Florence-2-large sample usage

```
import sys
print(sys.executable)
```

```
/home/students/anaconda3/envs/florencenv/bin/python
```

```
!pip install ollama
```

```
Collecting ollama
  Using cached ollama-0.4.4-py3-none-any.whl.metadata (4.7 kB)
Requirement already satisfied: httpx<0.28.0,>=0.27.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from ollama) (0.27.0)
Collecting pydantic<3.0.0,>=2.9.0 (from ollama)
  Downloading pydantic-2.10.4-py3-none-any.whl.metadata (29 kB)
Requirement already satisfied: anyio in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from httpx<0.28.0,>=0.27.0->ollama) (4.6.2)
Requirement already satisfied: certifi in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from httpx<0.28.0,>=0.27.0->ollama) (2024.12.14)
Requirement already satisfied: httpcore==1.* in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from httpx<0.28.0,>=0.27.0->ollama) (1.0.2)
Requirement already satisfied: idna in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from httpx<0.28.0,>=0.27.0->ollama) (3.7)
Requirement already satisfied: sniffio in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from httpx<0.28.0,>=0.27.0->ollama) (1.3.0)
Requirement already satisfied: h11<0.15,>=0.13 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from httpcore==1.*->httpx<0.28.0,>=0.27.0->ollama) (0.14.0)
Collecting annotated-types>=0.6.0 (from pydantic<3.0.0,>=2.9.0-
>ollama)
  Using cached annotated_types-0.7.0-py3-none-any.whl.metadata (15 kB)
Collecting pydantic-core==2.27.2 (from pydantic<3.0.0,>=2.9.0->ollama)
  Downloading pydantic_core-2.27.2-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (6.6 kB)
Requirement already satisfied: typing-extensions>=4.12.2 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from pydantic<3.0.0,>=2.9.0->ollama) (4.12.2)
Requirement already satisfied: exceptiongroup>=1.0.2 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from anyio->httpx<0.28.0,>=0.27.0->ollama) (1.2.0)
Using cached ollama-0.4.4-py3-none-any.whl (13 kB)
Downloading pydantic-2.10.4-py3-none-any.whl (431 kB)
Downloading pydantic_core-2.27.2-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (2.0 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 2.0/2.0 MB 446.0 kB/s eta
0:00:001m451.7 kB/s eta 0:00:01
a
Successfully installed annotated-types-0.7.0 ollama-0.4.4 pydantic-
2.10.4 pydantic-core-2.27.2
```

```
!pip list
```

```
Package                   Version
------------------------- -----------
annotated-types           0.7.0
anyio                     4.6.2
argon2-cffi               21.3.0
argon2-cffi-bindings      21.2.0
asttokens                 2.0.5
async-lru                 2.0.4
attrs                     24.3.0
Babel                     2.11.0
backcall                  0.2.0
beautifulsoup4            4.12.3
bleach                    6.2.0
Brotli                    1.0.9
certifi                   2024.12.14
cffi                      1.17.1
charset-normalizer        3.3.2
comm                      0.2.1
debugpy                   1.6.7
decorator                 5.1.1
defusedxml                0.7.1
exceptiongroup            1.2.0
executing                 0.8.3
fastjsonschema            2.20.0
h11                       0.14.0
httpcore                  1.0.2
httpx                     0.27.0
idna                      3.7
importlib_metadata        8.5.0
ipykernel                 6.29.5
ipython                   8.15.0
ipywidgets                8.1.5
jedi                      0.19.2
Jinja2                    3.1.4
json5                     0.9.25
jsonschema                4.23.0
jsonschema-specifications 2023.7.1
jupyter                   1.0.0
jupyter_client            8.6.0
jupyter-console           6.6.3
```

```
jupyter_core              5.7.2
jupyter-events            0.10.0
jupyter-lsp               2.2.0
jupyter_server            2.14.1
jupyter_server_terminals  0.4.4
jupyterlab                4.2.5
jupyterlab-pygments       0.1.2
jupyterlab_server         2.27.3
jupyterlab_widgets        3.0.13
MarkupSafe                2.1.3
matplotlib-inline         0.1.6
mistune                   2.0.4
nbclient                  0.8.0
nbconvert                 7.16.4
nbformat                  5.10.4
nest-asyncio              1.6.0
notebook                  7.2.2
notebook_shim             0.2.3
ollama                    0.4.4
overrides                 7.4.0
packaging                 24.2
pandocfilters             1.5.0
parso                     0.8.4
pexpect                   4.8.0
pickleshare               0.7.5
pip                       24.2
platformdirs              3.10.0
ply                       3.11
prometheus_client         0.21.0
prompt-toolkit            3.0.43
psutil                    5.9.0
ptyprocess                0.7.0
pure-eval                 0.2.2
pycparser                 2.21
pydantic                  2.10.4
pydantic_core             2.27.2
Pygments                  2.15.1
PyQt5                     5.15.10
PyQt5-sip                 12.13.0
PySocks                   1.7.1
python-dateutil           2.9.0.post0
python-json-logger        3.2.1
pytz                      2024.1
PyYAML                    6.0.2
pyzmq                     26.2.0
qtconsole                 5.6.0
QtPy                      2.4.1
referencing               0.30.2
requests                  2.32.3
```

```
rfc3339-validator         0.1.4
rfc3986-validator         0.1.1
rpds-py                   0.10.6
Send2Trash                1.8.2
setuptools                75.1.0
sip                       6.7.12
six                       1.16.0
sniffio                   1.3.0
soupsieve                 2.5
stack-data                0.2.0
terminado                 0.17.1
tinycss2                  1.2.1
tomli                     2.0.1
tornado                   6.4.2
traitlets                 5.14.3
typing_extensions         4.12.2
urllib3                   2.2.3
wcwidth                   0.2.5
webencodings              0.5.1
websocket-client          1.8.0
wheel                     0.44.0
widgetsnbextension        4.0.13
zipp                      3.21.0
```

```
!pip install transformers
```

```
Collecting transformers
  Using cached transformers-4.47.1-py3-none-any.whl.metadata (44 kB)
Collecting filelock (from transformers)
  Using cached filelock-3.16.1-py3-none-any.whl.metadata (2.9 kB)
Collecting huggingface-hub<1.0,>=0.24.0 (from transformers)
  Using cached huggingface_hub-0.27.0-py3-none-any.whl.metadata (13
kB)
Collecting numpy>=1.17 (from transformers)
  Using cached numpy-2.0.2-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (60 kB)
Requirement already satisfied: packaging>=20.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from transformers) (24.2)
Requirement already satisfied: pyyaml>=5.1 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from transformers) (6.0.2)
Collecting regex!=2019.12.17 (from transformers)
  Downloading regex-2024.11.6-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (40 kB)
Requirement already satisfied: requests in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from transformers) (2.32.3)
Collecting tokenizers<0.22,>=0.21 (from transformers)
  Using cached tokenizers-0.21.0-cp39-abi3-
```

```
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (6.7 kB)
Collecting safetensors>=0.4.1 (from transformers)
  Downloading safetensors-0.4.5-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (3.8 kB)
Collecting tqdm>=4.27 (from transformers)
  Using cached tqdm-4.67.1-py3-none-any.whl.metadata (57 kB)
Collecting fsspec>=2023.5.0 (from huggingface-hub<1.0,>=0.24.0-
>transformers)
  Using cached fsspec-2024.12.0-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from huggingface-hub<1.0,>=0.24.0->transformers) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->transformers) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->transformers) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->transformers) (2.2.3)
Requirement already satisfied: certifi>=2017.4.17 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->transformers) (2024.12.14)
Using cached transformers-4.47.1-py3-none-any.whl (10.1 MB)
Using cached huggingface_hub-0.27.0-py3-none-any.whl (450 kB)
Using cached numpy-2.0.2-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (19.5 MB)
Downloading regex-2024.11.6-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (780 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 780.9/780.9 kB 3.7 MB/s eta
0:00:00[31m5.4 MB/s eta 0:00:01
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (436 kB)
Using cached tokenizers-0.21.0-cp39-abi3-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.0 MB)
Using cached tqdm-4.67.1-py3-none-any.whl (78 kB)
Using cached filelock-3.16.1-py3-none-any.whl (16 kB)
Downloading fsspec-2024.12.0-py3-none-any.whl (183 kB)
Installing collected packages: tqdm, safetensors, regex, numpy,
fsspec, filelock, huggingface-hub, tokenizers, transformers
Successfully installed filelock-3.16.1 fsspec-2024.12.0 huggingface-
hub-0.27.0 numpy-2.0.2 regex-2024.11.6 safetensors-0.4.5 tokenizers-
0.21.0 tqdm-4.67.1 transformers-4.47.1
```

```
!pip install Pillow
```

```
Collecting Pillow
  Using cached pillow-11.0.0-cp39-cp39-
manylinux_2_28_x86_64.whl.metadata (9.1 kB)
Using cached pillow-11.0.0-cp39-cp39-manylinux_2_28_x86_64.whl (4.4
MB)
Installing collected packages: Pillow
Successfully installed Pillow-11.0.0
```

```
!pip install torch
```

```
Collecting torch
  Downloading torch-2.5.1-cp39-cp39-manylinux1_x86_64.whl.metadata (28
kB)
Requirement already satisfied: filelock in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (3.16.1)
Requirement already satisfied: typing-extensions>=4.8.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (4.12.2)
Collecting networkx (from torch)
  Using cached networkx-3.2.1-py3-none-any.whl.metadata (5.2 kB)
Requirement already satisfied: jinja2 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (3.1.4)
Requirement already satisfied: fsspec in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (2024.12.0)
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch)
  Using cached nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch)
  Using cached nvidia_cuda_runtime_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch)
  Using cached nvidia_cuda_cupti_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch)
  Using cached nvidia_cudnn_cu12-9.1.0.70-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch)
  Using cached nvidia_cublas_cu12-12.4.5.8-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cufft-cu12==11.2.1.3 (from torch)
  Using cached nvidia_cufft_cu12-11.2.1.3-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-curand-cu12==10.3.5.147 (from torch)
  Using cached nvidia_curand_cu12-10.3.5.147-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cusolver-cu12==11.6.1.9 (from torch)
  Using cached nvidia_cusolver_cu12-11.6.1.9-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cusparse-cu12==12.3.1.170 (from torch)
  Using cached nvidia_cusparse_cu12-12.3.1.170-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
```

```
Collecting nvidia-nccl-cu12==2.21.5 (from torch)
  Using cached nvidia_nccl_cu12-2.21.5-py3-none-
manylinux2014_x86_64.whl.metadata (1.8 kB)
Collecting nvidia-nvtx-cu12==12.4.127 (from torch)
  Using cached nvidia_nvtx_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.7 kB)
Collecting nvidia-nvjitlink-cu12==12.4.127 (from torch)
  Using cached nvidia_nvjitlink_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting triton==3.1.0 (from torch)
  Downloading triton-3.1.0-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (1.3 kB)
Collecting sympy==1.13.1 (from torch)
  Using cached sympy-1.13.1-py3-none-any.whl.metadata (12 kB)
Collecting mpmath<1.4,>=1.1.0 (from sympy==1.13.1->torch)
  Using cached mpmath-1.3.0-py3-none-any.whl.metadata (8.6 kB)
Requirement already satisfied: MarkupSafe>=2.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from jinja2->torch) (2.1.3)
Downloading torch-2.5.1-cp39-cp39-manylinux1_x86_64.whl (906.5 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 906.5/906.5 MB 2.6 MB/s eta
0:00:00m eta 0:00:01[36m0:00:10
anylinux2014_x86_64.whl (363.4 MB)
Using cached nvidia_cuda_cupti_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl (13.8 MB)
Using cached nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl (24.6 MB)
Using cached nvidia_cuda_runtime_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl (883 kB)
Using cached nvidia_cudnn_cu12-9.1.0.70-py3-none-
manylinux2014_x86_64.whl (664.8 MB)
Using cached nvidia_cufft_cu12-11.2.1.3-py3-none-
manylinux2014_x86_64.whl (211.5 MB)
Using cached nvidia_curand_cu12-10.3.5.147-py3-none-
manylinux2014_x86_64.whl (56.3 MB)
Using cached nvidia_cusolver_cu12-11.6.1.9-py3-none-
manylinux2014_x86_64.whl (127.9 MB)
Using cached nvidia_cusparse_cu12-12.3.1.170-py3-none-
manylinux2014_x86_64.whl (207.5 MB)
Using cached nvidia_nccl_cu12-2.21.5-py3-none-manylinux2014_x86_64.whl
(188.7 MB)
Using cached nvidia_nvjitlink_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl (21.1 MB)
Using cached nvidia_nvtx_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl (99 kB)
Using cached sympy-1.13.1-py3-none-any.whl (6.2 MB)
Downloading triton-3.1.0-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (209.5 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 209.5/209.5 MB 2.7 MB/s eta
```

```
0:00:00m eta 0:00:01[36m0:00:03
pmath-1.3.0-py3-none-any.whl (536 kB)
Installing collected packages: mpmath, triton, sympy, nvidia-nvtx-
cu12, nvidia-nvjitlink-cu12, nvidia-nccl-cu12, nvidia-curand-cu12,
nvidia-cufft-cu12, nvidia-cuda-runtime-cu12, nvidia-cuda-nvrtc-cu12,
nvidia-cuda-cupti-cu12, nvidia-cublas-cu12, networkx, nvidia-cusparse-
cu12, nvidia-cudnn-cu12, nvidia-cusolver-cu12, torch
Successfully installed mpmath-1.3.0 networkx-3.2.1 nvidia-cublas-cu12-
12.4.5.8 nvidia-cuda-cupti-cu12-12.4.127 nvidia-cuda-nvrtc-cu12-
12.4.127 nvidia-cuda-runtime-cu12-12.4.127 nvidia-cudnn-cu12-9.1.0.70
nvidia-cufft-cu12-11.2.1.3 nvidia-curand-cu12-10.3.5.147 nvidia-
cusolver-cu12-11.6.1.9 nvidia-cusparse-cu12-12.3.1.170 nvidia-nccl-
cu12-2.21.5 nvidia-nvjitlink-cu12-12.4.127 nvidia-nvtx-cu12-12.4.127
sympy-1.13.1 torch-2.5.1 triton-3.1.0
```

```
!pip install matplotlib
```

```
Collecting matplotlib
  Downloading matplotlib-3.9.4-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (11 kB)
Collecting contourpy>=1.0.1 (from matplotlib)
  Using cached contourpy-1.3.0-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (5.4 kB)
Collecting cycler>=0.10 (from matplotlib)
  Using cached cycler-0.12.1-py3-none-any.whl.metadata (3.8 kB)
Collecting fonttools>=4.22.0 (from matplotlib)
  Downloading fonttools-4.55.3-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (165 kB)
Collecting kiwisolver>=1.3.1 (from matplotlib)
  Using cached kiwisolver-1.4.7-cp39-cp39-
manylinux_2_12_x86_64.manylinux2010_x86_64.whl.metadata (6.3 kB)
Requirement already satisfied: numpy>=1.23 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from matplotlib) (2.0.2)
Requirement already satisfied: packaging>=20.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from matplotlib) (24.2)
Requirement already satisfied: pillow>=8 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from matplotlib) (11.0.0)
Collecting pyparsing>=2.3.1 (from matplotlib)
  Using cached pyparsing-3.2.0-py3-none-any.whl.metadata (5.0 kB)
Requirement already satisfied: python-dateutil>=2.7 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from matplotlib) (2.9.0.post0)
Collecting importlib-resources>=3.2.0 (from matplotlib)
  Using cached importlib_resources-6.4.5-py3-none-any.whl.metadata
(4.0 kB)
Requirement already satisfied: zipp>=3.1.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
```

```
(from importlib-resources>=3.2.0->matplotlib) (3.21.0)
Requirement already satisfied: six>=1.5 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from python-dateutil>=2.7->matplotlib) (1.16.0)
Downloading matplotlib-3.9.4-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (8.3 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 8.3/8.3 MB 356.4 kB/s eta
0:00:00[36m0:00:01m eta 0:00:02
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (321 kB)
Using cached cycler-0.12.1-py3-none-any.whl (8.3 kB)
Downloading fonttools-4.55.3-cp39-cp39-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (4.6 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 4.6/4.6 MB 374.8 kB/s eta
0:00:001m376.8 kB/s eta 0:00:01
portlib_resources-6.4.5-py3-none-any.whl (36 kB)
Using cached kiwisolver-1.4.7-cp39-cp39-
manylinux_2_12_x86_64.manylinux2010_x86_64.whl (1.6 MB)
Using cached pyparsing-3.2.0-py3-none-any.whl (106 kB)
Installing collected packages: pyparsing, kiwisolver, importlib-
resources, fonttools, cycler, contourpy, matplotlib
Successfully installed contourpy-1.3.0 cycler-0.12.1 fonttools-4.55.3
importlib-resources-6.4.5 kiwisolver-1.4.7 matplotlib-3.9.4 pyparsing-
3.2.0
```

```
!pip install timm einops
```

```
Requirement already satisfied: timm in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(1.0.12)
Requirement already satisfied: einops in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(0.8.0)
Requirement already satisfied: torch in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from timm) (2.5.1)
Requirement already satisfied: torchvision in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from timm) (0.20.1)
Requirement already satisfied: pyyaml in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from timm) (6.0.2)
Requirement already satisfied: huggingface_hub in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from timm) (0.27.0)
Requirement already satisfied: safetensors in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from timm) (0.4.5)
Requirement already satisfied: filelock in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from huggingface_hub->timm) (3.16.1)
```

```
Requirement already satisfied: fsspec>=2023.5.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from huggingface_hub->timm) (2024.12.0)
Requirement already satisfied: packaging>=20.9 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from huggingface_hub->timm) (24.2)
Requirement already satisfied: requests in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from huggingface_hub->timm) (2.32.3)
Requirement already satisfied: tqdm>=4.42.1 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from huggingface_hub->timm) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from huggingface_hub->timm) (4.12.2)
Requirement already satisfied: networkx in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (3.2.1)
Requirement already satisfied: jinja2 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (3.1.4)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (12.4.127)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127
in /home/students/anaconda3/envs/florencenv/lib/python3.9/site-
packages (from torch->timm) (12.4.127)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (12.4.127)
Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (9.1.0.70)
Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (12.4.5.8)
Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (11.2.1.3)
Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (10.3.5.147)
Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (11.6.1.9)
Requirement already satisfied: nvidia-cusparse-cu12==12.3.1.170 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (12.3.1.170)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in
```

```
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (12.4.127)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (12.4.127)
Requirement already satisfied: triton==3.1.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (3.1.0)
Requirement already satisfied: sympy==1.13.1 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch->timm) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from sympy==1.13.1->torch->timm) (1.3.0)
Requirement already satisfied: numpy in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torchvision->timm) (2.0.2)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torchvision->timm) (11.0.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from jinja2->torch->timm) (2.1.3)
Requirement already satisfied: charset-normalizer<4,>=2 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->huggingface_hub->timm) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->huggingface_hub->timm) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->huggingface_hub->timm) (2.2.3)
Requirement already satisfied: certifi>=2017.4.17 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from requests->huggingface_hub->timm) (2024.12.14)
```

```
from transformers import AutoProcessor, AutoModelForCausalLM
from PIL import Image
import requests
import copy
import torch
%matplotlib inline
```

```
!pip install torch torchvision torchaudio
```

```
Requirement already satisfied: torch in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
```

```
(2.5.1)
Requirement already satisfied: torchvision in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(0.20.1)
Requirement already satisfied: torchaudio in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(2.5.1)
Requirement already satisfied: filelock in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (3.16.1)
Requirement already satisfied: typing-extensions>=4.8.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (4.12.2)
Requirement already satisfied: networkx in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (3.2.1)
Requirement already satisfied: jinja2 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (3.1.4)
Requirement already satisfied: fsspec in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (2024.12.0)
Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (12.4.127)
Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127
in /home/students/anaconda3/envs/florencenv/lib/python3.9/site-
packages (from torch) (12.4.127)
Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (12.4.127)
Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (9.1.0.70)
Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (12.4.5.8)
Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (11.2.1.3)
Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (10.3.5.147)
Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (11.6.1.9)
Requirement already satisfied: nvidia-cusparse-cu12==12.3.1.170 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (12.3.1.170)
```

```
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (12.4.127)
Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (12.4.127)
Requirement already satisfied: triton==3.1.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (3.1.0)
Requirement already satisfied: sympy==1.13.1 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torch) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from sympy==1.13.1->torch) (1.3.0)
Requirement already satisfied: numpy in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torchvision) (2.0.2)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from torchvision) (11.0.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages
(from jinja2->torch) (2.1.3)
```

```
model_id = 'microsoft/Florence-2-large'
model = AutoModelForCausalLM.from_pretrained(model_id,
trust_remote_code=True, torch_dtype='auto').eval().cuda()
processor = AutoProcessor.from_pretrained(model_id,
trust_remote_code=True)
```

```
/home/students/anaconda3/envs/florencenv/lib/python3.9/site-packages/
timm/models/layers/__init__.py:48: FutureWarning: Importing from
timm.models.layers is deprecated, please import via timm.layers
  warnings.warn(f"Importing from {__name__} is deprecated, please
import via timm.layers", FutureWarning)
```

