

BAI

智源学者成果展示——机器学习方向

作者 颜水成（依图科技）

2020年6月

项目负责人：颜水成

Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution. ICCV-2019

自然图像中的信息是由不同频率的信号所构成，那些在空间上变换剧烈的细节与纹理等通常是由高频信号构成，而变化缓慢的结构与全局信息则是由低频信号构成。同样，在卷积神经网络中，卷积层中的特征图也可以看作是由不同频率的信号所构成的。在这项工作中，我们提出对卷积层中的特征按频率分解，并为此设计了一种新的算子，称为八分音符卷积 (OctConv)，目的是通过将变化缓慢的低频信号存储在分辨率较低的特征图中来减少空间冗余和计算开销，来更有效的处理不同频段的信息。不同于以往的多尺度的方法，八分音符卷积是一个即插即用的通用卷积算子，可以替换现有网络中的卷积层，并且无需修改网络结构。同一般卷积一样，八分音符卷积也可以使用“分组卷积”以及“深度卷积”等技术来进一步减少特征图之间的链接以达到稀疏和加速的目的。

实验证明，在二维图像识别以及三维视频分析中，八分音符卷积均可以稳定有效的提升卷积神经网络的识别精度并明显减少显存占用与计算开销。在 ImageNet 标准测试集上，八分音符卷积改造后的 ResNet-152 仅用 22.2 GFLOPs 就达到了 82.9% 的高识别率。

Beijing Academy of Artificial Intelligence



微信关注
北京智源人工智能研究院