Task 4.1 - P

Computer Vision

Saurabh Dharmadhikari

Saurabh.dharma01@gmail.com

Detecting objects with

Coordinates using

Azure computer vision

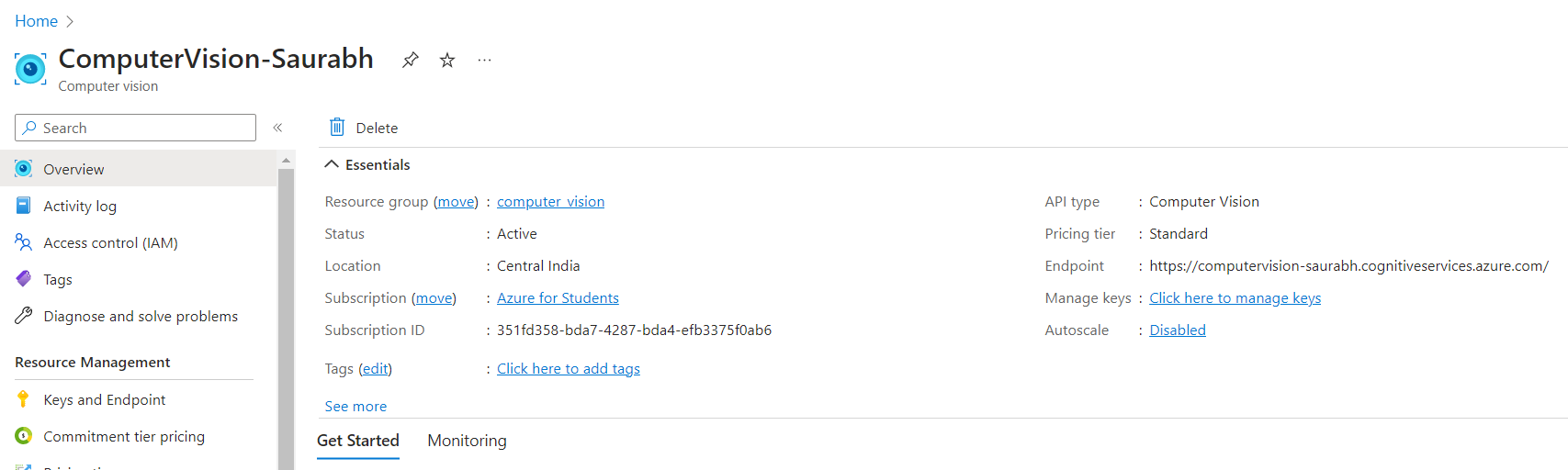
|  |  |
| --- | --- |
| Question 1. **Please explain cell by cell of your code from reading a local image to object detection, drawing a bounding box around different object. To complete this task, you need to provide the screenshot of your code and explain cell by cell of the code and explain what sort of API is being used.** | 2 |

Table of Content:

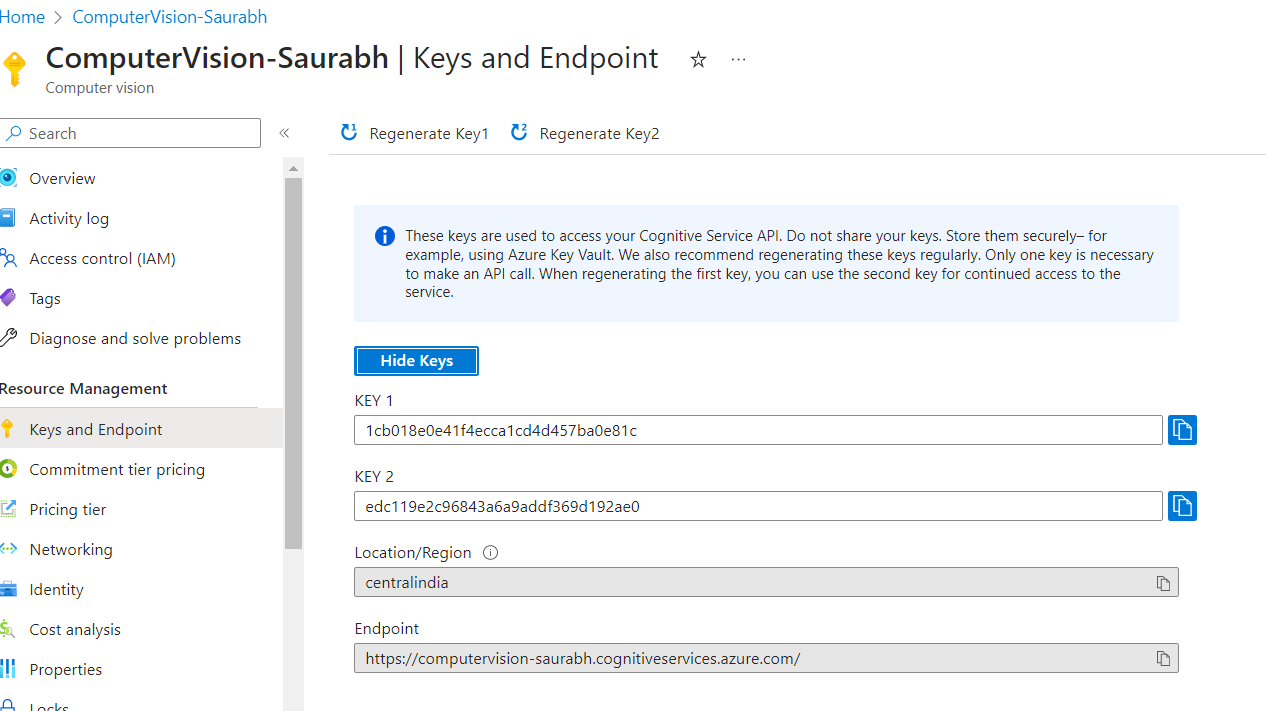
1. **Please explain cell by cell of your code from reading a local image to object detection, drawing a bounding box around different object. To complete this task, you need to provide the screenshot of your code and explain cell by cell of the code and explain what sort of API is being used.**