define the prediction function

```
def run_example(task_prompt, text_input=None):
    if text_input is None:
        prompt = task_prompt
    else:
        prompt = task_prompt + text_input
    inputs = processor(text=prompt, images=image,
return_tensors="pt").to('cuda', torch.float16)
```

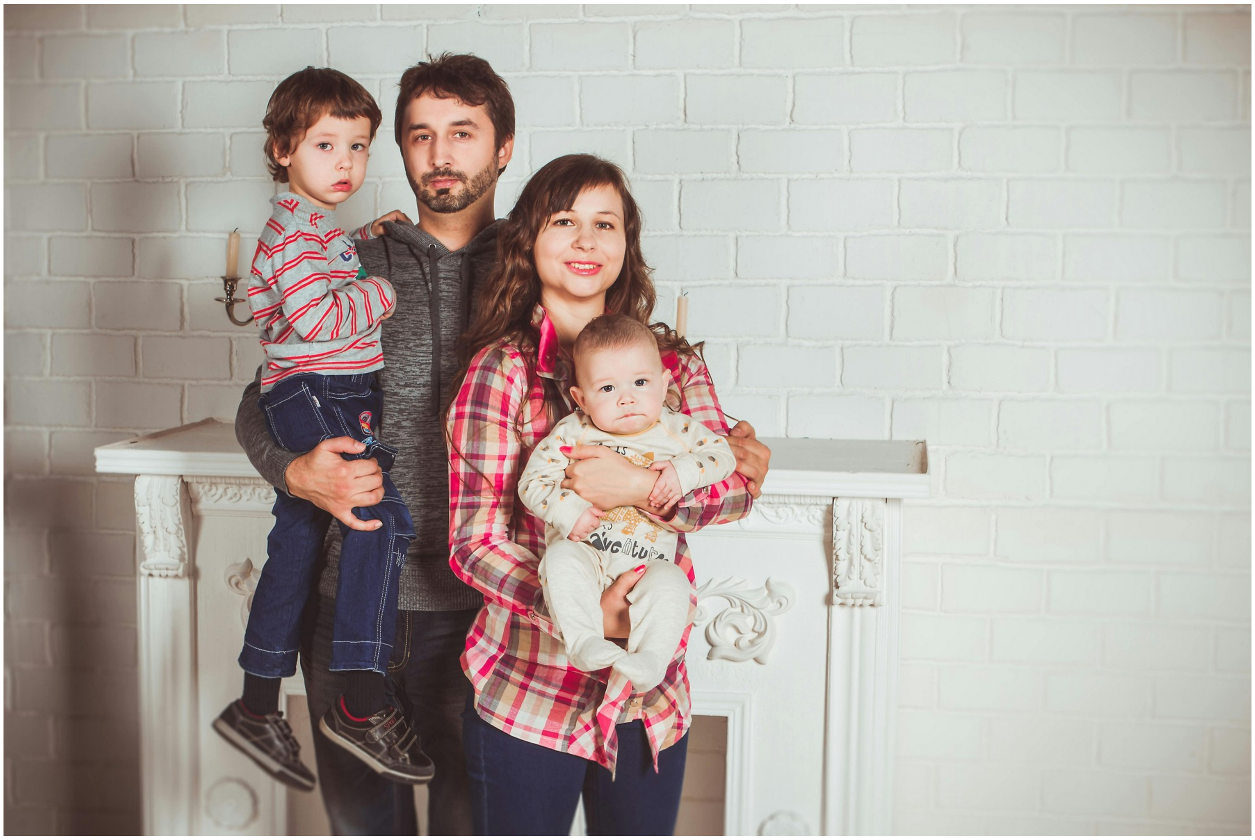
```
    generated_ids = model.generate(
      input_ids=inputs["input_ids"].cuda(),
      pixel_values=inputs["pixel_values"].cuda(),
      max_new_tokens=1024,
      early_stopping=False,
      do_sample=False,
      num_beams=3,
    )
    generated_text = processor.batch_decode(generated_ids,
skip_special_tokens=False)[0]
    parsed_answer = processor.post_process_generation(
        generated_text,
        task=task_prompt,
        image_size=(image.width, image.height)
    )

    return parsed_answer
```

init image

```
#url = "https://huggingface.co/datasets/huggingface/documentation-
images/resolve/main/transformers/tasks/car.jpg?download=true"
#image = Image.open(requests.get(url, stream=True).raw)
from PIL import Image

# Path to the local image
image_path = "/home/students/Downloads/family.jpg"

# Open the image
image = Image.open(image_path)

# Display the image (optional)
image.show()
```

```
huggingface/tokenizers: The current process just got forked, after
parallelism has already been used. Disabling parallelism to avoid
deadlocks...
To disable this warning, you can either:
	- Avoid using `tokenizers` before the fork if possible
	- Explicitly set the environment variable
TOKENIZERS_PARALLELISM=(true | false)
```

```
image
```



Run pre-defined tasks without additional inputs

Caption

```
task_prompt = '<CAPTION>'
run_example(task_prompt)

{'<CAPTION>': 'A family posing for a picture in front of a
fireplace.'}

task_prompt = '<DETAILED_CAPTION>'
run_example(task_prompt)

{'<DETAILED_CAPTION>': 'The image shows a family of four posing for a
portrait in front of a fireplace. The man and woman are standing on
the floor, each holding a baby in their arms, while the two children
stand on either side of them. In the background, there is a candle
stand with a candle on it, and a wall behind them.'}

task_prompt = '<MORE_DETAILED_CAPTION>'
run_example(task_prompt)

{'<MORE_DETAILED_CAPTION>': 'The image shows a family of four posing
for a photo in front of a white brick wall. The family consists of a
man, a woman, and two young children. The man is standing on the left
side of the image, holding one of the children in his arms. The woman
is holding a baby in her arms, while the other two children are
standing on either side of her. The child on the right side is wearing
a red and white striped shirt and blue jeans, and the child in the
middle is holding the baby in a white onesie. All four family members
are smiling and looking at the camera.'}
```

