others walk quickly. Some people are gathered near the benches, while others are spread out across the room.

Ahead of you, there's a large train platform where trains arrive and depart. You can see the train tracks stretching out into the distance.

To your right, there's a row of windows that let in natural light, and above them, more flags are waving. The walls of the station are made of bricks, which cover most of the area.

Overall, the atmosphere is bustling with activity as people move about, preparing for their journeys or waiting for their trains to arrive.

street



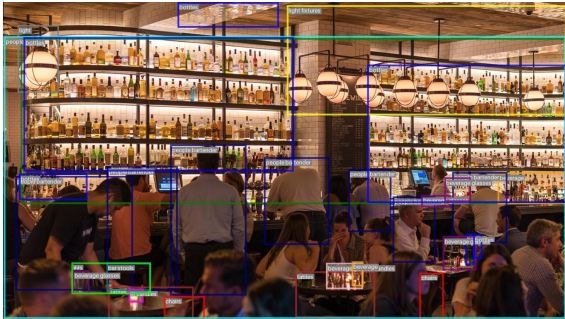
In this bustling urban street scene, tall buildings line the sides, stretching from the left to the right of the image. A variety of pedestrians can be seen, some walking alone, others in small groups or crowds. There are also several bicycles scattered throughout the scene, with one bicycle parked near a vehicle. A few handbags and backpacks are visible, carried by individuals walking past. A stroller is present, likely with a child in tow. A few trees can be seen, adding a touch of greenery to the urban environment. Signs are present, providing direction and information to the pedestrians. The street itself is the main thoroughfare, with vehicles and pedestrians sharing the space.

You are standing on a bustling city street surrounded by tall buildings that stretch up towards the sky. The sidewalk beneath your feet stretches out in front of you, lined with trees on either side. As you look around, you notice a variety of people going about their daily business. There are pedestrians walking along the sidewalk, some pushing strollers or bicycles. You see bicycles parked along the curb, some near the buildings and others closer to the street. Cars are driving by, some parked along the curb while others are moving down the road. You notice several signs posted along the street, some attached to buildings and others standing alone. The atmosphere is lively, with people going about their daily routines amidst the hustle and bustle of city life.

Here is a descriptive paragraph of the busy street scene:

You are standing in front of a bustling street with a sidewalk running along its edge. The sidewalk is lined with trees and has several signs posted on it. To your left, you see a row of buildings, including a large one that takes up most of the block. On the street itself, there are many people walking in different directions, some carrying bags or wearing hats. You notice bicycles parked along the curb, and there are also cars driving by. Above the sidewalk, there are signs mounted on poles, and more signs are posted on the buildings. The atmosphere is lively, with many people going about their daily business.

bar



In this bar scene, there are several objects present. At the top of the scene, there are two light fixtures that provide illumination. Below them, there are two bar stools positioned on the right side of the scene. On the left side, there are two tables, one near the top and another closer to the bottom. On the tables, there are several bottles, glasses, and shelves. The shelves are located at the bottom of the scene, and they contain various bottles.

In the center of the scene, there is a bartender who is standing behind a bar counter. The bartender is wearing a uniform and is preparing drinks for the customers. There are two beverage glasses on the counter, and a beverage is being poured into one of them. There are also several beverage glasses on the tables, and some of them have candles placed in them.

On the right side of the scene, there are three people sitting on bar stools. They are enjoying their drinks and engaging in conversation. The bartender is attending to their needs, taking their orders and serving them their beverages.

In summary, this bar scene features a bartender, customers, bar stools, tables, shelves, and various beverages and glasses. The scene is well-lit, and the atmosphere is lively and social.

You are standing in a lively bar atmosphere with many customers seated and standing around you. There are several bottles placed on shelves or tables throughout the room. Above you, there are two chandeliers hanging from the ceiling, one near the front of the room and another towards the back. The lighting in the room is well-distributed, with fixtures placed at various heights. You notice several patrons sitting on stools or chairs, some of whom appear to be bartenders attending to customers. There are also chairs scattered around the room, some of which are occupied by customers. The overall atmosphere is one of socializing and relaxation.

Here is a descriptive paragraph of the crowded bar scene:

You are standing in a bustling bar with a long counter stretching across the room. The counter is lined with bottles and glasses, and there are several menus displayed on it. Above the counter, there are three light fixtures hanging from the ceiling. To your left, you see a group of people gathered around the counter, while to your right, there's another group of people seated at tables. The tables are scattered throughout the room, with some near the walls and others in the center of the room. You notice several menus placed on the tables, likely belonging to the patrons. There are also several glasses and bottles scattered around the room, some on the counter and others on the tables. The atmosphere is lively, with many people socializing and enjoying their drinks.