We first create a computer vision resource on Azure.

Below is the overview of our resource. Location is central India and other details.

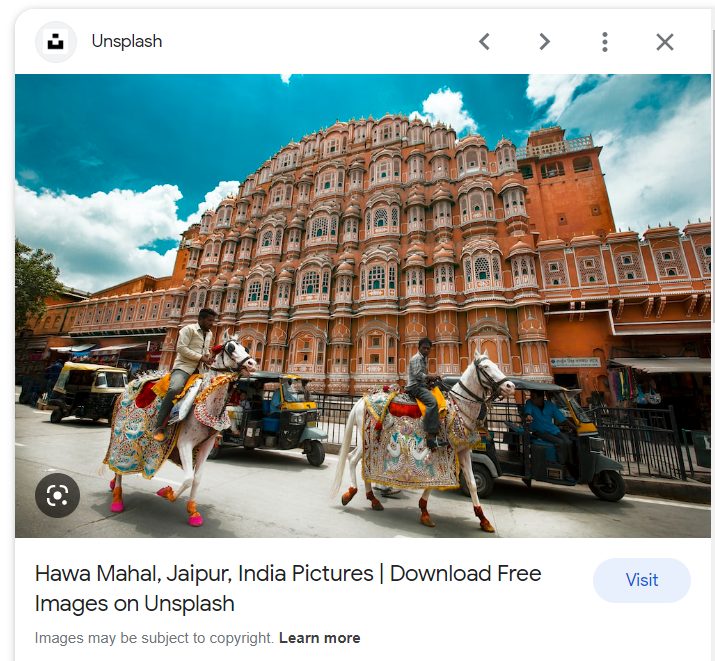


We can find keys from the left side of the screen in Keys and Endpoint.



We have created computer vision on azure to use these cognitive services using key and endpoint to later use it in python.

Below is the image that we are going to use in computer vision to detect, describe and find tags.



We are using a photograph by Siva (2016).

### Web address from where the image address was taken[¶](http://localhost:8888/notebooks/Object%20Detection.ipynb#Web-address-from-where-the-image-address-was-taken)

<https://www.google.com/search?q=unsplash+hawa+mahal&tbm=isch&ved=2ahUKEwiXxL_I3bH-AhXt-3MBHQoWBT0Q2-cCegQIABAA&oq=unsplash+hawa+mahal&gs_lcp=CgNpbWcQAzoECCMQJzoICAAQgAQQsQM6BQgAEIAEOggIABCxAxCDAToECAAQAzoHCAAQigUQQzoKCAAQigUQsQMQQzoLCAAQgAQQsQMQgwE6BwgAEBgQgARQjGRYpbUBYIC4AWgAcAB4AIAB7QWIAfQgkgEOMC4xNy4wLjEuMi4wLjGYAQCgAQGqAQtnd3Mtd2l6LWltZ8ABAQ&sclient=img&ei=zqg9ZJelOu33z7sPiqyU6AM&bih=746&biw=1536&rlz=1C1UEAD_enIN996IN996#imgrc=6mhniEGQhG9JHM>

