深度之眼
deepshare.net

Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning

导师：余老师

深度之眼
deepshare.net

第二课：论文精读

The second lesson: the paper details

Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning

Inception-v4, Inception-ResNet，残差连接对模型训练的影响

作者：Christian Szegedy , Sergey Ioffe，Vincent Vanhoucke

单位：Google Inc.

发表会议及时间：AAAI 2017

深度之眼
deepshare.net

目录

上节回顾

Review in the previous lesson

上节回顾

Review in the previous lesson

深度之眼
deepshare.net

知识回顾

研究背景：Inception-v1、v2、v3和ResNet

研究成果及意义：超越ResNet，top-5 error低至3.1%

本文摘要：研究背景1，研究背景2，提出问题，本文成果1，本文成果2，本文成果3

本文图表： 18张图展示3个模型具体的结构，4个模型的精度对比


深度之眼
deepshare.net
Inception-V4
Architecture of Inception-V4

Inception-V4

Architecture of Inception-V4

Inception-v4

Inception-v4可分为六大模块分别是：

Stem、Inception-A、B、C、Reduction-A、B

每个模块都有针对性的设计，模型总共76层

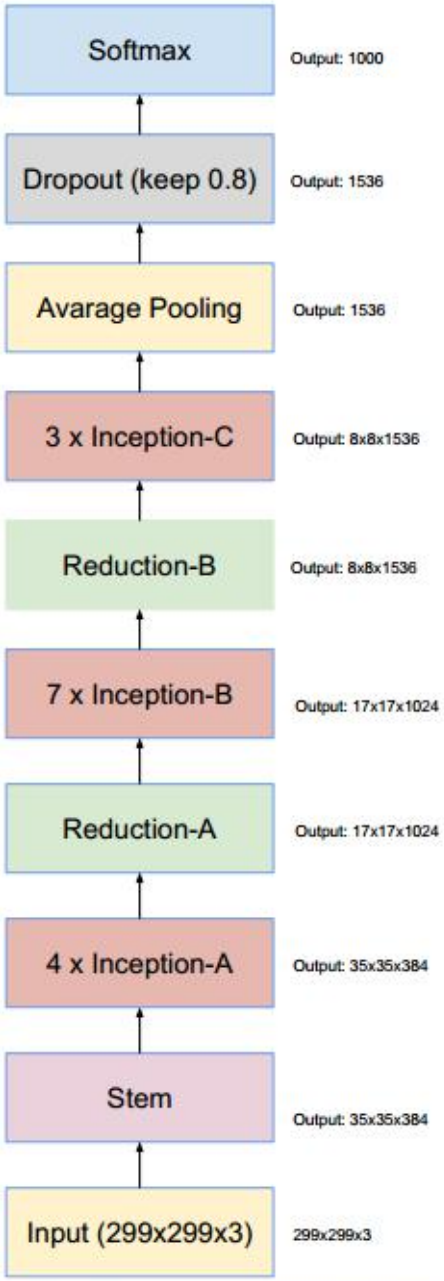


Figure 9. The overall schema of the Inception-v4 network. For the detailed modules, please refer to Figures 3, 4, 5, 6, 7 and 8 for the detailed structure of the various components.

Inception-V4

Architecture of Inception-V4

Inception-v4

Inception-v4可分为六大模块分别是：

Stem、Inception-A、B、C、Reduction-A、B

Stem（9层）：3个3*3卷积堆叠；高效特征图下降策略（参见InceptionV3）；非对称分解卷积

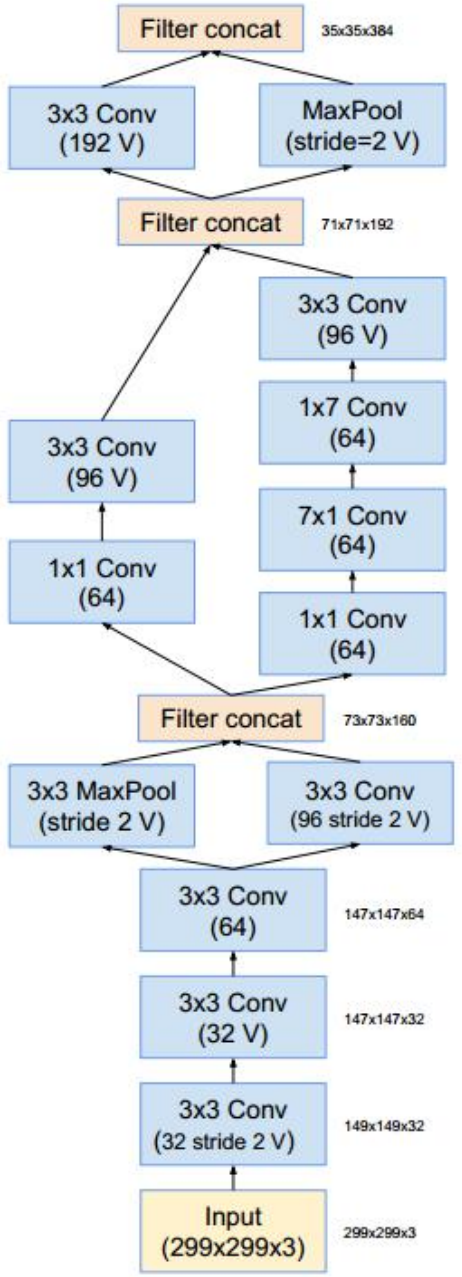


Figure 3. The schema for stem of the pure Inception-v4 and Inception-ResNet-v2 networks. This is the input part of those networks. Cf. Figures 9 and 15

Inception-V4

Architecture of Inception-V4

Inception-v4

Inception-v4可分为六大模块分别是：

Stem（9层）、Inception-A、B、C、Reduction-A、B

Inception-A（3层）：标准的Inception module

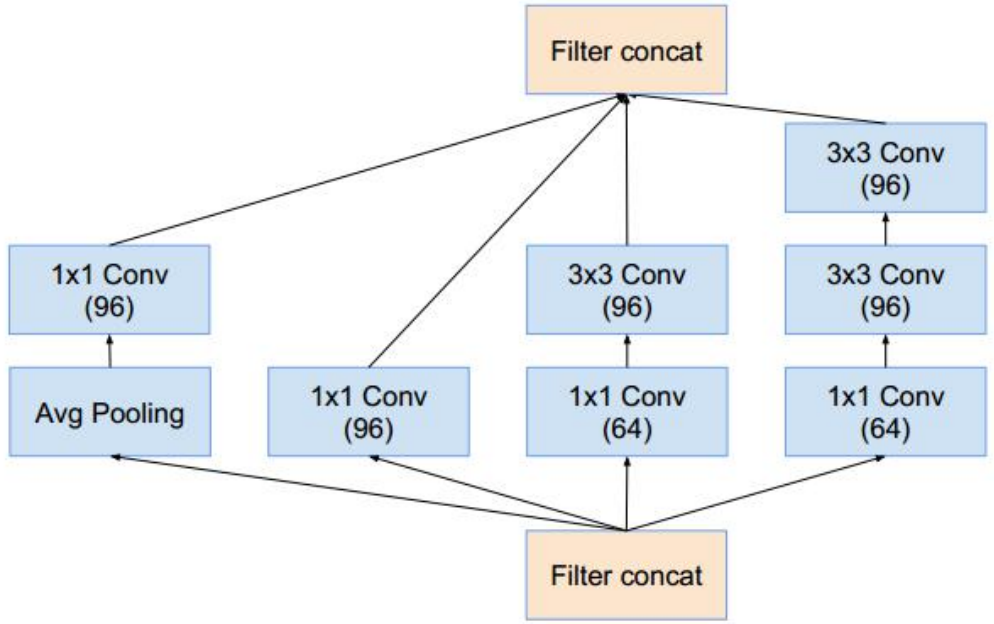


Figure 4. The schema for 35 × 35 grid modules of the pure Inception-v4 network. This is the Inception-A block of Figure 9.

