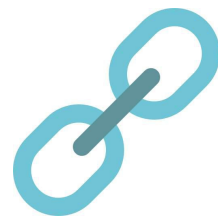TensorFlow

PYTORCH

bilibili: 霹雳吧啦Wz

深度学习-图像处理篇

作者：神秘的wz

GoogLeNet详解

GoogLeNet在2014年由Google团队提出，斩获当年ImageNet竞赛中Classification Task (分类任务) 第一名。

Going deeper with convolutions

Christian Szegedy
Google Inc.

Wei Liu
University of North Carolina, Chapel Hill

Yangqing Jia
Google Inc.

Pierre Sermanet
Google Inc.

Scott Reed
University of Michigan

Dragomir Anguelov
Google Inc.

Dumitru Erhan
Google Inc.

**Vincent Vanhoucke**
Google Inc.

Andrew Rabinovich
Google Inc.

网络中的亮点：

➤ 引入了**Inception**结构（融合不同尺度的特征信息）
➤ 使用**1x1**的卷积核进行降维以及映射处理
➤ 添加两个辅助分类器帮助训练
➤ 丢弃全连接层，使用平均池化层（大大减少模型参数）

GoogLeNet详解

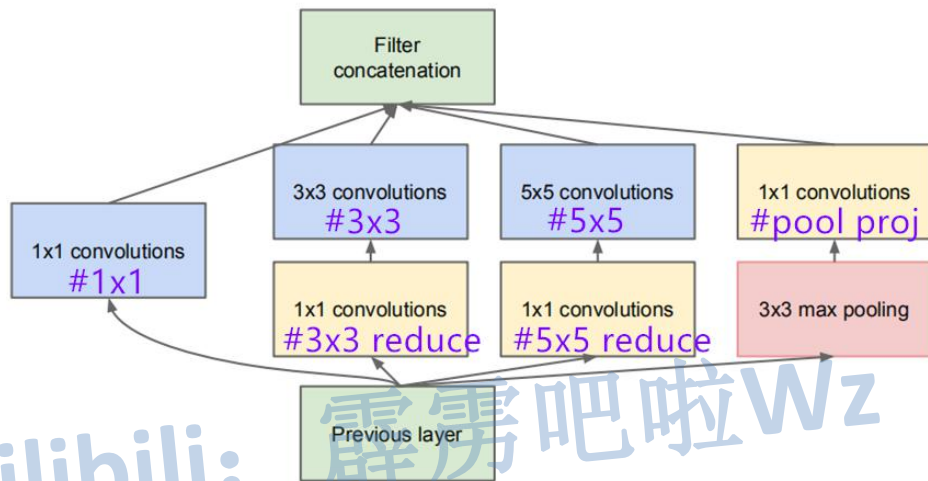| type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								
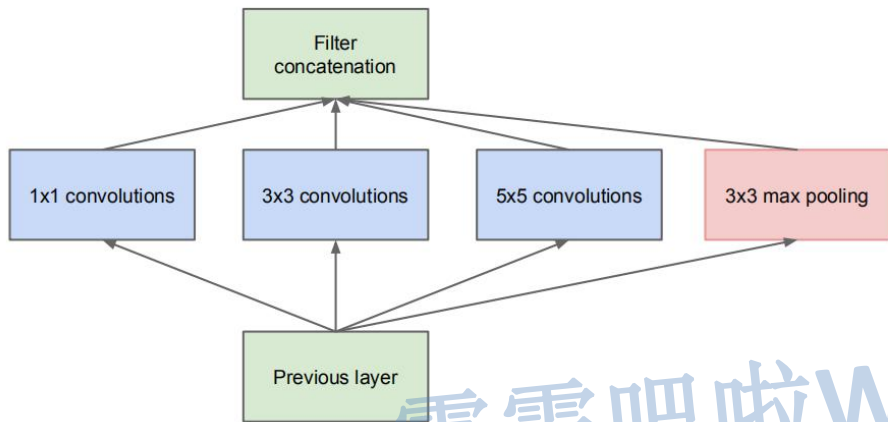
Inception结构

(a) Inception module, naïve version

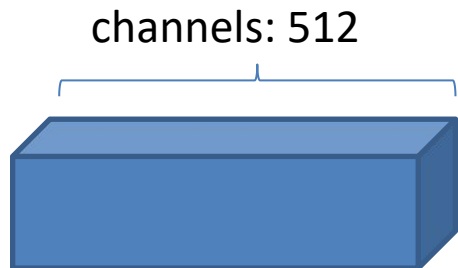(b) Inception module with dimension reductions

