Description of DOAI2019 Task1 & Task2

In the DOAI2019 competition, we propose a novel multi-category rotation detector for small, cluttered and rotated objects, which is designated to address the following issues: 1) small object: a sampling fusion network (SF-Net) is devised that incorporates feature fusion and finer anchor sampling; 2) noisy background: a supervised multi-dimensional attention network (MDA-Net) is developed which consists of pixel attention network and channel attention network to suppress the noise and highlight foreground. 3) cluttered and dense objects in arbitrary orientation: an angle sensitive network is devised by introducing an angle related parameter for estimation. Combing these three techniques as a whole, our approach achieves state-of-the-art performance on a public remote sensing benchmarks DOTA.

We first give an overview of our two-stage method as sketched in Fig. 1. In the first stage, the feature map is expected to contain more feature information and less noise by adding SF-Net and MDA-Net. For positional sensitivity of the angle parameters, this stage still regresses the horizontal box. By the improved five-parameter regression and the rotation nonmaximum-suppression (R-NMS) operation for each proposal in the second stage, we can obtain the final detection results under arbitrary rotations.

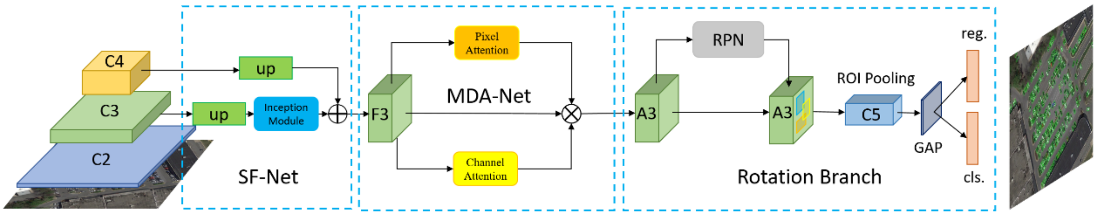

Fig. 1. Pipeline of our method

We use the pretrained ResNet-101 model for initialization. For DOTA, the model is trained by 600k iterations in total, and the learning rate changes during the 200k and 400k iterations from 3e-4 to 3e-6. Besides, weight decay and momentum are 0.0001 and 0.9, respectively. We employ MomentumOptimizer as optimizer and data augmentation is performed such as random image flip and rotation during training. It also should be noted that we shield the angle parameter in the code in HBB task. Meanwhile, we subtracted the mean value [103.939, 116.779, 123.68] which comes from ImageNet.

For parameter setting, we set the base anchor size to 256, and the anchor scales setting from 2^{-4} to 2^{1}. Since the multi-categories objects in DOTA have different shapes, we set anchor ratios to [1/1, 1/2, 1/3, 1/4, 1/5, 1/6, 1/7, 1/9]. These settings ensure that each ground-truth can be assigned with positive samples. When IoU > 0.7, the anchor is assigned as a positive sample, and as a negative sample if IoU < 0.3. Besides, due to the sensitivity between angle and IoU in the large aspect ratio rectangle, the two thresholds in the second stage are all set to 0.4, respectively. For training, the mini-batch size in two stages is 512.

Finally, we apply image pyramid, model ensemble, and other small tricks to achieve 74.7% and 78.4% mAP on OBB and HBB task, respectively. There are some reference code and paper for more details:

1) https://github.com/DetectionTeamUCAS/R2CNN_Faster-RCNN_Tensorflow
2) https://github.com/yangxue0827/R2CNN_FPN_Tensorflow
3) https://arxiv.org/abs/1811.07126