Inception-V4

Architecture of Inception-V4

Inception-v4

Inception-v4可分为六大模块分别是：

Stem（9层）、Inception-A（3*4层）、B、C、Reduction-A、B

Reduction-A（3层）：采用3个分支，其中卷积核的参数K，l，m，n分别为192， 224，256，384

Network	k	l	m	n
Inception-v4	192	224	256	384
Inception-ResNet-v1	192	192	256	384
Inception-ResNet-v2	256	256	384	384

Table 1. The number of filters of the Reduction-A module for the three Inception variants presented in this paper. The four numbers in the colums of the paper parametrize the four convolutions of Figure 7

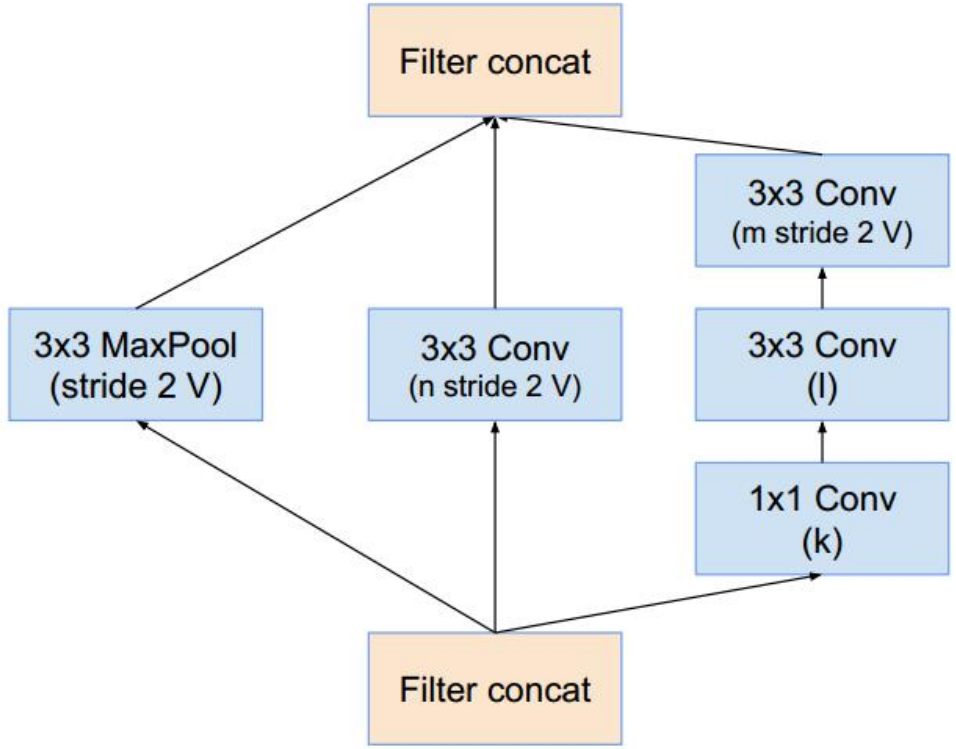


Figure 7. The schema for 35×35 to 17×17 reduction module. Different variants of this blocks (with various number of filters) are used in Figure 9, and 15 in each of the new Inception(-v4, -ResNet-v1, -ResNet-v2) variants presented in this paper. The k, l, m, n numbers represent filter bank sizes which can be looked up in Table 1.

Inception-V4

Architecture of Inception-V4

Inception-v4

Inception-v4可分为六大模块分别是：

Stem（9层）、Inception-A（3*4层）、B、C、Reduction-A（3层）、B

Inception-B（5层）：非对称卷积操作部分，参考Inception-v3

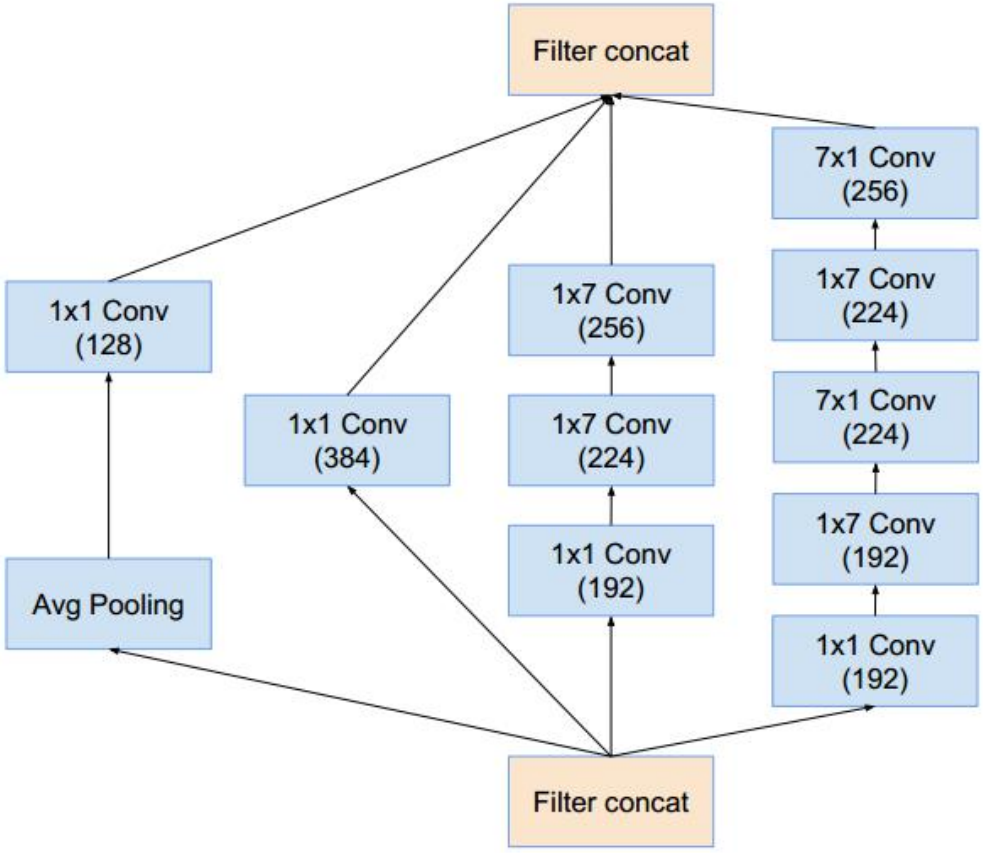


Figure 5. The schema for 17 × 17 grid modules of the pure Inception-v4 network. This is the Inception-B block of Figure 9.

