



Scenario	User Inputs	Desired Outputs
Location Recognition [2.1]	<p>Make an educated guess about the name of the landmark shown in the image. Think step by step, and then output your answer in the last line. Choose one of the options below as your answer:</p> <p>A. Arizona State Capitol B. California State Capitol C. Nebraska State Capitol</p> 	<p>[Reasoning Omitted] Based on the architectural style, the dome, and the layout of the grounds, the image most closely resembles Nebraska State Capitol</p>
Image Captioning [2.2]	<p>Generate a caption for the image in one sentence. Make sure to follow the following instructions: 1. Describe all the important parts of the remote sensing image.</p>  <p>[Instructions Omitted]</p>	<p>Aerial view of an airport terminal with adjacent aircraft, taxiways, and parking areas.</p>
Land Use & Land Cover Classification [2.3]	<p>You are given a satellite image and a list of land usage types or object names. Classify the image into one of the following options. Choose the best option that describes the given image. A list of possible options:</p>  <p>[Options Omitted]</p>	<p>The image appears to show a collection of lined up containers in a specific pattern, which is characteristic of a shipping yard. Therefore, the best option to describe the given image is: 28. Shipping Yard ✓</p>
Object Localization [3.1]	<p>You are given an 800 x 800 satellite image. Identify the extent of the object in the description below in the format of [xmin, ymin, xmax, ymax], where the top-left coordinate is (xmin, ymin) and the bottom-right coordinate is (xmax, ymax). You should answer the extent without further explanation. Description: The gray windmill in the middle</p> 	<p>[233, 383, 376, 542]</p>
Object Counting [3.2]	<p>Count the number of trees in the given image to the best of your ability. Output your count only without any further explanation.</p> 	<p>134</p>
Change Detection [4]	<p>You are given two satellite images taken before and after a natural disaster. The first image was taken before the natural disaster. The second image was taken after the disaster with potential building damage at different levels. [Descriptions of damage levels] Count the number of buildings in each damage category. Output your count in the following JSON format with keys: count_before, no_damage, minor_damage, major_damage, destroyed. You don't have to give extra explanations.</p> 	<pre>```json {"count_before": 75, "no_damage": 2, "minor_damage": 73, "major_damage": 0, "destroyed": 0} ```</pre>