

✓ Object Detection with DETr

Created by Facebook Research, DETr is an object detection model that replaces the object detection pipeline with a Transformer. Learn more about [DETr here](#) at the official repository..

[+ Code](#) [+ Text](#)

```
import torch
import torchvision
from torchvision import transforms as T
from torchvision.utils import draw_bounding_boxes # bounding box drawing function.
from PIL import Image
import matplotlib.pyplot as plt

CLASSES = [
    'N/A', 'person', 'bicycle', 'car', 'motorcycle', 'airplane', 'bus',
    'train', 'truck', 'boat', 'traffic light', 'fire hydrant', 'N/A',
    'stop sign', 'parking meter', 'bench', 'bird', 'cat', 'dog', 'horse',
    'sheep', 'cow', 'elephant', 'bear', 'zebra', 'giraffe', 'N/A', 'backpack',
    'umbrella', 'N/A', 'N/A', 'handbag', 'tie', 'suitcase', 'frisbee', 'skis',
    'snowboard', 'sports ball', 'kite', 'baseball bat', 'baseball glove',
    'skateboard', 'surfboard', 'tennis racket', 'bottle', 'N/A', 'wine glass',
    'cup', 'fork', 'knife', 'spoon', 'bowl', 'banana', 'apple', 'sandwich',
    'orange', 'broccoli', 'carrot', 'hot dog', 'pizza', 'donut', 'cake',
    'chair', 'couch', 'potted plant', 'bed', 'N/A', 'dining table', 'N/A',
    'N/A', 'toilet', 'N/A', 'tv', 'laptop', 'mouse', 'remote', 'keyboard',
    'cell phone', 'microwave', 'oven', 'toaster', 'sink', 'refrigerator', 'N/A',
    'book', 'clock', 'vase', 'scissors', 'teddy bear', 'hair drier',
    'toothbrush'
] # COCO Classes
# These variables are for loading from pytorch hub
MODEL_REPO = "facebookresearch/detr:main" # The model repo
MODEL_NAME = "detr_resnet50" # The model
```

```
# Load an image of around 10 cars.
!wget -O car.jpg https://hips.hearstapps.com/hmg-prod.s3.amazonaws.com/images/10best-cars-group-cropped-1542126037.jpg
```

```
→ --2022-08-11 14:30:54-- https://hips.hearstapps.com/hmg-prod.s3.amazonaws.com/images/10best-cars-group-cropped-1542126037.jpg
Resolving hips.hearstapps.com (hips.hearstapps.com)... 151.101.0.155, 151.101.64.155, 151.101.128.155, ...
Connecting to hips.hearstapps.com (hips.hearstapps.com)|151.101.0.155|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 2178423 (2.1M) [image/jpeg]
Saving to: 'car.jpg'

car.jpg          100%[=====] 2.08M  --.-KB/s   in 0.04s

2022-08-11 14:30:54 (50.6 MB/s) - 'car.jpg' saved [2178423/2178423]
```

```
im = Image.open("car.jpg")
im
```



```
def box_cxcywh_to_xyxy(x):
    """
    This is to convert the bounding boxes from DETr format to xy coordinate pairs
    """
    x_c, y_c, w, h = x.unbind(1)
    b = [(x_c - 0.5 * w), (y_c - 0.5 * h),
          (x_c + 0.5 * w), (y_c + 0.5 * h)]
    return torch.stack(b, dim=1)

def rescale_boxes(out_bbox, size):
    """
    Rescale the bounding boxes to the original image height and width so they can be overlaid.
    """
    img_w, img_h = size
    b = box_cxcywh_to_xyxy(out_bbox)
    b = b * torch.tensor([img_w, img_h, img_w, img_h], dtype=torch.float32)
    return b

transforms = T.Compose([
    T.Resize(800), # Resizes too 800px while maintaining the aspect ratio
    T.ToTensor(),
    T.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225]) # Imagenet normalization
])

img = transforms(im).unsqueeze(0)

model = torch.hub.load(MODEL_REPO, MODEL_NAME, pretrained=True)
```

```
↳ Downloading: "https://github.com/facebookresearch/detr/zipball/main" to /root/.cache/torch/hub/main.zip
/usr/local/lib/python3.7/dist-packages/torchvision/models/_utils.py:209: UserWarning: The parameter 'pretrained' is deprecated since
  f"The parameter '{pretrained_param}' is deprecated since 0.13 and will be removed in 0.15, "
/usr/local/lib/python3.7/dist-packages/torchvision/models/_utils.py:223: UserWarning: Arguments other than a weight enum or `None` + 
  warnings.warn(msg)
Downloading: "https://download.pytorch.org/models/resnet50-0676ba61.pth" to /root/.cache/torch/hub/checkpoints/resnet50-0676ba61.pt
100% 97.8M/97.8M [00:02<00:00, 33.4MB/s]
```

```
model = model.eval()
100% |██████████| 100.02~00.00, 43.2MB/s
```

```
out = model(img)
```

```
↳ /root/.cache/torch/hub/facebookresearch_detr_main/models/position_encoding.py:41: UserWarning: __floordiv__ is deprecated, and its t
  dim_t = self.temperature ** (2 * (dim_t // 2) / self.num_pos_feats)
```

```
probs = out["pred_logits"].softmax(-1)[0, :, :-1] # processing output to return bounding boxes and labels
keep = probs.max(-1).values > 0.7 # only predictions with 0.7 probability or higher
bboxes_scaled = rescale_boxes(out['pred_boxes'][0, keep], im.size)
```

```
labels = [CLASSES[i] for i in probs[keep].argmax(1)] # This is a good sign, the model detected many cars
labels
```

```
↳ ['car', 'car', 'car', 'car', 'car', 'car', 'car', 'car', 'car', 'car']
```

```
tensor_img = torch.tensor((T.ToTensor()(im) * 255), dtype=torch.uint8) # draw_bounding_boxes requires images in this format.
```

```
↳ /usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: UserWarning: To copy construct from a tensor, it is recommended to u
  """Entry point for launching an IPython kernel.
```

```
bb = draw_bounding_boxes(tensor_img, boxes=bboxes_scaled, width=6) # this function will automatically overlay bounding boxes over the ir
# imported at beginning of file
```

```
# draw figure
fig = plt.figure(figsize=(14, 8))
plt.imshow(bb.permute(1, 2, 0)) # convert it into a image that matplotlib can display
```

```
↳ <matplotlib.image.AxesImage at 0x7fd98fb827d0>
```

