Cotta网络实现暗夜语义分割任务

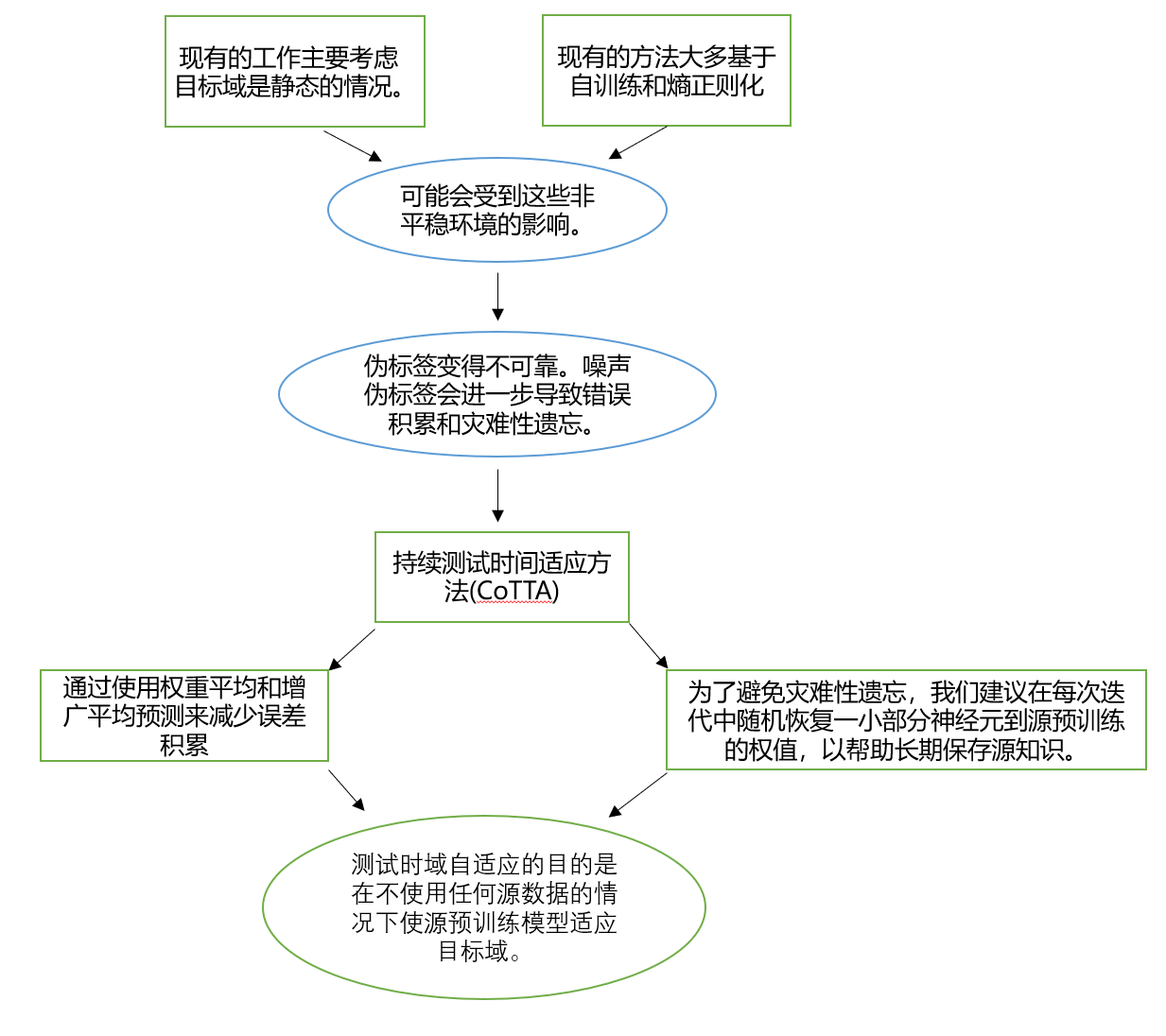


图1：选择Cotta的原因

1.RefineNet模型

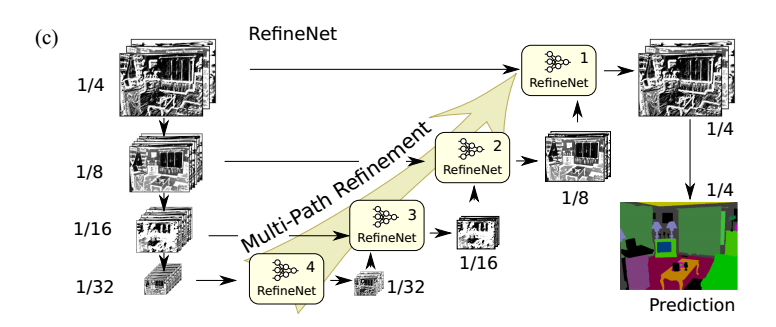


图2：RefineNet的整体图

低分辨率特征指的是在图像或视频处理过程中，将图像或视频的分辨率降低，从而得到的较粗糙的高级特征表示。可以提供大概和抽象的信息细粒度的低级特征通常是指在图像或视频处理中较低层次的特征表达，比如边缘、纹理、颜色等。这些特征具有较高的空间分辨率，可以提供更加详细和精确的信息。

