

Our baseline of object detection algorithm is based on RetinaNet. It is a one-stage detection algorithm. We further design a simple one-stage object detector called Attention Feature Pyramid Networks (AFPN). It draws on instructive feature channel information.

In the Challenge of Detecting objects in Aerial Images. All experiments are based on Keras as the deep learning framework. All models are trained on NVIDIA GTX1080 GPU.

We evaluate the performance of our proposed AFPN for geospatial object detection on a large-scale Dataset for Object detection in Aerial images (DOTA), which consists 2806 images from different sensors and platforms. There are 16 object categories, such as plane, basketball diamond, bridge, ground track field, small vehicle, large vehicle, ship, tennis court, basketball court, storage tank, soccer ball field, roundabout, harbor, swimming pool, helicopter, and container crane. Half of the origin images were randomly selected as the training set, 1/6 as the validation set, and 1/3 as the testing set. We split origin images into 1200\*1200 sub-images with an overlap of 200.

We did data augmentation training, such as used multi-scale testing. The result is shown in Table 1:

<i>objects</i>	<i>mAP</i>
<i>plane</i>	0.794
<i>Baseball-diamond</i>	0.812
<i>bridge</i>	0.545
<i>Ground-track-field</i>	0.635
<i>Small-vehicle</i>	0.646
<i>large-vehicle</i>	0.707
<i>ship</i>	0.861
<i>Tennis-court</i>	0.907
<i>Basketball-court</i>	0.737
<i>Storage-tank</i>	0.743
<i>Soccer-ball-field</i>	0.458
<i>roundabout</i>	0.719
<i>harbor</i>	0.737
<i>Swimming-pool</i>	0.755
<i>helicopter</i>	0.551
<i>Container-crane</i>	0.195

<i>mAP</i>	0.675
------------	-------

Table 1 The results on test set of objects in the Challenge.