

# Benchmark

---

berechnung der inferenz-fps für unterschiedl. anzahl an infer requests

## Verwendung

- In *models\_dir* pfad zu konvertierten openvino models angeben, modelle haben folgende struktur:

```
└─ model_folder
   └─ frozen_inference_graph.bin
   └─ frozen_inference_graph.xml
   └─ exported_model (optional)
```

falls vorhanden, wird *exported\_model* verwendet, ansonsten wird *exec\_model* aus den *.xml* und *.bin* files erzeugt und exportiert.

- In *test\_image* pfad zu inferierender bilddatei angeben
- *iterationen* gibt an wie oft das bild inferiert werden soll

### benchmark.py

starten über

```
python3 benchmark.py
```

- model über nummer auswählen
- anzahl der inferenz requests eingeben

### benchmark\_auto.py

In *models* liste mit zu inferierenden modellen (in *models\_dir* enthalten sind) festlegen.

zB:

```
models = ['ssd_mobilenet_v2',
          'ssd_inception_v2',
          'faster_rcnn_inception_v2']
```

starten über

```
python3 benchmark_auto.py
```

werden dann jeweils für 1 bis 4 inferenz-requests automatisch inferiert.

`infer_async.py`

wird von *benchmark.py* und *benchmark\_auto.py* für die asynchrone inferenz verwendet.