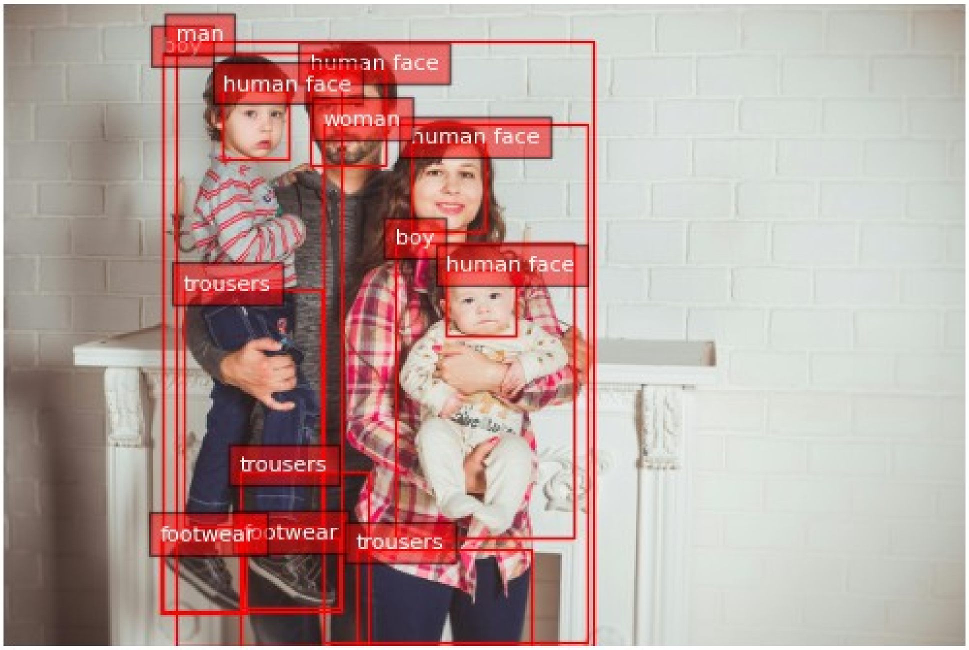
Object detection

OD results format: {'<OD>': { 'bboxes': [[x1, y1, x2, y2], ...], 'labels': ['label1', 'label2', ...] } }

```
task_prompt = '<OD>'
results = run_example(task_prompt)
print(results)
```

```
{'<OD>': {'bboxes': [[852.7680053710938, 254.01600646972656,
1822.176025390625, 3271.10400390625], [2107.2958984375,
1283.904052734375, 3071.52001953125, 2866.751953125],
[1288.2239990234375, 2870.2080078125, 1806.6240234375,
3246.912109375], [842.4000244140625, 2880.576171875, 1308.9599609375,
3278.01611328125], [1651.10400390625, 347.3280029296875,
2060.639892578125, 865.72802734375], [2190.239990234375,
741.31201171875, 2594.592041015625, 1218.239990234375],
[1184.5439453125, 454.4640197753906, 1537.0560302734375,
831.1680297851562], [2387.23193359375, 1429.0560302734375,
2755.2958984375, 1785.0240478515625], [925.343994140625,
191.80799865722656, 3175.199951171875, 3450.816162109375],
[966.8159790039062, 1532.736083984375, 1718.4959716796875,
2804.544189453125], [1272.6719970703125, 2510.7841796875,
1967.3280029296875, 3454.27197265625], [1899.93603515625,
2928.9599609375, 2843.424072265625, 3454.27197265625],
[1723.6800537109375, 644.5440063476562, 3149.280029296875,
3430.080078125]], 'labels': ['boy', 'boy', 'footwear', 'footwear',
'human face', 'human face', 'human face', 'human face', 'man',
'trousers', 'trousers', 'trousers', 'woman']}}
```

```
import matplotlib.pyplot as plt
import matplotlib.patches as patches
def plot_bbox(image, data):
   # Create a figure and axes
    fig, ax = plt.subplots()

    # Display the image
    ax.imshow(image)

    # Plot each bounding box
    for bbox, label in zip(data['bboxes'], data['labels']):
        # Unpack the bounding box coordinates
        x1, y1, x2, y2 = bbox
        # Create a Rectangle patch
        rect = patches.Rectangle((x1, y1), x2-x1, y2-y1, linewidth=1,
edgecolor='r', facecolor='none')
        # Add the rectangle to the Axes
        ax.add_patch(rect)
        # Annotate the label
        plt.text(x1, y1, label, color='white', fontsize=8,
bbox=dict(facecolor='red', alpha=0.5))

    # Remove the axis ticks and labels
    ax.axis('off')

    # Show the plot
    plt.show()

plot_bbox(image, results['<OD>'])
```

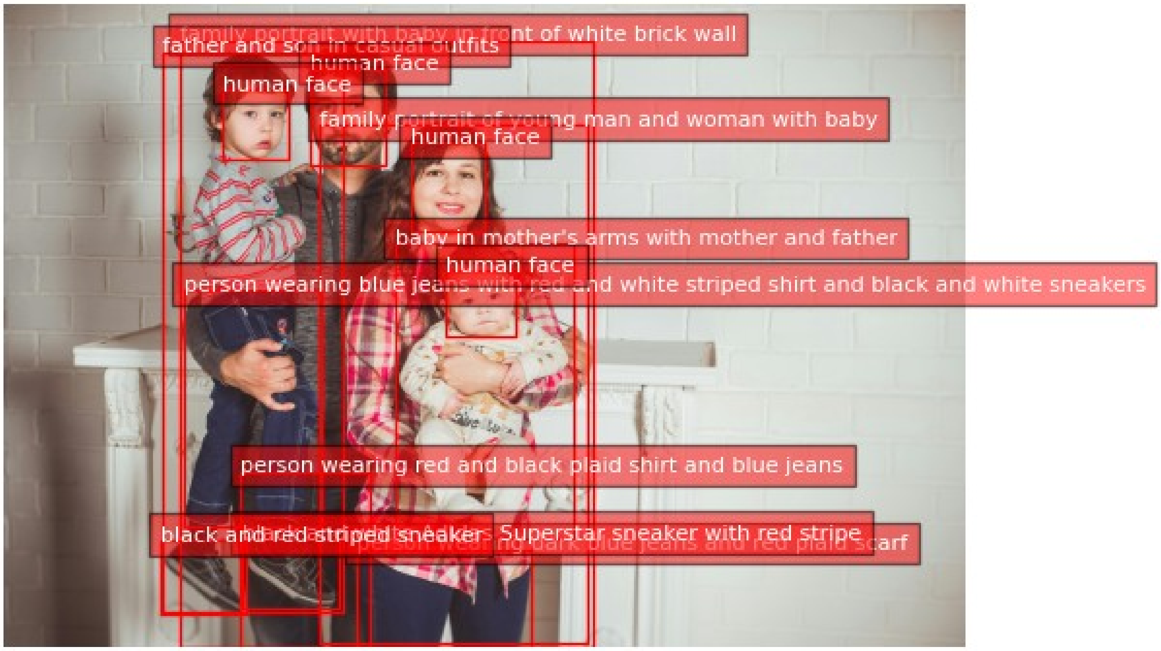


Dense region caption

Dense region caption results format: {'<DENSE_REGION_CAPTION>': {'bboxes': [[x1, y1, x2, y2], ...], 'labels': ['label1', 'label2', ...]}}

```
task_prompt = '<DENSE_REGION_CAPTION>'
results = run_example(task_prompt)
print(results)
```

```
{'<DENSE_REGION_CAPTION>': {'bboxes': [[946.0800170898438,
191.80799865722656, 3175.199951171875, 3450.816162109375],
[1702.9439697265625, 644.5440063476562, 3144.095947265625,
3430.080078125], [857.9520263671875, 254.01600646972656,
1827.3599853515625, 3271.10400390625], [2107.2958984375,
1283.904052734375, 3071.52001953125, 2866.751953125],
[966.8159790039062, 1532.736083984375, 1718.4959716796875,
2804.544189453125], [1277.85595703125, 2510.7841796875,
1972.511962890625, 3454.27197265625], [1899.93603515625,
2928.9599609375, 2843.424072265625, 3454.27197265625],
[1651.10400390625, 347.3280029296875, 2060.639892578125,
865.72802734375], [2190.239990234375, 741.31201171875,
2594.592041015625, 1218.239990234375], [1283.0400390625,
2870.2080078125, 1806.6240234375, 3246.912109375], [842.4000244140625,
2880.576171875, 1308.9599609375, 3278.01611328125], [1184.5439453125,
454.4640197753906, 1537.0560302734375, 831.1680297851562],
[2387.23193359375, 1429.0560302734375, 2755.2958984375,
1785.0240478515625]], 'labels': ['family portrait with baby in front
of white brick wall', 'family portrait of young man and woman with
baby', 'father and son in casual outfits', "baby in mother's arms with
mother and father", 'person wearing blue jeans with red and white
striped shirt and black and white sneakers', 'person wearing red and
black plaid shirt and blue jeans', 'person wearing dark blue jeans and
red plaid scarf', 'human face', 'human face', 'black and white Adidas
Superstar sneaker with red stripe', 'black and red striped sneaker',
'human face', 'human face']}}

plot_bbox(image, results['<DENSE_REGION_CAPTION>'])
```

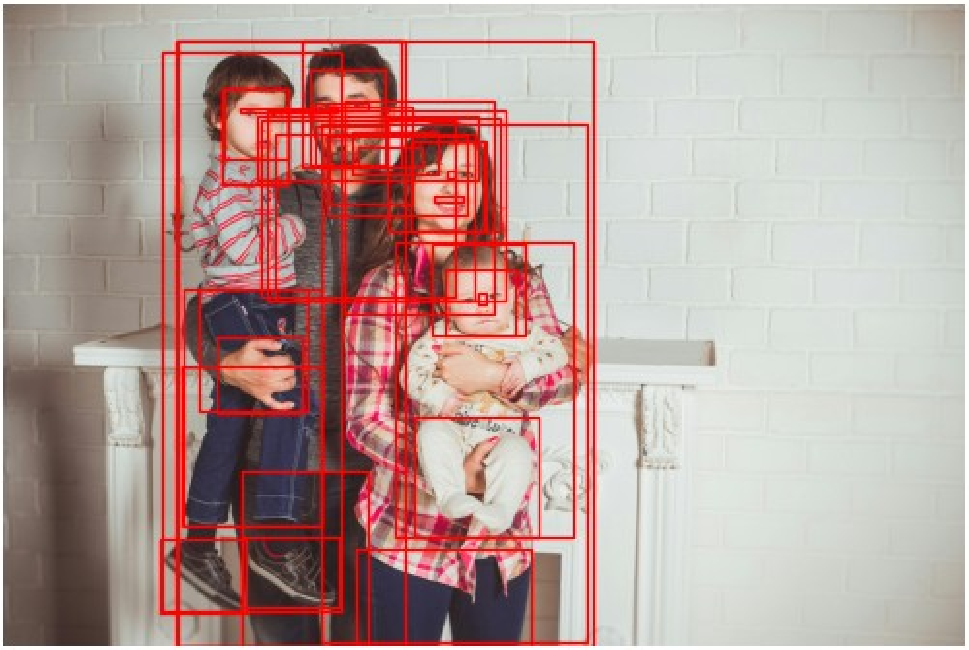

family portrait with baby in front of white brick wall
father and son in casual outfits
human face
human face
family portrait of young man and woman with baby
human face
baby in mother's arms with mother and father
human face
person wearing blue jeans with red and white striped shirt and black and white sneakers
person wearing red and black plaid shirt and blue jeans
black and red striped sneaker
Superstar sneaker with red stripe
person wearing dark blue jeans and red plaid scarf

Region proposal

Region proposal results format: {'' : {'bboxes': [[x1, y1, x2, y2], ...], 'labels': ['', '', ...]}}

```
task_prompt = '<REGION_PROPOSAL>'
results = run_example(task_prompt)
print(results)
```

```
{'<REGION_PROPOSAL>': {'bboxes': [[940.89599609375,
195.26400756835938, 3175.199951171875, 3450.816162109375],
[1718.4959716796875, 641.0880126953125, 3149.280029296875,
3433.5361328125], [930.5280151367188, 195.26400756835938,
2164.320068359375, 3450.816162109375], [852.7680053710938,
254.01600646972656, 1816.9920654296875, 3271.10400390625],
[2107.2958984375, 1283.904052734375, 3071.52001953125,
2866.751953125], [966.8159790039062, 1529.280029296875,
1718.4959716796875, 2804.544189453125], [961.6320190429688,
1950.912109375, 1718.4959716796875, 2804.544189453125],
[1288.2239990234375, 2510.7841796875, 1967.3280029296875,
3454.27197265625], [1899.93603515625, 2928.9599609375,
2843.424072265625, 3454.27197265625], [2210.97607421875,
2220.47998046875, 2879.7119140625, 2866.751953125],
[1604.447998046875, 191.80799865722656, 2143.583984375,
869.1840209960938], [2065.823974609375, 616.89599609375,
2677.535888671875, 1221.696044921875], [1054.9439697265625,
1529.280029296875, 1635.552001953125, 2192.83203125],
[2304.2880859375, 1287.3599853515625, 2812.320068359375,
1785.0240478515625], [1651.10400390625, 347.3280029296875,
2060.639892578125, 865.72802734375], [2190.239990234375,
741.31201171875, 2594.592041015625, 1218.239990234375],
[1288.2239990234375, 2870.2080078125, 1806.6240234375,
3246.912109375], [842.4000244140625, 2880.576171875, 1308.9599609375,
3278.01611328125], [1148.2559814453125, 1788.4801025390625,
1609.6319580078125, 2196.2880859375], [2190.239990234375,
741.31201171875, 2599.776123046875, 1218.239990234375],
[2382.048095703125, 1425.60009765625, 2755.2958984375,
1785.0240478515625], [1184.5439453125, 454.4640197753906,
1531.8719482421875, 972.864013671875], [1184.5439453125,
457.9200134277344, 1531.8719482421875, 972.864013671875],
[1189.72802734375, 457.9200134277344, 1531.8719482421875,
831.1680297851562], [1677.0240478515625, 658.3680419921875,
2065.823974609375, 876.0960083007812], [2205.7919921875,
1073.0880126953125, 1754.7840576171875, 551.2319946289062],
[1879.199951171875, 682.5599975585938, 2345.760009765625,
914.1119995117188], [2475.360107421875, 1031.615966796875,
2314.656005859375, 1059.2640380859375], [2501.280029296875,
1131.840087890625, 1744.416015625, 706.7520141601562],
[1915.488037109375, 768.9600219726562, 2252.447998046875,
876.0960083007812], [2392.416015625, 934.8480224609375,
2433.887939453125, 889.9200439453125], [2563.488037109375,
948.6719970703125, 1365.9840087890625, 603.072021484375],
[1459.2960205078125, 692.9280395507812, 2537.568115234375,
1594.944091796875], [2641.248046875, 1670.97607421875,
1837.72802734375, 516.6719970703125], [1967.3280029296875,
571.968017578125, 1677.0240478515625, 527.0400390625],
[1796.2559814453125, 585.7919921875, 1272.6719970703125,
565.0560302734375], [1386.719970703125, 620.3519897460938,
2439.072021484375, 1570.7520751953125], [2558.303955078125,
1619.135986328125, 2599.776123046875, 1550.0159912109375],
[2708.639892578125, 1598.4000244140625, 1423.008056640625,
568.5120239257812]], 'labels': ['', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '', '', '', '', '', '', '', '',
'', '', '', '', '', '', '', '', '', '', '', '', '']}}

plot_bbox(image, results['<REGION_PROPOSAL>'])
```



