

第5章 VGG原理 与实战

A THE STATE OF THE PARTY OF THE

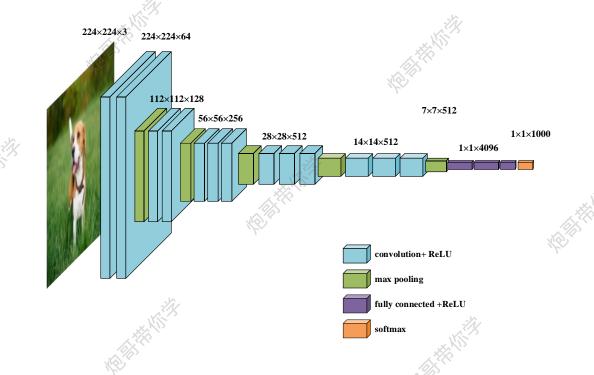
A CHANGE

720

VGG网络诞生背景

VGGNet是牛津大学计算机视觉组(Visual Geometry Group)和谷歌 DeepMind 一起研究出来的深度卷积神经网络,因而冠名为 VGG。VGG是一种被广泛使用的卷积神经网络结构,其在在2014年的 ImageNet 大规模视觉识别挑战(ILSVRC -2014)中获得了亚军,不是VGG不够强,而是对手太强,因为当年获得冠军的是GoogLeNet。

通常人们说的VGG是指VGG-16(13层卷积层+3层全连接层)。虽然其屈居亚军,但是由于其规律的设计、简洁可堆叠的卷积块,且在其他数据集上都有着很好的表现,从而被人们广泛使用,从这点上还是超过了GoogLeNet。VGG和之前的AlexNet相比,深度更深,参数更多(1.38亿),效果和可移植性更好。





VGG网络结构

VGGNet有6种不同结构,我们以通常所说的VGG-16(即图D列)为例。

如图,可以发现,VGG中卷积层是通过block块状形式相连的,block内的卷积层结构相同;block外,block之间通过maxpool连接。图中VGG-16中的一个vgg-block块:

conv3-256 conv3-256

其中conv3-256表示: 这是一个卷积层, 卷积核尺寸为3×3,通道数为256。

	X,'\'1'			
	ConvNet Co	onfiguration		
A-LRN	δ, B	С	D	E
11 weight	13 weight	16 weight	16 weight	19 weight
layers	layers	layers	layers	layers
i	nput (224×2	24 RGB image	e)	
conv3-64	conv3-64	conv3-64	conv3-64	conv3-64
î LRN	conv3-64	conv3-64	conv3-64	conv3-64
			-½v	
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128
	conv3-128	conv3-128	conv3-128	conv3-128
	max		26	
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
		conv1-256	conv3-256	conv3-256
				conv3-256
	max	pool		
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
		conv1-512	conv3-512	conv3-512
	vici),			conv3-512
	max	pool		
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
		conv1-512	conv3-512	conv3-512
-1/2				conv3-512
()	max	pool	-1/-	
(12				
			* 17 / 13	
	FC-	1000	400	
	soft.	may		
	11 weight layers in conv3-64 LRN conv3-128 conv3-256 conv3-256 conv3-256 conv3-512 conv3-512	A-LRN 13 weight layers input (224 × 2) conv3-64 conv3-64 conv3-64 conv3-64 conv3-128 conv3-128 conv3-128 conv3-256 conv3-256 conv3-256 conv3-256 conv3-512 c	11 weight 13 weight 16 weight layers layers 16 weight layers 16 weight layers 16 weight layers 16 weight layers 16 weight layers 16 weight layers 16 weight layers 16 weight layers 16 weight layers 16 weight 16 weigh	A-LRN

Table 2: Number of parameters (in millions).

Natavanle	A.A-LRN	D	C	D.	E
Network	A,A-LKN	Ь		ט	E
Number of parameters	133	133	134	138	144



VGG网络结构

经典卷积神经网络的基本组成部分是下面的这个序列:

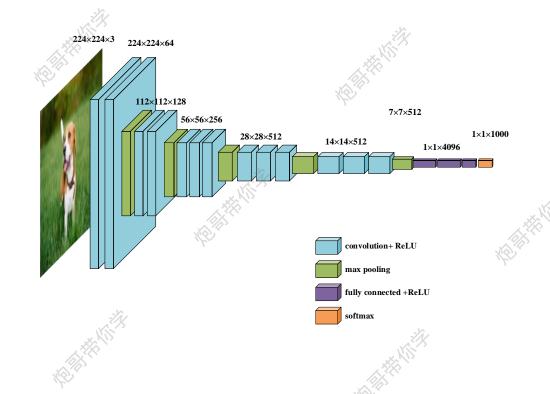
- (1)带填充以保持分辨率的卷积层;
- (2)非线性激活函数,如ReLU;
- (3)池化层,最大池化层。