RefineNet以递归的方式用细粒度的低级特征细化低分辨率(粗)特征，以生成高分辨率的特征映射。

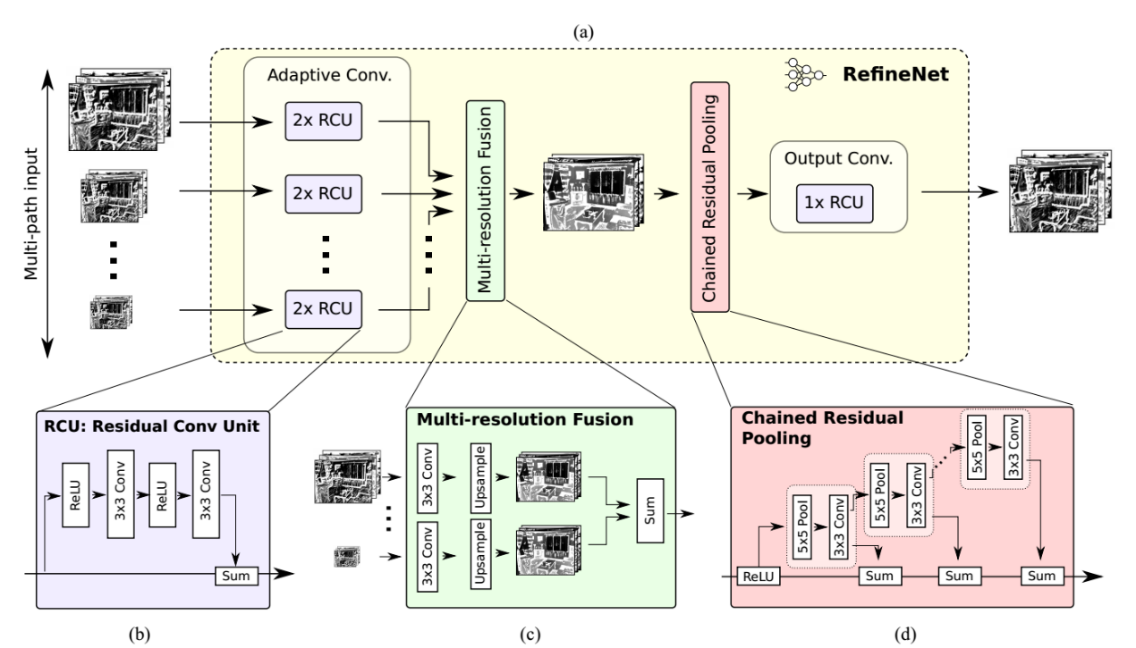


图3：RefineNet的内部结构图

通常RefineNet的输入是两个，一个来自于ResNet，另一个来自上一级的RefineNet。

每个RefineNet块的第一部分由一个自适应卷积集组成，主要用于为我们的任务微调预训练的ResNet权重。第二部分首先应用卷积进行输入自适应，生成相同特征维度(输入中最小的一个)的特征映射，然后对所有(较小的)特征映射进行上采样，使其达到输入的最大分辨率。最后，对所有特征映射进行求和融合。第三部分在从大图像区域捕获背景内容。它能够有效地汇集具有多个窗口大小的特征，并使用可学习的权重将它们融合在一起。最后一部分输出，特征维度不变。