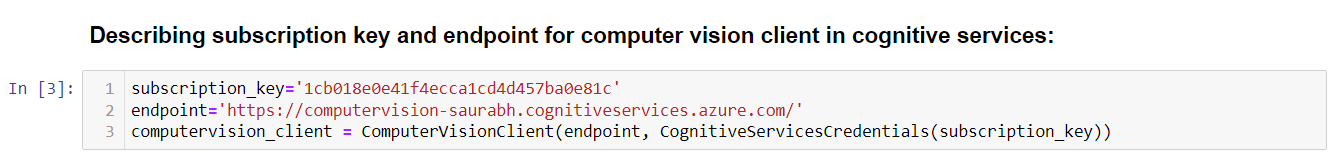
Courtesy Unsplash images.

Importing libraries:



We first import all the necessary libraries on jupyter notebook needed for computer vision. This includes CommputerVisionClient, OperatioStatusCodes, CognitiveServiceCredentials from azure.

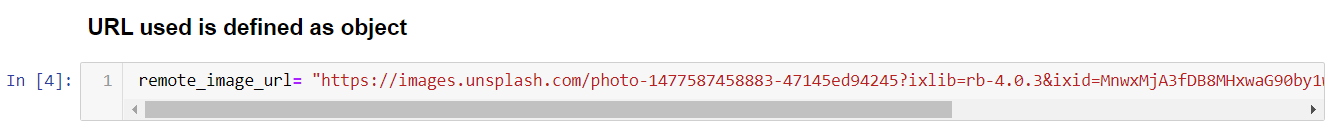
Also, Image from PIL.



We create an object subscription\_key to pass KEY1 that we created in Azure for computer vision resource.

Similarly, we pass endpoint details created in azure computer vision under object endpoint as shown above.

We connect to azure using endpoint and KEY.

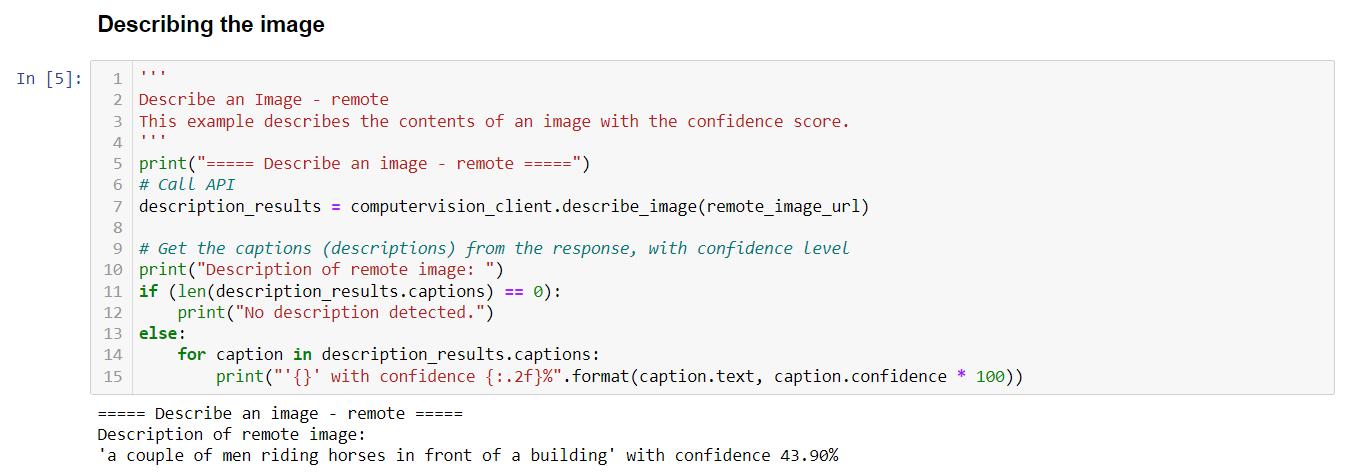


#### Image address

<https://images.unsplash.com/photo-1477587458883-47145ed94245?ixlib=rb-4.0.3&ixid=MnwxMjA3fDB8MHxzZWFyY2h8M3x8aGF3YSUyMG1haGFsJTJDJTIwamFpcHVyJTJDJTIwaW5kaWF8ZW58MHx8MHx8&w=1000&q=80>

We create an object which has a web address of an image.

