

## 每日小结

	周一	周二	周三	周四	周五
早	跑 attention, 上课	跑 attention	跑 CBAM, 上课	上课	跑 GPT
中	跑 CBAM, 上课	论文阅读	论文阅读	跑代码, 上课	跑 GPT
晚	上课	上课		GPT 学习	跑 GPT

注：简单表述当前时间段工作，如看文献 1，整理数据等

## 科研详情

## 文献阅读

## 文献1

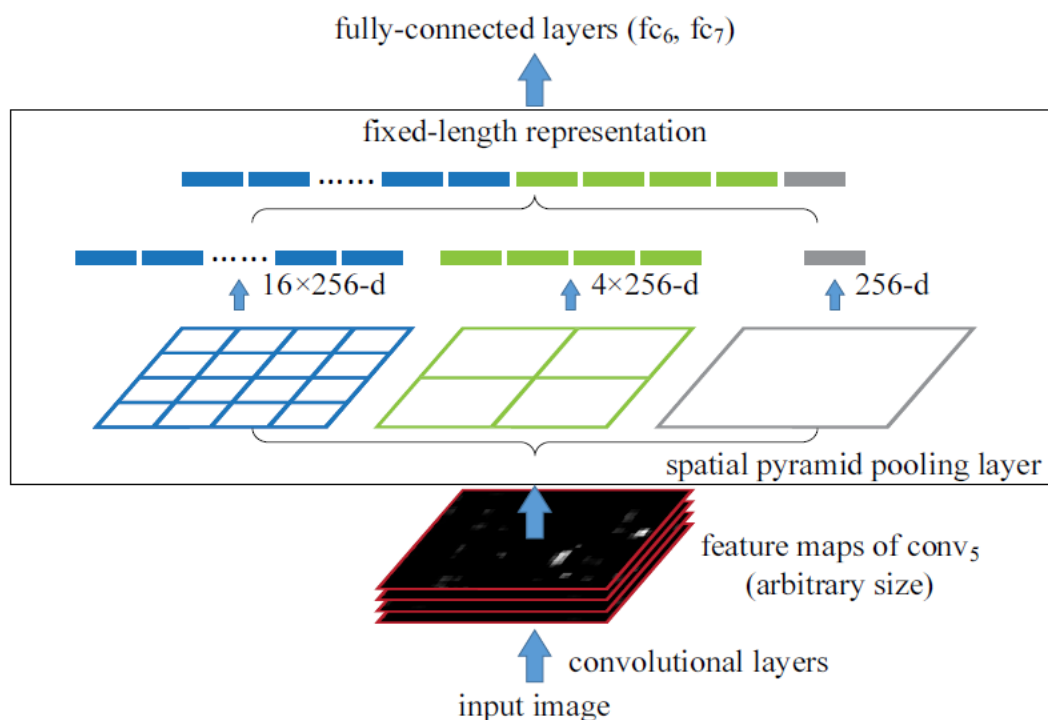
题目： Learning Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition

作者： Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun

出处： arxiv1506

## 方法：

文章以 AlexNet 为例，讨论 SPP layer。AlexNet 包括五层卷积层和两层全连接层，在最后一层卷积层后面加上 SPP 层，如图所示：



金字塔池化是分等级的，每一级的大小都用  $n \times n$  bins 来描述，意思是经过该等级的池化操作后，输出特征图的大小就是  $n \times n$ 。那么如何做到这一点呢，因为我们池化都是用滑动窗口来实现的，这涉及到滑动窗口大小和步长。金字塔池化的精髓就是滑动窗口的大小和步长是正比于输入图像的大小的，这样就能够产生大小（长度）的输出特征。

$$\text{window size } win = \lfloor a/n \rfloor$$

$$\text{stride str} = \lfloor a/n \rfloor$$

因此，不管 conv5 特征图的尺寸是多少，经过 SPP 后都产生 15\*256 的输出，然后送入全连接

层中进行下一步的分类等操作。卷积层的输出特征图的尺寸会根据输入图像的尺寸发生变化，因此，加入了 SPP 模块，SPPNet 就可以处理任意尺寸输入的图片进行训练和测试。