如果只有一条输入路径(如图2(c)中的RefineNet-4)，则输入路径将直接穿过第二、三部分，不做任何改变。如果是最后一个RefineNet，则连接softmax函数生成最后预测。

2.cotta

(1)结构

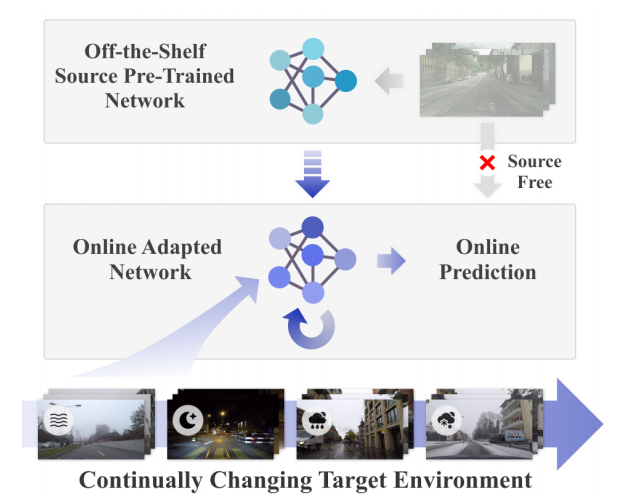


图4：cotta的整体结构

目标数据（Zurich）是按顺序提供的，并且来自不断变化的环境。使用现成的RefineNet来初始化目标网络。模型基于当前目标数据在线更新，并以在线方式给出预测结果。目标网络的自适应不依赖于任何源数据。

（2）数据流向

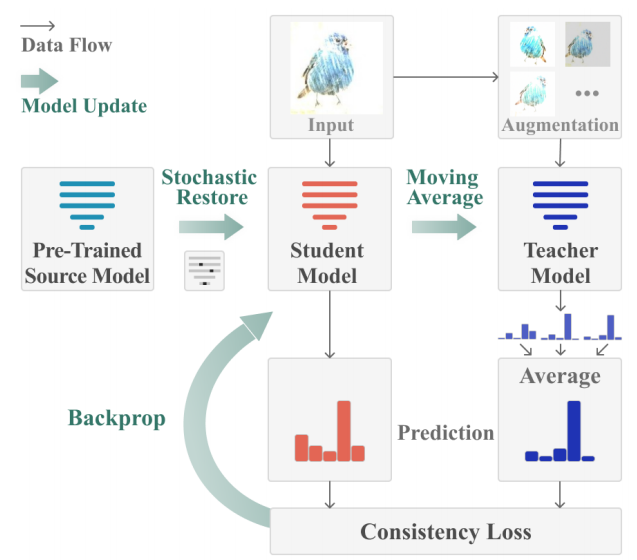
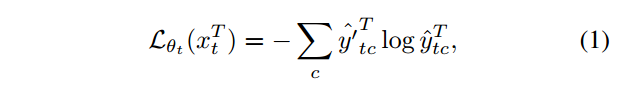