Inception-V4

Architecture of Inception-V4

Inception-v4

Inception-v4可分为六大模块分别是：

Stem（9层）、Inception-A（3*4层）、B（5*7）、C、Reduction-A（3层）、B

Reduction-B（4层）：非对称卷积操作部分，参考Inception-v3

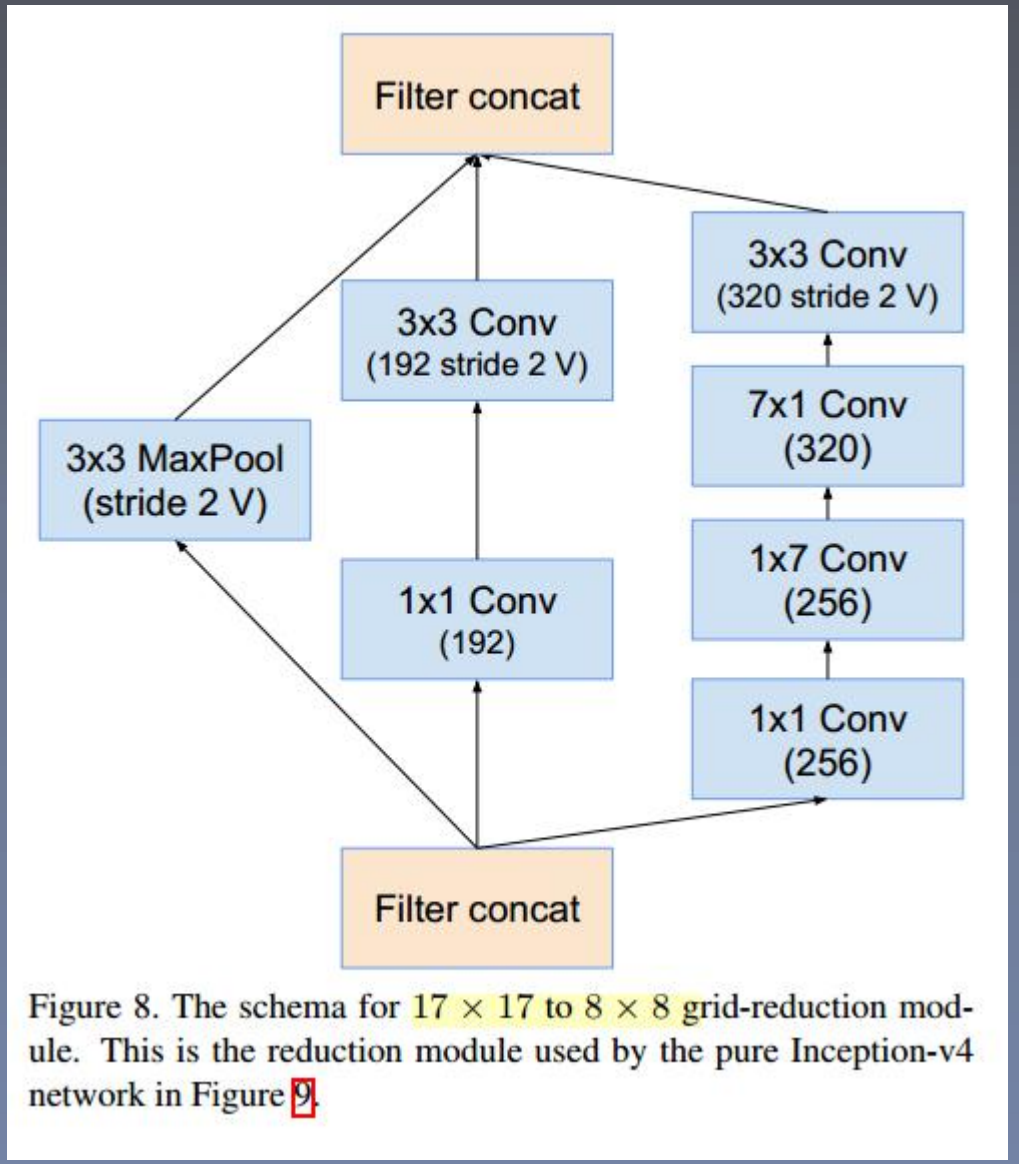


Figure 8. The schema for 17 × 17 to 8 × 8 grid-reduction module. This is the reduction module used by the pure Inception-v4 network in Figure 9.

Inception-V4

Architecture of Inception-V4

Inception-v4

Inception-v4可分为六大模块分别是：

Stem（9层）、Inception-A（3*4层）、B（5*7）、C、Reduction-A（3层）、B（4层）

Inception-C（4层） ：结构参考自Inception-v3

Inception-v4总共9+3*4+5*7+4*3+3+4+1 = 76层

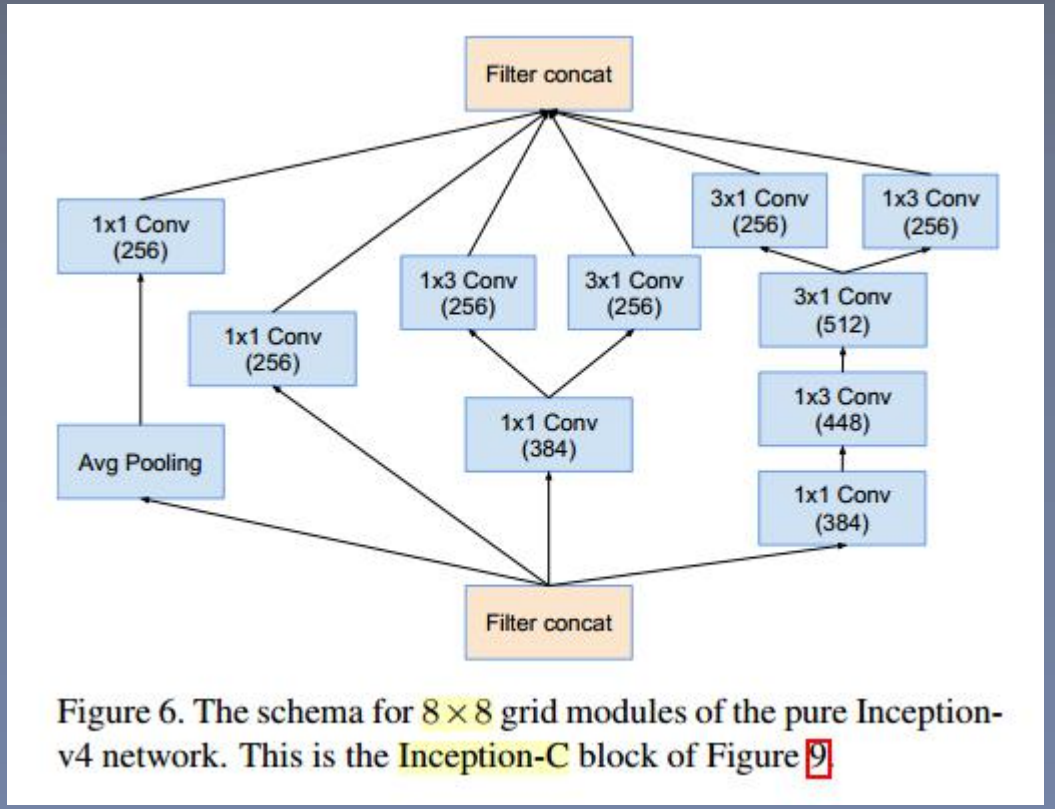


Figure 6. The schema for 8 × 8 grid modules of the pure Inception-v4 network. This is the Inception-C block of Figure 9.

深度之眼
deepshare.net

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

将ResNet中的residual connection思想加到Inception中

根据Stem和卷积核数量的不同，设计出了Inception-ResNet-V1和V2

六大模块分别是：

Stem、Inception-ResNetA、B、C、Reduction-A、B

两个模型的六大模块分别对应

图14， 图10， 图7， 图11， 图12， 图13

图3， 图16， 图7， 图17， 图18， 图19

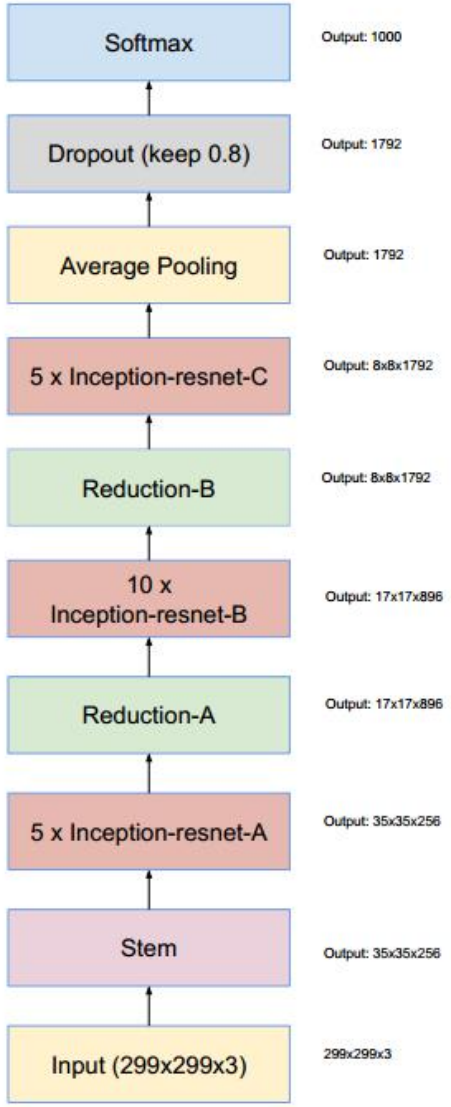


Figure 15. Schema for Inception-ResNet-v1 and Inception-ResNet-v2 networks. This schema applies to both networks but the underlying components differ. Inception-ResNet-v1 uses the blocks as described in Figures 14, 10, 7, 11, 12 and 13. Inception-ResNet-v2 uses the blocks as described in Figures 3, 16, 7, 17, 18 and 19. The output sizes in the diagram refer to the activation vector tensor shapes of Inception-ResNet-v1.

Inception-ResNe

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

Stem模块:

V1无分支（7层）；

V2与Inception-V4相同（9层）

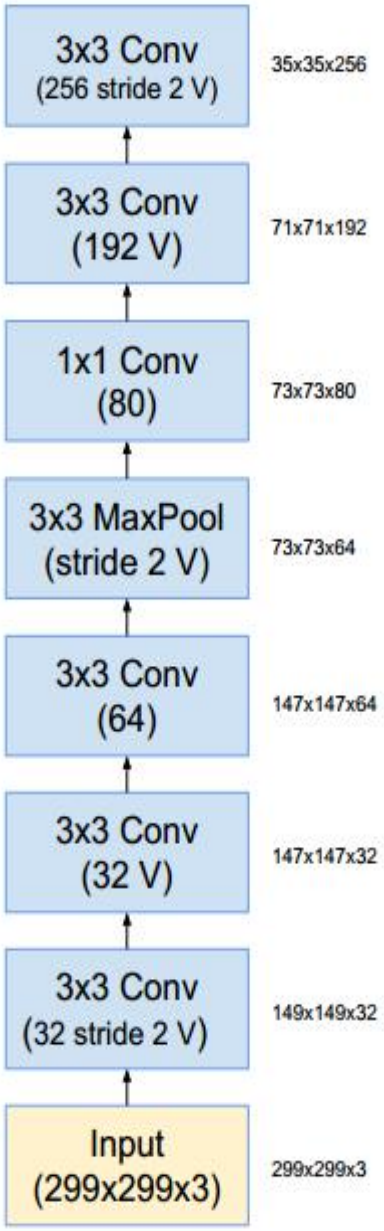


Figure 14. The stem of the Inception-ResNet-v1 network.

