Large Field and High Resolution: Detecting

Needle in Haystack

Experiments - Summary

# Hyperparameters

The models for the experiments were trained using the default hyperparameters for the Mask-RCNN ResNet101 model. Those are detailed on the GitHub repository of the Mask-RCNN project [1]. Only *model\_multI\_mixed* (described below) was trained using a different hyperparameter configuration which differs from the original in the following way:

|  |  |  |
| --- | --- | --- |
| Hyperparameter | Original value | Value used for training |
| STEPS | 60000, 80000 | 180000, 240000 |
| MAX\_ITER | 90000 | 270000 |

# Notation

1. The following is notation which will be referred to in order to address different model types throughout this document. All models are based on Mask-RCNN ResNet101, as described in *Hyperparameters*:

* Single channel model – A model which takes as input a single image of uniform effective resolution. It has been trained almost exclusively on images of constant effective resolution. “Almost exclusively” addresses the fact that FOV of CH2 and CH3 (described in *Datasets*) can sometimes contain images of varying sizes.
* Multi-channel model – A model which takes as input single images of various sizes. It has been trained collectively on all FOV Channels (CH1, CH2 and CH3 – described in Datasets).
* Multi-channel model stacked – A model which takes as input stacked images from all FOVs (CH1, CH2, CH3).

1. The following notation will be used to describe the train/test configuration of a particular model type (*1.*):

* Model Type | Training Set | Testing Set

# Datasets

All datasets were generated to contain 95% of the provided images for training and 5% for validation/testing. All datasets which consist of images sized differently than the original COCO Dataset images have ground truths cropped to the respective size of the dataset images (this includes bounding box adjustments as well as segmentation adjustments).

Several datasets based on the COCO 2017 Dataset [2] were prepared in order to test different input configurations for the various model types.

The generated Datasets used for **training** are as follows:

* CH1 (E.g *Fig. 1*) – Consists of images of fixed size – 60x60 – cropped from the middle of the corresponding full-resolution COCO images.
* CH2 (E.g *Fig. 2*) – Generally consists of images of fixed size – 934x934 – extracted from the middle of the COCO full-resolution image as described in the *Process of extraction*. There are several instances in the CH2 dataset where a particular image is of different size (not 934x934) due to the natural variability of image sizes in the COCO Dataset, hence the word “generally” at the beginning.
  + Process of extraction – images for the CH2 dataset were initially cropped from the middle of the full-resolution COCO images and sampled at a uniform sampling rate, decreasing their effective resolution to X% of the original cropped image.
* CH3 (E.g *Fig. 3*) – Consists of images of variable size, same as the original COCO full-resolution images.
  + Process of extraction – images for the CH3 dataset are duplicates of the full-resolution COCO images, sampled at a uniform sampling rate, decreasing their effective resolution to Y% of the original COCO full-resolution image
* CH1\_interp – Consists of images of fixed size – 60x60 – sampled at a uniform sampling rate from the middle of the full-resolution COCO images. The sampling rate decreases the effective resolution of the region to Z% of the original 60x60 samples.
* Multi – Consists of all CH1, CH2 and CH3 images combined and fed one-by-one into the model
* Multi stacked – Consists of all CH1, CH2 and CH3 images, combined as described in the *Process of extraction*
  + Process of extraction – The corresponding images from each channel have been stacked together along the color-channel axis, effectively resulting in a 9-channel image input (each image individually has 3 color channels – R, G, B). Due to the different sizes of the images in CH1, CH2 and CH3, all images in CH1 and CH2 are padded with 0-s to achieve the corresponding CH3 image size.

All **testing** datasets have been generated with 5% of the total number of images from the COCO dataset. The rest were used for **testing**, as described above.

Generated datasets used for **testing** are as follows:

* CH1\_val – images extracted thorough the same procedure as CH1 images
* CH2\_val – images extracted thorough the same procedure as CH2 images
* CH3\_val – images extracted thorough the same procedure as CH3 images
* CH1\_interp\_val – images extracted thorough the same procedure as CH1\_interp images
* CH1\_val\_0.8 – images extracted thorough the same procedure as CH1 images. The ground truth for the CH1\_val\_0.8 dataset, however, has been modified to contain annotations only for objects with area more than 80% contained within the FOV of each image (Fig. 4 for more information).
* CH1\_val\_0.7 – images extracted thorough the same procedure as CH1 images. The ground truth for the CH1\_val\_0.8 dataset, however, has been modified to contain annotations only for objects with area more than 70% contained within the FOV of each image (Fig. 4 for more information).
* CH1\_val\_0.6 – images extracted thorough the same procedure as CH1 images. The ground truth for the CH1\_val\_0.8 dataset, however, has been modified to contain annotations only for objects with area more than 60% contained within the FOV of each image (Fig. 4 for more information).
* CH2\_val\_0.6 – images extracted thorough the same procedure as CH2 images. The ground truth for the CH2\_val\_0.8 dataset, however, has been modified to contain annotations only for objects with area more than 60% contained within the FOV of each image (Fig. 4 for more information).
* CH3\_val\_0.6 – images extracted thorough the same procedure as CH3 images. The ground truth for the CH3\_val\_0.8 dataset, however, has been modified to contain annotations only for objects with area more than 60% contained within the FOV of each image (Fig. 4 for more information).
* CH1\_val\_0.6\_no\_empt – all images from CH1\_val\_0.6 which contain at least 1 annotation

# References

[1] (Online) Mask-RCNN Project; <https://github.com/facebookresearch/maskrcnn-benchmark/blob/master/configs/e2e_mask_rcnn_R_101_FPN_1x.yaml>

[2] (Online) COCO Dataset; <http://cocodataset.org/>