

不需要进行图像预处理

VGG

VGG组成规律: 连续使用数个相同的填充为1、窗口形状为3x3的卷积层后接上一个步幅为2、窗口形状为2x2的最大池化层。卷积层保持输入的高和宽不变, 而池化层则对其减半。

VGG的名字来源于作者所在的**V**isual **G**eometry **G**roup实验室。该网络是在ILSVRC 2014上的相关工作, 主要工作是证明了增加网络的深度能够在一定程度上影响网络最终的性能。VGG有两种结构, 分别是VGG16和VGG19, 两者并没有本质上的区别, 只是网络深度不一样。提出了可以通过重复使用简单的基础块来构建深度模型的思路。

原理:在VGG中使用3个3x3卷积核代替7x7(ZFNet) 卷积核, 使用2个3x3卷积核来代替5x5(LeNet)卷积核, 在保证具有相同感知野的条件下, 提升了网络的深度, 减小了网络参数, 在一定程度上提升了神经网络的效果。

ResNet

2015年, 何凯明等人提出的残差网络(ResNet)在ImageNet图像识别大赛 夺冠, 并深刻影响了后来的深度学习网络的设计。ResNets解决深度神经网络的“退化”问题。

使用了残差链接:

- 设输入为 x , 如果期望的潜在映射为 $H(x)$, 可以拟合的函数为 $F(x)$ 。与其让 $F(x)$ 直接学习潜在的映射, 不如去学习残差 $H(x)-x$, 即 $F(x):=H(x)-x$, 这样原本的前向路径上就变成了 $F(x)+x$, 用 $F(x)+x$ 来拟合 $H(x)$ 。
- 残差映射在实际中更容易优化。
- 残差块中的输入可以通过跨层的数据线(短路链接, Shortcut)更快的向前传播。

ResNet采用残差块和Shortcut identity mapping, 取得了极其好的效果。