PhD

Feature Pyramid Networks for Object Detection ax170419

2018/12/05 2:08 PM

Not sure about this but the main idea seems to be to combine the semantically high-level features present in later layers with the low-level features from earlier layers;

since higher level features are also low-resolution, they are upsampled before being added to the lower level features passed to 1X1 convolutions to reduce their feature dimensionality;

The point of all of this seems to be to approximate the extraction of features from multiple scales while having to process only a single scale and using extrapolation instead;

The output of the FPN is assert a feature maps of different sizes corresponding to different levels of the pyramid

all layers of the base network whose output maps are of the same size are said to be the same as stage of the network and one pyramid level is attached to each stage by taking the output of the last layer of that stage and processing it further – this is called the bottom of pathway

the top-down pathway uses upsampling to increase the size of upper level feature maps to make them the same size as one coming from the bottom up pathway and then adds them up element wise through lateral connections;

the top-down feature map is low-resolution but it is supposed to have thematically stronger features while the bottom up 1 is higher resolution so this combining thing is supposed to get the best of both worlds