Open world Object Detection

Vadlamudi Karthikeya- IMT2021504

Asrith- IMT2021509

1. Problem Statement and Introduction

Open world object detection

Existing object detection methods encounter significant challenges when they are confronted with the task of detecting unknown objects, particularly within dynamic environments. These challenges stem from the inherent limitations imposed by closed-world assumptions, which restrict the detection capabilities of current systems. In such closed-world scenarios, the detection models are trained and tested on a predefined set of object classes, leading to a lack of adaptability and robustness when faced with novel or unforeseen objects. Consequently, in dynamic environments where the composition of objects may evolve or include previously unseen entities, the effectiveness of traditional object detection approaches diminishes. Addressing these limitations is crucial for enhancing the reliability and flexibility of object detection systems, thereby enabling their deployment in real-world scenarios where the presence of unknown objects is commonplace.

2. Problem Motivation

In real-world scenarios, such as autonomous driving, surveillance systems, and industrial automation, there is a growing demand for object detection systems that can effectively adapt to encountering new or previously unseen objects. However, existing Object Detection with Objectness (OWOD) methods often fall short in this regard. These methods typically categorise all unknown objects into a single class, lacking the capability for fine-grained recognition. This limitation poses significant challenges, particularly in dynamic environments where the composition of objects may change rapidly or include novel entities.

The motivation for our project stems from the pressing need for object detection systems that can robustly detect and accurately classify unknown objects in open-world scenarios. By 'open-world,' we refer to environments where the types and characteristics of objects are diverse, constantly evolving, and may include objects not encountered during the training phase. Conventional OWOD approaches struggle to cope with this variability, resulting in suboptimal performance and limited practical applicability.

Therefore, our project aims to address this critical gap by developing novel techniques and algorithms that enable object detection systems to adapt seamlessly to novel objects in open-world settings. By enhancing the system's ability for fine-grained recognition and classification of unknown objects, we aspire to empower these systems with greater versatility, reliability, and real-world utility across a wide range of applications.

3. Literature Survey

To extend we have also read these following papers on open world object:

1) UC-OWOD: Unknown-Classified Open World Object Detection.
   2207.11455 (arxiv.org)

2)LoCalization and IdentificAtion Cascade Detection Transformer for Open-World Object Detection.
Ma_CAT_LoCalization_and_IdentificAtion_Cascade_Detection_Transformer_for_Open-World_Object_CVPR_2023_paper.pdf (thecvf.com)

3)DetectingEverythingintheOpenWorld:TowardsUniversalObjectDetection.
2303.11749 (arxiv.org)

4)Towards Open World Object Detection.
2103.02603 (arxiv.org)

4. Summary of the Papers

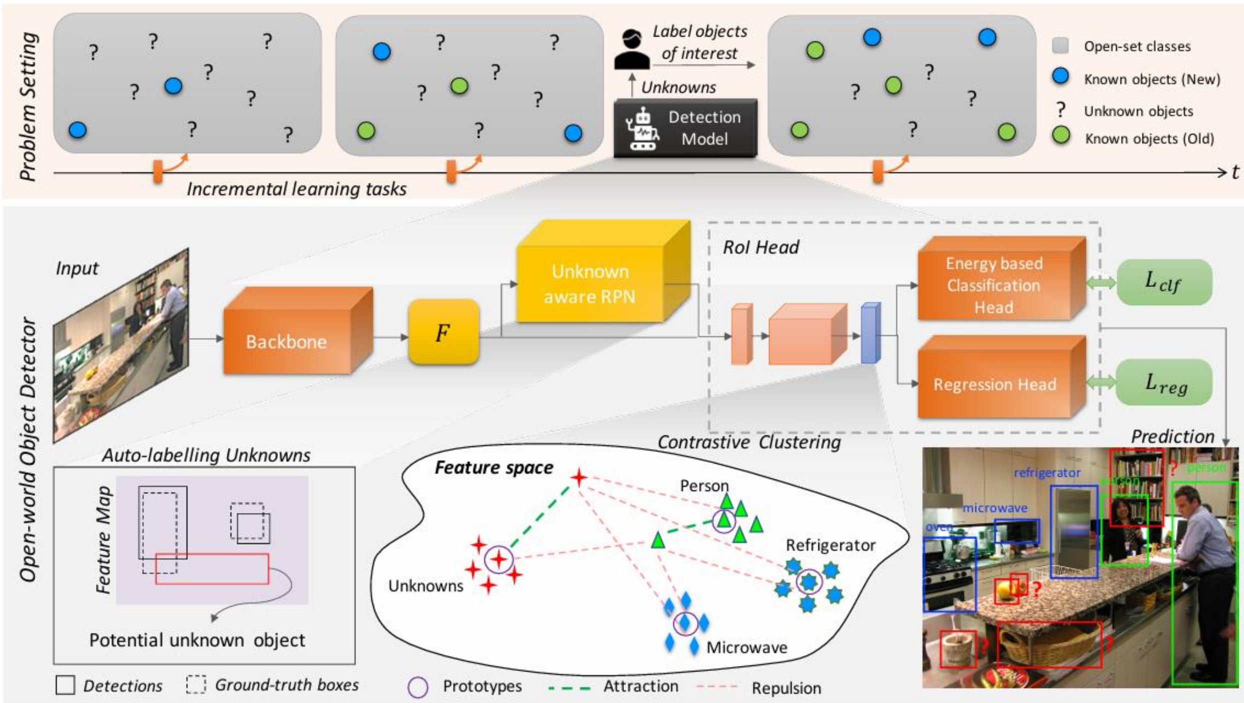
Uni Detector:
It is a novel object detection model designed to address the challenge of detecting unknown objects in dynamic environments. Unlike traditional methods, it employs a unified approach that allows for fine-grained recognition of novel objects, enhancing adaptability and robustness in open-world scenarios. By incorporating advanced techniques for object detection and classification, UniDetector aims to significantly improve the performance and practical applicability of object detection systems across various real-world applications.

