深蓝学院
shenlanxueyuan.com

# 深度学习：从理论到实践

## 第一章：深度学习理论（三）
## ——卷积神经网络（下）

# 课程大纲



- ✓ 卷积神经网络的设置
- ✓ 卷积神经网络的训练
- ✓ 卷积神经网络的应用现状

# 课程大纲

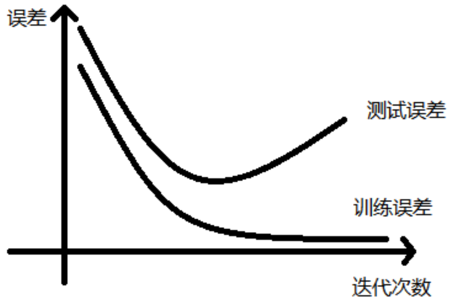
深蓝学院
shenlanxueyuan.com

- ✓ 卷积神经网络的设置
- ✓ 卷积神经网络的训练
- ✓ 卷积神经网络的应用现状

# 网络设置

## 1. 层的类别（a）：必需

- ✓ Data层：接收输入数据
- ✓ 卷积层：核心部分
- ✓ 非线性层：Sigmoid，tanh，ReLU……
- ✓ 池化层（下采样层）：max，average
- ✓ Loss层：用于误差反传（训练、验证），交叉熵、均方误差等
- ✓ Accuracy层：计算最终输出的指标（测试）



# 网络设置

## 1. 层的类别（b）：常用

- ✓ 全连接层（内积层）：相当于多层感知器
- ✓ Dropout层：降低节点之间关联性带来的影响
- ✓ Batch normalization层：“归一化”操作，可学习（有参数）
- ✓ Softmax层：相当于多类的Sigmoid

BN层的原理解读：http://blog.csdn.net/hjimce/article/details/50866313
BN学习笔记：http://blog.csdn.net/hjimce/article/details/50866313
为什么加入BN层效果好？https://www.zhihu.com/question/38102762

# 网络设置

## 1. 层的类别（b）：常用

**Input:** Values of $x$ over a mini-batch: $\mathcal{B} = \{x_{1\ldots m}\}$;
Parameters to be learned: $\gamma$, $\beta$

**Output:** $\{y_i = \mathrm{BN}_{\gamma,\beta}(x_i)\}$

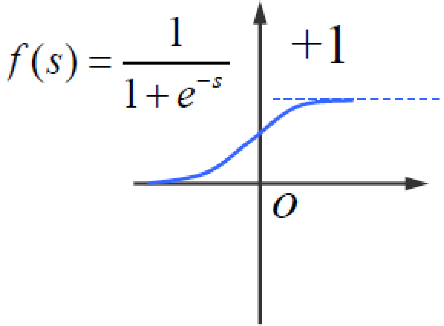
$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \text{// mini-batch mean}$$

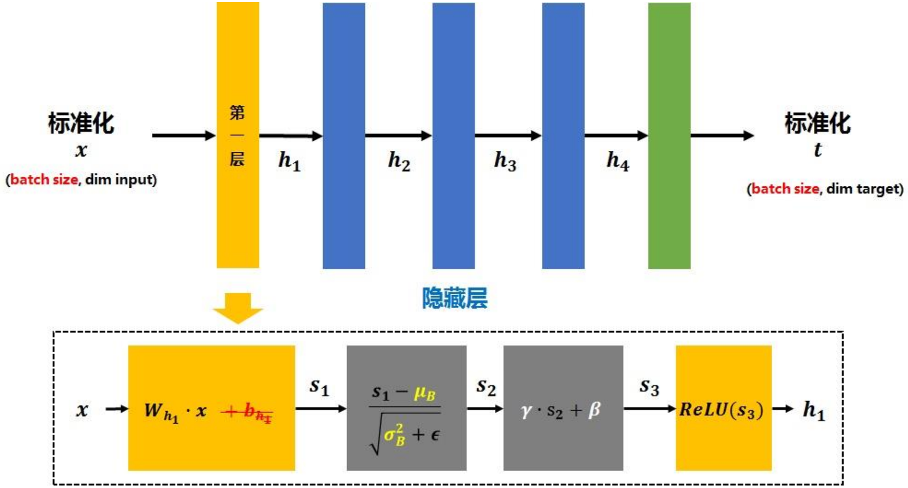
$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_{\mathcal{B}})^2 \qquad \text{// mini-batch variance}$$

