**INTRO TO ARTIFICIAL INTELLIGENCE**

**Phase 2: Term project mid-status report**

Shivapriya Pillalamarri Harsha Vardhan Mahankali

### **RESEARCH QUESTION**

### Object detection is a fundamental problem in computer vision that requires the identification and localization of objects within an image or video. The task is twofold: classification and localization. Classification involves determining the category or class of each detected object, such as identifying whether an object is a car, dog, or person. Localization focuses on identifying the spatial position of the object within the image, typically represented by bounding box coordinates that enclose the detected object.

1. **INTRODUCTION & IMPORTANCE OF PROBLEM SOLVING**

In the medical imaging field, object detection is essential for accurately detecting cancers, abnormalities, and other vital situations in X-rays, MRIs, and CT scans. Similar to this, object detection is essential to autonomous cars' safety and dependability since it enables real-time recognition of barriers, other cars, pedestrians, and road signs, thereby facilitating navigation and decision-making. The technology is equally vital in surveillance and security systems, where it improves public safety by facilitating automated monitoring, facial identification in congested areas, and intruder detection. Innovative solutions like virtual try-ons, inventory management, and visual search are made possible by object detection in e-commerce and retail, which enhances user experience and operational effectiveness. Furthermore, object detection is used in a variety of industries, including manufacturing, agriculture, and environmental monitoring, to track wildlife, identify crops, and monitor machinery, respectively. Across many industries, real-time item detection may boost production, save lives, and avoid accidents. Additionally, it advances artificial intelligence, opening the door for increasingly sophisticated and self-sufficient devices that can help people with daily activities.

1. **RELATED LITERATURE**

In [1], focuses on enhancing object detection through the Single Shot Detector (SSD) algorithm. The research introduces improvements to the standard SSD model by incorporating depth-wise separable convolutions and better spatial resolution techniques to optimize accuracy and processing speed. In [2], discusses the Histogram of Oriented Gradients (HOG) method for human detection. HOG is a feature descriptor used in computer vision and image processing, emphasizing the distribution of intensity gradients in localized image regions. The study concludes that HOG is highly effective for detecting objects like pedestrians, offering robustness against variations in lighting and pose. In [3], examines advancements in object detection techniques using deep learning, focusing on improvements in algorithm efficiency and accuracy. It highlights innovations in combining neural network architectures, such as CNNs, with real-time detection frameworks like YOLO and SSD. The study provides insights into overcoming challenges related to detecting small objects and occlusions while ensuring scalability for real-world applications.

Our proposed solution is different from the above discussed papers is that we have included Histogram of Oriented Gradients(HOG),Singe Shot Detector(SSD),Convolutional Neural Networks (CNN),YOLO(You Only Look Once). By combining these algorithms, the strengths of each are leveraged to create a robust and versatile object detection solution:

* **HOG** provides baseline features for initial analysis and feature engineering.
* **SSD** offers real-time detection with good precision for medium-complexity scenarios.
* **CNN** ensures high-quality feature extraction and scalability for advanced tasks.
* **YOLO** adds efficiency and simplicity, excelling in real-time applications with minimal latency.

This integrated approach ensures a better and more accurate solution for object detection, addressing challenges like varying object scales, occlusion, and the need for real-time performance in diverse application domains.

1. Q. Shuai and X. Wu, "Object detection system based on SSD algorithm," 2020 International Conference on Culture-oriented Science & Technology (ICCST), Beijing, China, 2020, pp. 141-144, doi: 10.1109/ICCST50977.2020.00033. <https://ieeexplore.ieee.org/document/9262816>
2. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177. https://ieeexplore.ieee.org/document/1467360
3. B. Balakrishnan, R. Chelliah, M. Venkatesan and C. Sah, "Comparative Study On Various Architectures Of Yolo Models Used In Object Recognition," 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2022, pp. 685-690, doi: 10.1109/ICCCIS56430.2022.10037635. <https://ieeexplore.ieee.org/document/10037635>
4. **ACHIEVEMENTS, CHALLENGES & SOLUTIONS**
5. Importing libararies- Importing libraries From Tensorflow.keras is successfully done.
6. Loading the data sets have been done. For ex, Minist dataset.
7. Creating the model with all algorithms and loading the datasets into the model.

**Occlusion and Cluttered Backgrounds**

**Challenge:**  
Objects partially obscured by other objects or blending into complex backgrounds are harder to detect accurately. This can lead to missed detections or false positives.

**Workarounds:**

* Training the model with occlusion-augmented datasets to improve robustness.
* Using region-based detection frameworks like Faster R-CNN, which perform better at handling occluded objects.

### ****Overlapping and Similar Objects****

**Challenge:**  
Objects with similar appearances or overlapping regions can confuse models, leading to misclassifications or multiple detections of the same object.

**Workarounds:**

* Applying non-maximum suppression (NMS) to eliminate redundant detections.
* Using models with refined bounding box regression and classification, such as Faster R-CNN or YOLOv5.
* Employing instance segmentation models like Mask R-CNN to better separate overlapping objects.

1. **GITHUB**

<https://github.com/Shivapriyapillalamarri/OBJECT-DETECTION-USING-DEEP-LEARNING->