**使用带PPM模块的U-Net网络框架解决语义分割中的遥感图像问题**

## **摘要**

## 本文是将我看的两篇论文《U-Net: Convolutional Networks for Biomedical Image Segmentation》《Pyramid Scene Parsing Network》中的网络框架运用于变化检测遥感图像领域中。通过实验数据，分析两个网络框架在遥感图像领域使用的优缺点。首先我将介绍我要使用的这两个网络框架模型，以及使用的优化器，学习率调整策略等等，我还将介绍我使用的数据集。最后通过过实验结果得出结论。

**一、U-Net**

**（一）网络结构**

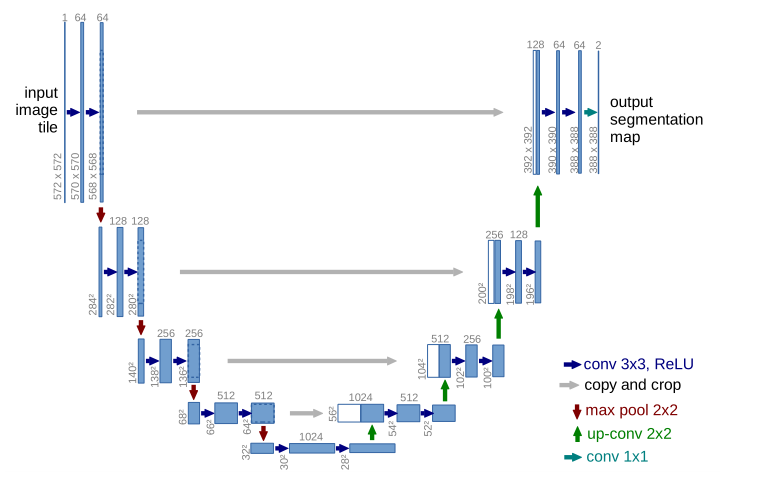


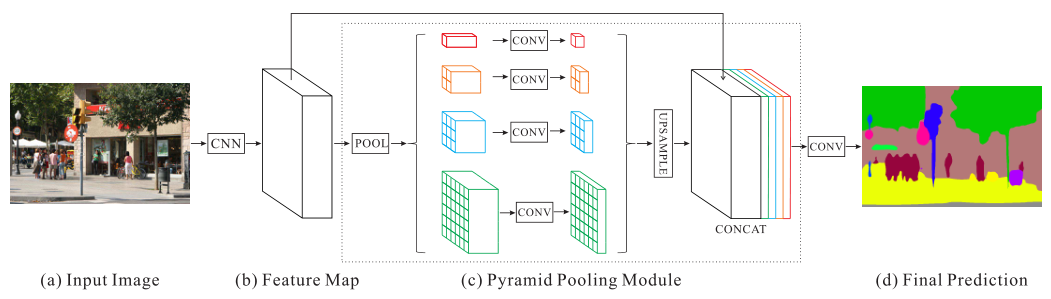
图1.1

1 左：特征提取部分：它是一个收缩网络，通过四个下采样，使图片尺寸减小，在这不断下采样的过程中，特征提取到的是浅层信息。具体过程是，输入图片然后经过两个卷积核（3x3后面紧跟着一个Relu）以论文原图为例：输入572x572，经过两个卷积核（大小为3x3）大小从572-570-568，然后经过一个Maxpool（2x2）图片尺寸变为284这即为一个完整的下采样，接下来三个也是如此。在下采样的过程中，通道数翻倍，例如图上的从64-128。2.中（copy and crop）：在论文中叫拼接，在U-Net有四个拼接操作。如上图所示：有人也叫Skip connect,这一操作的目的是为了融合特征信息，使深层和浅层的信息融合起来，在拼接的时候要注意，不仅图片大小要一致（故要crop,是为了使图片大小一致）而且特征的维度（channels）也要才一样，才可以拼接。3.右（上采样部分 up-conv）也叫扩张网络，图片尺寸变大，提取的是深层信息，使用了四个上采样，在上采样的过程中，图片的通道数是减半的，与左部分的特征提取通道数的变化相反。在上采样的过程融合了左边的浅层的信息即拼接了左边的特征。