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

$$y_i \leftarrow \gamma \widehat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i) \qquad \text{// scale and shift}$$

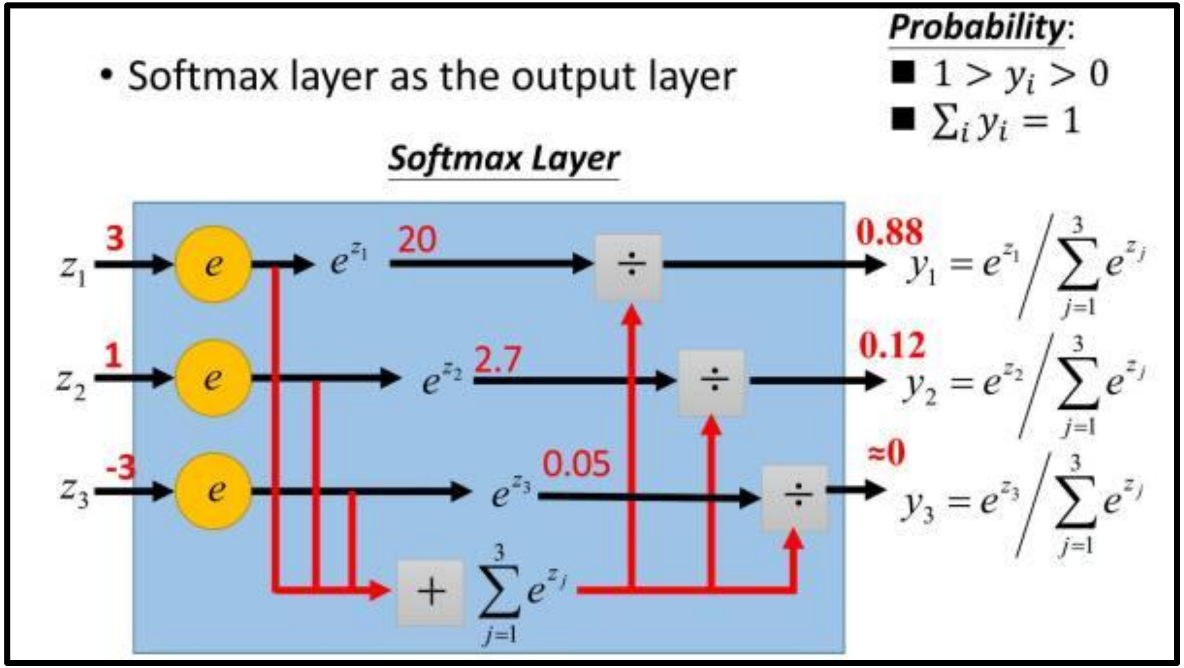
**Algorithm 1:** Batch Normalizing Transform, applied to activation $x$ over a mini-batch.