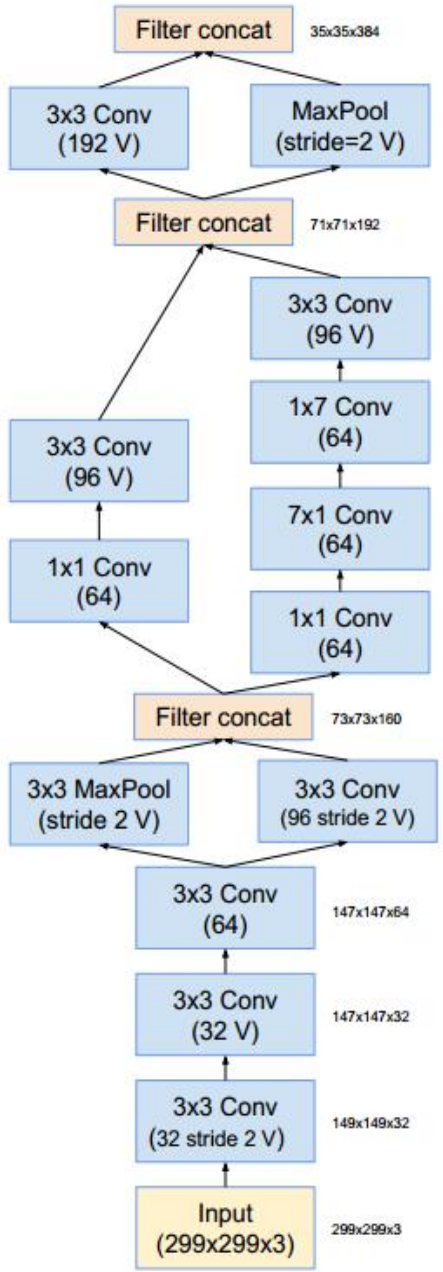


Figure 3. The schema for stem of the pure Inception-v4 and Inception-ResNet-v2 networks. This is the input part of those networks. Cf. Figures 9 and 15

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

Inception-ResNet-A模块（4层）：

均处理35*35大小的特征图

V1卷积核数量少

V2卷积核数量多

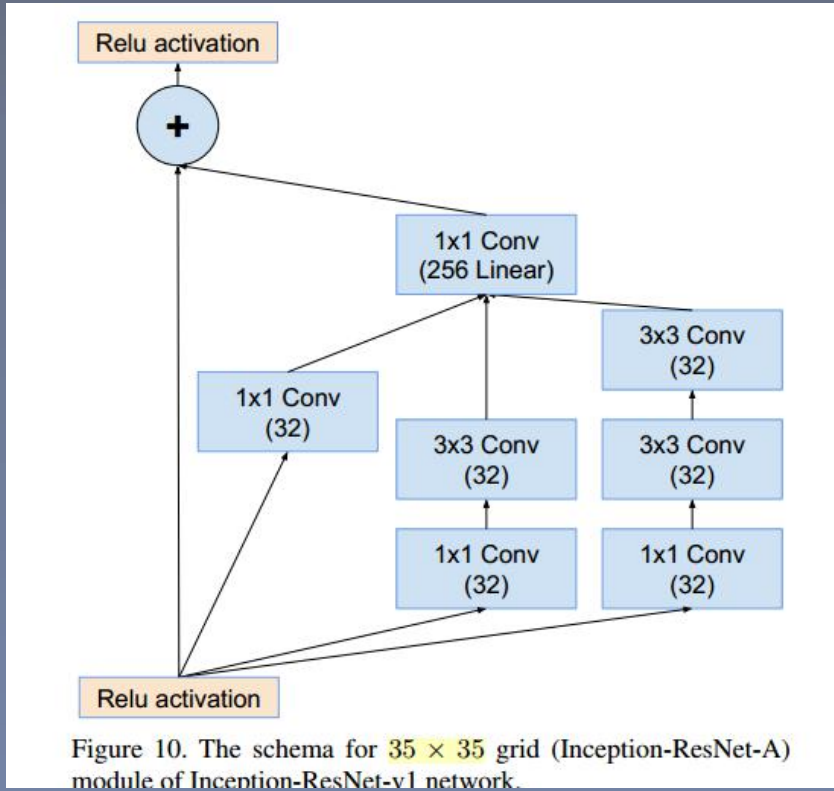


Figure 10. The schema for 35 × 35 grid (Inception-ResNet-A) module of Inception-ResNet-v1 network.

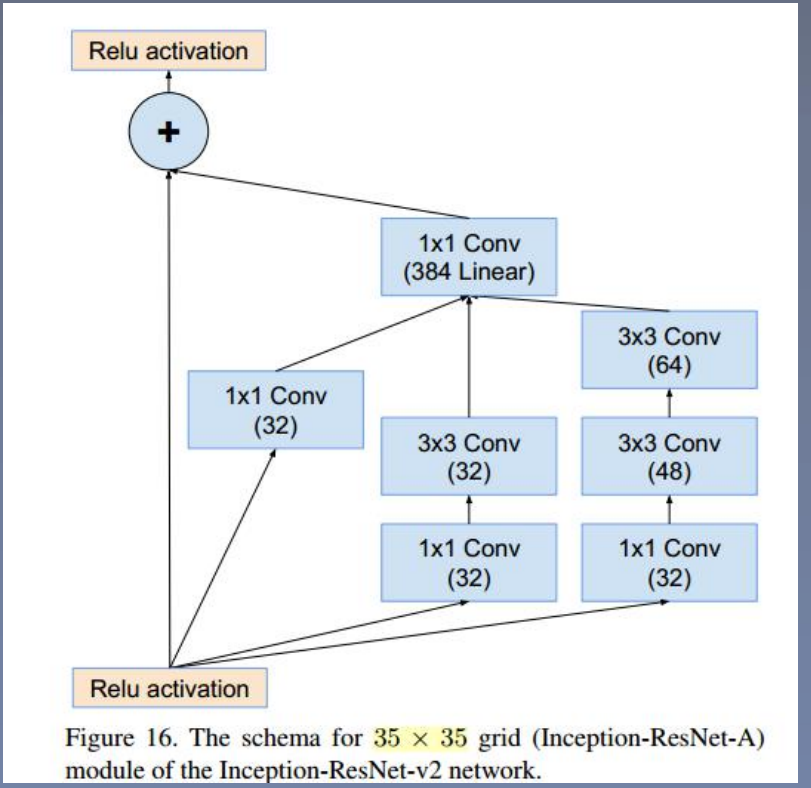


Figure 16. The schema for 35 × 35 grid (Inception-ResNet-A) module of the Inception-ResNet-v2 network.

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

Reduction-A模块（3层）：

将35*35大小的特征图降低至17*17

Inception-V4和两个Inception-ResNet都一样

Network	k	l	m	n
Inception-v4	192	224	256	384
Inception-ResNet-v1	192	192	256	384
Inception-ResNet-v2	256	256	384	384

Table 1. The number of filters of the Reduction-A module for the three Inception variants presented in this paper. The four numbers in the colums of the paper parametrize the four convolutions of Figure 7

