TF-Vision Model Garden

⚠ Disclaimer: All datasets hyperlinked from this page are not owned or distributed by Google. The dataset is made available by third parties. Please review the terms and conditions made available by the third parties before using the data.

Table of Contents

Introduction

TF-Vision modeling library for computer vision provides a collection of baselines and checkpoints for image classification, object detection, and segmentation.

Image Classification

ResNet models trained with vanilla settings

▶ Details

ResNet-RS models trained with various settings

▶ Details

Vision Transformer (ViT)

▶ Details

Object Detection and Instance Segmentation

Common Settings and Notes

▶ Details

COCO Object Detection Baselines

RetinaNet (ImageNet pretrained)

▶ Details

RetinaNet (Trained from scratch)

▶ Details

Mobile-size RetinaNet (Trained from scratch):

▶ Details

Instance Segmentation Baselines

Mask R-CNN (Trained from scratch)

▶ Details

Cascade RCNN-RS (Trained from scratch)

▶ Details

Semantic Segmentation

- We support DeepLabV3 and DeepLabV3+ architectures, with Dilated ResNet backbones.
- Backbones are pre-trained on ImageNet.

PASCAL-VOC

▶ Details

CITYSCAPES

▶ Details

Video Classification

Common Settings and Notes

▶ Details

Kinetics-400 Action Recognition Baselines

▶ Details

Kinetics-600 Action Recognition Baselines

▶ Details