Describing the image:



The above code is for describing the image. Thus, azure describes the image as a couple of men riding horses in front of a building' with confidence 43.90%’. Which means it rightly describes that men are riding horses in front of a building with a confidence of 43.90% but there is a lot more happening in the picture like tuk tuk rikshaws are also on the street. Azure cognitive services chose to describe the most relevant event.

Confidence is not very high for image description.

Categories in the image:



In the above code we are categorizing this image from the web. So we use analyse this image.

Categories and confidence:

Building: 57.42%

Outdoor: 0.78%

Outdoor street: 14.84%

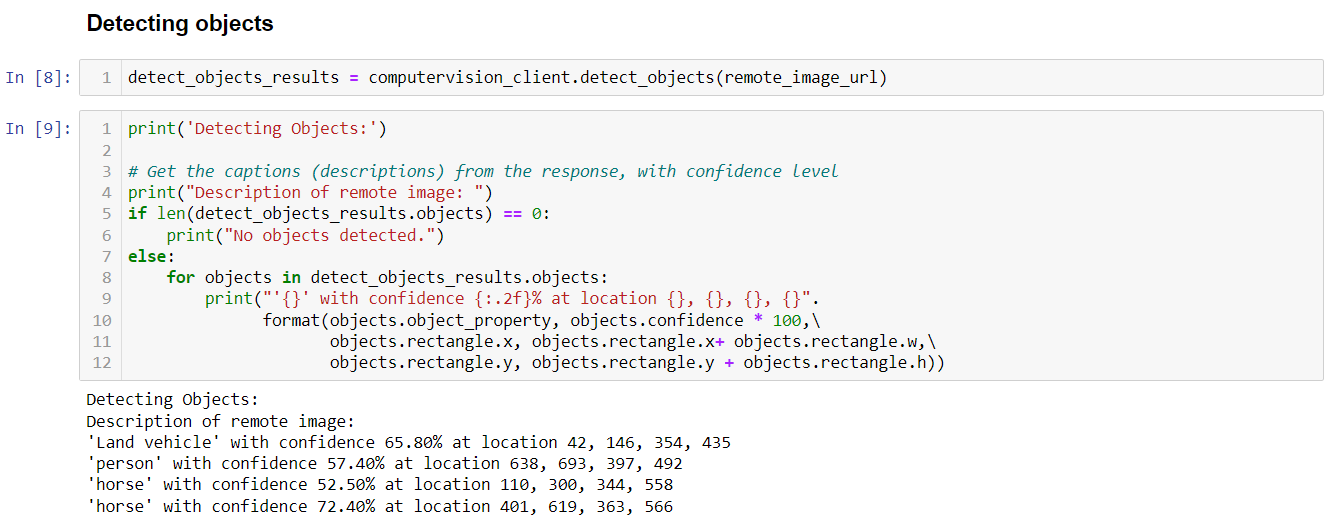
We observe that building has a good confidence level but outdoor has a very little confidence level.

Let us see how many tags can be associated with this image:

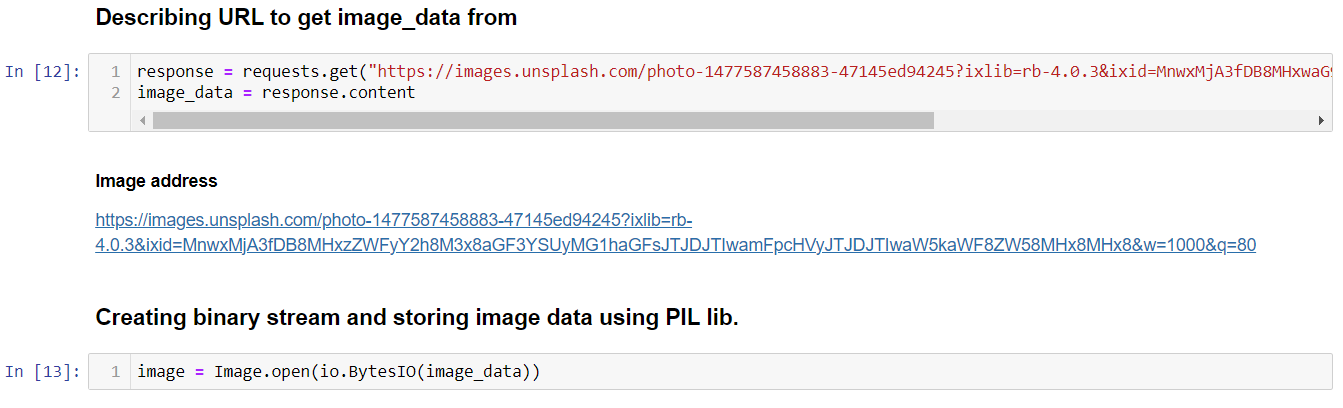


We see that outdoor has been associated by azure with the highest confidence of 99.56% followed by sky and building. Lowest confidence for tag given by azure are city and people with confidence of 61.45% and 64.57% respectively.