注意：每个分支所得的特征矩阵高和宽必须相同
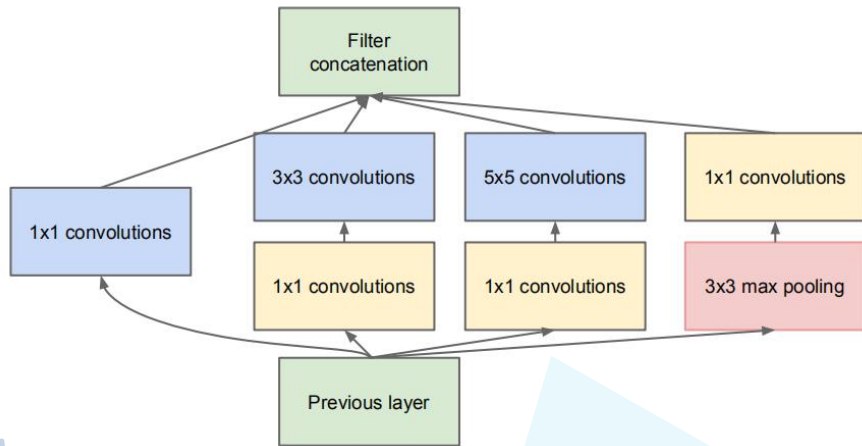
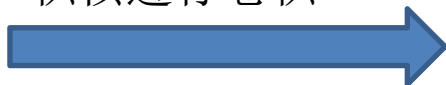Inception结构

(b) Inception module with dimension reductions

注意：每个分支所得的特征矩阵高和宽必须相同

GoogLeNet详解

不使用1x1卷积核降维

channels: 512

使用64个5x5的卷积核进行卷积

$$5 \times 5 \times 512 \times 64 = 819\,200$$

channels: 512

channels: 24

使用24个1x1的卷积核进行卷积

使用64个5x5的卷积核进行卷积

$$1 \times 1 \times 512 \times 24 = 12\,288$$

$$5 \times 5 \times 24 \times 64 = 38\,400$$

$$12\,288 + 38\,400 = 50\,688$$

使用1x1卷积核降维

GoogLeNet详解

辅助分类器（Auxiliary Classifier）

The exact structure of the extra network on the side, including the auxiliary classifier, is as follows:

- An average pooling layer with 5×5 filter size and stride 3, resulting in an $4 \times 4 \times 512$ output for the (4a), and $4 \times 4 \times 528$ for the (4d) stage.
- A 1×1 convolution with 128 filters for dimension reduction and rectified linear activation.
- A fully connected layer with 1024 units and rectified linear activation.
- A dropout layer with 70% ratio of dropped outputs.
- A linear layer with softmax loss as the classifier (predicting the same 1000 classes as the main classifier, but removed at inference time).

bilibili：霹雳吧啦Wz
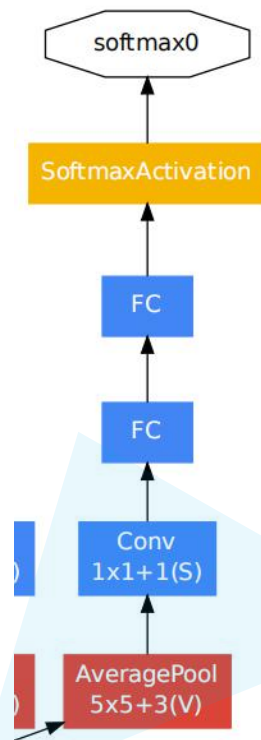
$$out_{size} = \left(in_{size} - F_{size} + 2P \right) / S + 1$$

GoogLeNet详解

GoogLeNet模型参数（不包含辅助分类器）

```
Total params: 6,994,392
Trainable params: 6,994,392
Non-trainable params: 0
```

VGGNet模型参数

```
Total params: 138,357,544
Trainable params: 138,357,544
Non-trainable params: 0
```

$$\frac{6\,994\,392}{138\,357\,544} \approx \frac{1}{20}$$

沟通方式

1.github

https://github.com/WZMIAOMIAO/deep-learning-for-image-processing

2.CSDN

https://blog.csdn.net/qq_37541097/article/details/103482003

3.bilibili

bilibili: 霹雳吧啦Wz

https://www.bilibili.com/video/av79436317

尽可能每周更新

感谢各位的观看！