(Zhao 等, 2018, p. 1) 我们注意到由于卷积滤波器的物理设计，卷积神经网络中的信息流被限制在局部邻域内，这限制了对复杂场景的整体理解。在本文中，我们提出了从点到面的空间注意网络（PSANet）来放松局部邻域的约束。特征图上的每个位置都通过自适应学习的注意力掩码与所有其他位置相连。此外，用于场景解析的双向信息传播被启用。其他位置的信息可以被收集以帮助预测当前位置，反之亦然，当前位置的信息可以被分发以帮助预测其他位置。我们提出的方法在各种有竞争力的场景解析数据集上取得了最佳性能，包括ADE20K、PASCAL VOC 2012和Cityscapes，证明了其有效性和通用性。

(Zhao 等, 2018, p. 2) 在本文中，我们提出了点式空间注意网络（PSANet），以灵活和自适应的方式聚集长距离的背景信息。特征图中的每个位置都通过自适应预测的注意力图与所有其他位置相连，从而收获附近和远处的各种信息。此外，我们设计了双向的信息传播路径，以全面了解复杂的场景。每个位置收集所有其他位置的信息以帮助预测自己，反之亦然，每个位置的信息可以分布到全球，帮助预测所有其他位置。最后，双向聚合的上下文信息与本地特征相融合，形成复杂场景的最终表示。我们提出的PSANet在三个最具竞争力的语义分割数据集，即ADE20K[48]、PASCAL VOC 2012[9]和Cityscapes[8]上取得了顶级性能。我们相信，所提出的明智的空间注意模块与双向信息传播范式一起，也能使其他密集预测任务受益。我们给出了所有的实施细节，并将代码和训练好的模型公开提供给社区  
  
我们的主要贡献有三个方面  
我们通过学习点-明智的位置敏感的上下文依赖和双向信息传播范式，实现了场景解析的长程上下文聚合  
我们提出了点式空间注意网络（PSANet），从特征图中的所有位置收集上下文信息。每个位置都通过自适应学习的注意力图谱与所有其他位置相连  
PSANet在各种有竞争力的场景解析数据集上取得了顶级性能，证明了其有效性和通用性。

(Zhao 等, 2018, p. 4) 为了捕捉上下文信息，特别是远距离的信息，信息聚合对场景解析具有重要意义[24,5,45,38]。在本文中，我们将信息聚合步骤表述为一种信息流，并提出从两个角度自适应地学习每个位置的像素级全局注意图，以聚合整个特征图的上下文信息

(Zhao 等, 2018, p. 7) 另一方面，我们将当前位置的信息分配给其他位置。在每个位置，我们预测当前位置的信息对其他位置的重要性。援助的产生与AIC类似。这个注意力图谱有助于分配信息，以便更好地进行预测。这两个地图以互补的方式编码了不同位置对之间的上下文依赖性，导致了信息传播的改善和对长距离上下文的强化利用。利用这两种不同的注意力的好处在实验中得到了体现。

(Zhao 等, 2018, p. 8) 我们的PSA模块是可扩展的，可以附加到FCN结构中的任何阶段。我们在图4中展示了我们的实例。给定一个输入图像I，我们通过FCN获得它的局部表示，即特征图X，这是PSA模块的输入。和[45]一样，我们采用ResNet[13]作为FCN的主干。我们提出的PSA模块被用来聚合来自本地表示的远距离上下文信息。它遵循ResNet中的第五阶段，这是FCN骨干网的最后阶段。第五阶段的特征在语义上更强。将它们聚集在一起会导致对长距离背景的更全面的表述。此外，第五阶段的特征图的空间大小较小，可以减少计算开销和内存消耗。参照[45]，我们也采用了同样的深度监督技术。如图4所示，除了主损失之外，还应用了一个辅助损失分支。

场景解析和语义分割。最近，基于CNN的方法[26,4,5,42,45,6]在场景解析和语义分割的任务中取得了显著的成功。FCN[26]是第一个用卷积层取代分类网络中的全连接层进行语义分割的方法。DeconvNet[29]和SegNet[1]采用了编码器-解码器结构，利用低层的信息来帮助完善分割掩码。扩张卷积[4,42]在特征图上应用跳过卷积来扩大网络的接受域。UNet[33]将低级层的输出与高级层的输出连接起来进行信息融合。DeepLab[4]和CRF-RNN[46]利用CRF进行场景解析的结构预测。DPN[25]使用MRF进行语义分割。LRR[11]和RefineNet[21]采用分步重建和细化来获得解析结果。PSPNet[45]通过金字塔集合策略实现了高性能。还有一些高效率的框架，如ENet[30]和ICNet[44]，用于自动驾驶等实时应用。语境信息聚合。语境信息对图像理解起着关键作用。扩张卷积[4,42]在经典卷积核内插入了扩张，以扩大CNN的接受域。全局集合被广泛地应用于各种基本分类骨干[19,35,36,13,14]，以收获全局表示的上下文信息。Liu等人提出了ParseNet[24]，利用全局池来聚合上下文信息进行场景解析。陈晓东等人开发了ASPP[5]模块，Zhao等人提出了PPM[45]模块来获取不同区域的上下文信息。在此基础上，我们提出了ReSeg[38]，利用RNN来获取长距离的上下文依赖信息。注意机制。注意机制在神经网络中被广泛使用。Mnih等人[28]学习了一个注意力模型，该模型可以自适应地选择一系列的区域或位置进行处理。Chen等人[7]学习了几个注意力掩码来融合来自不同分支的特征图或预测。Vaswani等人[37]学习了一个用于机器翻译的自我注意力模型。Wang等人[40]通过计算特征图中每个空间点之间的相关矩阵得到了注意力掩码。我们的基于点的注意力掩码与上述研究不同。具体来说，通过我们的PSA模块学习的掩码是自适应的，对位置和类别信息敏感。PSA学习为每个单独的点自适应地和具体地汇总上下文信息。