Name: Nayan Biramane

Roll No: 21

Aim: To Detecting and Recognizing Objects

Objective:
Object Detection and recognition techniques HOG descriptor The Scale issues The location issue Non-maximum (or non-maxima) suppression vector machine people detection

Theory :

Object detection and recognition Techniques:

Object detection finds and identifies things in images, and it's one of the biggest accomplishments of deep learning and image processing. One of the common approaches to creating localizations for objects is with the help of bounding boxes. You can train an object detection model to identify and detect more than one specific object, so it's versatile.

Object detection models are usually trained to detect the presence of specific objects. The constructed models can be used in images, videos, or real-time operations. Even before the deep learning methodologies and modern-day image processing technologies, object detection had a high scope of interest. Certain methods (like SIFT and HOG with their feature and edge extraction techniques) had success with object detection, and there were relatively few other competitors in this field.

HOG descriptors:

The Histogram of Oriented Gradients is one of the oldest methods of object detection. It was first introduced in 1986. Despite some developments in the upcoming decade, the approach did not gain a lot of popularity until 2005 when it started being used in many tasks related to computer vision. HOG uses a feature extractor to identify objects in an image.

The feature descriptor used in HOG is a representation of a part of an image where we extract only the most necessary information while disregarding anything else. The function of the feature descriptor is to convert the overall size of the image into the form of an array or feature vector. In HOG, we use the gradient orientation procedure to localize the most critical parts of an image.

The scale issue:

The scale issue in computer vision, particularly in object detection and recognition tasks, refers to the challenge of dealing with objects appearing at different sizes in images or scenes. This variation in the size of objects can occur due to factors such as distance from the camera, viewing angle, focal length, and the resolution of the imaging device. Addressing the scale issue is crucial because objects can look drastically different depending on their scale, and detecting or recognizing them accurately across different scales is a fundamental problem in computer vision.

The Location issue:

The "location issue" in computer vision refers to the challenge of accurately determining the position or location of objects in images or scenes. Object detection and recognition algorithms not only need to identify objects but also provide information about where these objects are located within the image. Addressing the location issue is critical for a wide range of applications, including robotics, augmented reality, surveillance, and autonomous vehicles.
Here are some aspects of the location issue:

1. Object Localization:
Object detection algorithms often need to not only classify objects but also localize them by drawing bounding boxes around the objects.
2. Precision and Recall:
Achieving a balance between precision and recall is essential for reliable object localization.
3. Occlusion and Overlapping Objects:
Objects in the real world can overlap or be partially occluded by other objects. Object detection algorithms must handle these cases to accurately determine the boundaries of individual objects, even when they are not fully visible in the image.
4. Object Pose and Orientation:
In some applications, understanding the pose or orientation of an object is crucial. This is particularly relevant in robotics, where the orientation of objects needs to be known for manipulation tasks.

Non-maximum(or Non-maxima)Suppression:
Non-maximum suppression (NMS) is a technique used in object detection to filter out multiple overlapping bounding boxes predicted by an algorithm. When an object detection model detects objects in an image, it often generates multiple bounding boxes around the same object. These bounding boxes can overlap significantly, especially in cases where the object is large or close to the camera. NMS is applied to remove redundant bounding boxes, ensuring that only the most accurate and highest-confidence bounding box is kept for each detected object.

Support vector machines:

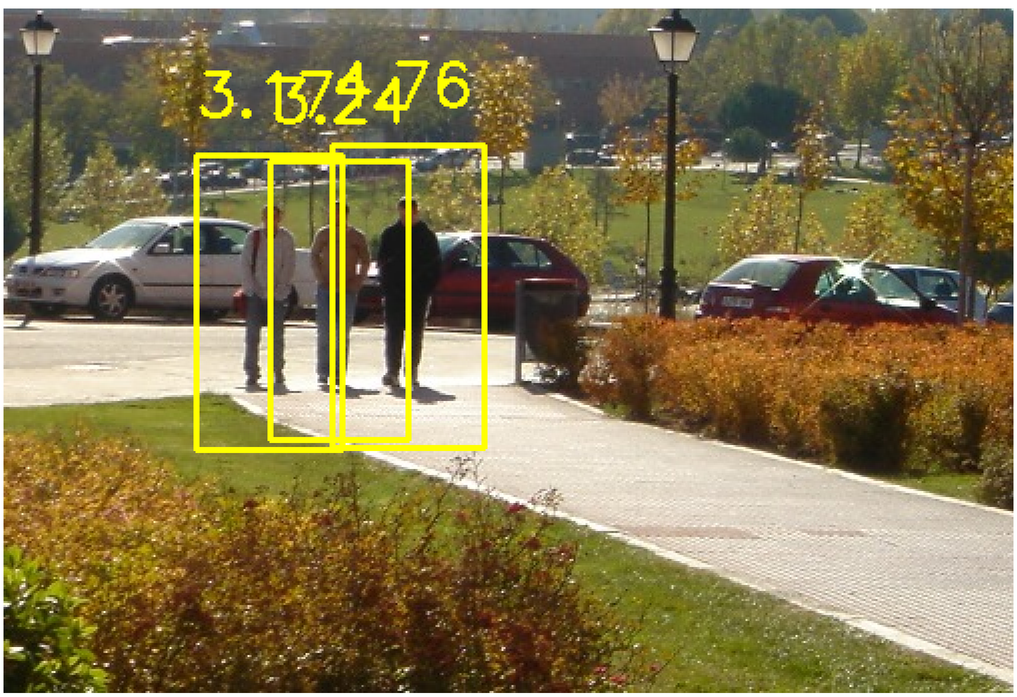
Support Vector Machines (SVM) is a supervised machine learning algorithm commonly used for classification and regression tasks. While it's not typically used for object detection on its own, SVMs can be part of the process in certain object detection pipelines, especially when dealing with binary classification tasks within the context of object detection.

Code:

```
import cv2
from google.colab.patches import cv2_imshow
image_path = '/content/humans.jpg'
img = cv2.imread(image_path)

def is_inside(i, o):
    ix, iy, iw, ih = i
    ox, oy, ow, oh = o
    return ix > ox and ix + iw < ox + ow and \
           iy > oy and iy + ih < oy + oh

hog = cv2.HOGDescriptor()
hog.setSVMDetector(cv2.HOGDescriptor_getDefaultPeopleDetector())

found_rects, found_weights = hog.detectMultiScale(
    img, winStride=(4, 4), scale=1.02, groupThreshold=1.9)

found_rects_filtered = []
found_weights_filtered = []
for ri, r in enumerate(found_rects):
    for qi, q in enumerate(found_rects):
        if ri != qi and is_inside(r, q):
            break
    else:
        found_rects_filtered.append(r)
        found_weights_filtered.append(found_weights[ri])

for ri, r in enumerate(found_rects_filtered):
    x, y, w, h = r
```

```
    cv2.rectangle(img, (x, y), (x + w, y + h), (0, 255, 255), 2)
    text = '%.2f' % found_weights_filtered[ri]
    cv2.putText(img, text, (x, y - 20), cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 255,
255), 2)

cv2_imshow(img)

cv2.waitKey(0)
cv2.destroyAllWindows()
```

Output:



Conclusion:

In conclusion, our experiment highlighted the importance of a multi-faceted approach to object detection and recognition. Utilizing techniques like HOG descriptors, addressing scale and location issues, implementing non-maximum suppression, and leveraging SVM for specific tasks, such as people detection, allowed us to create a robust and accurate object detection system. However, it's crucial to note that the field of computer vision continues to evolve rapidly, with deep learning methods currently dominating object detection tasks due to their ability to learn intricate features directly from raw data, making them the go-to choice for most real-world applications.