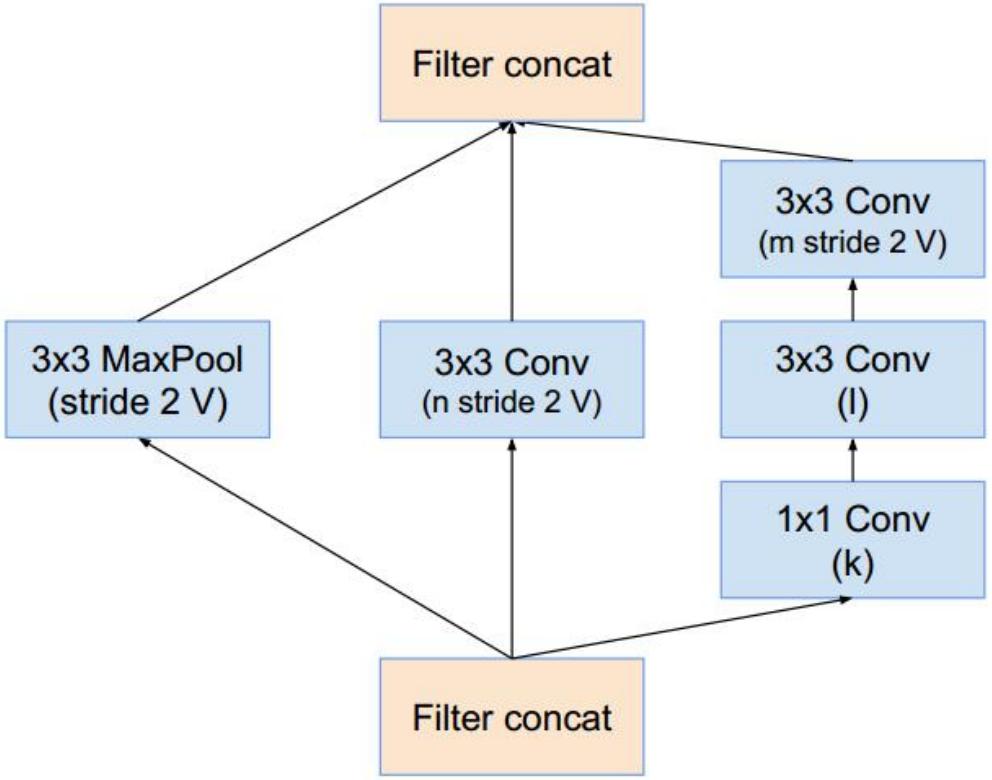


Figure 7. The schema for 35×35 to 17×17 reduction module. Different variants of this blocks (with various number of filters) are used in Figure 9, and 15 in each of the new Inception(-v4, -ResNet-v1, -ResNet-v2) variants presented in this paper. The k, l, m, n numbers represent filter bank sizes which can be looked up in Table 1.

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

Inception-ResNet-B模块（4层）：

处理17*17大小的特征图

V1卷积核数量少

V2卷积核数量多

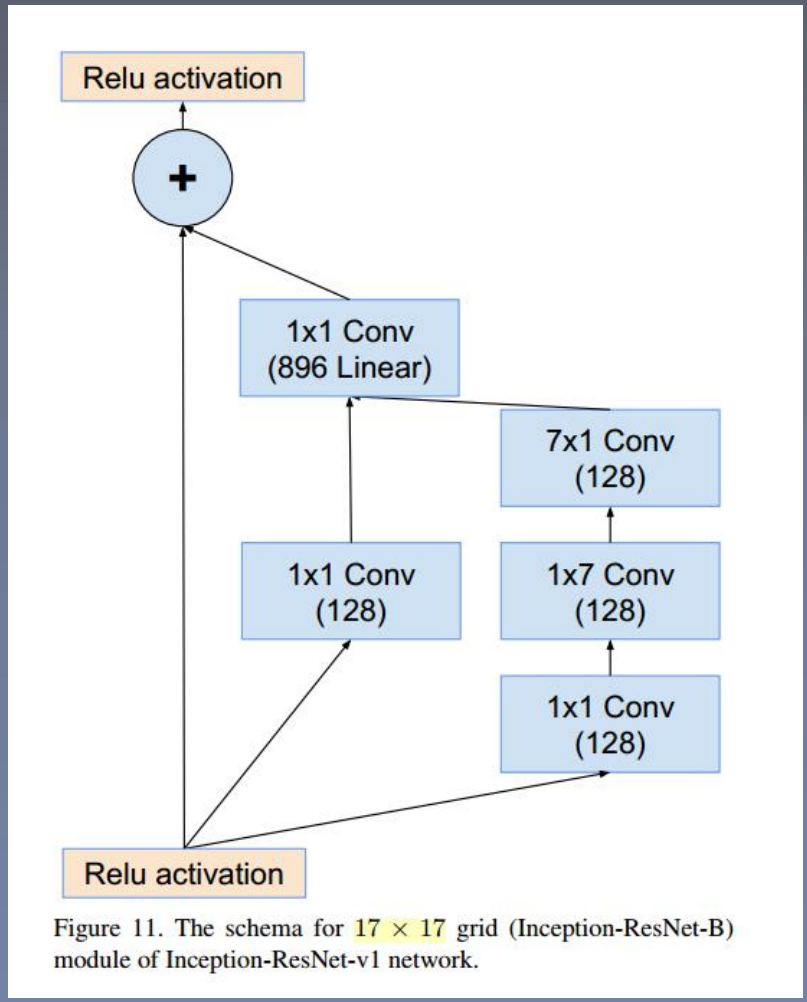


Figure 11. The schema for 17 × 17 grid (Inception-ResNet-B) module of Inception-ResNet-v1 network.

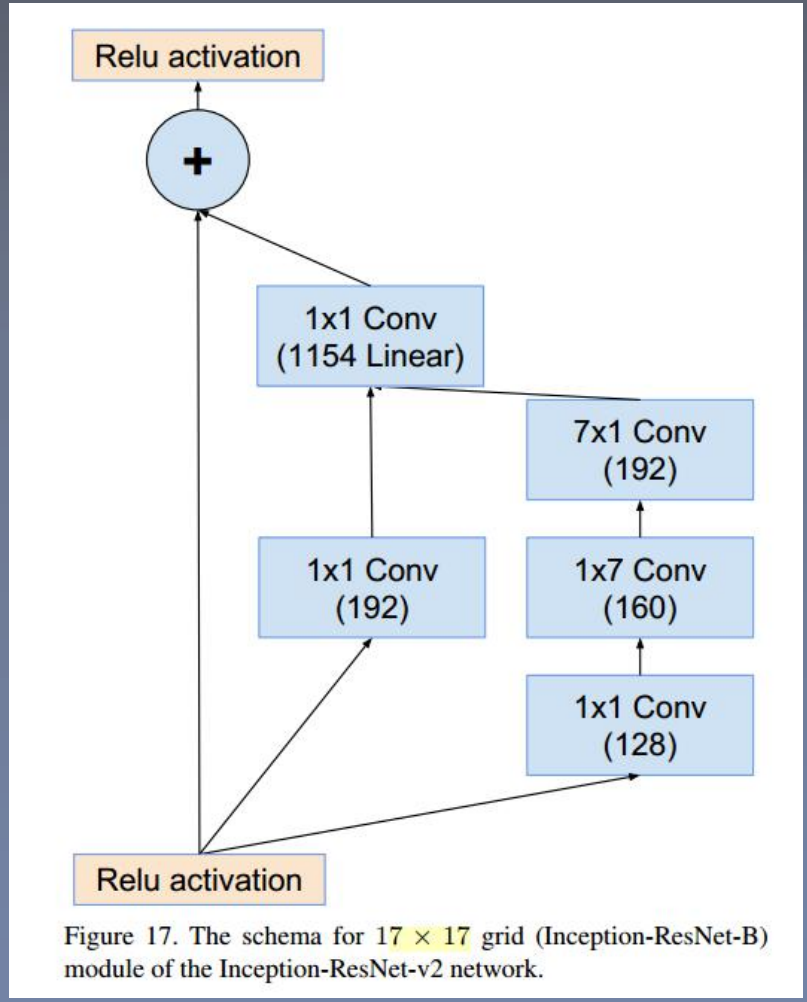


Figure 17. The schema for 17 × 17 grid (Inception-ResNet-B) module of the Inception-ResNet-v2 network.

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

Reduction-B模块（3层）：

将17*17大小的特征图降低至7*7

图12：V1

图18：V2，文中描述为 the wider Inception-ResNet-v1

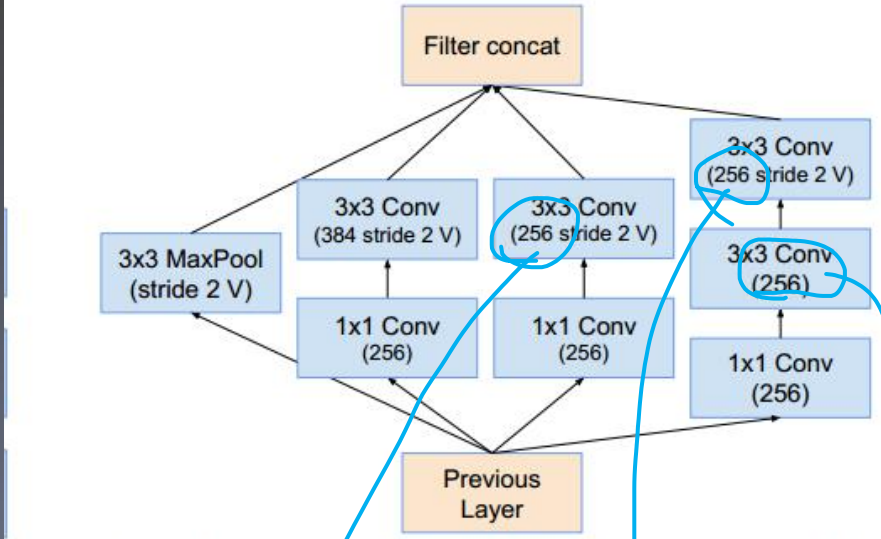


Figure 12. "Reduction-B" 17×17 to 8×8 grid-reduction module. This module used by the smaller Inception-ResNet-v1 network in Figure 15.