**（二）U-Net结构优势**

网络的高层（就是U型的上部分），获取了图形的细节信息（因为这时候图片很大，很多细节可以得以保留）。在网络的底层（U型的下部分），获取了图形的低频信息（感受野很大，便于获取大的轮廓信息）。然后用过skip connection将各个层次的信息保留。使得整个网络可以很好的记住图片的所有信息。

**二、PSPNet**

**（一）网络结构**

在PSPNet中，首先将输入进来的图像使用主干提取网络（ResNet101）提取特征层，一部分将直接连接在后方，另一部分将进入PSP模块中，PSP模块会将输入进来的特征层划分成6x6，3x3，2x2，1x1的网格，对应了图片中的绿色、蓝色、橙色、红色的的输出

其中：

红色：将输入进来的特征层整个进行平均池化。

橙色：将输入进来的特征层划分为2×2个子区域，然后对每个子区域进行平均池化。

蓝色：将输入进来的特征层划分为3×3个子区域，然后对每个子区域进行平均池化。

绿色：将输入进来的特征层划分为6×6个子区域，然后对每个子区域进行平均池化。

最后将这4次平均池化的结果进行上采样，将这四次上采样之后的结果和黑框部分进行堆叠，其中主干提取网络进行初步特征提取，PSP模块进行加强特征提取。然后利用堆叠之后的

结果进行卷积和上采样，对每一个像素点进行分类获得最终的预测结果。输入图像的尺寸和输出图像的尺寸相同。

（二）PSPNet结构优势

1.多尺度特征融合。由于在深层网络中，高层特征包含了更多的语义和较少的位置信息。结合多尺度特征可以提高性能.

2.基于结构进行预测。例如使用条件随机场(CRF)进行后处理，对分割结果进行细化。

**三、Deep Supervision for ResNet-Based FCN**

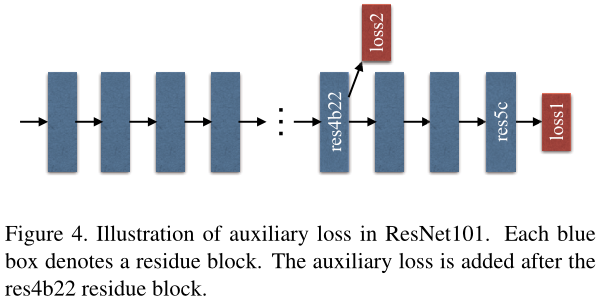


图3为深度监督ResNet101模型的例子

除了使用Softmax loss来训练最终分类器的的主分支外，在第四阶段后再使用另一个分类器，即res4b22残差块

这两个loss同时传播，通过各自前面所有层

辅助loss有助于优化学习过程，主loss仍是主要的优化方向。增加权重，以平衡辅助loss

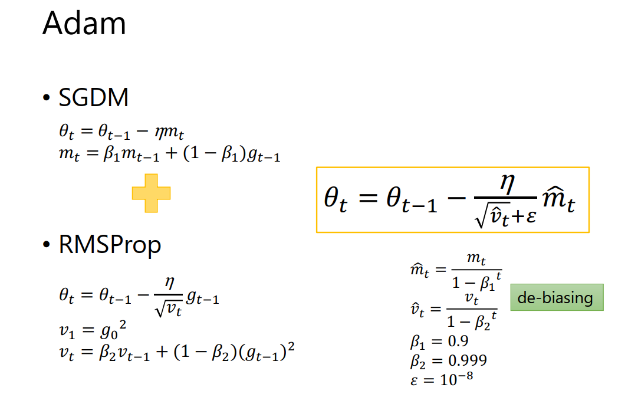
在测试阶段，放弃辅助分支，只使用优化好的主分支来进行最终的预测。这种对基于ResNet的FCN进行深度监督的训练策略在不同的实验环境下是非常有用的，与预训练的ResNet模型也可以很好结合。体现了这种学习策略的普遍性。

