

背景

ResNeXt 首先提出了增加子模块的拓扑结构的数量(cardinality)比增加深度和宽度更加有效。DenseNet 设计的密集的特征重用策略也可以看作是增加子模块的拓扑数量。但是 DenseNet 的设计结构使得其在进行梯度反传更新参数的时候,出现了大量梯度信息被重复用来更新不同层的参数,导致了不同模块的参数是使用相同的梯度信息进行更新。

对于图像分类和检测任务来说,一些关键的信息被图像级的真值框信息给屏蔽掉了,而且网络从图像级标签中学习时效果并不好。

方法

①提出了跨阶段分层网络结构

提出该网络框架结构的主要目的是为了实现更加丰富的梯度组合,避免相同的梯度信息用于更新不同的网络层。

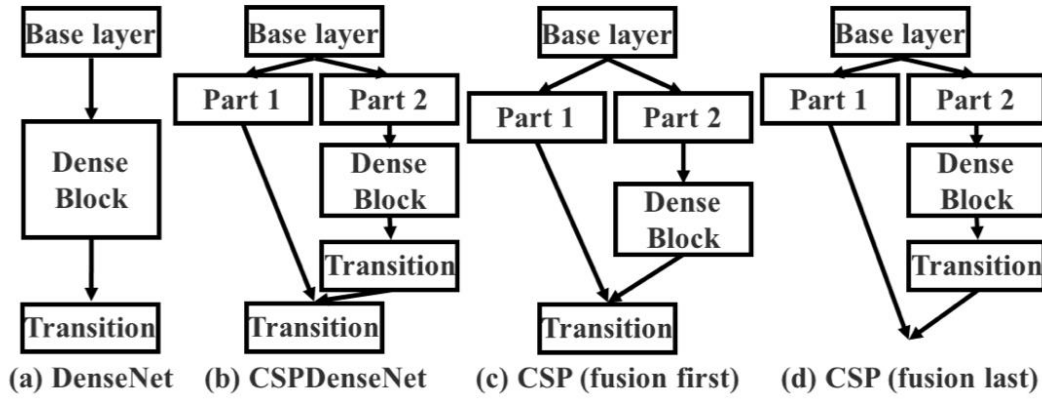


图 1 不同类型的特征融合策略

如上图中的(b)所示,将基础层的特征图 x_0 分成两路 $[x'_0, x''_0]$, x'_0 直接连接到该阶段的终点, x''_0 输入到 dense block 中,随后 dense block 的输出 $[x''_0, x_1, x_2, \dots, x_k]$ 再经过一个过渡层(transition layer),然后和 x'_0 进行 concatenate 操作最后再经过一个过渡层。

$$\begin{aligned} x_k &= w_k * [x_0'', x_1, \dots, x_{k-1}] \\ x_T &= w_T * [x_0'', x_1, \dots, x_k] \\ x_U &= w_U * [x_0', x_T] \end{aligned} \quad (7)$$

$$\begin{aligned} w'_k &= f_k(w_k, \{g_0'', g_1, \dots, g_{k-1}\}) \\ w'_T &= f_T(w_T, \{g_0'', g_1, \dots, g_k\}) \\ w'_U &= f_U(w_U, \{g_0', g_T\}) \end{aligned} \quad (8)$$

绿色框中的第一个公式是 dense block 的梯度信息,绿色框中的第二个公式是没有经过 dense block 的另一路梯度信息,可以看出两路梯度信息没有重复的

部分，通过这样分割梯度流的方式，避免了过多的重复梯度。

除了将原特征图信息分流之外，还将原 **densenet** 中的过渡层也进行了分割处理，提出了一种特征融合的策略。图 1(c)是先将两路的特征进行 **concatenate**，再进行转换，如果采用这种方式，会出现梯度重复利用的现象（**densenet** 就是先 **concat** 再转换）。图 1(d)是先将 **denseblock** 的输出先转换，再与第一路的特征图进行 **concat**，这样的策略提高了梯度组合的差异，使得梯度信息不再被重用，降低计算量。

②精确融合模型

精确融合模型是基于 **YOLOv3**，能够更好地聚合特征金字塔，**YOLOv3** 将特征图划分为三个不同尺度（13、26、52）的 **grid cell** 进行预测，每个 **grid cell** 上产生 3 个 **anchor box**，网络基于 **anchor box** 进行预测 **bounding box**，每个 **ground truth** 框都有一个与之匹配的预测框，且每个真值框都对应一个匹配度(IOU 值)最高的 **anchor box**。如果 **anchor box** 的大小与其对应的 **grid cell** 大小相等，那么对于第 s 尺度的网格来说，它里面的 **bounding box** 框的尺寸就是介于 $s+1$ 尺度与 $s-1$ 尺度之间。精确融合模型将三个尺度的特征组合了起来，可以为每一个 **anchor box** 获取一个合适的 **grid cell** 的尺寸。

由于特征金字塔，不同尺寸特征之家进行 **concat** 操作的计算量庞大，因此采用 **maxout** 技术解决这一问题。

总结

Partial dense block 的作用更像是“引流”，由于基础特征图通道数多，通过增加分支，将原来基础特征图通道数量一半输入到 **dense block** 中，另一半由于不经过 **dense block**，这样直接减少了计算量。

补充

✍ 假如设计了一个模块或者从别的论文里引进了一个模块，直接加进来发现没有什么效果，不要直接放弃，可以像这篇论文一样尝试不同的融合方式，例如先融合再怎样，或者先各自怎样再融合，或者两者直接串行

✍ 假如设计了一个模块或者引入的模块发现有效果，同样可以设计多组融合方式、多组不同的参数来增加工作量以及论文字数。