# 网络设置

## 1. 层的类别（b）：常用



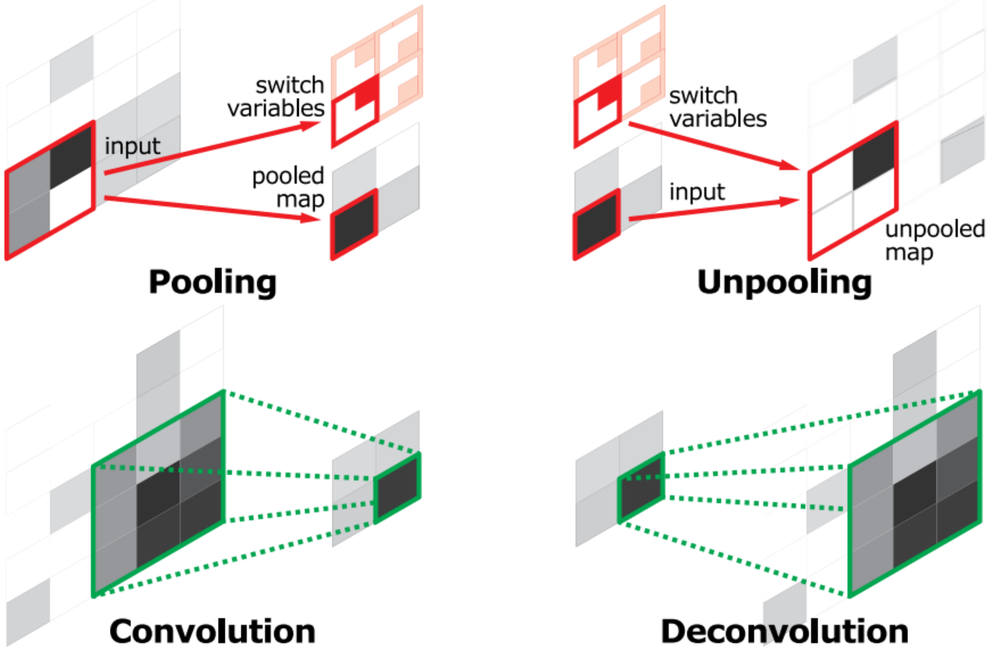

$$y = \frac{e^{z_1}}{\sum_{j=1}^{2} e^{z_j}} = \frac{1}{1+e^{z_2-z_1}} \triangleq \frac{1}{1+e^{z}}$$

# 网络设置

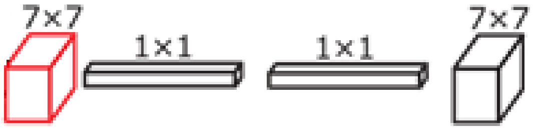
## 1. 层的类别（c）：图像分割网络必需

✓ 反卷积层（不是数学意义上的反/去卷积运算）：

✓ 上采样层：



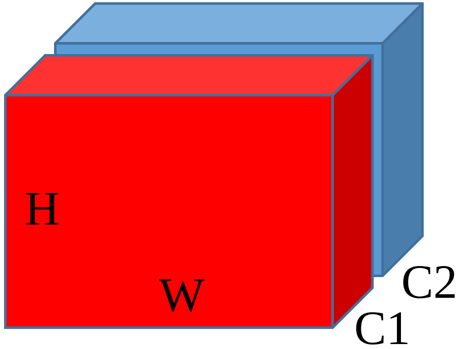

**思考：**从1*1到7*7的卷积是如何实现的？



# 网络设置

## 1. 层的类别（d）：可选

- ✓ Crop层：对图像或者特征图进行裁剪
- ✓ Elementwise层：1）相当于matlab的点乘，用于施加权重，类似于“门（gate）”的功能；2）不同分支的结果进行融合（SUM、MAX）
- ✓ Concatenate层：特征图在channel维度进行拼接
- ✓ ……



# 网络设置

## 2. 层的设置

- ✓ 卷积层（可连续多层）与池化层交替
- ✓ 池化层不可过多（考虑特征图尺寸）
- ✓ 卷积层->BN层->非线性层 （为了保证激励函数输入的分布）
- ✓ 全连接层->Softmax层->Loss层通常在网络最后
- ✓ 层数特别多的网络通常由基本模块重复构成：GoogLeNet中的Inception module、ResNet中的residual module等

# 网络设置

深蓝学院
shenlanxueyuan.com

## 3. 层的数目：举例

- ✓ AlexNet：8层
- ✓ VGG：16层、19层两个版本
- ✓ Conv-Deconv Net：大约是VGG-16的两倍
- ✓ ResNet：152层，1000+层等多个版本
- ✓ One-hundred-layer-Tiramisu： 100层左右
- ✓ ……