辅助损失项用于训练过程中。辅助损失项有0.4的权重，以用来平衡最终损失和辅助损失。在测试时，则会放弃辅助损失。这是一种针对深度网络训练的深度监督训练策略。

简单来说就是两个loss一起传播，使用不同的权重，共同优化参数

**四、Adam**

Adam是SGDM和RMSProp的结合，它基本解决了之前提到的梯度下降的一系列问题，比如随机小样本、自适应学习率、容易卡在梯度较小点等问题，2015年提出。如下

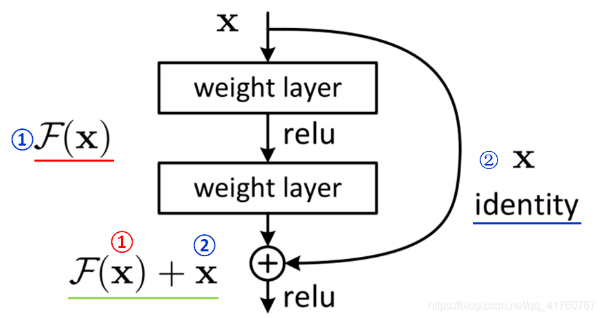


由上可见，mt即为动量，根号vt即为自适应学习率。加入了两个衰减系数β1和β2。刚开始所需动量比较大，后面模型基本稳定后，逐步减小对动量的依赖。自适应学习率同样也会随迭代次数逐渐衰减。𝜀则是防止除数为0，仅仅是数学计算考虑。

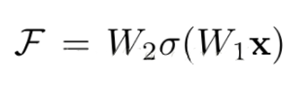
**五、ResNet残差网络**

**（一）ResNet网络结构**

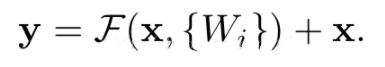
ResNet在2015年被提出，在ImageNet比赛classification任务上获得第一名，因为它“简单与实用”并存，之后很多方法都建立在ResNet50或者ResNet101的基础上完成的，检测，分割，识别等领域里得到广泛的应用。它使用了一种连接方式叫做“shortcut connection”，顾名思义，shortcut就是“抄近道”的意思，下面是这个resnet的网络结构：



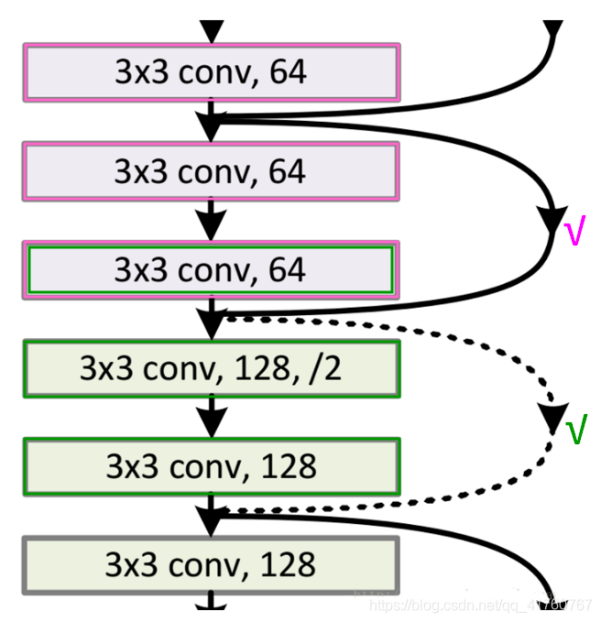
它对每层的输入做一个reference（X）, 学习形成残差函数， 而不是学习一些没有reference（X）的函数。这种残差函数更容易优化，能使网络层数大大加深。在上图的残差块中它有二层，如下表达式，其中σ代表非线性函数ReLU。