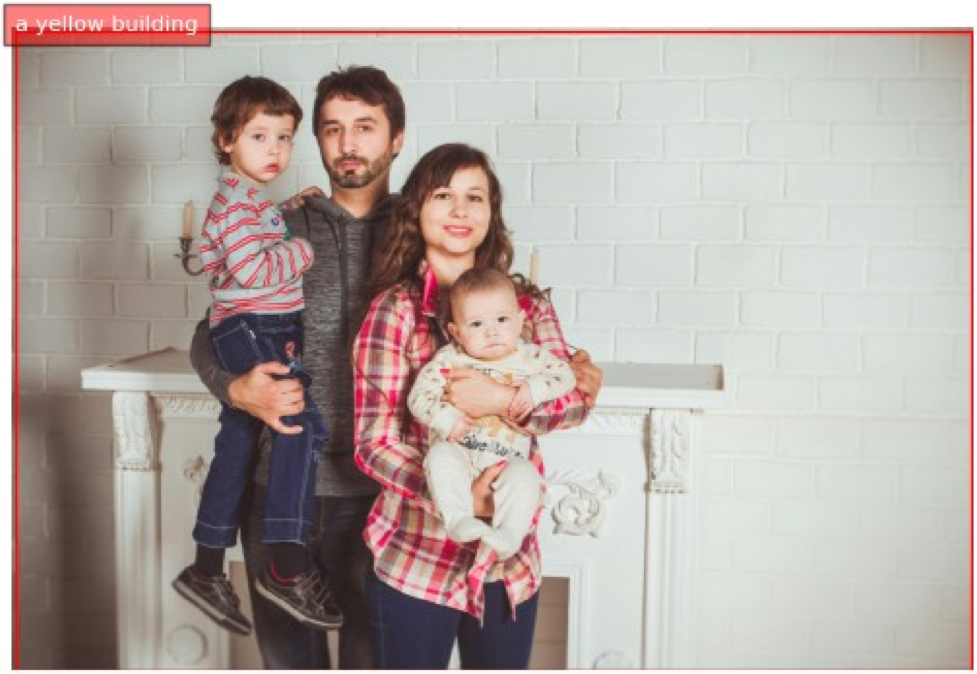
Run pre-defined tasks that requires additional inputs

Phrase Grounding

Phrase grounding results format: {'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[x1, y1, x2, y2], ...], 'labels': ['', '', ...]}}

```
task_prompt = '<CAPTION_TO_PHRASE_GROUNDING>'
results = run_example(task_prompt, text_input="A green car parked in
front of a yellow building.")
print(results)

{'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[18.144001007080078,
15.552000045776367, 5171.0400390625, 3450.816162109375]], 'labels':
['a yellow building']}}
```

```
plot_bbox(image, results['<CAPTION_TO_PHRASE_GROUNDING>'])
```