# 网络设置



## 4. 每层特征图数目（卷积核数目）

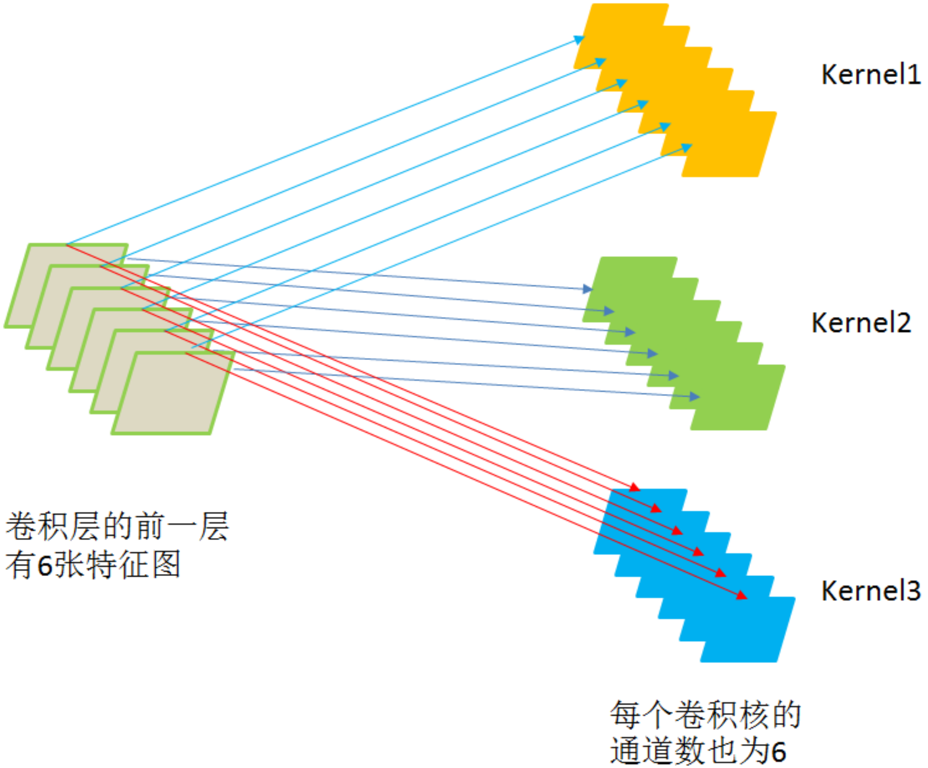
- ✓ 整体上随着网络层数的加深而增加
- ✓ 通常每次进行池化，特征图数目相应增加
- ✓ 池化操作时，前后两层特征图一一对应（即数目不变）
- ✓ 过多：参数数目巨大，需要训练数据多，容易过拟合
- ✓ 过少：网络建模能力差
- ✓ 通常设为8的倍数（或者2的幂）