启发：

1. CNN 由于全连接层的存在，需要固定尺寸的输入图像，这限制了输入图像的纵横比和尺度，当前方法主要通过裁剪或变形将输入图像变换到固定尺寸，这样会导致图像内容的丢失或失真，从而影响图像识别的准确性。
2. SPP 网络的输出特征长度固定，与输入无关。
3. SPP 的多级池化提取多尺度特征，使网络对图像变形更加鲁棒。
4. SPP 可以处理不同尺寸的输入图像，提升网络的性能。

## 文献2

题目：CBAM: Convolutional Block Attention Module

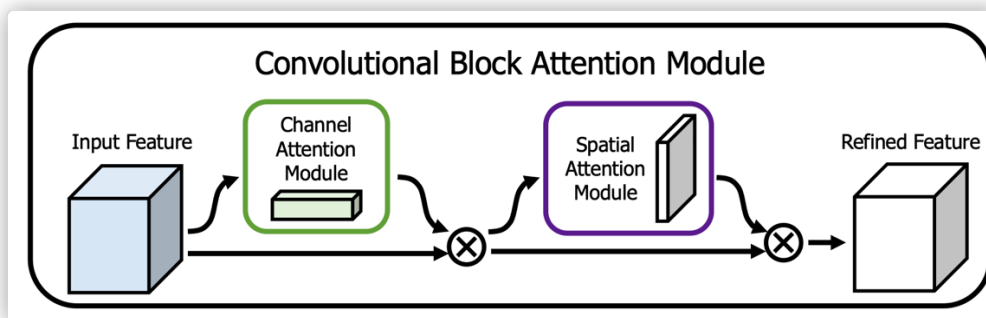
作者：Sanghyun Woo, Jongchan Park, Joon-Young Lee, In So Kweon;

出处：Proceedings of the European Conference on Computer Vision (ECCV), 2018

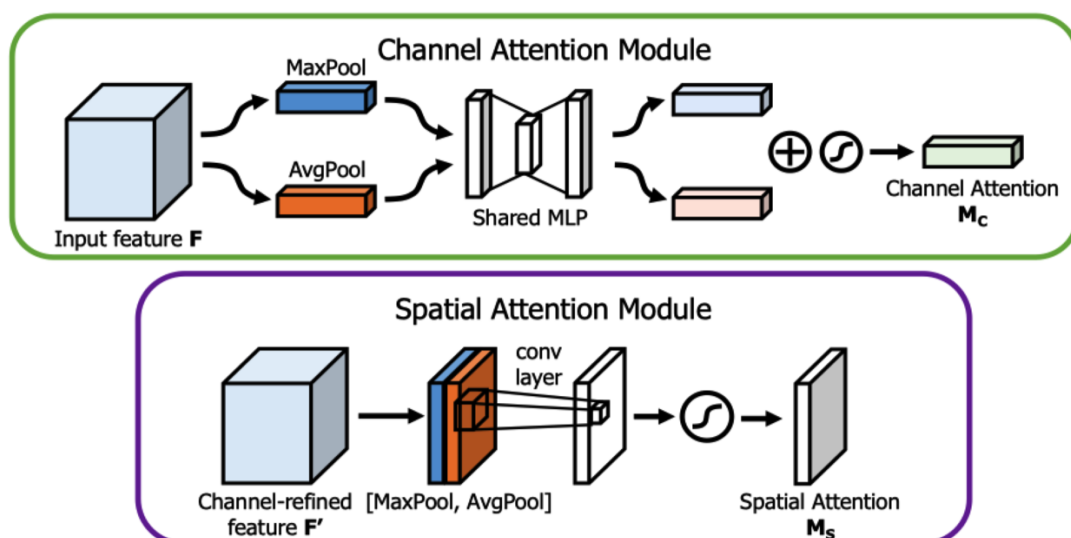
方法：

提出了卷积块注意模块（CBAM），用于前馈卷积神经网络的简单而有效的注意模块。给定一个中间特征图，模块沿两个单独的维度（通道和空间）顺序推断注意力图，然后将注意力图乘以输入特征图以进行自适应特征细化。因为 CBAM 是一个轻量级的通用模块，它可以无缝集成到任何 CNN 架构中，开销可以忽略不计，并且可以与基础 CNN 一起进行端到端训练。