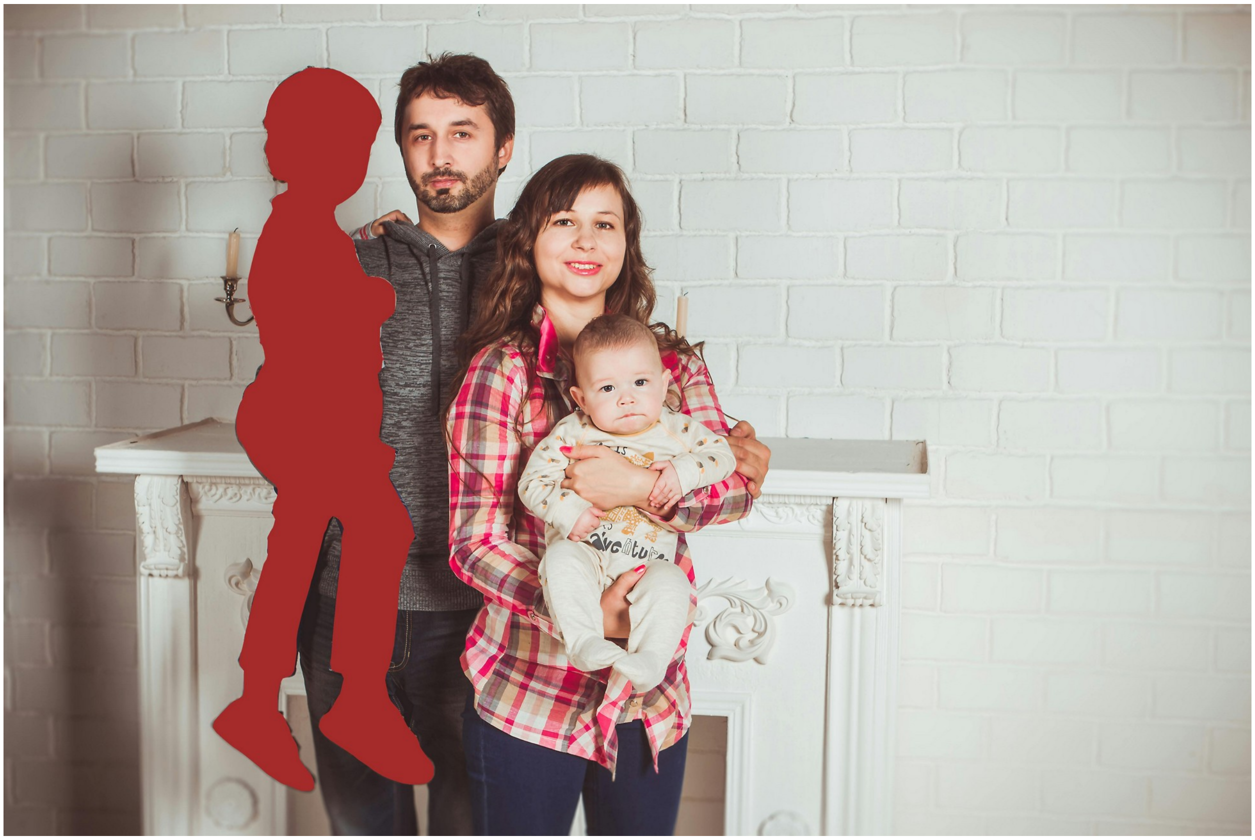
Referring expression segmentation

Referring expression segmentation results format: {'<REFERRING_EXPRESSION_SEGMENTATION>': {'Polygons': [[[polygon]], ...], 'labels': ['', '', ...]}}, one object is represented by a list of polygons. each polygon is [x1, y1, x2, y2, ..., xn, yn]

```
task_prompt = '<REFERRING_EXPRESSION_SEGMENTATION>'
results = run_example(task_prompt, text_input="peoples")
print(results)
```

```
{'<REFERRING_EXPRESSION_SEGMENTATION>': {'polygons':
[[[1132.7039794921875, 350.78399658203125, 1179.3599853515625,
309.31201171875, 1210.4639892578125, 292.0320129394531,
1241.5679931640625, 278.2080078125, 1267.488037109375,
264.3840026855469, 1283.0400390625, 264.3840026855469,
1365.9840087890625, 271.2960205078125, 1397.0880126953125,
281.66400146484375, 1417.823974609375, 292.0320129394531,
1443.7440185546875, 305.85601806640625, 1469.6640625,
323.1360168457031, 1511.135986328125, 357.6960144042969,
1537.0560302734375, 392.2560119628906, 1552.6080322265625,
419.9040222167969, 1562.9759521484375, 440.6400146484375,
1568.1600341796875, 492.4800109863281, 1557.7919921875,
516.6719970703125, 1547.4239501953125, 540.864013671875,
1542.239990234375, 565.0560302734375, 1526.68798828125,
589.2479858398438, 1521.5040283203125, 616.89599609375,
1516.3199462890625, 648.0, 1511.135986328125, 672.1920166015625,
1505.9520263671875, 706.7520141601562, 1495.583984375,
734.4000244140625, 1485.216064453125, 755.1360473632812,
1459.2960205078125, 786.239990234375, 1428.1920166015625,
810.4320068359375, 1402.27197265625, 827.7120361328125,
1381.5360107421875, 838.0800170898438, 1371.16796875, 865.7280273437 5,
1376.35205078125, 893.3760375976562, 1386.719970703125,
917.5680541992188, 1402.27197265625, 938.3040161132812,
1423.008056640625, 955.5840454101562, 1443.7440185546875,
972.864013671875, 1454.112060546875, 993.6000366210938,
1459.2960205078125, 1017.7920532226562, 1464.47998046875,
1045.4400634765625, 1474.8480224609375, 1069.632080078125,
1485.216064453125, 1090.3680419921875, 1495.583984375,
1111.10400390625, 1511.135986328125, 1131.840087890625,
1557.7919921875, 1135.2960205078125, 1583.7120361328125, 1145.6640625,
1609.6319580078125, 1166.4000244140625, 1620.0, 1187.135986328125,
1625.1839599609375, 1207.8720703125, 1625.1839599609375,
1259.7120361328125, 1620.0, 1283.904052734375, 1599.2640380859375,
1304.6400146484375, 1573.343994140625, 1325.3759765625,
1562.9759521484375, 1346.112060546875, 1562.9759521484375,
1411.7760009765625, 1568.1600341796875, 1439.424072265625,
1573.343994140625, 1477.4400634765625, 1568.1600341796875,
1518.9119873046875, 1552.6080322265625, 1536.1920166015625,
1557.7919921875, 1570.7520751953125, 1562.9759521484375,
1594.944091796875, 1573.343994140625, 1615.6800537109375,
1573.343994140625, 1691.7120361328125, 1568.1600341796875,
1729.72802734375, 1562.9759521484375, 1760.83203125, 1557.7919921875,
1798.8480224609375, 1568.1600341796875, 1819.5841064453125,
1588.89599609375, 1829.9520263671875, 1609.6319580078125,
1840.320068359375, 1625.1839599609375, 1850.68798828125,
1625.1839599609375, 1878.3360595703125, 1620.0, 1899.072021484375,
1614.8160400390625, 1923.2640380859375, 1599.2640380859375, 1944.0,
1599.2640380859375, 1968.1920166015625, 1609.6319580078125,
1988.9281005859375, 1620.0, 2009.6640625, 1630.3680419921875,
2026.944091796875, 1640.7359619140625, 2047.6800537109375,
1651.10400390625, 2068.416015625, 1666.656005859375,
2089.152099609375, 1677.0240478515625, 2109.887939453125,
1682.2080078125, 2127.16796875, 1692.5760498046875, 2147.904052734375,
1697.760009765625, 2168.64013671875, 1702.9439697265625,
2189.3759765625, 1708.1280517578125, 2210.112060546875,
1697.760009765625, 2230.84814453125, 1687.3919677734375,
2248.128173828125, 1682.2080078125, 2268.864013671875,
1677.0240478515625, 2289.60009765625, 1671.8399658203125,
2310.336181640625, 1666.656005859375, 2331.072021484375,
1656.2879638671875, 2351.80810546875, 1651.10400390625,
2372.5439453125, 1645.9200439453125, 2396.736083984375,
1640.7359619140625, 2448.576171875, 1635.552001953125,
2521.152099609375, 1630.3680419921875, 2562.6240234375,
1625.1839599609375, 2614.464111328125, 1620.0, 2662.84814453125,
1614.8160400390625, 2693.9521484375, 1609.6319580078125,
2718.14404296875, 1604.447998046875, 2738.880126953125,
1599.2640380859375, 2759.615966796875, 1588.89599609375,
2780.35205078125, 1583.7120361328125, 2808.0, 1588.89599609375,
2832.192138671875, 1594.0799560546875, 2859.840087890625,
1599.2640380859375, 2880.576171875, 1609.6319580078125,
2897.85595703125, 1625.1839599609375, 2915.135986328125,
1640.7359619140625, 2932.416015625, 1651.10400390625,
2953.152099609375, 1661.4720458984375, 2970.43212890625,
1671.8399658203125, 2991.16796875, 1682.2080078125, 3008.447998046875,
1697.760009765625, 3025.72802734375, 1713.31201171875,
3043.008056640625, 1723.6800537109375, 3063.744140625,
1728.864013671875, 3084.47998046875, 1739.2320556640625,
3101.760009765625, 1754.7840576171875, 3122.49609375,
1775.52001953125, 3146.68798828125, 1785.8880615234375,
3167.424072265625, 1785.8880615234375, 3201.984130859375,
1775.52001953125, 3222.719970703125, 1754.7840576171875,
3236.544189453125, 1728.864013671875, 3240.0, 1682.2080078125, 3240.0,
1656.2879638671875, 3236.544189453125, 1635.552001953125,
3229.632080078125, 1614.8160400390625, 3222.719970703125,
1594.0799560546875, 3215.80810546875, 1573.343994140625,
3205.440185546875, 1557.7919921875, 3195.072021484375,
1537.0560302734375, 3184.7041015625, 1516.3199462890625,
3167.424072265625, 1495.583984375, 3153.60009765625,
1474.8480224609375, 3139.776123046875, 1454.112060546875,
3122.49609375, 1433.3759765625, 3108.672119140625, 1412.6400146484375,
3094.84814453125, 1397.0880126953125, 3084.47998046875,
1376.35205078125, 3070.656005859375, 1355.615966796875, 3056.83203125,
1334.8800048828125, 3039.552001953125, 1319.3280029296875,
3022.27197265625, 1308.9599609375, 3001.5361328125, 1314.14404296875,
2973.88818359375, 1329.696044921875, 2953.152099609375,
1350.4320068359375, 2939.328125, 1365.9840087890625,
2918.592041015625, 1376.35205078125, 2897.85595703125,
1386.719970703125, 2880.576171875, 1397.0880126953125,
2863.296142578125, 1402.27197265625, 2842.56005859375,
1407.4560546875, 2821.823974609375, 1412.6400146484375,
2801.088134765625, 1402.27197265625, 2780.35205078125,
1376.35205078125, 2773.440185546875, 1360.800048828125,
2763.072021484375, 1355.615966796875, 2735.424072265625,
1360.800048828125, 2707.776123046875, 1365.9840087890625,
2638.656005859375, 1371.16796875, 2569.5361328125, 1376.35205078125,
2521.152099609375, 1381.5360107421875, 2465.85595703125,
1386.719970703125, 2424.384033203125, 1391.904052734375, 2376.0,
1397.0880126953125, 2317.248046875, 1402.27197265625,
2275.776123046875, 1402.27197265625, 2230.84814453125,
1407.4560546875, 2199.744140625, 1412.6400146484375,
2179.008056640625, 1402.27197265625, 2154.816162109375,
1386.719970703125, 2134.080078125, 1365.9840087890625, 2127.16796875,
1350.4320068359375, 2144.447998046875, 1345.248046875,
2161.72802734375, 1340.06396484375, 2179.008056640625,
1329.696044921875, 2199.744140625, 1324.511962890625,
2220.47998046875, 1319.3280029296875, 2241.216064453125,
1314.14404296875, 2261.9521484375, 1308.9599609375, 2282.68798828125,
1303.7760009765625, 2303.424072265625, 1298.592041015625,
2320.7041015625, 1293.407958984375, 2341.43994140625,
1288.2239990234375, 2358.719970703125, 1283.0400390625, 2376.0,
1277.85595703125, 2393.280029296875, 1272.6719970703125,
2414.01611328125, 1267.488037109375, 2431.296142578125,
1262.303955078125, 2448.576171875, 1257.1199951171875,
2465.85595703125, 1251.93603515625, 2483.135986328125, 1246.751953125,
2503.8720703125, 1241.5679931640625, 2521.152099609375,
1236.384033203125, 2538.43212890625, 1231.199951171875, 2559.16796875,
1226.0159912109375, 2579.904052734375, 1220.83203125,
2604.095947265625, 1215.64794921875, 2638.656005859375,
1215.64794921875, 2693.9521484375, 1210.4639892578125,
2735.424072265625, 1205.280029296875, 2769.984130859375,
1194.9119873046875, 2787.26416015625, 1179.3599853515625,
2794.176025390625, 1158.6240234375, 2801.088134765625,
1148.2559814453125, 2818.3681640625, 1143.072021484375,
2842.56005859375, 1137.8880615234375, 2870.2080078125,
1137.8880615234375, 2935.8720703125, 1158.6240234375,
2953.152099609375, 1168.991943359375, 2970.43212890625,
1174.176025390625, 2987.712158203125, 1184.5439453125, 3004.9921875,
1189.72802734375, 3022.27197265625, 1194.9119873046875,
3039.552001953125, 1205.280029296875, 3056.83203125, 1215.64794921875,
3074.112060546875, 1220.83203125, 3094.84814453125,
1226.0159912109375, 3132.864013671875, 1236.384033203125,
3150.14404296875, 1251.93603515625, 3170.880126953125,
1272.6719970703125, 3195.072021484375, 1283.0400390625,
3215.80810546875, 1288.2239990234375, 3233.088134765625,
1288.2239990234375, 3257.280029296875, 1272.6719970703125,
3267.648193359375, 1236.384033203125, 3267.648193359375,
1215.64794921875, 3260.736083984375, 1200.095947265625,
3253.823974609375, 1184.5439453125, 3246.912109375,
1163.8079833984375, 3240.0, 1148.2559814453125, 3233.088134765625,
1132.7039794921875, 3222.719970703125, 1111.968017578125,
3208.89599609375, 1091.2320556640625, 3195.072021484375,
1060.1280517578125, 3167.424072265625, 1034.2080078125,
3146.68798828125, 1013.4719848632812, 3129.408203125,
992.7360229492188, 3112.128173828125, 972.0, 3098.30419921875,
951.2639770507812, 3084.47998046875, 930.5280151367188,
3067.2001953125, 909.7919921875, 3049.920166015625, 883.8720092773438,
3025.72802734375, 868.3200073242188, 3004.9921875, 873.5040283203125,
2980.800048828125, 889.0560302734375, 2960.06396484375,
909.7919921875, 2939.328125, 935.7119750976562, 2911.68017578125,
956.447998046875, 2894.400146484375, 972.0, 2887.488037109375,
992.7360229492188, 2873.6640625, 997.9199829101562, 2849.47216796875,
997.9199829101562, 2769.984130859375, 982.3679809570312,
2749.248046875, 977.1840209960938, 2731.968017578125,
982.3679809570312, 2704.320068359375, 987.552001953125,
2687.0400390625, 992.7360229492188, 2669.760009765625,
997.9199829101562, 2645.568115234375, 1003.10400390625,
2614.464111328125, 1008.2880249023438, 2593.72802734375,
1013.4719848632812, 2572.9921875, 1018.656005859375,
2552.256103515625, 1023.8400268554688, 2528.06396484375,
1029.0240478515625, 2503.8720703125, 1034.2080078125,
2483.135986328125, 1039.3919677734375, 2462.400146484375,
1044.5760498046875, 2445.1201171875, 1049.760009765625,
2427.840087890625, 1054.9439697265625, 2410.56005859375,
1060.1280517578125, 2393.280029296875, 1065.31201171875, 2376.0,
1070.4959716796875, 2358.719970703125, 1075.6800537109375,
2341.43994140625, 1080.864013671875, 2324.16015625,
1091.2320556640625, 2306.880126953125, 1096.416015625,
2289.60009765625, 1096.416015625, 2213.568115234375,
1106.7840576171875, 2192.83203125, 1117.1519775390625,
2175.552001953125, 1122.3360595703125, 2161.72802734375,
1127.52001953125, 2140.991943359375, 1117.1519775390625,
2120.256103515625, 1117.1519775390625, 2089.152099609375,
1122.3360595703125, 2071.8720703125, 1132.7039794921875,
2054.592041015625, 1137.8880615234375, 2037.31201171875,
1137.8880615234375, 2016.5760498046875, 1127.52001953125,
1995.840087890625, 1106.7840576171875, 1978.56005859375,
1086.0479736328125, 1961.280029296875, 1070.4959716796875, 1944.0,
1049.760009765625, 1923.2640380859375, 1034.2080078125,
1902.528076171875, 1023.8400268554688, 1885.248046875,
1013.4719848632812, 1867.968017578125, 1003.10400390625,
1850.68798828125, 992.7360229492188, 1829.9520263671875,
977.1840209960938, 1809.216064453125, 972.0, 1795.39208984375,
966.8159790039062, 1774.656005859375, 966.8159790039062,
1715.904052734375, 972.0, 1691.7120361328125, 977.1840209960938,
1674.4320068359375, 982.3679809570312, 1657.152099609375,
992.7360229492188, 1639.8720703125, 997.9199829101562,
1619.135986328125, 1003.10400390625, 1601.8560791015625,
1023.8400268554688, 1581.1199951171875, 1044.5760498046875,
1567.2960205078125, 1054.9439697265625, 1546.56005859375,
1060.1280517578125, 1529.280029296875, 1070.4959716796875, 1512.0,
1080.864013671875, 1494.7200927734375, 1086.0479736328125,
1473.9840087890625, 1080.864013671875, 1439.424072265625,
1075.6800537109375, 1422.14404296875, 1065.31201171875,
1404.864013671875, 1060.1280517578125, 1356.47998046875,
1054.9439697265625, 1335.7440185546875, 1049.760009765625,
1318.4639892578125, 1044.5760498046875, 1301.18408203125,
1034.2080078125, 1283.904052734375, 1029.0240478515625,
1266.6240234375, 1023.8400268554688, 1245.8880615234375,
1018.656005859375, 1225.1519775390625, 1013.4719848632812,
1204.416015625, 1013.4719848632812, 1152.5760498046875,
1018.656005859375, 1128.384033203125, 1023.8400268554688,
1107.6480712890625, 1029.0240478515625, 1080.0, 1034.2080078125,
1059.2640380859375, 1039.3919677734375, 1041.9840087890625,
1044.5760498046875, 1024.7039794921875, 1049.760009765625,
1007.4240112304688, 1054.9439697265625, 990.14404296875,
1060.1280517578125, 972.864013671875, 1070.4959716796875,
955.5840454101562, 1075.6800537109375, 938.3040161132812,
1080.864013671875, 921.0240478515625, 1091.2320556640625,
903.7440185546875, 1101.5999755859375, 886.4640502929688,
1111.968017578125, 869.1840209960938, 1117.1519775390625,
851.9039916992188, 1111.968017578125, 827.7120361328125,
1117.1519775390625, 810.4320068359375, 1137.8880615234375,
793.1520385742188, 1158.6240234375, 786.239990234375,
1179.3599853515625, 772.416015625, 1179.3599853515625,
744.7680053710938, 1163.8079833984375, 734.4000244140625,
1137.8880615234375, 730.9440307617188, 1122.3360595703125,
720.5759887695312, 1111.968017578125, 703.2960205078125,
1106.7840576171875, 686.0159912109375, 1101.5999755859375,
668.7360229492188, 1096.416015625, 648.0, 1091.2320556640625,
630.7200317382812, 1086.0479736328125, 609.9840087890625,
1091.2320556640625, 575.4240112304688, 1096.416015625,
558.14404296875, 1096.416015625, 530.4960327148438,
1086.0479736328125, 513.2160034179688, 1080.864013671875,
495.9360046386719, 1086.0479736328125, 468.28802490234375,
1091.2320556640625, 451.0080261230469, 1096.416015625,
423.3600158691406, 1101.5999755859375, 402.6240234375,
1111.968017578125, 381.88800048828125, 1122.3360595703125,
364.6080017089844]]], 'labels': ['']}}
```

```
from PIL import Image, ImageDraw, ImageFont
import random
import numpy as np
colormap =
['blue','orange','green','purple','brown','pink','gray','olive','cyan'
,'red',

'lime','indigo','violet','aqua','magenta','coral','gold','tan','skyblu
e']
def draw_polygons(image, prediction, fill_mask=False):
    """
    Draws segmentation masks with polygons on an image.

    Parameters:
    - image_path: Path to the image file.
    - prediction: Dictionary containing 'polygons' and 'labels' keys.
                  'polygons' is a list of lists, each containing
vertices of a polygon.
                  'labels' is a list of labels corresponding to each
polygon.
    - fill_mask: Boolean indicating whether to fill the polygons with
color.
    """
    # Load the image

    draw = ImageDraw.Draw(image)


    # Set up scale factor if needed (use 1 if not scaling)
    scale = 1

    # Iterate over polygons and labels
    for polygons, label in zip(prediction['polygons'], 
prediction['labels']):
        color = random.choice(colormap)
        fill_color = random.choice(colormap) if fill_mask else None

        for _polygon in polygons:
            _polygon = np.array(_polygon).reshape(-1, 2)
            if len(_polygon) < 3:
                print('Invalid polygon:', _polygon)
                continue

            _polygon = (_polygon * scale).reshape(-1).tolist()

            # Draw the polygon
            if fill_mask:
                draw.polygon(_polygon, outline=color, fill=fill_color)
            else:
                draw.polygon(_polygon, outline=color)

            # Draw the label text
            draw.text((_polygon[0] + 8, _polygon[1] + 2), label, 
fill=color)

    # Save or display the image
    #image.show()  # Display the image
    display(image)

output_image = copy.deepcopy(image)
draw_polygons(output_image, 
results['<REFERRING_EXPRESSION_SEGMENTATION>'], fill_mask=True)
```