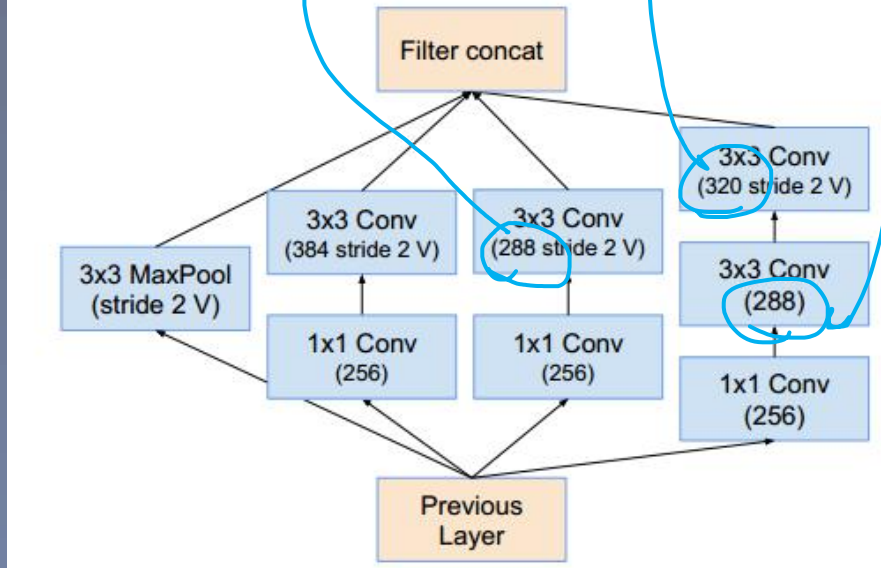


Figure 18. The schema for 17×17 to 8×8 grid-reduction module. Reduction-B module used by the wider Inception-ResNet-v1 network in Figure 15.

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

Inception-ResNet-C模块（4层）：

处理8*8大小的特征图

V1卷积核数量少

V2卷积核数量多

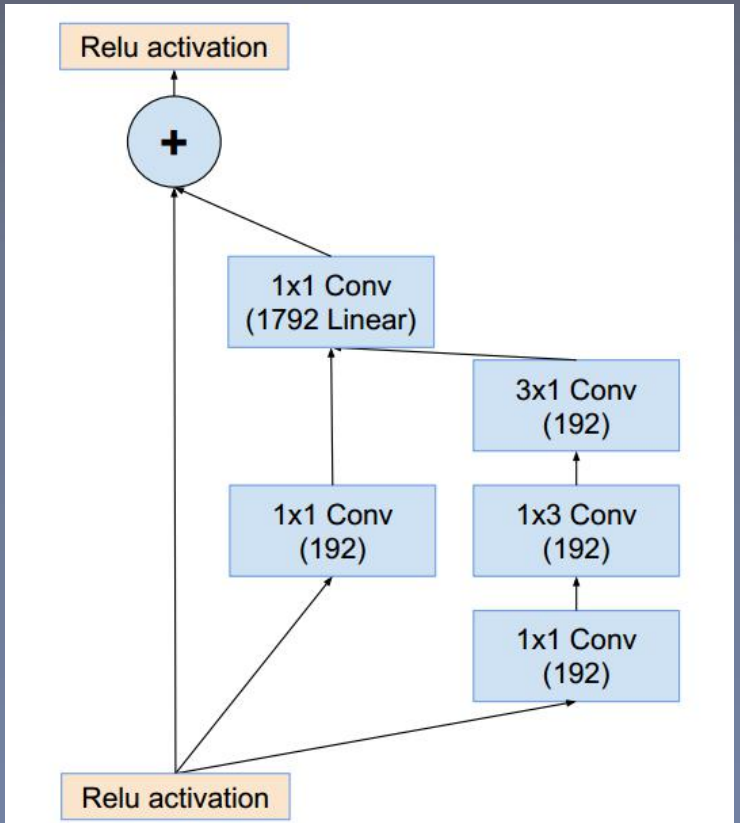


Figure 13. The schema for 8×8 grid (Inception-ResNet-C) module of Inception-ResNet-v1 network.

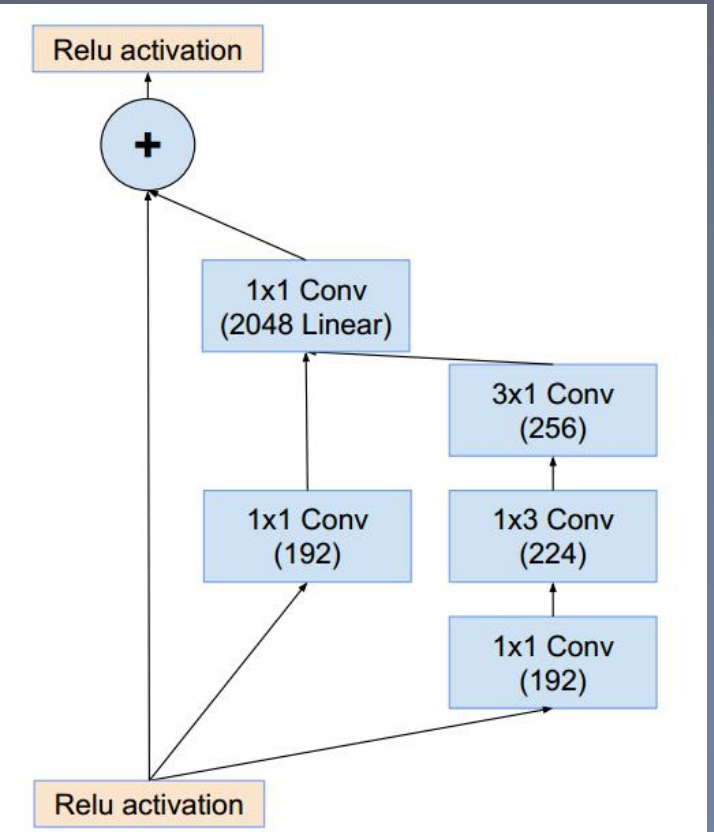


Figure 19. The schema for 8×8 grid (Inception-ResNet-C) module of the Inception-ResNet-v2 network.

Inception-ResNet

Architecture of Inception-Resnet

Inception-ResNet V1 、V2

将ResNet中的residual connection思想加到Inception中

根据Stem和卷积核数量的不同，设计出了Inception-ResNet-V1和V2

Inception-ResNet-V1共 7+5*4+3+10*4+3+5*4+1=94层

Inception-ResNet-V2共 9+5*4+3+10*4+3+5*4+1=96层

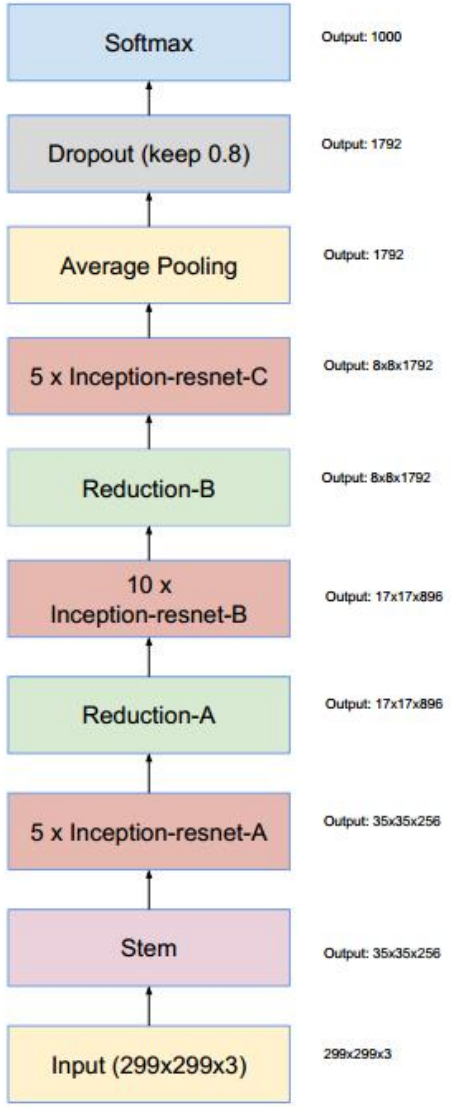


Figure 15. Schema for Inception-ResNet-v1 and Inception-ResNet-v2 networks. This schema applies to both networks but the underlying components differ. Inception-ResNet-v1 uses the blocks as described in Figures 14, 10, 7, 11, 12 and 13. Inception-ResNet-v2 uses the blocks as described in Figures 3, 16, 7, 17, 18 and 19. The output sizes in the diagram refer to the activation vector tensor shapes of Inception-ResNet-v1.