Detecting Objects:

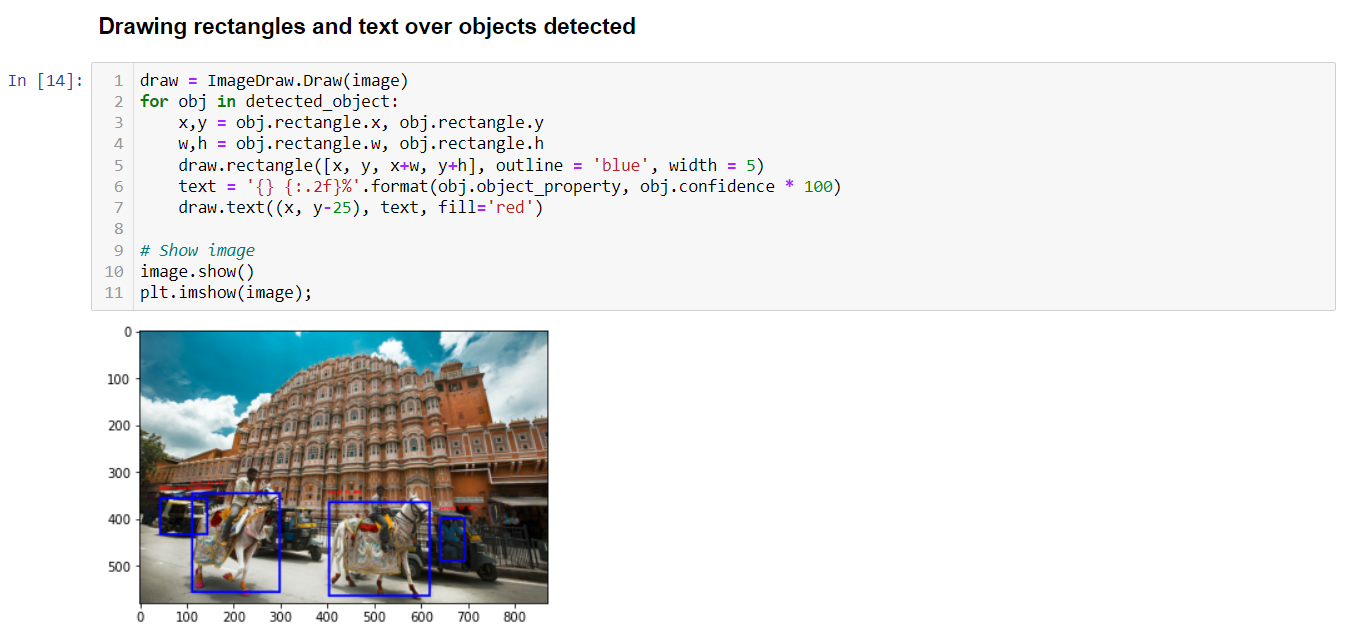


We can see above objects detected with confidence and their coordinates.



We created binary stream using PIL library and stored data as in image.

Detected Objects:



We can see that objects have been detected but they are not very clear.

Let us enlarge the image and adjust font size to have a clear view.



Thus, we can see clearly now that a total of four objects are detected by our machine. Above we can see rectangles drawn on each of the object deteected and their names as detected with the confidence level of each one of them.

|  |  |  |
| --- | --- | --- |
| Object | Confidence | Coordinates |
| Land vehicle | 65.80% | 42, 146, 354, 435 |
| Person | 57.40% | 638, 693, 397, 492 |
| Horse | 52.50% | 110, 300, 344, 558 |
| Horse | 72.40% | 401, 619, 363, 566 |

**References:**

Siva, Aditya. (2016). Hawa Mahal Road, Jaipur, India. [Photograph]. Retrieved from https://images.unsplash.com/photo-1477587458883-47145ed94245?ixlib=rb-4.0.3&ixid=MnwxMjA3fDB8MHxzZWFyY2h8M3x8aGF3YSUyMG1haGFsJTJDJTIwamFpcHVyJTJDJTIwaW5kaWF8ZW58MHx8MHx8&w=1000&q=80 (Accessed April 19, 2023).