该模块有两个顺序子模块：通道和空间。中间特征图通过我们的模块（CBAM）在深度网络的每个卷积块上自适应地细化：

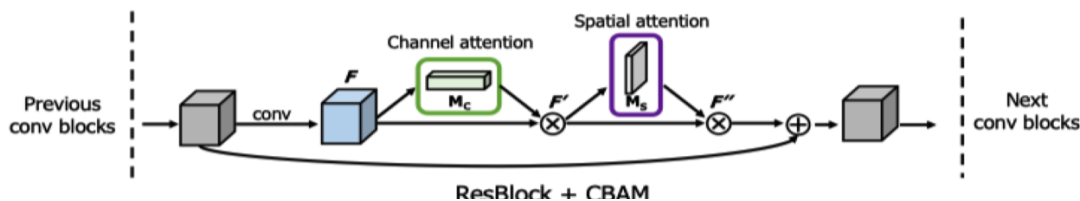


通道注意力模块:通过利用特征的通道间关系来生成通道注意力图。



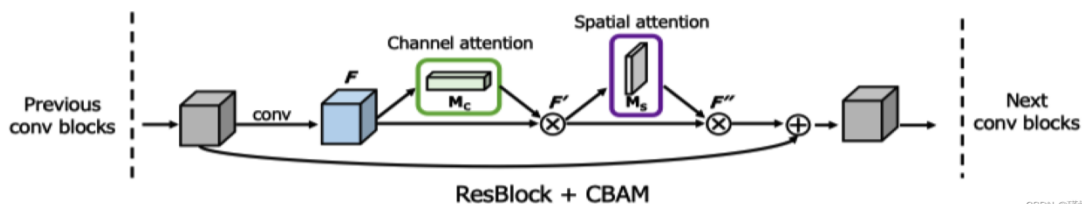
空间注意力模块:利用特征的空间间关系生成空间注意力图。与通道注意力不同，空间注意力集中在“哪里”是一个信息部分，这是对通道注意力的补充。

CBAM 与 Res Net 中的 Res Block 集成：此图显示模块在集成到 Res Block 时的确切位置，在每个块的卷积输出上应用 CBAM：



启发:

1. 通道注意力中使用最大池化特征和平均池化特征，从而产生比 SE 更好的注意力
2. 主要学习了一下本文 CBAM 的使用细节和代码，从而考虑在 DPN 的什么位置加入 Spatiotemporal self-attention 或者 CBAM 更好，例如：