图5：结构图/数据流向

在给定目标数据XTt和模型fθt的情况下，在自训练框架下，我们的目标是通过最小化预测值y^Tt = fθt (XTt)与伪标签的交叉熵一致性来进行优化。

学生和教师预测之间的交叉熵损失：



在给定的情况下，yˆ′Ttc表示教师模型生成的伪标签中类别c在时间步长t出现的概率，而yˆTtc表示主模型（学生模型）在时间步长t对类别c的预测值。

**权重平均伪标签**

在时间步长t = 0时，将教师网络初始化为与源预训练网络相同。

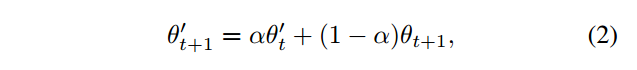
在时间步长t处，伪标签是由加权平均教师模型生成的，

即yˆ′Tt = fθ’(XTt)，

其中fθ’表示加权平均教师模型，XTt为输入的目标数据。教师模型会根据输入的目标数据XTt生成预测值yˆ′Tt，作为该时间步长t处的伪标签。

**EMA**

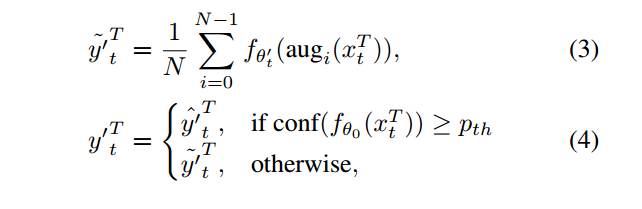
使用学生权重通过指数移动平均更新教师模型的权重:



其中α是平滑因子。

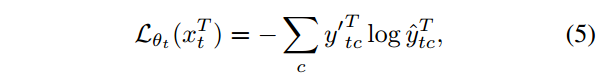
**增广平均伪标签**

在不断变化的环境下，测试分布可能发生巨大变化，这可能使增强策略无效。在这里，我们考虑了测试时间的域移，并通过预测置信度近似域差。仅在域差较大时进行增广，以减少误差累积。



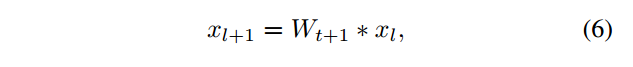
其中，y˜′Tt是来自教师模型的增强平均预测，yˆ′Tt是来自教师模型的直接预测，conf(fθ0(XTt))是源预训练模型对当前输入XTt的预测置信度，pth是置信度阈值。**我们假设较低的置信度表示较大的领域差距，较高的置信度表示较小的领域差距**。因此，当置信度较高且大于阈值时，我们直接使用yˆ′Tt作为伪标签，不使用任何增值。当置信度较低时，我们额外应用N个随机增广来进一步提高伪标签质量。

最后又通过改进的标签代入公式5来更新学生模型：



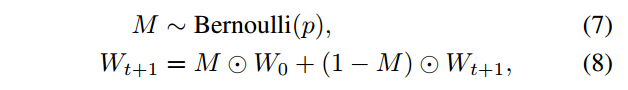
**随机修复**

考虑在时间步长为t时，基于方程1的**梯度更新**后的学生模型fθ内的卷积层:



其中\*表示卷积运算，xl和xl+1表示该层的输入和输出，Wt+1表示**可训练的卷积滤波器**。

本文提出的随机恢复方法通过以下方式对权值W进行更新:



其中⊙表示逐元素的乘法。p是一个小的恢复概率，M是与Wt+1形状相同的掩模张量，Bernoulli表示伯努利分布。掩码张量决定Wt+1中的哪个元素要恢复到源权重W0。如果M中对应位置的值是1，那么Wt+1中的对应元素将从源权重W0中恢复；如果M中对应位置的值是0，那么Wt+1中的对应元素将保持不变。

3.算法

（1）初始化：准备源预训练模型refinent，由refinenet初始化老师网络。

（2）输入：每个时间步长t，当前数据流xt

（3）训练： 1.增广xt，由式4从教师网络得到权值和增广平均伪标签。

2.通过公式5中的一致性损失更新学生网络。

3.通过方程2中的移动平均指数（EMA）更新教师网络。

4.用公式8随机还原学生网络。

（4）输出：更新的学生网络和更新的老师网络，还有最终的老师网络的预测。

（5）评估