# 网络设置

## 4. 每层特征图数目

- ✓ 每个卷积核的通道数与前一层特征图的通道数一致
- ✓ 后一层的特征图数目取决于卷积核的数目

卷积层的前一层，有多少个特征图，卷积核的通道就有多少个。



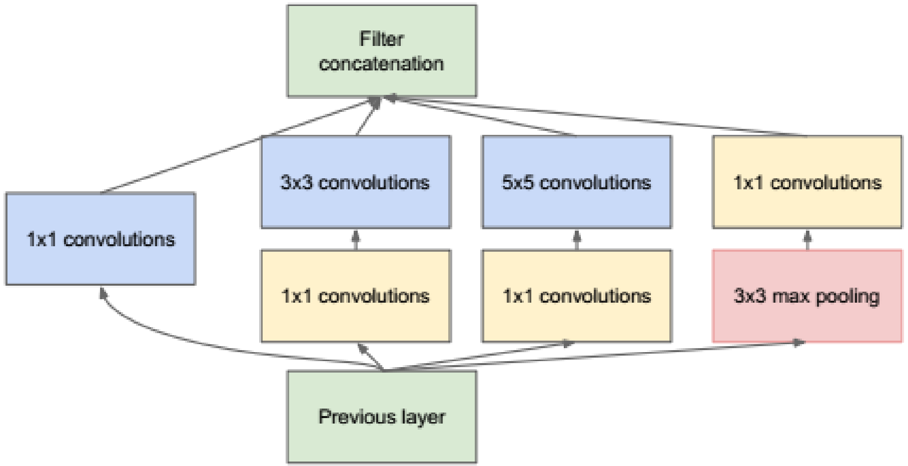
# 网络设置

## 5. 卷积核尺寸（a）

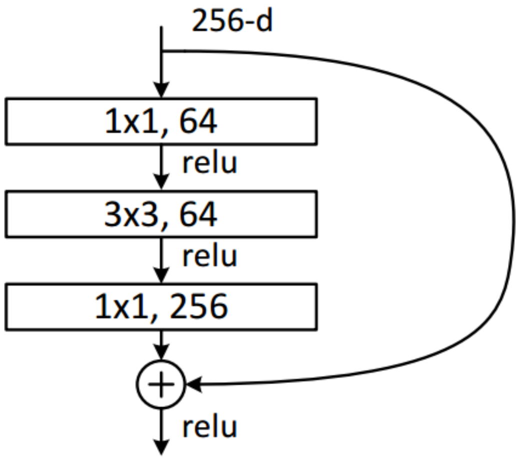
- ✓ 考虑图像尺寸和目标尺度
- ✓ 考虑高层节点的实际感受野
- ✓ 为了获得特定尺寸的输出特征图而设定：如7*7->1*1
- ✓ 3*3最为常用
- ✓ 多种尺寸混合使用，检测不同尺度、不同类型的特征
- ✓ 1*1卷积：不需限制输入尺寸；改变特征图数目（如降维）

# 网络设置

## 5. 卷积核尺寸（b）：举例




GoogLeNet中的Inception module



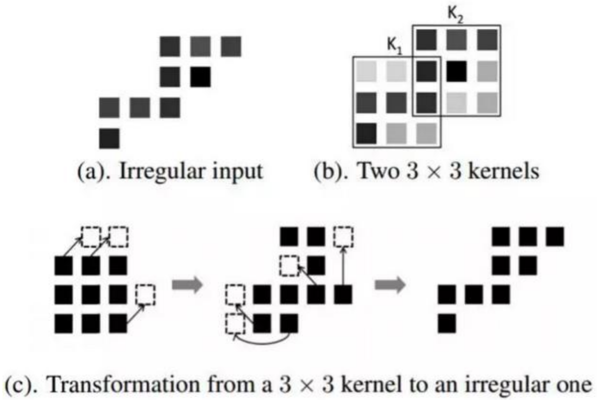

ResNet中的residual module

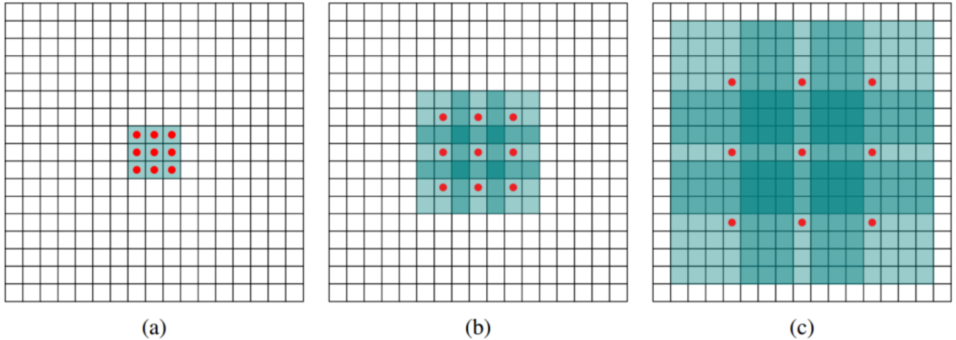

C. Szegedy et al., Going deeper with convolutions. CVPR 2015.
K. He et al., Deep Residual Learning for Image Recognition. ICCV 2015.

# 网络设置

## 6. 卷积核形状

- ✓ 规则k*k正方形
- ✓ 不规则形状：对元素位置也进行学习
- ✓ “离散”孔洞：可以增加高层节点的实际感受野（同时参数数目不变）




F. Yu and V. Koltun, Multi-Scale Context Aggregation by Dilated Convolutions. ICLR 2016.
J. Ma et al., Irregular Convolution Neural Networks. arXiv 2017.