激活值缩放

Scaling of the Residuals

激活值缩放

Scaling of the Residuals

Activation Scaling

为了让模型训练稳定，在残差模块中对残差进行数值大小的缩放，通常乘以0.1至0.3之间的一个数

这个操作不是必须的！

不是 Even where the scaling was not strictly necessary, it never seemed to harm the final accuracy, but it helped to stabilize the training.

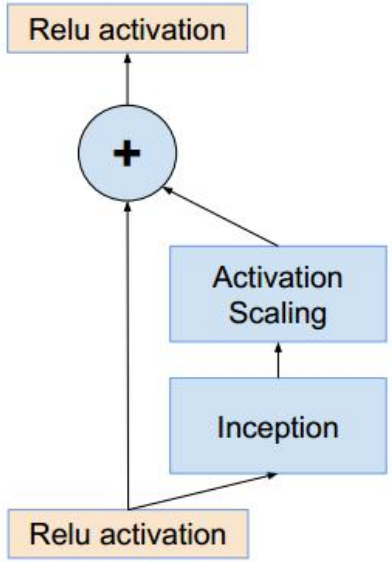


Figure 20. The general schema for scaling combined Inception-resnet moduels. We expect that the same idea is useful in the general resnet case, where instead of the Inception block an arbitrary subnetwork is used. The scaling block just scales the last linear activations by a suitable constant, typically around 0.1.

实验结果及分析

Results and Discussion

实验结果及分析

Results and Discussion

single-crop 对比各模型top-1、top-5 error

Network	Top-1 Error	Top-5 Error
BN-Inception [6]	25.2%	7.8%
Inception-v3 [15]	21.2%	5.6%
Inception-ResNet-v1	21.3%	5.5%
Inception-v4	20.0%	5.0%
Inception-ResNet-v2	19.9%	4.9%

Table 2. Single crop - single model experimental results. Reported on the non-blacklisted subset of the validation set of ILSVRC 2012.

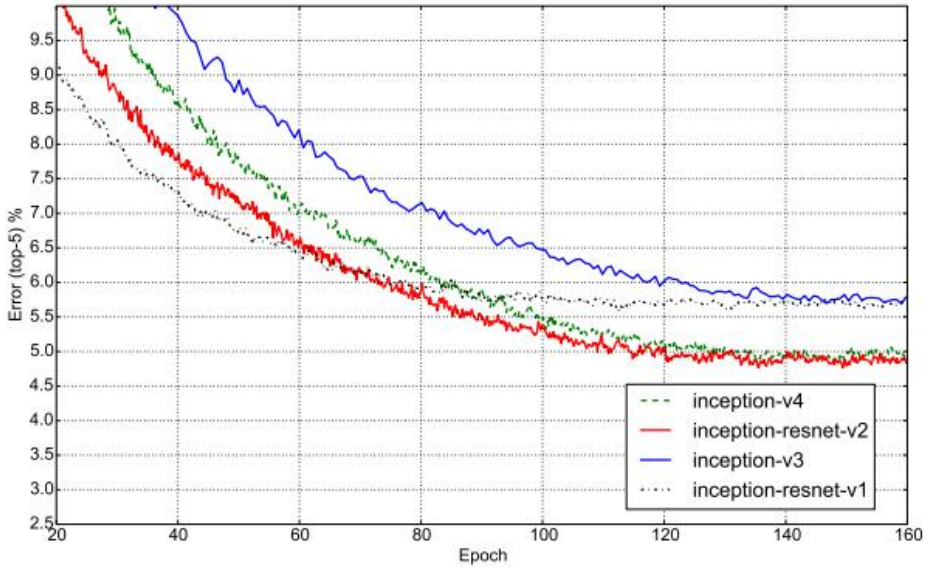


Figure 25. Top-5 error evolution of all four models (single model, single crop). Showing the improvement due to larger model size. Although the residual version converges faster, the final accuracy seems to mainly depend on the model size.

top-5 error

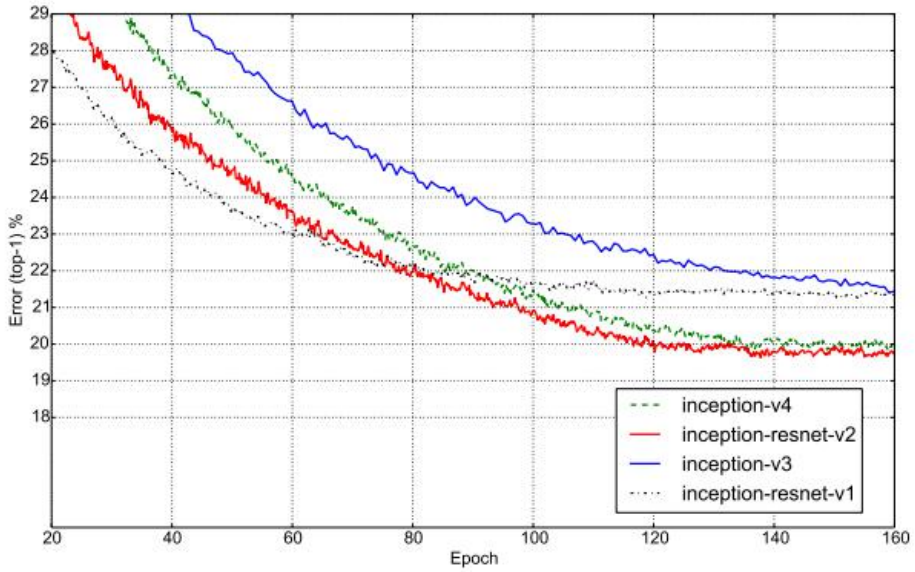


Figure 26. Top-1 error evolution of all four models (single model, single crop). This paints a similar picture as the top-5 evaluation.

top-1 error

实验结果及分析

Results and Discussion

10-crop、144-crop以及模型融合对比各模型

各种crop方式以及模型融合对比，从实验中可看出

- Inception-ResNet-V2优于Inception-V4优于Inception-ResNet-V1
- 144crop优于12crop
- 模型集成中采用最优的模型进行集成Inception-v4 + 3个Inception-ResNet-v2
- 本文方法优于ResNet论文方法3.1% VS 3.6%

Network	Crops	Top-1 Error	Top-5 Error
ResNet-151 [5]	10	21.4%	5.7%
Inception-v3 [15]	12	19.8%	4.6%
Inception-ResNet-v1	12	19.8%	4.6%
Inception-v4	12	18.7%	4.2%
Inception-ResNet-v2	12	18.7%	4.1%

Table 3. 10/12 crops evaluations - single model experimental results. Reported on the all 50000 images of the validation set of ILSVRC 2012.

Network	Crops	Top-1 Error	Top-5 Error
ResNet-151 [5]	dense	19.4%	4.5%
Inception-v3 [15]	144	18.9%	4.3%
Inception-ResNet-v1	144	18.8%	4.3%
Inception-v4	144	17.7%	3.8%
Inception-ResNet-v2	144	17.8%	3.7%

Table 4. 144 crops evaluations - single model experimental results. Reported on the all 50000 images of the validation set of ILSVRC 2012.

Network	Models	Top-1 Error	Top-5 Error
ResNet-151 [5]	6	–	3.6%
Inception-v3 [15]	4	17.3%	3.6%
Inception-v4 + 3× Inception-ResNet-v2	4	16.5%	3.1%

Table 5. Ensemble results with 144 crops/dense evaluation. Reported on the all 50000 images of the validation set of ILSVRC 2012. For Inception-v4(+Residual), the ensemble consists of one pure Inception-v4 and three Inception-ResNet-v2 models and were evaluated both on the validation and on the test-set. The test-set performance was 3.08% top-5 error verifying that we don't overfit on the validation set.

