Name: Verdad, Jane Benneth Dione
Year & Section: BSCS IS 4B

Date: October 2024
Instructor: Sir Mark Bernardino

Object Detection and Recognition Documentation

Task 1: HOG (Histogram of Oriented Gradients) Object Detection

The Histogram of Oriented Gradients (HOG) technique is used for feature extraction. The process begins by converting the image to grayscale, after which the HOG descriptor is applied to compute gradient orientations within local regions. Key parameters, including the number of orientations, pixels per cell, and cells per block, are stated to configure the HOG computation. Based on observations, the HOG visual representation emphasizes edges and gradients, which are essential for detecting shape-related information within an image. This is useful in object detection tasks, especially for identifying objects with distinct edges and structures. The HOG visualization is rescaled to enhance contrast, allowing for clearer interpretation. The final output displays two images: the original and the HOG image.

Task 2: YOLO (You Only Look Once) Object Detection

The YOLO (You Only Look Once) v8 model, a deep-learning-based object detection method, is used to load a pre-trained YOLOv8 model for this task. The process begins by loading an image. Then applying the YOLO model to detect objects, where each detected object is marked with a bounding box, labeled with the class ID, and annotated with its confidence level. A confidence threshold of 0.5 is set to filter out low-confidence detections, thereby enhancing detection accuracy. Based on observations, YOLO enables real-time detection by processing the entire image at once, eliminating the need for region proposals required by traditional models. The confidence threshold refines the output by excluding potential false positives. The resulting image displays a bounding box around the detected object, with each box labeled by class ID and confidence level, showcasing YOLO's balance of speed and accuracy, that makes it ideal for real-time object detection scenarios.

Task 3: SSD (Single Shot MultiBox Detector) with TensorFlow

The SSD MobileNet V2 model is loaded using TensorFlow's save model feature. This model performs single-shot object detection by processing images through a convolutional neural network. For the preparation, the input image is resized and preprocessed to match the model's required input size of 300x300 pixels. Only detections with a confidence threshold above 0.5 are visualized, with bounding boxes and labels displaying class information and confidence scores. Based on my observations, SSD is effective for fast object detection with reasonable accuracy. The use of a lightweight backbone, MobileNet V2, keeps the model efficient and well-suited for mobile applications. The requirement for specific image formats and sizes emphasizes the need for preprocessing in some deep-learning-based object detection models. The output image displays a bounding box and class label for the detected object, demonstrating the SSD MobileNet model's balance of accuracy and speed, that offers a reliable detection while maintaining efficiency.

Task 4: Traditional vs. Deep Learning Object Detection Comparison

In this part, we compare traditional (HOG-SVM) and deep learning-based (YOLO) object detection methods using a single image. For the HOG-SVM method, a pre-trained SVM with HOG descriptors is applied to detect objects, often used for human detection. For YOLO, the YOLOv8 model is utilized for object detection, just like in the previous task. Processing times are recorded for both methods to assess their speed, and bounding boxes are drawn on the images for visualization. Based on the observations, YOLO processes the image faster than HOG-SVM, and in terms of detection quality, YOLO provides more robust detections with confidence scores, while HOG-SVM is more limited to simpler, edge-based detections. This comparison underscores YOLO's strengths in both speed and detection quality, making it the preferred choice for real-time applications. The results show two images displaying detections from each method, with YOLO exceeding HOG-SVM in both speed and accuracy.