然后通过一个shortcut，和第2个ReLU，获得输出y。



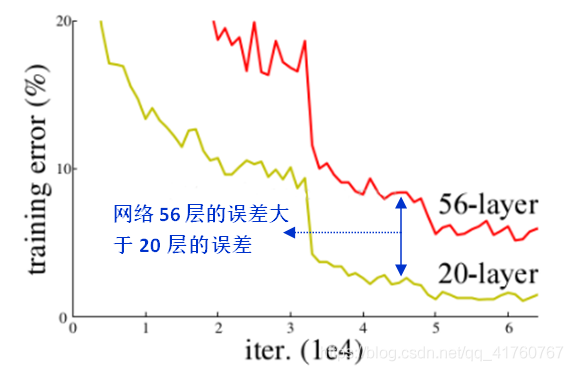
然而实验证明x已经足够了，不需要再搞个维度变换，除非需求是某个特定维度的输出，如是将通道数翻倍，如下图所示：



由上图，我们可以清楚的看到“实线”和“虚线”两种连接方式， 实线的Connection部分 (第一个粉色矩形和第三个粉色矩形) 都是执行3x3x64的卷积，他们的channel个数一致，所以采用计算方式：

Y = F(x) + x，虚线的Connection部分 (第一个绿色矩形和第三个绿色矩形) 分别是3x3x64和3x3x128的卷积操作，他们的channel个数不同(64和128)，所以采用计算方式： y=F(x)+Wx 。其中W是卷积操作，用来调整x的channel维度。

在计算机视觉里，网络的深度是实现网络好的效果的重要因素，输入特征的“等级”随增网络深度的加深而变高。然而在网络深度不断加深的情况下，梯度弥散/爆炸成为训练深层次的网络的障碍，导致导致网络无法收敛。虽然，归一初始化，各层输入归一化，使得可以收敛的网络的深度提升为原来的十倍。虽然网络收敛了，但网络却开始退化 （增加网络层数却导致更大的误差）， 如下图所示：



由上图可知，在一个浅层网络的基础上叠加y=x的层（称identity mappings，恒等映射），可以让网络随深度增加而不退化。这反映了多层非线性网络无法逼近恒等映射网络。

但是，在深度学习中我们希望有更好性能的网络，而网络不退化则不是我们的目的。在 RestNet网络中学习的残差函数是F(x) = H(x) - x, 这里如果F(x) = 0, 那么就是上面提到的恒等映射（H(x) = x）。事实上，RestNet是“shortcut connections”的在connections是在恒等映射下的特殊情况，它没有引入额外的参数和计算的复杂度。 假如优化目标函数是逼近一个恒等映射, 而不是0映射（F(x) = 0）或者说恒等映射，那么学习找到对恒等映射的扰动会比重新学习一个映射函数要容易。

**五、学习率调整策略**

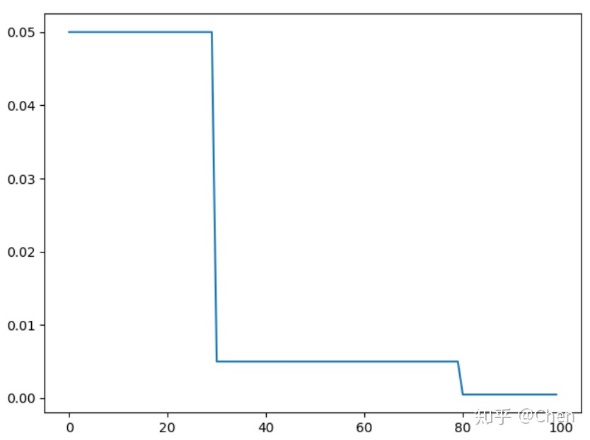
（一）**MultiStep LR**

更新规律：

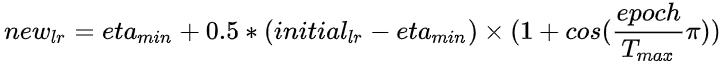


其中，new表示得到的新学习率；initial表示初始学习率；y表示更新lr的乘法因子milestones是递增的list，存放要更新lr的epoch。bisect-right返回把epoch插入排序好的列表milestones式的位置。