而一个VGG块与之类似,由一系列卷积层组成,后面再加上用于空间下采样的最大池化层。

而一个VGG块与之类似,由一系列卷积层组成,后面再加上用于空间下采样的最大池化层。

VGG特点:

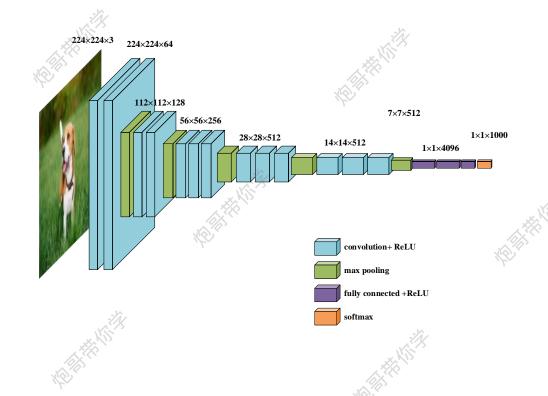
vgg-block内的卷积层都是同结构的 池化层都得上一层的卷积层特征缩减一半 深度较深,参数量够大 较小的filter size/kernel size





VGG网络结构

- 1、vgg-block内的卷积层都是同结构的意味着输入和输出的尺寸一样,且卷积层可以堆叠复用,其中的实现是通过统一的size为3×3的kernel size + stride1 + padding1实现。
- 2、maxpool层将前一层(vgg-block层)的特征缩减一半 使得尺寸缩减的很规整,从224-112-56-28-14-7。其中是通过pool size2 + stride2实现。
- 3、深度较深,参数量够大较深的网络层数使得训练得到的模型分类效果优秀,但是较大的参数对训练和模型保存提出了更大的资源要求。
- 4、较小的filter size/kernel size **这里全局的kernel size都为3×3,相比以前的网络模型来说,尺寸足够小。





第1个vgg block层:

- (1) 输入为224×224×3, 卷积核数量为64个; 卷积核的尺寸大小为3×3×3; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为224×224×64的特征图输出。
- (2) 输入为224×224×64, 卷积核数量为64个; 卷 积核的尺寸大小为3×3×64; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为224×224×64的特征图输出。
- (3) 输入为224×224×64, 池化核为2×2, 步幅为2 (stride = 2) 后得到尺寸为112×112×64的池化层的特征图输出。

	V.	ConvNet C	onfiguration		
Α	A-LRN	В	C	D	E
11 weight	11 weight	13 weight	16 weight	16 weight	19 weight
layers	layers	layers	layers	layers	layers
		nput (224×2			
conv3-64	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64
	LRN	conv3-64	conv3-64	conv3-64	conv3-64
,×//> `			pool	THE STATE OF THE S	
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128
1/60		conv3-128	conv3-128	conv3-128	conv3-128
			pool 🎇		
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
			conv1-256	conv3-256	conv3-256
					conv3-250
		V . V .	pool		
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
			conv1-512	conv3-512	conv3-512
		1/20			conv3-512
			pool		*
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
	-14		conv1-512	conv3-512	conv3-512
	<i>:</i> 4%			10	conv3-512
***	12		pool	7.1%	
			4096	75/12	
KO, C			4096	65	
1			1000	N/O	·
		soft	-max	1	



第2个vgg block层:

- (1) 输入为112×112×64, 卷积核数量为128个; 卷积核的尺寸大小为3×3×64; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape 为112×112×128的特征图输出。
- (2) 输入为112×112×128, 卷积核数量为128个; 卷积核的尺寸大小为3×3×128; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape 为112×112×128的特征图输出。
- (3) 输入为112×112×128, 池化核为2×2, 步幅为2 (stride = 2) 后得到尺寸为56×56×128的池化层的特征图输出。

	J. 4/2>				. 25/25. `			
	1/6,	ConvNet Co	onfiguration		4/5			
A	A-LRN	В	С	D	₩E			
11 weight	11 weight	13 weight	16 weight	16 weight	19 weight			
layers	layers	layers	layers	layers	layers			
input (224×224 RGB image)								
conv3-64	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64			
XX (II)	LRN	conv3-64	conv3-64	conv3-64	conv3-64			
			pool 👸	1/12.				
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128			
		conv3-128	conv3-128	conv3-128	conv3-128			
			pool					
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256			
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256			
			conv1-256	conv3-256	conv3-256			
		-1/2/			conv3-256			
		1.17	pool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512			
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512			
	1/6	*	conv1-512	conv3-512	conv3-512			
	•				conv3-512			
			pool		,			
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512			
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512			
://x			conv1-512	conv3-512	conv3-512			
X Kis				-//	conv3-512			
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\		max	pool					
			4096					
1			4096					
			1000					
		soft-	-max					



第3个vgg block层:

- (1) 输入为56×56×128, 卷积核数量为256个; 卷积核的尺寸大小为3×3×128; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为56×56×256的特征图输出。
- (2) 输入为56×56×256, 卷积核数量为256个; 卷积核的尺寸大小为3×3×256; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为56×56×256的特征图输出。
- (3) 输入为56×56×256, 卷积核数量为256个; 卷积核的尺寸大小为3×3×256; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为56×56×256的特征图输出。
- (4) 输入为56×56×256, 池化核为2×2, 步幅为2 (stride = 2) 后得到尺寸为28×28×256的池化层的特征 图输出。



第4个vgg block层:

- (1) 输入为28×28×256, 卷积核数量为512个; 卷积核的尺寸大小为3×3×256; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为28×28×512的特征图输出。
- (2) 输入为28×28×512, 卷积核数量为512个; 卷积核的尺寸大小为3×3×512; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为28×28×512的特征图输出。
- (3) 输入为28×28×512, 卷积核数量为512个; 卷积核的尺寸大小为3×3×512; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为28×28×512的特征图输出。
- (4) 输入为28×28×512, 池化核为2×2, 步幅为2 (stride = 2) 后得到尺寸为14×14×512的池化层的特征 图输出。

	4(4),		-		\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\		
	.7-	ConvNet Co	onfiguration		10), '		
A	A-LRN	В	C	D	E		
11 weight	11 weight	13 weight	16 weight	16 weight	19 weight		
layers	layers	layers	layers	layers	layers		
$224 \times 224 \text{ RGB image}$							
conv3-64	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64		
	LRN	conv3-64	conv3-64	conv3-64	conv3-64		
>		max					
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128		
		conv3-128	conv3-128	conv3-128	conv3-128		
			pool				
	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256		
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256		
		10	conv1-256	conv3-256	conv3-256		
		7/1/			conv3-256		
			pool				
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512		
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512		
	1/2		conv1-512	conv3-512	conv3-512		
					conv3-512		
			pool				
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512		
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512		
W. S.			conv1-512	cony3-512	conv3-512		
XXX XXX					conv3-512		
4/6			pool	XXXXXX			
\$\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\			4096 <i>(</i>)				
			4096				
		FC-	1000				
•			max				



第5个vgg block层:

- (1) 输入为14×14×512, 卷积核数量为512个; 卷积核的尺寸大小为3×3×512; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为14×14×512的特征图输出。
- (2) 输入为14×14×512, 卷积核数量为512个; 卷积核的尺寸大 小为3×3×512; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为14×14×512的特征图输出。
- (3) 输入为14×14×512, 卷积核数量为512个; 卷积核的尺寸大小为3×3×512; 步幅为1 (stride = 1), 填充为1 (padding=1); 卷积后得到shape为14×14×512的特征图输出。
- (4) 输入为14×14×512, 池化核为2×2, 步幅为2 (stride = 2) 后得到尺寸为7×7×512的池化层的特征图输出。该层后面还隐藏了flatten操作, 通过展平得到7×7×512=25088个参数后与之后的全连接层相连。
- **第1~3层全连接层:** 第1~3层神经元个数分别为4096, 4096, 1000。其中前两层在使用relu后还使用了Dropout对神经元随机失活, 最后一层全连接层用softmax输出1000个分类。

	1		_		WALL TO THE PARTY OF THE PARTY			
		ConvNet C	onfiguration		iko			
Α	A-LRN	В	С	D	Е			
11 weight	11 weight	13 weight	16 weight	16 weight	19 weight			
layers	layers	layers	layers	layers	layers			
input (224 × 224 RGB image)								
conv3-64	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64			
55	LRN	conv3-64	conv3-64	conv3-64	conv3-64			
			pool					
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128			
		conv3-128	conv3-128	conv3-128	conv3-128			
			pool					
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256			
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256			
		.7/X	conv1-256	conv3-256	conv3-256			
		* KI)			conv3-256			
	6	_	pool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512			
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512			
			conv1-512	conv3-512	conv3-512			
					conv3-512			
			pool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512			
conv3-512	conv3-512	conv3-512	conv3-512	cony3-512	conv3-512			
****			conv1-512	conv3-512	conv3-512			
6/5				the first	conv3-512			
			pool					
1			4096					
			4096					
		FC-	1000					
		soft-	-max					



VGGNet总结

VGGNet总结

VGGNet通过在传统卷积神经网络模型(AlexNet)上的拓展, 发现除了较为复杂的模型结构的设计(如GoogLeNet)外,深度对 于提高模型准确率很重要,VGG和之前的AlexNet相比,深度更 深,参数更多(1.38亿),效果和可移植性更好,且模型设计的简 洁而规律,从而被广泛使用。还有一些特点总结如下:

- 1、小尺寸的filter(3×3)不仅使参数更少,效果也并不弱于大尺寸filter如5×5
- 2、块的使用导致网络定义的非常简洁。使用块可以有效地设计 复杂的网络。
- 3、AlexNet中的局部响应归一化作用不大

