[net]

batch=64 每batch个样本更新一次参数。

subdivisions=8 如果内存不够大，将batch分割为subdivisions个子batch，每个子batch的大小为batch/subdivisions。

在darknet代码中，会将batch/subdivisions命名为batch。

height=416 input图像的高

width=416 Input图像的宽

channels=3 Input图像的通道数

momentum=0.9 动量

decay=0.0005 权重衰减正则项，防止过拟合

angle=0 通过旋转角度来生成更多训练样本

saturation = 1.5 通过调整饱和度来生成更多训练样本

exposure = 1.5 通过调整曝光量来生成更多训练样本

hue=.1 通过调整色调来生成更多训练样本

learning\_rate=0.0001 初始学习率

max\_batches = 45000 训练达到max\_batches后停止学习

policy=steps 调整学习率的policy，有如下policy：CONSTANT, STEP, EXP, POLY, STEPS, SIG, RANDOM

steps=100,25000,35000 根据batch\_num调整学习率

scales=10,.1,.1 学习率变化的比例，累计相乘

[convolutional]

batch\_normalize=1 是否做BN

filters=32 输出多少个特征图

size=3 卷积核的尺寸

stride=1 做卷积运算的步长

pad=1 如果pad为0,padding由 padding参数指定。如果pad为1，padding大小为size/2

activation=leaky 激活函数：

logistic，loggy，relu，elu，relie，plse，hardtan，lhtan，linear，ramp，leaky，tanh，stair

[maxpool]

size=2 池化层尺寸

stride=2 池化步进

[convolutional]

batch\_normalize=1

filters=64

size=3

stride=1

pad=1

activation=leaky

[maxpool]

size=2

stride=2

......

......

#######

[convolutional]

batch\_normalize=1

size=3

stride=1

pad=1

filters=1024

activation=leaky

[convolutional]

batch\_normalize=1

size=3

stride=1

pad=1

filters=1024

activation=leaky

[route] the route layer is to bring finer grained features in from earlier in the network

layers=-9

[reorg] the reorg layer is to make these features match the feature map size at the later layer.

The end feature map is 13x13, the feature map from earlier is 26x26x512.

The reorg layer maps the 26x26x512 feature map onto a 13x13x2048 feature map

so that it can be concatenated with the feature maps at 13x13 resolution.

stride=2

[route]

layers=-1,-3

[convolutional]

batch\_normalize=1

size=3

stride=1

pad=1

filters=1024

activation=leaky

[convolutional]

size=1

stride=1

pad=1

filters=125 region前最后一个卷积层的filters数是特定的，计算公式为filter=num\*(classes+5)

5的意义是5个坐标，论文中的tx,ty,tw,th,to

activation=linear

[region]

anchors = 1.08,1.19, 3.42,4.41, 6.63,11.38, 9.42,5.11, 16.62,10.52 预选框，可以手工挑选，

也可以通过k means 从训练样本中学出

bias\_match=1

classes=20 网络需要识别的物体种类数

coords=4 每个box的4个坐标tx,ty,tw,th

num=5 每个grid cell预测几个box,和anchors的数量一致。当想要使用更多anchors时需要调大num，且如果调大num后训练时Obj趋近0的话可以尝试调大object\_scale

softmax=1 使用softmax做激活函数

jitter=.2 通过抖动增加噪声来抑制过拟合

rescore=1 暂理解为一个开关，非0时通过重打分来调整l.delta（预测值与真实值的差）

object\_scale=5 栅格中有物体时，bbox的confidence loss对总loss计算贡献的权重

noobject\_scale=1 栅格中没有物体时，bbox的confidence loss对总loss计算贡献的权重

class\_scale=1 类别loss对总loss计算贡献的权重

coord\_scale=1 bbox坐标预测loss对总loss计算贡献的权重

absolute=1

thresh = .6

random=0 random为1时会启用Multi-Scale Training，随机使用不同尺寸的图片进行训练。