torch**.**optim**.**lr\_scheduler**.**MultiStepLR(optimizer,milestones,gamma**=**0.1,last\_epoch**=-**1)

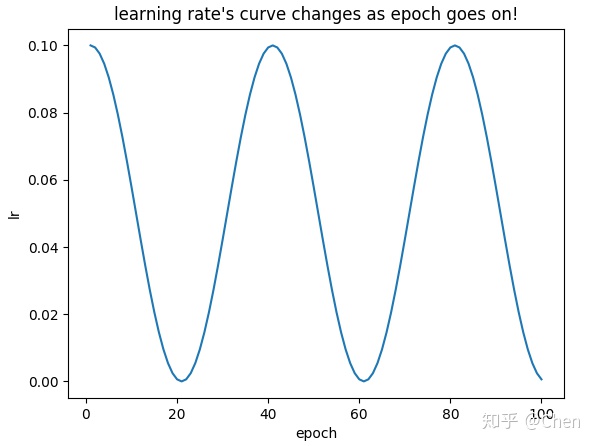
**（二）Cosine Annealing LR**

让学习率随epoch的变化图类似于cos，更新策略：



其中，余弦退火学习率中LR的变化是周期性的，T\_max是周期的1/2；eta\_min(float)表示学习率的最小值，默认为0；last\_epoch(int)代表上一个epoch数，该变量用来指示学习率是否需要调整。当last\_epoch符合设定的间隔时，就会对学习率进行调整。当为-1时，学习率设为初始值。

torch**.**optim**.**lr\_scheduler**.**CosineAnnealingLR(optimizer,T\_max,eta\_min**=**0,last\_epoch**=-**1)

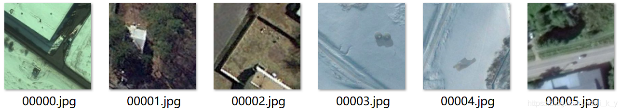


**六、SVCD数据集**

该数据集具有三种类型：没有对象相对移动的合成图像，对象相对移动较小的合成图像，随季节变化的真实遥感图像（由Google Earth获得）。 随季节变化的遥感图像具有16000个图像集，图像尺寸为256x256像素（10000个训练集以及3000个测试和验证集），空间分辨率为3至100 cm / px。

数据集中包含生成图像（具有随机背景和随机不相交的几何图形（正方形，圆形，矩形，三角形）的RGB图像对）和真实图像。具体介绍在论文的 EXPERIMENTS 部分。