# 网络设置

## 6. 网络参数初始化

- ✓ 从头训练/预训练（pre-train）：
  - 正态分布随机数
  - xavier、MSRA等
  - 不可全零！（思考：为什么？）

X. Glorot and Y. Bengio, Understanding the difficulty of training deep feedforward neural networks. AISTATS 2010.

K. He et al., Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. ICCV 2015.

- ✓ 基于已训练好的网络进行微调（fine-tune）

跨数据集、跨任务微调的可行性，
证实了卷积神经网络学到特征的有效性和通用性

# 课程大纲

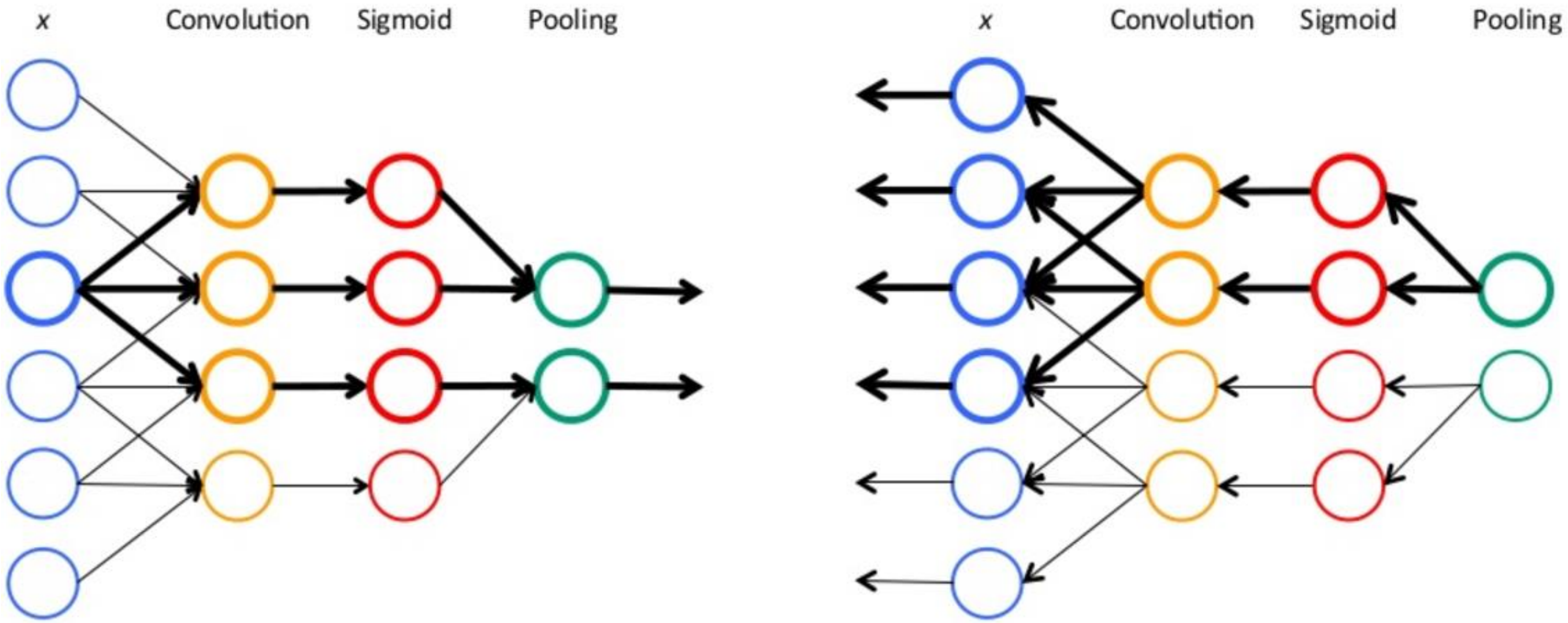
深蓝学院
shenlanxueyuan.com

- ✓ 卷积神经网络的设置
- ✓ 卷积神经网络的训练
- ✓ 卷积神经网络的应用现状

# 网络训练

深蓝学院 shenlanxueyuan.com

## 与多层感知器类似，采用误差反向传播（BP）算法



# 网络训练

深蓝学院
shenlanxueyuan.com

## 卷积神经网络的BP算法推导

- ✓ 核心部分：卷积层、池化层
- ✓ 重要部分：BN层（见参考文献）
- ✓ 与多层感知器相同部分：非线性层、全连接层和损失层等
- ✓ 可“偷懒”部分：反卷积层、上采样层