region to segmentation

with additional region as inputs, format is '<loc_x1><loc_y1><loc_x2><loc_y2>', [x1, y1, x2, y2] is the quantized corrdinates in [0, 999].

```
task_prompt = '<REGION_TO_SEGMENTATION>'
results = run_example(task_prompt,
text_input="<loc_702><loc_575><loc_866><loc_772>")
print(results)
```

```
{'<REGION_TO_SEGMENTATION>': {'polygons': [[[3636.575927734375,
1961.280029296875, 3641.760009765625, 1957.8240966796875,
3646.944091796875, 1957.8240966796875, 3652.1279296875,
1954.3680419921875, 3657.31201171875, 1954.3680419921875,
3662.49609375, 1954.3680419921875, 3667.679931640625,
1957.8240966796875, 3672.864013671875, 1957.8240966796875,
3678.048095703125, 1957.8240966796875, 3683.23193359375,
1961.280029296875, 3688.416015625, 1961.280029296875,
3693.60009765625, 1964.736083984375, 3698.783935546875,
1964.736083984375, 3703.968017578125, 1968.1920166015625,
3709.152099609375, 1968.1920166015625, 3714.3359375,
1968.1920166015625, 3719.52001953125, 1971.6480712890625,
3724.7041015625, 1971.6480712890625, 3729.887939453125,
1971.6480712890625, 3735.072021484375, 1975.10400390625,
3740.256103515625, 1975.10400390625, 3745.43994140625,
1975.10400390625, 3750.6240234375, 1978.56005859375, 3755.80810546875,
1978.56005859375, 3760.991943359375, 1978.56005859375,
3766.176025390625, 1982.0159912109375, 3771.360107421875,
1982.0159912109375, 3776.5439453125, 1985.4720458984375,
3781.72802734375, 1985.4720458984375, 3786.912109375,
1985.4720458984375, 3792.095947265625, 1988.9281005859375,
3797.280029296875, 1988.9281005859375, 3802.464111328125,
1988.9281005859375, 3807.64794921875, 1992.384033203125,
3812.83203125, 1992.384033203125, 3818.01611328125, 1992.384033203125,
3823.199951171875, 1995.840087890625, 3828.384033203125,
1995.840087890625, 3833.568115234375, 1999.2960205078125,
3838.751953125, 1999.2960205078125, 3843.93603515625,
1999.2960205078125, 3849.1201171875, 1999.2960205078125,
3854.303955078125, 2002.7520751953125, 3854.303955078125,
2002.7520751953125, 3859.488037109375, 2006.2080078125,
3864.672119140625, 2006.2080078125, 3869.85595703125, 2006.2080078125,
3875.0400390625, 2009.6640625, 3880.22412109375, 2009.6640625,
3885.407958984375, 2009.6640625, 3890.592041015625,
2013.1199951171875, 3890.592041015625, 2013.1199951171875,
3895.776123046875, 2016.5760498046875, 3900.9599609375,
2016.5760498046875, 3906.14404296875, 2016.5760498046875, 3911.328125,
2016.5760498046875, 3916.511962890625, 2020.0321044921875,
3921.696044921875, 2020.0321044921875, 3926.880126953125,
2020.0321044921875, 3932.06396484375, 2023.488037109375,
3937.248046875, 2023.488037109375, 3942.43212890625,
2026.944091796875, 3947.615966796875, 2026.944091796875,
3952.800048828125, 2026.944091796875, 3957.984130859375,
2030.4000244140625, 3963.16796875, 2030.4000244140625,
3968.35205078125, 2030.4000244140625, 3973.5361328125,
2033.8560791015625, 3978.719970703125, 2033.8560791015625,
3983.904052734375, 2037.31201171875, 3989.087890625, 2037.31201171875,
3994.27197265625, 2037.31201171875, 3999.4560546875, 2040.76806640625,
4004.639892578125, 2040.76806640625, 4009.823974609375,
2040.76806640625, 4015.008056640625, 2044.2239990234375,
4020.19189453125, 2044.2239990234375, 4025.3759765625,
2044.2239990234375, 4030.56005859375, 2047.6800537109375,
4035.743896484375, 2047.6800537109375, 4040.927978515625,
2051.135986328125, 4046.112060546875, 2051.135986328125,
4051.2958984375, 2051.135986328125, 4056.47998046875,
2054.592041015625, 4061.6640625, 2054.592041015625, 4066.847900390625,
2054.592041015625, 4072.031982421875, 2058.048095703125,
4077.216064453125, 2058.048095703125, 4082.39990234375,
2058.048095703125, 4087.583984375, 2061.504150390625,
4092.76806640625, 2061.504150390625, 4097.9521484375,
2061.504150390625, 4103.13623046875, 2064.9599609375,
4103.13623046875, 2064.9599609375, 4108.31982421875, 2068.416015625,
4113.50390625, 2068.416015625, 4118.68798828125, 2068.416015625,
4123.8720703125, 2071.8720703125, 4129.05615234375, 2071.8720703125,
4134.240234375, 2071.8720703125, 4139.423828125, 2075.328125,
4144.60791015625, 2075.328125, 4149.7919921875, 2075.328125,
4154.97607421875, 2078.783935546875, 4160.16015625, 2078.783935546875,
4165.34423828125, 2078.783935546875, 4170.52783203125,
2082.239990234375, 4175.7119140625, 2082.239990234375,
4180.89599609375, 2082.239990234375, 4186.080078125,
2085.696044921875, 4191.26416015625, 2085.696044921875,
4196.4482421875, 2085.696044921875, 4201.6318359375,
2089.152099609375, 4206.81591796875, 2089.152099609375, 4212.0,
2089.152099609375, 4217.18408203125, 2089.152099609375,
4222.3681640625, 2089.152099609375, 4227.55224609375,
2092.608154296875, 4232.73583984375, 2092.608154296875,
4237.919921875, 2096.06396484375, 4243.10400390625, 2096.06396484375,
4248.2880859375, 2096.06396484375, 4253.47216796875, 2099.52001953125,
4258.65625, 2099.52001953125, 4263.83984375, 2099.52001953125,
4269.02392578125, 2102.97607421875, 4274.2080078125, 2102.97607421875,
4279.39208984375, 2102.97607421875, 4284.576171875, 2106.43212890625,
4289.76025390625, 2106.43212890625, 4294.94384765625,
2106.43212890625, 4300.1279296875, 2109.887939453125,
4305.31201171875, 2109.887939453125, 4310.49609375, 2109.887939453125,
4315.68017578125, 2113.343994140625, 4315.68017578125,
2113.343994140625, 4320.86376953125, 2113.343994140625,
4320.86376953125, 2116.800048828125, 4326.0478515625,
2116.800048828125, 4331.23193359375, 2116.800048828125,
4336.416015625, 2120.256103515625, 4341.60009765625,
2120.256103515625, 4346.7841796875, 2120.256103515625,
4351.9677734375, 2123.712158203125, 4357.15185546875,
2123.712158203125, 4362.3359375, 2123.712158203125, 4367.52001953125,
2127.16796875, 4372.7041015625, 2127.16796875, 4377.88818359375,
2127.16796875, 4383.07177734375, 2130.6240234375, 4388.255859375,
2130.6240234375, 4393.43994140625, 2130.6240234375, 4398.6240234375,
2134.080078125, 4403.80810546875, 2134.080078125, 4408.9921875,
2134.080078125, 4414.17578125, 2137.5361328125, 4419.35986328125,
2137.5361328125, 4424.5439453125, 2137.5361328125, 4429.72802734375,
2140.991943359375, 4429.72802734375, 2140.991943359375,
4434.912109375, 2140.991943359375, 4440.09619140625,
2144.447998046875, 4445.27978515625, 2144.447998046875,
4450.4638671875, 2144.447998046875, 4455.64794921875,
2147.904052734375, 4455.64794921875, 2147.904052734375, 4460.83203125,
2147.904052734375, 4466.01611328125, 2151.360107421875,
4466.01611328125, 2151.360107421875, 4471.2001953125,
2151.360107421875, 4476.3837890625, 2154.816162109375,
4476.3837890625, 2154.816162109375, 4481.56787109375,
2154.816162109375, 4486.751953125, 2158.27197265625, 4486.751953125,
2158.27197265625, 4491.93603515625, 2158.27197265625, 4497.1201171875,
2161.72802734375, 4497.1201171875, 2161.72802734375, 4502.30419921875,
2161.72802734375, 4507.48779296875, 2161.72802734375, 4512.671875,
2165.18408203125, 4512.671875, 2165.18408203125, 4517.85595703125,
2165.18408203125, 4523.0400390625, 2168.64013671875, 4523.0400390625,
2168.64013671875, 4528.22412109375, 2168.64013671875, 4533.408203125,
2172.095947265625, 4533.408203125, 2172.095947265625, 4538.591796875,
2175.552001953125, 4538.591796875, 2175.552001953125,
4543.77587890625, 2179.008056640625, 4543.77587890625,
2179.008056640625, 4548.9599609375, 2179.008056640625,
4554.14404296875, 2182.464111328125, 4554.14404296875,
2182.464111328125, 4559.328125, 2182.464111328125, 4564.51220703125,
2182.464111328125, 4569.69580078125, 2185.920166015625,
4569.69580078125, 2185.920166015625, 4574.8798828125, 2189.3759765625,
4574.8798828125, 2189.3759765625, 4580.06396484375, 2192.83203125,
4580.06396484375, 2192.83203125, 4585.248046875, 2196.2880859375,
4585.248046875, 2196.2880859375, 4590.43212890625, 2199.744140625,
4590.43212890625, 2199.744140625, 4595.6162109375, 2199.744140625,
4600.7998046875, 2203.199951171875, 4600.7998046875,
2203.199951171875, 4605.98388671875, 2206.656005859375,
4605.98388671875, 2206.656005859375, 4611.16796875, 2210.112060546875,
4611.16796875, 2210.112060546875, 4616.35205078125, 2213.568115234375,
4616.35205078125, 2213.568115234375, 4621.5361328125,
2217.024169921875, 4621.5361328125, 2217.024169921875,
4626.72021484375, 2220.47998046875, 4626.72021484375,
2220.47998046875, 4631.90380859375, 2223.93603515625,
4631.90380859375, 2223.93603515625, 4637.087890625, 2227.39208984375,
4637.087890625, 2227.39208984375, 4642.27197265625, 2230.84814453125,
4642.27197265625, 2230.84814453125, 4647.4560546875,
2234.303955078125, 4647.4560546875, 2234.303955078125,
4652.64013671875, 2237.760009765625, 4652.64013671875,
2237.760009765625, 4657.82421875, 2241.216064453125, 4657.82421875,
2241.216064453125, 4663.0078125, 2244.672119140625, 4663.0078125,
2244.672119140625, 4668.19189453125, 2248.128173828125,
4668.19189453125, 2248.128173828125, 4673.3759765625, 2251.583984375,
4673.3759765625, 2251.583984375, 4678.56005859375, 2255.0400390625,
4678.56005859375, 2255.0400390625, 4683.744140625, 2258.49609375,
4683.744140625, 2258.49609375, 4688.92822265625, 2261.9521484375,
4688.92822265625, 2261.9521484375, 4694.11181640625,
2265.407958984375, 4694.11181640625, 2265.407958984375,
4699.2958984375, 2268.864013671875, 4699.2958984375,
2268.864013671875, 4704.47998046875, 2272.320068359375,
4704.47998046875, 2272.320068359375, 4709.6640625, 2275.776123046875,
4709.6640625, 2275.776123046875, 4714.84814453125, 2279.232177734375,
4714.84814453125, 2282.68798828125, 4714.84814453125,
2282.68798828125, 4720.0322265625, 2286.14404296875, 4720.0322265625,
2286.14404296875, 4725.2158203125, 2289.60009765625, 4725.2158203125,
2289.60009765625, 4730.39990234375, 2293.05615234375,
4730.39990234375, 2293.05615234375, 4735.583984375, 2296.511962890625,
4735.583984375, 2296.511962890625, 4740.76806640625,
2299.968017578125, 4740.76806640625, 2299.968017578125,
4745.9521484375, 2303.424072265625, 4745.9521484375,
2303.424072265625, 4751.13623046875, 2306.880126953125,
4751.13623046875, 2306.880126953125, 4756.31982421875,
2310.336181640625, 4756.31982421875, 2310.336181640625, 4761.50390625,
2313.7919921875, 4761.50390625, 2313.7919921875, 4766.68798828125,
2317.248046875, 4766.68798828125, 2317.248046875, 4771.8720703125,
2320.7041015625, 4771.8720703125, 2320.7041015625, 4777.05615234375,
2324.16015625, 4777.05615234375, 2324.16015625, 4782.240234375,
2327.615966796875, 4782.240234375, 2327.615966796875, 4787.423828125,
2331.072021484375, 4787.423828125, 2331.072021484375,
4792.60791015625, 2334.528076171875, 4792.60791015625,
2334.528076171875, 4797.7919921875, 2337.984130859375,
4797.7919921875, 2337.984130859375, 4802.97607421875,
2341.43994140625, 4802.97607421875, 2341.43994140625, 4808.16015625,
2344.89599609375, 4808.16015625, 2344.89599609375, 4813.34423828125,
2348.35205078125, 4813.34423828125, 2348.35205078125,
4818.52783203125, 2351.80810546875, 4818.52783203125,
2351.80810546875, 4823.7119140625, 2355.26416015625, 4823.7119140625,
2355.26416015625, 4828.89599609375, 2358.719970703125,
4828.89599609375, 2358.719970703125, 4834.080078125,
2362.176025390625, 4834.080078125, 2362.176025390625,
4839.26416015625, 2365.632080078125, 4844.4482421875,
2365.632080078125, 4844.4482421875, 2369.088134765625,
4844.4482421875, 2372.5439453125, 4844.4482421875, 2372.5439453125,
4849.6318359375, 2376.0, 4849.6318359375, 2376.0, 4854.81591796875,
2379.4560546875, 4854.81591796875, 2379.4560546875, 4860.0,
2382.912109375, 4860.0, 2382.912109375, 4865.18408203125,
2386.3681640625, 4865.18408203125, 2386.3681640625, 4870.3681640625,
2389.823974609375, 4870.3681640625, 2389.823974609375,
4875.5522460937 5, 2393.280029296875, 4875.5522460937 5,
2393.280029296875, 4880.73583984375, 2396.736083984375,
4880.73583984375, 2396.736083984375, 4885.919921875,
2396.736083984375, 4885.919921875, 2400.192138671875, 4885.919921875,
2400.192138671875, 4891.10400390625, 2400.192138671875,
4891.10400390625, 2403.64794921875, 4891.10400390625,
2403.64794921875, 4896.2880859375, 2403.64794921875, 4896.2880859375,
2407.10400390625, 4896.2880859375, 2407.10400390625, 4901.47216796875,
2407.10400390625, 4901.47216796875, 2410.56005859375,
4901.47216796875, 2410.56005859375, 4906.65625, 2414.01611328125,
4906.65625, 2414.01611328125, 4911.83984375, 2417.47216796875,
4911.83984375, 2417.47216796875, 4917.02392578125, 2420.927978515625,
4917.02392578125, 2420.927978515625, 4922.2080078125,
2420.927978515625, 4922.2080078125, 2424.384033203125,
4922.2080078125, 2424.384033203125, 4927.39208984375,
2427.840087890625, 4927.39208984375, 2427.840087890625,
4932.576171875, 2431.296142578125, 4932.576171875, 2431.296142578125,
4937.76025390625, 2434.751953125, 4937.76025390625, 2434.751953125,
4942.94384765625, 2438.2080078125, 4942.94384765625, 2438.2080078125,
4948.1279296875, 2441.6640625, 4948.1279296875, 2441.6640625,
4953.31201171875, 2445.1201171875, 4953.31201171875, 2445.1201171875,
4958.49609375, 2445.1201171875, 4958.49609375, 2448.576171875,
4958.49609375, 2448.576171875, 4963.68017578125, 2448.576171875,
4963.68017578125, 2452.031982421875, 4963.68017578125,
2452.031982421875, 4968.8642578125, 2455.488037109375,
4968.8642578125, 2455.488037109375, 4974.0478515625,
2458.944091796875, 4974.0478515625, 2458.944091796875,
4979.23193359375, 2462.400146484375, 4979.23193359375,
2462.400146484375, 4984.416015625, 2465.85595703125, 4984.416015625,
2465.85595703125, 4989.60009765625, 2469.31201171875,
4989.60009765625, 2469.31201171875, 4994.7841796875, 2472.7680664062 5,
4994.7841796875, 2472.76806640625, 4999.9677734375, 2476.22412109375,
4999.9677734375, 2476.22412109375, 5005.1518554687 5, 2476.22412109375,
5005.15185546875, 2479.68017578125, 5005.15185546875,
2479.68017578125, 5010.3359375, 2483.135986328125, 5010.3359375,
2483.135986328125, 5015.52001953125, 2486.592041015625,
5015.52001953125, 2486.592041015625, 5020.7041015625,
2490.048095703125, 5020.7041015625, 2490.048095703125,
5025.88818359375, 2490.048095703125, 5031.07177734375,
2490.048095703125, 5031.07177734375, 2493.504150390625,
5031.07177734375, 2493.504150390625, 5036.255859375, 2496.9599609375,
5036.255859375, 2496.9599609375, 5041.43994140625, 2500.416015625,
5041.43994140625, 2500.416015625, 5046.6240234375, 2503.8720703125,
5046.6240234375, 2503.8720703125, 5051.80810546875, 2507.328125,
5051.80810546875, 2507.328125, 5056.9921875, 2510.7841796875,
5056.9921875, 2510.7841796875, 5062.17578125, 2514.239990234375,
5062.17578125, 2514.239990234375, 5067.35986328125, 2517.696044921875,
5067.35986328125, 2517.696044921875, 5072.5439453125,
2521.152099609375, 5072.5439453125, 2521.152099609375,
5077.72802734375, 2521.152099609375, 5077.72802734375,
2524.608154296875, 5077.72802734375, 2524.608154296875,
5082.912109375, 2528.06396484375, 5082.912109375, 2528.06396484375,
5088.09619140625, 2528.06396484375, 5088.09619140625,
2531.52001953125, 5088.09619140625, 2531.52001953125,
5093.27978515625, 2534.97607421875, 5093.27978515625,
2534.97607421875, 5098.4638671875, 2538.43212890625, 5098.4638671875,
2538.43212890625, 5103.64794921875, 2541.88818359375,
5103.64794921875, 2541.88818359375, 5108.83203125, 2541.88818359375,
5108.83203125, 2545.343994140625, 5108.83203125, 2545.343994140625,
5114.01611328125, 2545.343994140625, 5114.01611328125,
2548.800048828125, 5114.01611328125, 2548.800048828125,
5119.2001953125, 2552.256103515625, 5119.2001953125,
2552.256103515625, 5124.3837890625, 2555.712158203125,
5124.3837890625, 2555.712158203125, 5129.56787109375, 2559.16796875,
5129.56787109375, 2559.16796875, 5134.751953125, 2559.16796875,
5134.751953125, 2562.6240234375, 5134.751953125, 2562.6240234375,
5139.93603515625, 2566.080078125, 5139.93603515625, 2566.080078125,
5145.1201171875, 2569.5361328125, 5145.1201171875, 2569.5361328125,
5150.30419921875, 2569.5361328125, 5150.30419921875, 2572.9921875,
5150.30419921875, 2572.9921875, 5155.48779296875, 2576.447998046875,
5155.48779296875, 2576.447998046875, 5160.671875, 2576.447998046875,
5160.671875, 2579.904052734375, 5160.671875, 2579.904052734375,
5165.85595703125, 2583.360107421875, 5165.85595703125,
2583.360107421875, 5171.0400390625, 2586.816162109375,
5171.0400390625, 2586.816162109375, 5176.22412109375,
2590.27197265625, 5176.22412109375, 2590.27197265625, 5181.408203125,
2614.464111328125, 5181.408203125, 2614.464111328125,
5176.22412109375, 2614.464111328125, 5176.22412109375,
2617.920166015625, 5176.22412109375, 2617.920166015625,
5171.0400390625, 2617.920166015625, 5171.0400390625, 2621.3759765625,
5171.0400390625, 2621.3759765625, 5165.85595703125, 2621.3759765625,
5165.85595703125, 2624.83203125, 5165.85595703125, 2624.83203125,
5160.671875, 2624.83203125, 5160.671875, 2628.2880859375, 5160.671875,
2628.2880859375, 5155.48779296875, 2628.2880859375, 5155.48779296875,
2631.744140625, 4942.94384765625, 2628.2880859375, 4942.94384765625,
2628.2880859375, 4937.76025390625, 2631.744140625, 4937.76025390625,
2631.744140625, 4932.576171875, 2628.2880859375, 4932.576171875,
2628.2880859375, 4927.39208984375, 2628.2880859375, 4927.39208984375,
2631.744140625, 4927.39208984375, 2631.744140625, 4922.2080078125,
2631.744140625, 4922.2080078125, 2490.048095703125, 4922.2080078125,
2490.048095703125, 4917.02392578125, 2490.048095703125, 4911.83984375,
2490.048095703125, 4911.83984375, 2493.504150390625, 4911.83984375,
2493.504150390625, 4906.65625, 2496.9599609375, 4906.65625,
2496.9599609375, 4901.47216796875, 2496.9599609375, 4901.47216796875,
2500.416015625, 4896.2880859375, 2500.416015625, 4891.10400390625,
2500.416015625, 4891.10400390625, 2503.8720703125, 4891.10400390625,
2503.8720703125, 4885.919921875, 2507.328125, 4885.919921875,
2507.328125, 4906.65625, 2514.239990234375, 4906.65625,
2514.239990234375, 4901.47216796875, 2514.239990234375,
4896.2880859375, 2514.239990234375, 4891.10400390625,
2514.239990234375, 4885.919921875, 2514.239990234375, 4885.919921875,
2510.7841796875, 4885.919921875, 2510.7841796875, 4896.2880859375,
2528.06396484375, 4896.2880859375, 2528.06396484375, 4891.10400390625,
2528.06396484375, 4885.919921875, 2528.06396484375, 4880.73583984375,
2528.06396484375, 4875.55224609375, 2528.06396484375, 4896.2880859375,
2496.9599609375, 4896.2880859375, 2496.9599609375, 4891.10400390625,
2500.416015625, 4885.919921875, 2517.696044921875, 4896.2880859375,
2490.048095703125]]], 'labels': ['']}}

output_image = copy.deepcopy(image)
draw_polygons(output_image, results['<REGION_TO_SEGMENTATION>'],
fill_mask=True)
```