**七、实验**

**（一）实验设计**

将PSPNet，U-Net运用于变化检测。

（1）数据集：SVCD（CDD）

（2）PSPNet1：resnet34，bs16，multistepLR，ep100

PSPNet2：resnet34，bs16，CosineAnnealingLR，ep60

UNet1：resnet34，bs16，multistepLR，ep100

UNet2：resnet34，bs16，CosineAnnealingLR，ep60

（3）主要做了4组对比试验：

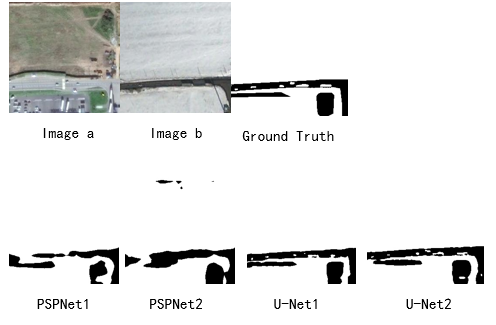
第一组：PSPNet1，PSPNet2 改变了学习率调整策略

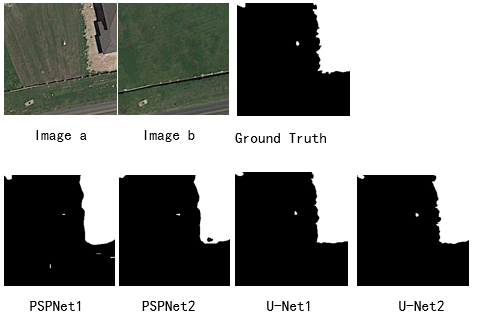
第二组：UNet1，UNet2 改变了学习率调整策略

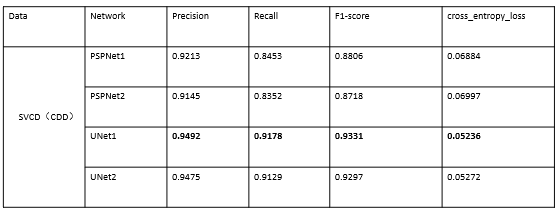
第三组：PSPNet1，UNet1 改变了网络框架

第四组：PSPNet2，UNet2 改变了网络框架

**（二）实验结果**







从实验结果来看，使用U-Net框架的效果要优于使用PSPNet框架的效果。

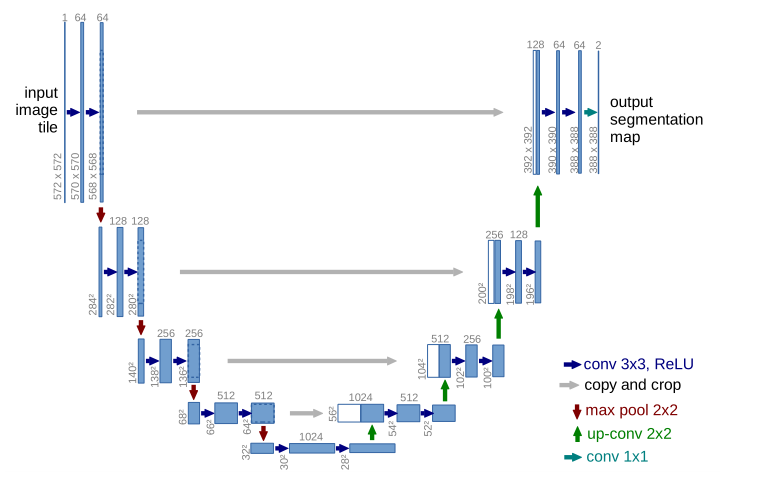
使用MultiStepLR学习率调整策略要优于CosineAnnealingLR学习率调整策略。

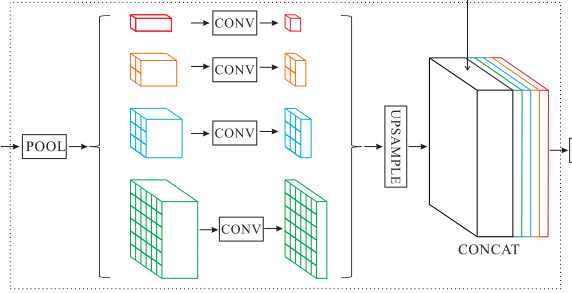
**（三）实验结果分析**

PSPNet 的上采样过程是在PPM模块中直接进行的8倍上采样。

U-Net在上采样的同时，网络架构进行了skip connection，有利于将下采样的各个阶段的信息在上采样过程中进行整合，就是在上采样的过程中，结合了各个层次的结构信息，从而使得预测的特征信息更为准确。