UC-OWOD: Unknown-Classified Open World Object Detection:
It employs methods that allow for the identification and categorization of previously unseen objects, thus enhancing the adaptability of the detection system. This involves training the model to differentiate between known and unknown objects, enabling it to accurately classify unfamiliar entities encountered in dynamic environments. Additionally, UC-OWOD continuously

learns and updates its knowledge base to improve its performance over time, ensuring robustness and reliability in open-world scenarios.

Towards Open World Object Detection;

A new methodology named ORE is introduced. It aims to identify unknown instances without explicit supervision and prevent forgetting of previous instances when new labels are introduced for model improvement without retraining from scratch. ORE utilises contrastive clustering in the latent space of object detectors to achieve clear discrimination between classes and facilitate incremental learning without overlapping representations. Auto-labelling based on the Region Proposal Network is proposed to address the challenge of annotating unknown instances, enabling the energy-based classification head to differentiate between known and unknown instances. The architectural overview of ORE utilises Faster R-CNN as the base detector, incorporating a class-agnostic Region Proposal Network and a classification head adapted for auto-labelling and unknown identification.The main idea of the model that it is trained to detect familiar objects and flag unknown ones. These unknown instances are reviewed by humans, helping the model learn and improve without needing to start training from scratch.



5. Code analysis

The paper which we have implemented is the 3rd paper based on ORE. So, we start off by using the Detectron2 model trained on COCO 2017 dataset to execute object detection within the given image. By using its pre-trained capabilities, the model quickly identifies and delineates various objects within the scene, spanning diverse categories such as vehicles, animals, and household items. However, acknowledging the inherent limitations of any model. So, after it finds everything it can, we take a closer look to see if there's anything it missed or if it got something wrong.

During this phase, each detected object undergoes examination, evaluating its classification and confidence scores. Objects deemed unidentified or misclassified by the model are marked as "unknown," setting the option for further refinement and enhancement of the model's capabilities. This adaptive approach not only ensures a comprehensive understanding of the image contents but also facilitates continual learning and improvement of the detection model.

Following the initial object detection phase, the process transitions to an evaluation stage using a specialised dataset focused on balloons. This curated dataset serves as a benchmark to assess the model's proficiency in discerning and categorising objects within a specific domain. By subjecting the model to this focused evaluation, we test its ability to identify balloons amidst varying backgrounds and environmental conditions.

Upon completion of the evaluation phase, the process proceeds to annotate all previously categorised "unknown" objects as balloons. This step underscores the model's adaptability and capacity to assimilate new knowledge, thereby refining its understanding of object categories. By extending its capabilities to encompass previously unrecognised objects, the model transcends its initial boundaries, embodying a more holistic understanding of the visual world.

The iterative process culminates in a comprehensive rendering of the image, where the balloons, once obscured amidst unidentified objects, now stand prominently alongside their initially recognized counterparts. This integration of identified objects and newly annotated balloons helps towards increasing precision and accuracy in object recognition and classification.