S. Ioffe and C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. ICML 2015.

# 网络训练

## 卷积与反卷积、下采样与上采样（以caffe为例）

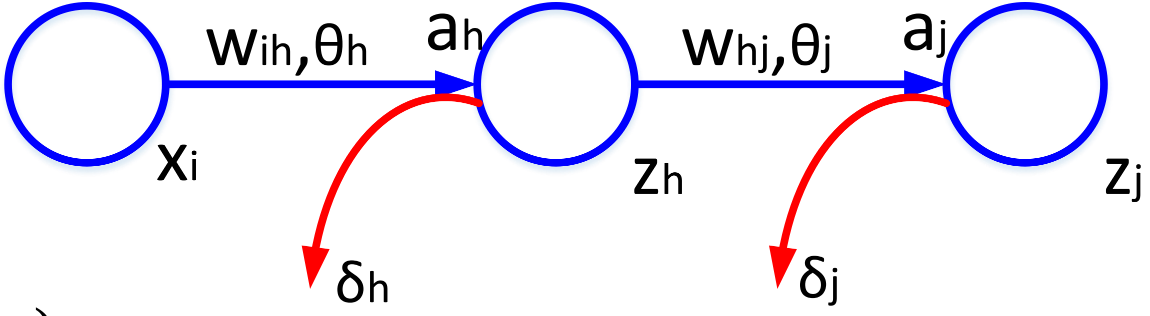
| 层名称 | 信号正向传播 | 误差反向传播 |
|---|---|---|
| 卷积 | forward_cpu() | backward_cpu() |
| 反卷积 | backward_cpu() | forward_cpu() |
| 池化（下采样） | H*W -> 1 | 1 -> H*W |
| 上采样 | 1 -> H*W | H*W -> 1 |

# 网络训练



## 卷积神经网络的BP算法推导——回顾

### 多层感知器的推导




$$a_h = \sum_{i=1}^{d} w_{ih} x_i + \theta_h, \quad z_h = f(a_h)$$

$$a_j = \sum_{h=1}^{n_H} w_{hj} z_h + \theta_j, \quad z_j = f(a_j)$$

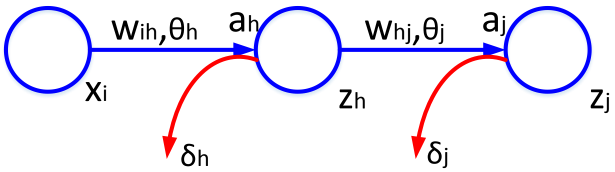
损失函数：

$$E(W) = \frac{1}{2} \sum_j (t_j - z_j)^2$$

# 网络训练

## 卷积神经网络的BP算法推导——回顾

### 多层感知器的推导




隐含层-输出层

$$\frac{\partial E}{\partial w_{hj}} = \frac{\partial E}{\partial a_j}\frac{\partial a_j}{\partial w_{hj}} = \delta_j z_h$$

$$\delta_j = -(t_j - z_j) f'(a_j)$$

$$\frac{\partial E}{\partial \theta_j} = \frac{\partial E}{\partial a_j}\frac{\partial a_j}{\partial \theta_j} = \delta_j$$

输入层-隐含层

$$\frac{\partial E}{\partial w_{ih}} = \frac{\partial E}{\partial a_h}\frac{\partial a_h}{\partial w_{ih}} = \delta_h x_i$$

$$\delta_h = \sum_{j=1}^{c} w_{hj}(\delta_j f'(a_h))$$

$$\frac{\partial E}{\partial \theta_h} = \frac{\partial E}{\partial a_h}\frac{\partial a_h}{\partial \theta_h} = \delta_h$$

# 网络训练

深蓝学院 shenlanxueyuan.com

## 卷积神经网络的BP算法推导——卷积

前向传播 $\mathbf{x}_j^\ell = f\left(\sum_i \mathbf{x}_i^{\ell-1} * \mathbf{k}_{ij}^\ell + b_j^\ell\right) = f(\mathbf{u}_j^\ell)$

梯度 $\frac{\partial E}{\partial \mathbf{k}_{ij}^\ell} = \sum_{u,v} (\boldsymbol{\delta}_j^\ell)_{uv} (\mathbf{p}_i^{\ell-1})_{uv} = \boxed{\tilde{\boldsymbol{\delta}}_j^\ell * \mathbf{x}_i^{\ell-1}}$