**八、新模型**

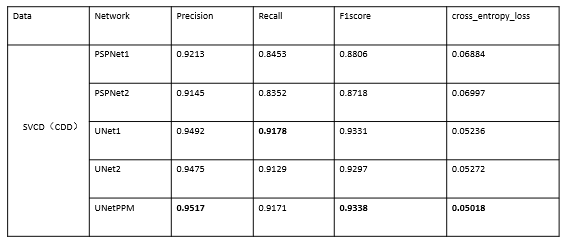


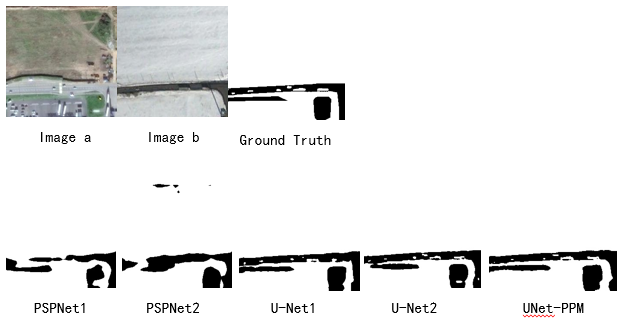


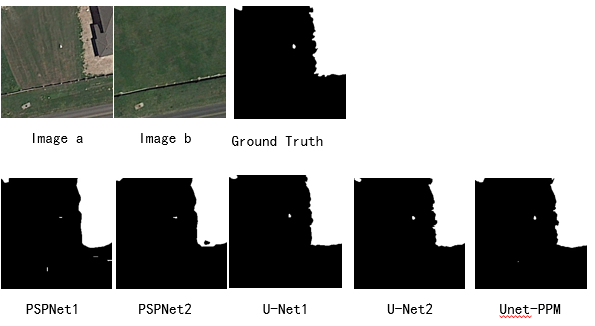
通过刚才对实验结果的分析，将PPM模块融入U-Net网络框架中。

将U-Net框架的底端插入PPM模块。进行新的实验。

**九、实验**







从实验结果上看，新的模型在SVCD数据集的实验结果比UNet，PSPNet的实验结果要好

**九、总结**

PSPNet框架的核心是PPM模块，PPM是以一种特殊的池化模型。可以有效增大感受野，增大全局信息的利用效率。PPM模块可以添加到许多网络框架中。

增强感受野的策略除PPM模块，还有如：SPP模块、ASPP模块、RFB模块等策略。

**十、参考文献**

（1）U-Net: Convolutional Networks for Biomedical Image Segmentation

（2）Pyramid Scene Parsing Network

（3）FULLY CONVOLUTIONAL SIAMESE NETWORKS FOR CHANGE DETECTION

（4）Unified Perceptual Parsing for Scene Understanding

（5）Deep Residual Learning for Image Recognition

（6）[1]刘国强,房胜,李哲.用于遥感图像变化检测的全尺度特征聚合网络[J/OL].北京航空航天大学学报:1-9[2021-11-07].https://doi.org/10.13700/j.bh.1001-5965.2021.0522.

（7）[1]刘雨婷,刘帆.基于卷积稀疏表示和NSCT的遥感图像融合[J/OL].太原理工大学学报:1-10[2021-11-07].http://kns.cnki.net/kcms/detail/14.1220.n.20211014.2200.002.html.

（8）[1]刘冠群,刘豪,王新,王威.一种基于深度学习的卫星遥感图像分割方法[J].湖南城市学院学报(自然科学版),2021,30(05):54-58.

（9）[1]刘高天,段锦,范祺,吴杰,赵言.基于改进RFBNet算法的遥感图像目标检测[J].吉林大学学报(理学版),2021,59(05):1188-1198.

（10）[1]董如婵,焦李成,赵进,沈维燕.一种深度融合机制的遥感图像目标检测技术[J/OL].西安电子科技大学学报:1-13[2021-11-07].http://kns.cnki.net/kcms/detail/61.1076.tn.20210827.1713.010.html.