Open vocabulary detection

open vocabulary detection can detect both objects and ocr texts.

results format:

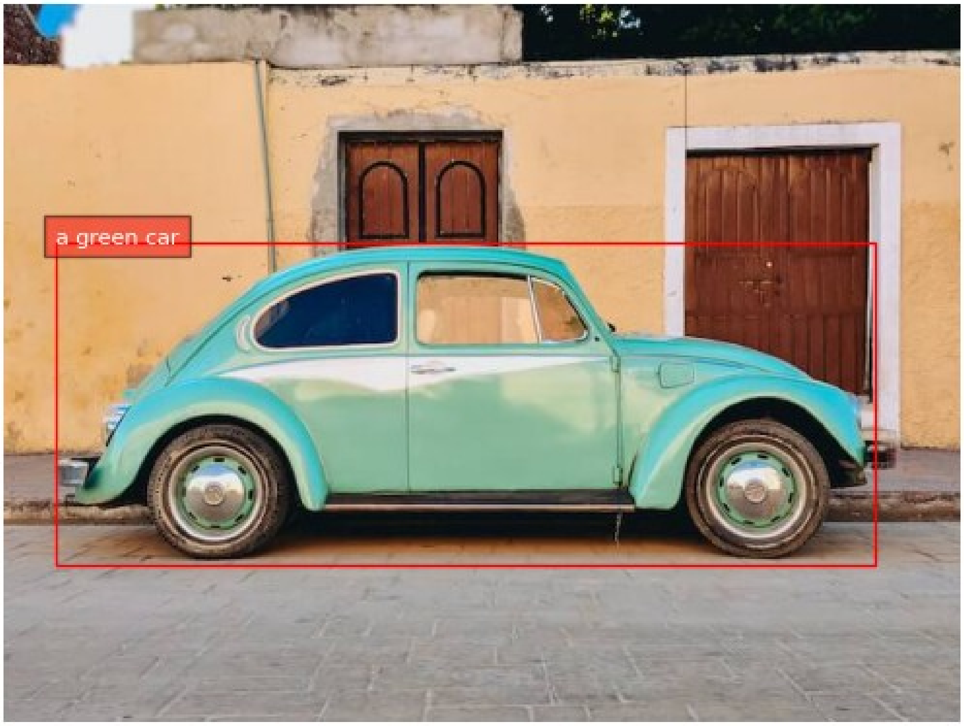
{ '<OPEN_VOCABULARY_DETECTION>': {'bboxes': [[x1, y1, x2, y2], [x1, y1, x2, y2], ...]], 'bboxes_labels': ['label_1', 'label_2', ..], 'polygons': [[[x1, y1, x2, y2, ..., xn, yn], [x1, y1, ..., xn, yn]], ...], 'polygons_labels': ['label_1', 'label_2', ...] }}

```
task_prompt = '<OPEN_VOCABULARY_DETECTION>'
results = run_example(task_prompt, text_input="a green car")
print(results)
```

```
{'<OPEN_VOCABULARY_DETECTION>': {'bboxes': [[34.23999786376953,
158.63999938964844, 582.0800170898438, 374.1600036621094]],
'bboxes_labels': ['a green car'], 'polygons': [], 'polygons_labels':
[]}}
```

```
def convert_to_od_format(data):
    """
    Converts a dictionary with 'bboxes' and 'bboxes_labels' into a
dictionary with separate 'bboxes' and 'labels' keys.

    Parameters:
```

```
    - data: The input dictionary with 'bboxes', 'bboxes_labels',
'polygons', and 'polygons_labels' keys.

    Returns:
    - A dictionary with 'bboxes' and 'labels' keys formatted for
object detection results.
    """
    # Extract bounding boxes and labels
    bboxes = data.get('bboxes', [])
    labels = data.get('bboxes_labels', [])

    # Construct the output format
    od_results = {
        'bboxes': bboxes,
        'labels': labels
    }

    return od_results

bbox_results  =
convert_to_od_format(results['<OPEN_VOCABULARY_DETECTION>'])

plot_bbox(image, bbox_results)
```



region to texts

```
task_prompt = '<REGION_TO_CATEGORY>'
results = run_example(task_prompt,
text_input="<loc_52><loc_332><loc_932><loc_774>")
print(results)
```

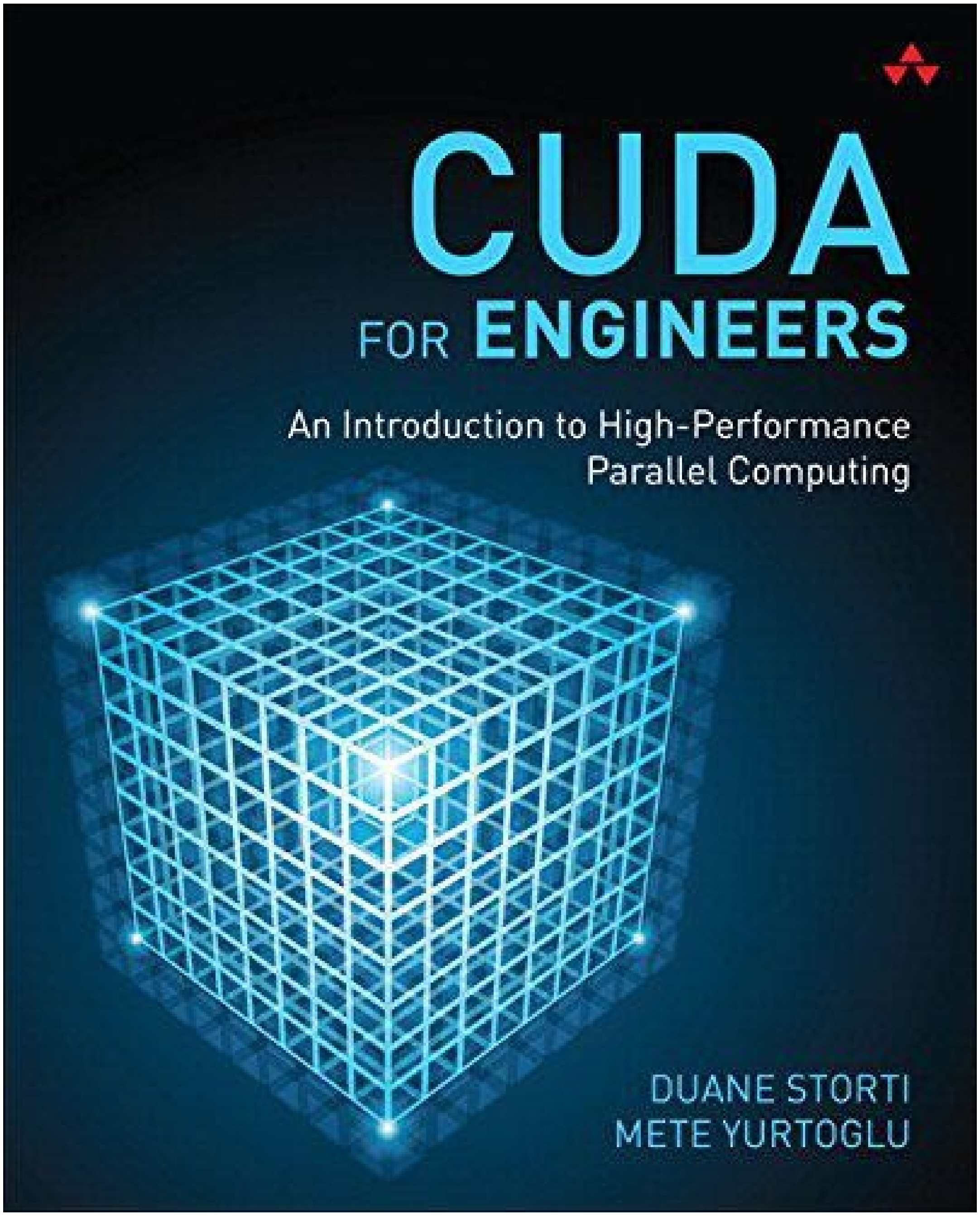
```
{'<REGION_TO_CATEGORY>': 'car<loc_52><loc_332><loc_932><loc_774>'}
```

```
task_prompt = '<REGION_TO_DESCRIPTION>'
results = run_example(task_prompt,
text_input="<loc_52><loc_332><loc_932><loc_774>")
print(results)
```

```
{'<REGION_TO_DESCRIPTION>': 'turquoise Volkswagen
Beetle<loc_52><loc_332><loc_932><loc_774>'}
```