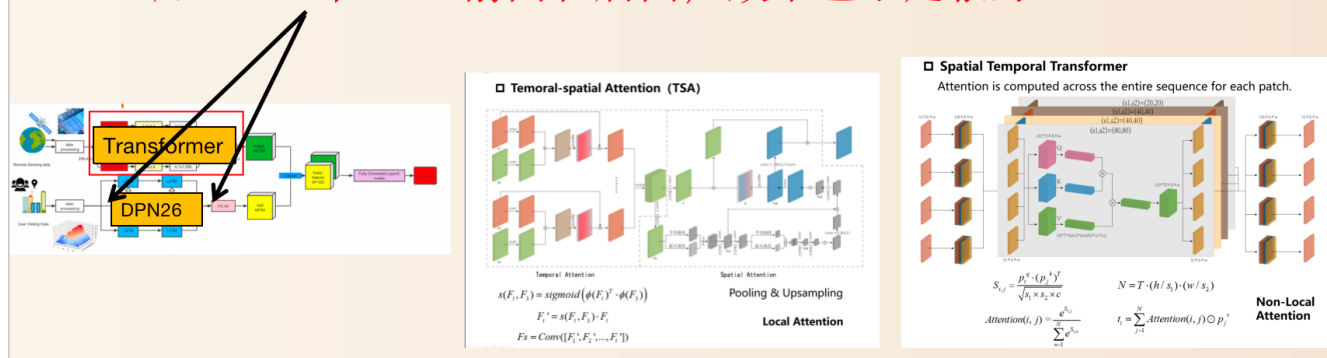


CSDN @理理研研

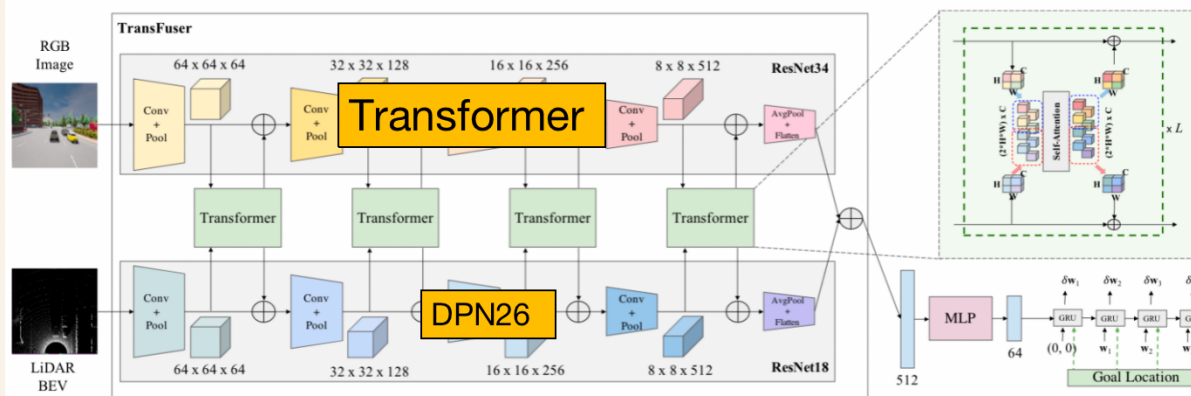
## 工作进展

1. 阅读文献；
2. URFC 实验结果：加 CBAM，效果也掉点，掉了 5 个点

- 加 CBAM 在 DPN 前面和后面，效果也不是很好

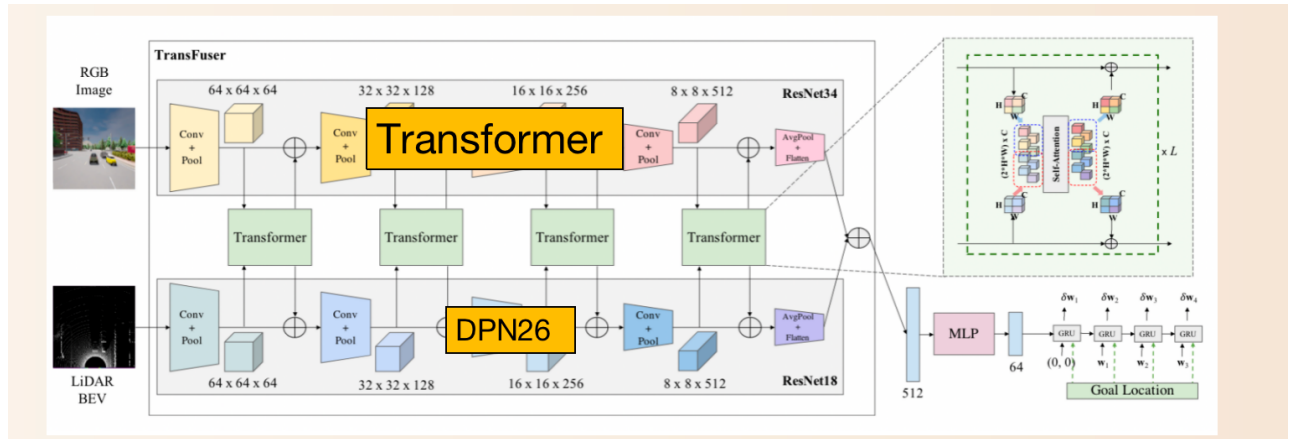


3. 在实现 GPT 结构的层间嵌入，代码没改完



## 下周计划

- 1: 觉得 **Spatial Attention** 代码放的位置可能有问题，准备放在 **DPN 内部**（DPN 有四层）。**CBAM** 准备也做一下相同的处理
- 2: 寻找 DPN+attention 论文
- 3: 实现 GPT 结构的层间嵌入代码



- 4: Spatial Attention代码的实现可能是有问题的，完全按照图像思路走的，还在找其他代码或者论文