论文总结

Summary of the paper

论文总结

Summary of the paper

关键点&创新点

- 将当下（2015-2016年）流行和实用的 residual connection技术，引入GoogLeNet系列中来，提出 Inception-ResNet结构网络

论文总结

Summary of the paper

启发点

1. 本文提出的Inception-ResNet结构依旧是在大量实验中得出的模型结构，这一点与Inception-V3论文一样，这也是Inception再也难以发展的原因，缺乏亮点

In fact we have tested bigger and wider Inception-ResNet variants and they performed very similarly on the ImageNet classification challenge [11] dataset (1. Introduction p4)

2. 图像识别任务中，没有残差连接也可训练深度卷积网络。但本论文汇中最深的网络仅95层（Inception-ResNet-V2），这一句话还是欠缺数据支撑的

The authors argue that residual connections are inherently necessary for training very deep convolutional models. Our findings do not seem to support this view, at least for image recognition.(2. Related Work p2)

论文总结

Summary of the paper

深度之眼
deepshare.net

启发点

3. 做了大量实验仅展示两个版本，这两个版本的计算量分别与Inception-v3和v4差不多，这也是实验部分为什么用Inception-ResNet-v1与Inception-v3比较，用Inception-ResNet-v2与Inception-v4比较的原因。

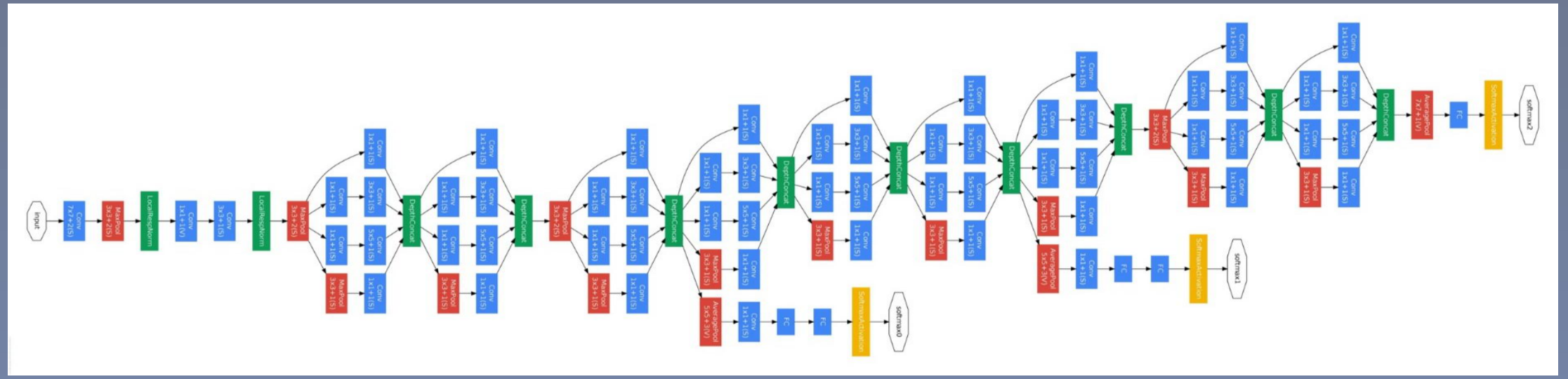
Only two of them are detailed here. The first one “Inception-ResNet-v1” roughly the computational cost of Inception-v3, while “Inception-ResNet-v2” matches the raw cost of the newly introduced Inception-v4 network. (3.2 Residual Inception Blocks)

论文总结

Summary of the paper

GoogLeNet系列论文回顾

- GoogLeNet-V1（2014-09）：借鉴NIN广泛应用1*1卷积，借鉴多尺度Gabor滤波器提出多尺度卷积的Inception结构，开启多尺度卷积，1*1卷积时代



论文总结

Summary of the paper

深度之眼
deepshare.net

GoogLeNet系列论文回顾

- GoogLeNet-V1（2014-09）：借鉴NIN广泛应用1*1卷积，借鉴多尺度Gabor滤波器提出多尺度卷积的Inception结构，开启多尺度卷积，1*1卷积时代
- GoogLeNet-V2（2015-03）：针对ICS（Internal Covariate Shift，内部协变量偏移）问题，提出Batch Normalization，加快整个深度学习的发展，让标准化层成为深度神经网络的标配

$$\widehat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}}$$

$$y_i \leftarrow \gamma \widehat{x}_i + \beta \equiv \mathrm{N}_{\gamma,\beta}(x_i)$$

论文总结

Summary of the paper

深度之眼
deepshare.net

GoogLeNet系列论文回顾

- GoogLeNet-V1（2014-09）：借鉴NIN广泛应用1*1卷积，借鉴多尺度Gabor滤波器提出多尺度卷积的Inception结构，开启多尺度卷积，1*1卷积时代
- GoogLeNet-V2（2015-03）：针对ICS（Internal Covariate Shift，内部协变量偏移）问题，提出Batch Normalization，加快整个深度学习的发展，让标准化层成为深度神经网络的标配
- GoogLeNet-V3（2015-12）：大量实验，总结模型设计准则，提出卷积分解、高效特征图分辨率、标签平滑技巧，开始了针对性模块的设计思路，也导致Inception系列越走越窄

论文总结

Summary of the paper

深度之眼
deepshare.net

GoogLeNet系列论文回顾

- GoogLeNet-V1（2014-09）：借鉴NIN广泛应用1*1卷积，借鉴多尺度Gabor滤波器提出多尺度卷积的Inception结构，开启多尺度卷积，1*1卷积时代
- GoogLeNet-V2（2015-03）：针对ICS（Internal Covariate Shift，内部协变量偏移）问题，提出Batch Normalization，加快整个深度学习的发展，让标准化层成为深度神经网络的标配
- GoogLeNet-V3（2015-12）：大量实验，总结模型设计准则，提出卷积分解、高效特征图分辨率、标签平滑技巧，开始了针对性模块的设计思路，也导致Inception系列越走越窄
- GoogLeNet-V4（2016-02） GoogLeNet-V4（2016-02）：从Inception-v3上看到Inception系列有些“走投无路”，于是引入当时最火最热的Residual connection思想进行改进

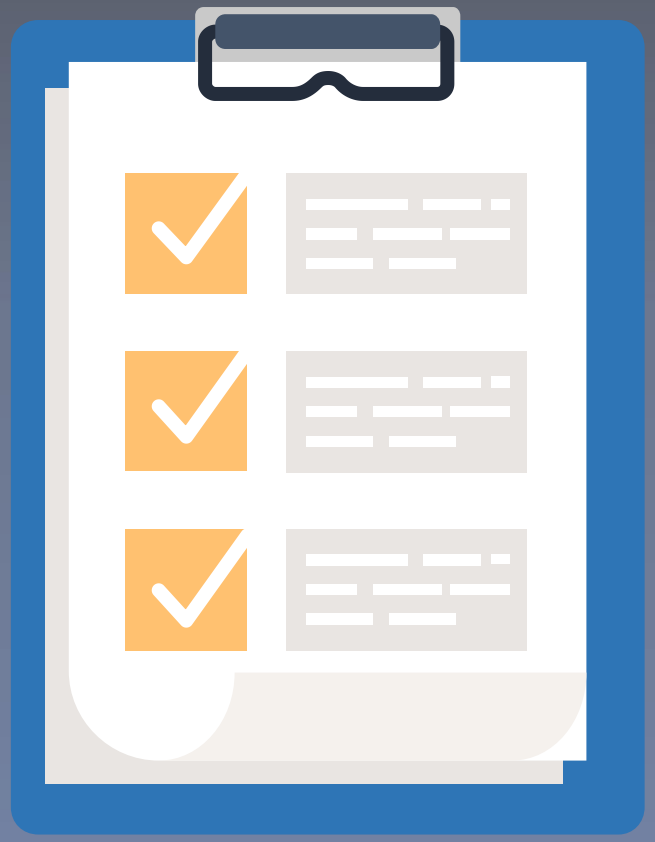
深度之眼
deepshare.net

本课回顾及下节预告

Review in the lesson and Preview of next lesson

本课回顾

Review in the lesson



深度之眼
deepshare.net

01 Inception-v4

Stem、Inception-A\B\C、Reduction-A\B 六大模块，
9+3*4+5*7+4*3+3+4+1 = 76层

02 Inception-ResNet网络结构

Inception-ResNet-V1共 7+5*4+3+10*4+3+5*4+1=94层

Inception-ResNet-V2共 9+5*4+3+10*4+3+5*4+1=96层

03 实验结果分析

Single crop、10-crop、144-crop及模型融合方式的top1及top5 error对比

04 论文总结

总结论文中创新点、关键点及启发点

深度之眼
deepshare.net

下节预告

Preview of next lesson

01 Inception-V4 模型推理

分析Stem、Inception-A\B\C、Reduction-A\B 六大模块

02 01 Inception-ResNet-V2 模型推理

分析Stem、Inception-A\B\C、Reduction-A\B 六大模块

结　语

读书之法,在循序而渐进,熟读而精思

——朱熹

深度之眼
deepshare.net

联系我们：

电话：18001992849

邮箱：service@deepshare.net

Q Q： 2677693114

公众号

客服微信