ocr related tasks

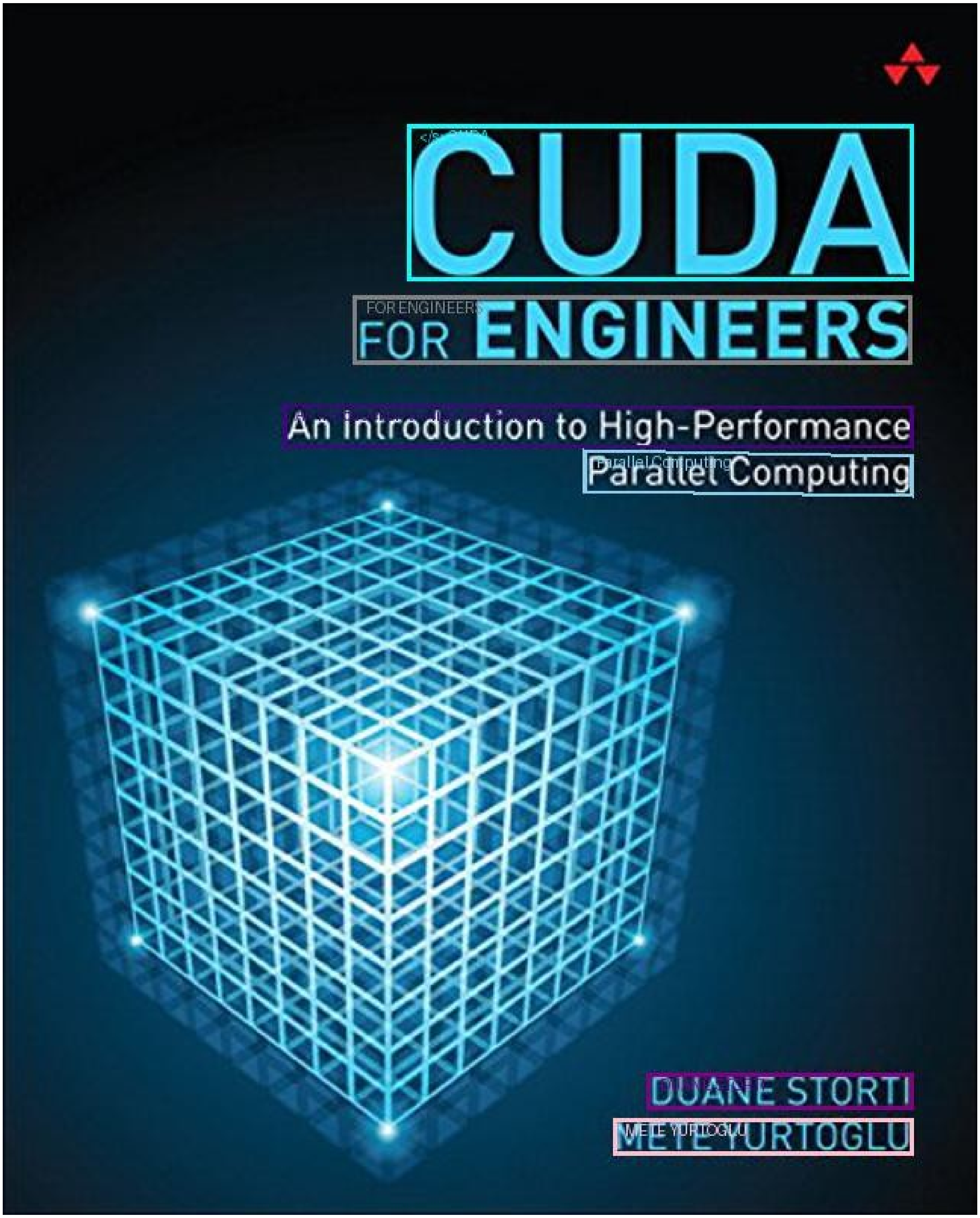
```
url = "http://ecx.images-amazon.com/images/I/51UUzBDAMsL.jpg?
download=true"
image = Image.open(requests.get(url, stream=True).raw).convert('RGB')

image
```



```python
task_prompt = '<OCR>'
run_example(task_prompt)
```

```
{'<OCR>': 'CUDA\nFOR ENGINEERS\nAn Introduction to High-Performance\
nParallel Computing\nDUANE STORTI\nMETE YURTOGLU\n'}
```

```python
task_prompt = '<OCR_WITH_REGION>'
results = run_example(task_prompt)
print(results)
# ocr results format
# {'OCR_WITH_REGION': {'quad_boxes': [[x1, y1, x2, y2, x3, y3, x4,
y4], ...], 'labels': ['text1', ...]}}

{'<OCR_WITH_REGION>': {'quad_boxes': [[167.0435028076172, 50.25,
376.20050048828125, 50.25, 376.20050048828125, 114.75,
167.0435028076172, 114.75], [144.8784942626953, 120.75,
375.3945007324219, 120.75, 375.3945007324219, 149.25,
144.8784942626953, 149.25], [115.86249542236328, 165.25,
376.20050048828125, 166.25, 376.20050048828125, 184.25,
115.86249542236328, 183.25], [239.9864959716797, 184.25,
376.20050048828125, 186.25, 376.20050048828125, 204.25,
239.9864959716797, 202.25], [266.1814880371094, 441.25,
376.20050048828125, 441.25, 376.20050048828125, 456.25,
266.1814880371094, 456.25], [252.0764923095703, 460.25,
376.20050048828125, 460.25, 376.20050048828125, 475.25,
252.0764923095703, 475.25]], 'labels': ['</s>CUDA', 'FOR ENGINEERS',
'An Introduction to High-Performance', 'Parallel Computing', 'DUANE
STORTI', 'METE YURTOGLU']}}

def draw_ocr_bboxes(image, prediction, scale=1):
    draw = ImageDraw.Draw(image)
    bboxes, labels = prediction['quad_boxes'], prediction['labels']
    for box, label in zip(bboxes, labels):
        color = random.choice(colormap)
        new_box = (np.array(box) * scale).tolist()
        draw.polygon(new_box, width=3, outline=color)
        draw.text((new_box[0]+8, new_box[1]+2),
                    "{}".format(label),
                    align="right",

                    fill=color)

    display(image)

output_image = copy.deepcopy(image)
w, h = output_image.size
scale = 800 / max(w, h)
new_output_image = output_image.resize((int(w * scale), int(h *
scale)))
draw_ocr_bboxes(new_output_image, results['<OCR_WITH_REGION>'],
scale=scale)
```



CUDA FOR ENGINEERS

An Introduction to High-Performance Parallel Computing

DUANE STORTI
METE YURTOGLU

Cascaded tasks

Caption + Phrase Grounding

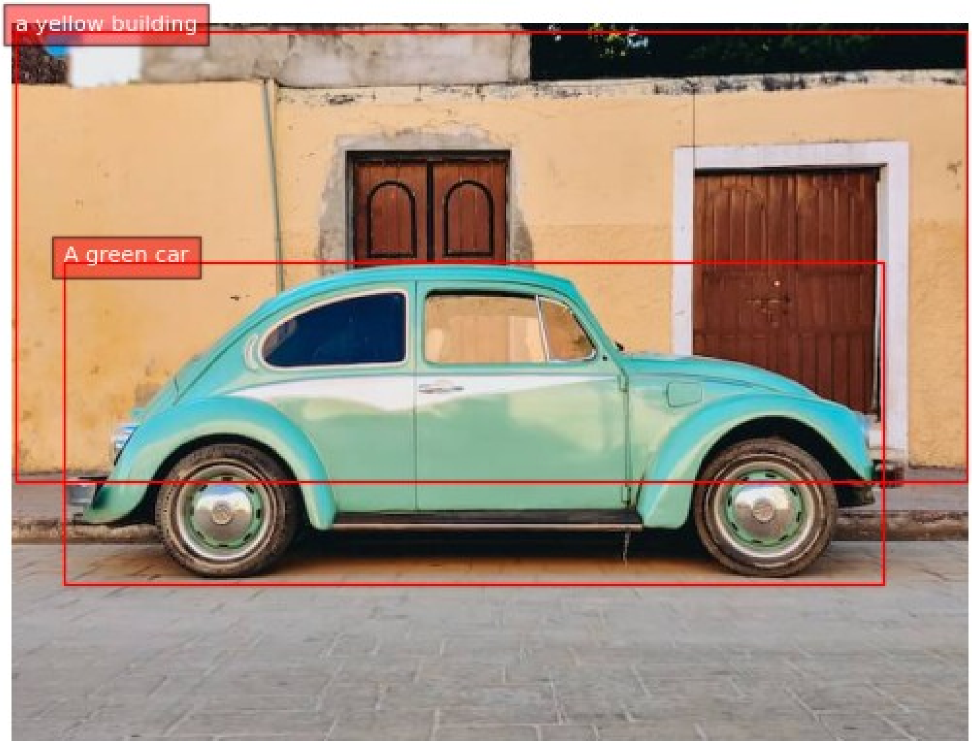
results format:

{ '<CAPTION': pure_text, {'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[x1, y1, x2, y2], ...], 'labels': ['', '', ...]}} }

```
url = "https://huggingface.co/datasets/huggingface/documentation-
images/resolve/main/transformers/tasks/car.jpg?download=true"
image = Image.open(requests.get(url, stream=True).raw)

task_prompt = '<CAPTION>'
results = run_example(task_prompt)
text_input = results[task_prompt]
task_prompt = '<CAPTION_TO_PHRASE_GROUNDING>'
results = run_example(task_prompt, text_input)
results['<CAPTION>'] = text_input

results
```

```
{'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[34.23999786376953,
    159.1199951171875,
    582.0800170898438,
    374.6399841308594],
   [1.5999999046325684,
    4.079999923706055,
    639.0399780273438,
    305.03997802734375]],
  'labels': ['A green car', 'a yellow building']},
 '<CAPTION>': 'A green car parked in front of a yellow building.'}
```

```
plot_bbox(image, results['<CAPTION_TO_PHRASE_GROUNDING>'])
```



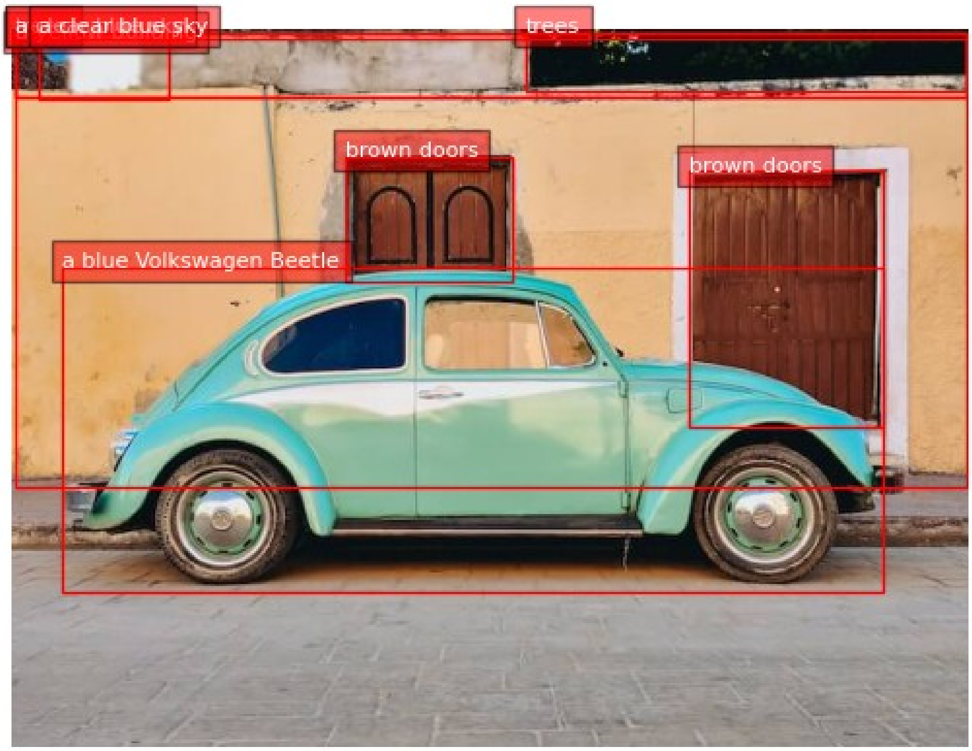
Detailed Caption + Phrase Grounding

results format:

{ '<DETAILED_CAPTION>': pure_text, {'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[x1, y1, x2, y2], ...], 'labels': ['', '', ...]}} }

```
task_prompt = '<DETAILED_CAPTION>'
results = run_example(task_prompt)
text_input = results[task_prompt]
task_prompt = '<CAPTION_TO_PHRASE_GROUNDING>'
results = run_example(task_prompt, text_input)
results['<DETAILED_CAPTION>'] = text_input

results
```

```
{'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[33.599998474121094,
    158.63999938964844,
    582.719970703125,
    375.1199951171875],
   [1.5999999046325684, 5.039999961853027, 639.0399780273438, 306.0],
   [452.79998779296875,
    94.31999969482422,
    582.0800170898438,
    265.67999267578125],
   [222.39999389648438,
    84.23999786376953,
    335.03997802734375,
    167.27999877929688],
   [2.240000009536743,
    1.1999999284744263,
    639.0399780273438,
    43.91999816894531],
   [343.3599853515625,
    1.1999999284744263,
    639.0399780273438,
    40.55999755859375],
   [2.240000009536743,
    1.1999999284744263,
    638.3999633789062,
    43.91999816894531],
   [18.239999771118164,
    1.1999999284744263,
    104.63999938964844,
    45.36000061035156]],
  'labels': ['a blue Volkswagen Beetle',
   'a yellow building',
   'brown doors',
   'brown doors',
   'trees',
   'trees',
   'a clear blue sky',
   'a clear blue sky']]},
 '<DETAILED_CAPTION>': 'The image shows a blue Volkswagen Beetle
parked in front of a yellow building with two brown doors, surrounded
by trees and a clear blue sky.'}

plot_bbox(image, results['<CAPTION_TO_PHRASE_GROUNDING>'])
```



More Detailed Caption + Phrase Grounding

results format:

{ '<MORE_DETAILED_CAPTION': pure_text, {'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[x1, y1, x2, y2], ...], 'labels': ['', '', ...]}} }

```
task_prompt = '<MORE_DETAILED_CAPTION>'
results = run_example(task_prompt)
text_input = results[task_prompt]
task_prompt = '<CAPTION_TO_PHRASE_GROUNDING>'
results = run_example(task_prompt, text_input)
results['<MORE_DETAILED_CAPTION>'] = text_input

results
```

```
{'<CAPTION_TO_PHRASE_GROUNDING>': {'bboxes': [[35.52000045776367,
    157.67999267578125,
    581.4400024414062,
    372.7200012207031],
   [1.5999999046325684,
    340.0799865722656,
    639.0399780273438,
    479.2799987792969],
   [454.0799865722656,
    93.83999633789062,
    580.7999877929688,
    263.2799987792969],
   [223.67999267578125,
    84.23999786376953,
    333.7599792480469,
    163.9199981689453],
   [36.79999923706055,
    157.67999267578125,
    580.1599731445312,
    372.239990234375],
   [454.0799865722656,
    93.83999633789062,
    581.4400024414062,
    263.2799987792969],
   [223.67999267578125,
    84.72000122070312,
    333.7599792480469,
    163.9199981689453],
   [164.16000366210938,
    177.36000061035156,
    265.2799987792969,
    232.0800018310547],
   [2.879999876022339,
    4.559999942779541,
    639.0399780273438,
    301.1999816894531],
   [19.520000457763672, 1.1999999284744263, 639.0399780273438, 42.0],
   [20.15999984741211,
    0.7199999690055847,
    102.07999420166016,
    42.959999084472656],
   [345.91998291015625,
    1.1999999284744263,
    639.0399780273438,
    37.68000030517578]],
  'labels': ['a vintage Volkswagen Beetle car',
   'a cobblestone street',
   'wooden doors',
   'wooden doors',
   'The car',
   'doors',
   'doors',
   'small window',
   'The building',
   'The sky',
   'The sky',
   'trees']},
 '<MORE_DETAILED_CAPTION>': 'The image shows a vintage Volkswagen
Beetle car parked on a cobblestone street in front of a yellow
building with two wooden doors. The car is painted in a bright
```

```
turquoise color and has a sleek, streamlined design. It has two doors
on either side of the car, one on top of the other, and a small window
on the front. The building appears to be old and dilapidated, with
peeling paint and crumbling walls. The sky is blue and there are trees
in the background.'}

plot_bbox(image, results['<CAPTION_TO_PHRASE_GROUNDING>'])
```