6. Code output for a test case

Input image which has been taken

Object identification done using instance segmentation

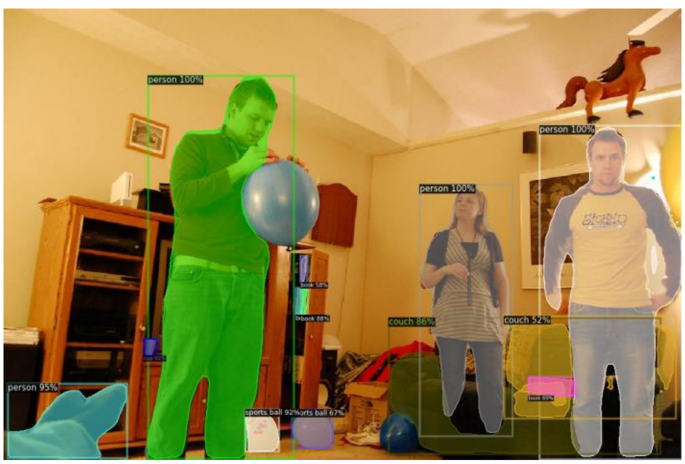


Image which unknown objects identified

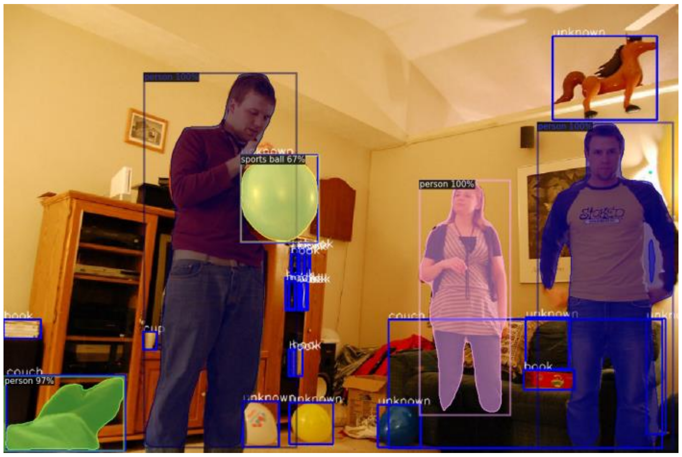
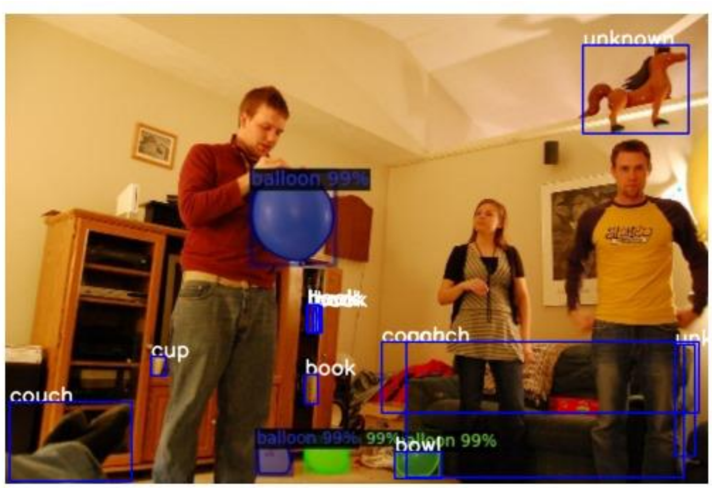


image detecting only balloons after training through its data set



output image which detects balloons now as it was trained on ballon data set

7. Experiments conducted and Results

So ran various Detection 2's inbuilt masking models Namely instance segmentation(uses RCNN) and panoptic segmentation(Fully Convolutional Panoptic Segmentation).these are the mAP(Mean average precision)% which we got.

Model	mAP (%)
mask_rcnn_R_101_C4_3x	38
mask_rcnn_R_50_C4_1x	32
mask_rcnn_R_50_FPN_1x	36

Model	mAP (%)
panoptic_fpn_R_101_3x	43
panoptic_fpn_R_50_1x	37

8.Conclusion

In this project, we explored the task of object detection and instance segmentation using various learning models implemented in Detectron2. Our investigation focused on evaluating the performance of various model architectures, including Mask R-CNN and Panoptic FPN, on custom datasets.

Through this we observed that different model architectures exhibit varying levels of accuracy and efficiency in detecting and segmenting objects. For instance, the mask_rcnn_R_101_C4_3x model demonstrated superior performance with a mean Average Precision (mAP) of 38%, highlighting its effectiveness in accurately identifying objects in complex scenes.

Furthermore, we investigated the impact of model architecture and training parameters on performance metrics such as mAP. Our results indicate that the choice of backbone network, training duration, and dataset characteristics significantly influence the overall performance of object detection and instance segmentation models.

Overall, our findings underscore the importance of careful model selection, parameter tuning, and dataset preparation in achieving optimal performance in object detection tasks.