Computer Vision Cheat Sheet

Image Classification

-> Predict the type or class of an object in an image.

Input: An image with a single object, such as a photograph.

Output: A class label (e.g. one or more integers that are mapped to class labels).

Object Localization

-> Object localization refers to identifying the location of one or more objects in an image and drawing abounding box around their extent.

Input: An image with one or more objects, such as a photograph.

Output: One or more bounding boxes (e.g. defined by a point, width, and height).

Object Detection

-> Locate the presence of objects with a bounding box and detect the classes of the located objects in these boxes

-> Object detection combines these two tasks(object localization and classification) and localizes and classifies one or more objects in an image.

Input: An image with one or more objects, such as a photograph.

Output: One or more bounding boxes (e.g. defined by a point, width, and height), and a class label for each bounding box.