

COMMON OBJECTS IN CONTEXT(COCO) PAPER SUMMARY

Abstract:-

COCO is the name of a sizable dataset for object detection segmentation, and captioning. Microsoft's Common Objects in Context dataset is now the most popular object detection dataset (COCO). It is widely used to assess how well computer vision algorithms work

Features Of COCO model:-

- Object segmentation
- Recognition in context
- Super pixel stuff segmentation
- 330K images (>200K labelled)
- 1.5 million object instances
- 80 object categories
- 91 stuff categories
- 5 captions per image
- 250,000 people with key points

COCO DATASET:-

The Microsoft Common Objects in Context (MS COCO) collection consists of 91 common object categories, with 82 of them having more than 5,000 annotated instances. There are 2,500,000 instances with labels and 328,000 images in the collection.

Compared to the well-known ImageNet dataset, COCO has fewer categories but more instances per category. It can be advantageous to learn complex object models with precise 2D localization. The dataset has a substantially greater instance count per category than the PASCAL VOC and SUN databases do.

COCO DATASET PREPARATION:-

The following 80 pre-trained objects are part of the object detection and tracking classes in the COCO dataset. Three values are written on each of the 17 pre-trained important spots in the COCO. The coordinates are represented by the x and y values, while the visibility of the key point is indicated by the v value. There are annotations on images of typical objects and situations in their natural settings.

These products are categorised, for instance, as chairs or bananas. In computer vision, labelling, also referred to as image annotation, is a common tactic. Other object recognition datasets focus on classifying images, segmenting objects at the semantic pixel level, and segmenting specific instances of objects. Localizing objects within their bounding boxes is also a focus of these datasets. There are numerous accessible perspectives for different item categories.

CHALLENGES FACED WHILE MS COCO DATASET PREPARATION:-

bias in data There is some bias in the data because more people with light complexion are represented in the sample. In the pictures, there are twice as many men as women, more people with light skin than those with dark skin, and even fewer persons with dark skin than women. A

number of the visual descriptions have racial overtones. This might lead to the socially unacceptable use of computer vision to characterise people. The researchers cite numerous studies showing fair-skinned individuals outperform dark-skinned individuals in tasks like pedestrian identification and facial recognition, and that picture caption systems developed using the COCO dataset work better for fair-skinned individuals. Prejudice is present in the photograph's setting. Lighter people appear indoors, whereas darker ones appear indoors with furniture in the background

RECENT DEVELOPMENTS IN THE COCO DATASET AND ITS MODEL:-

In 2014, the MS COCO dataset's initial version was made public. There are 164K photos total, divided into 83K training images, 41K validation images, and 41K test images. A second test set of 81K photos was released in 2015, including the 40K fresh photographs and all of the test images from the past.