1. **What type of object detectors are there?**

There are mainly two types of state-of-the-art object detectors. On one hand, we have two-stage detectors, such as Faster R-CNN (Region-based Convolutional Neural Networks) or Mask R-CNN, that (i) use a Region Proposal Network to generate regions of interests in the first stage and (ii) send the region proposals down the pipeline for object classification and bounding-box regression. Such models reach the highest accuracy rates, but are typically slower. On the other hand, we have single-stage detectors, such as YOLO (You Only Look Once) and SSD (Singe Shot MultiBox Detector), that treat object detection as a simple regression problem by taking an input image and learning the class probabilities and bounding box coordinates. Such models reach lower accuracy rates, but are much faster than two-stage object detectors.

1. **What are some SOTA methods available?**

For object detection some of the SOTA methods available are :

2.1 Integral Channel Features (ICF)

In integral channel features , we try to combine multiple registered image channels computed by linear and nonlinear transformations. Those eight transformations, including color, gradient, edge, gradient histogram, difference of Gaussian, thresholding, and the absolute value of Gaussian, are all translationally invariant, meaning that they only need to be computed once. Training those features by cascading AdaBoost classifier, the whole method is fast and effective on pedestrain detection.

2.2 Discriminatively Trained Part Based Models (DPM)

The hardest part for object detection is that there are lots of variances. Those are arose from illumination, viewpoint, non-rigid deformation, occlusion, and intraclass variability. The deformable parts model is trying to capture those variances. It assumes an object is constructed by its parts. Thus, the detector will first found a match by coarser (at half resolution) root filter, and then using its part models to fine-tune the result. It uses HOG features on pyramid levels before filtering, and linear SVM for training to find the different part locations of an object.

2.3 Rich Feature Hierarchies for Convolutional Neural Networks (RCNN)

Recently, many detection algorithms using selective search as region proposals, to avoid exhaustive sliding window method. In addition, Convolutional Neural Networks (CNN) is successful in ImageNet classification challenge. Since the progress of DPM model is now restrictive, the RCNN paper is trying to do detection in another way. CNN can learn a diverse set of features from experience. However, it is computational complex, has many parameters, and hard to tell its meaning of the whole mechanism. In RCNN paper, they use region proposal to reduce two order in the number of image windows. All proposals are then warped to 224 by 224-pixel size as inputs to CNN. For CNN, the experiments also show that the number of parameters can be reduced to 6%, by using pool 5 instead of the last layer with only few percentage drops on accuracy, though the reason is still a mystery.

1. **What are the datasets that are publicly available?**

There are various datasets publicly available. We can find most of them on Kaggle.

Links are mentioned below:

<https://www.kaggle.com/andrewmvd/helmet-detection>

<https://www.kaggle.com/abhishek4273/helmet-dataset>

<https://www.kaggle.com/andrewmvd/hard-hat-detection>

1. **Can we create our own dataset?**

Yes. We can create our own dataset. We can collect images from these datasets and what we can do is either we can draw the bounding boxes and label them using many labelling software like LabelImg.

1. **What are some software or techniques that we can use to label the dataset?**

There are various software techniques for Labelling the dataset. Some of them I will list below.

5.1 LABELIMG

5.2 SUPERVISE.LY

5.3 LABELBOX

1. **Has someone done similar work in this area?**

Yes, various Machine Learning and Computer Vison Experts have done contributions to this subject.

1. **What are some open-source codebases that people reading the document can refer to?**

People can refer to Software Documentation present online. Tensorflow’s online documentation is available. Various Object Detection Model’s github repositories are present.

1. **Provide more links to blogs, articles, videos, etc. to learn more about the topic.**

<https://medium.com/tektorch-ai/best-image-labeling-tools-for-computer-vision-393e256be0a0>

<https://ieeexplore.ieee.org/abstract/document/6726989/>

<https://patents.google.com/patent/US6965312B2/en>

<https://www.kaggle.com/andrewmvd/helmet-detection>

<https://www.kaggle.com/abhishek4273/helmet-dataset>

<https://www.kaggle.com/andrewmvd/hard-hat-detection>