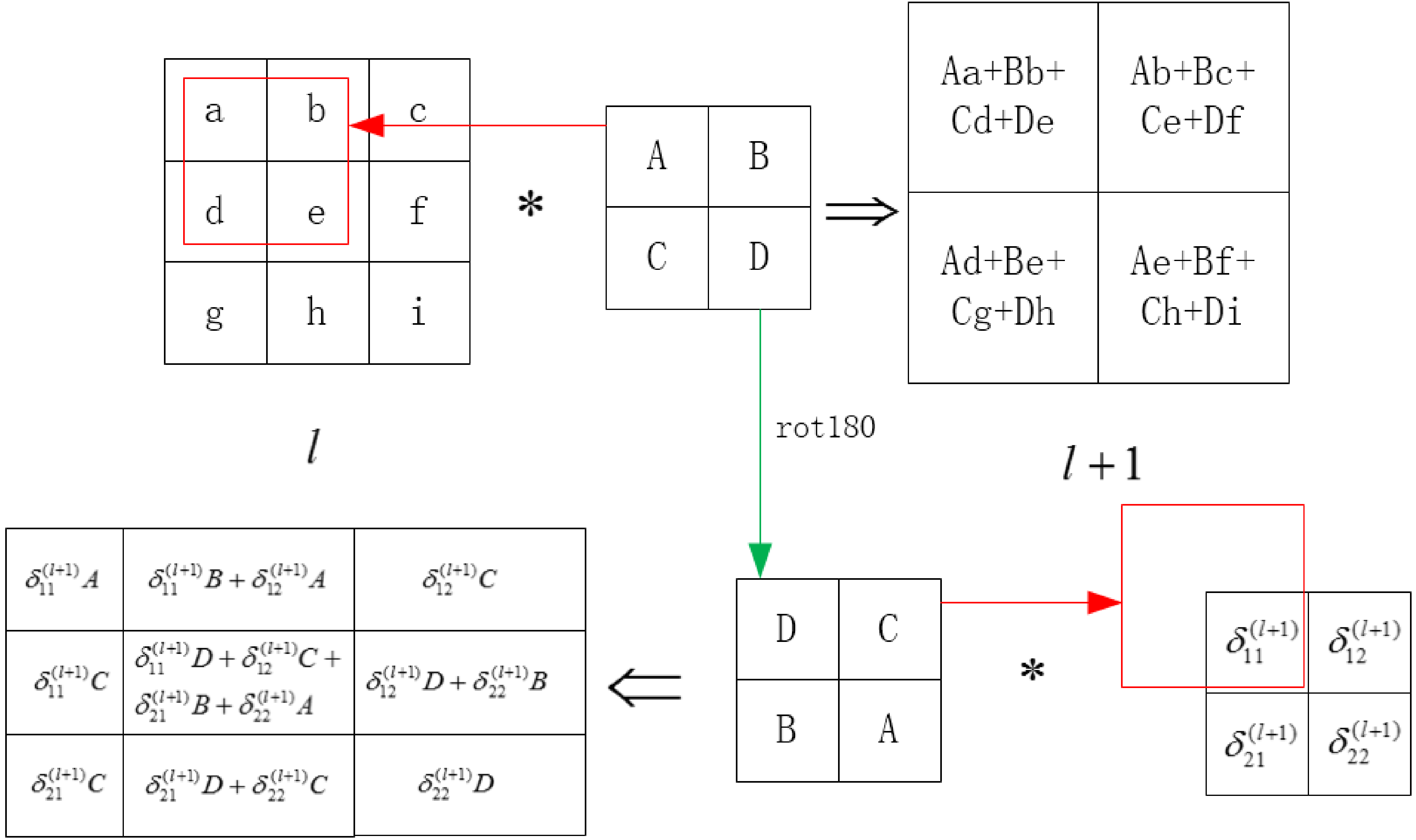
$\tilde{\boldsymbol{\delta}}_j^\ell$ 是$\boldsymbol{\delta}_j^\ell$旋转180度

$$\frac{\partial E}{\partial b_j^\ell} = \sum_{u,v} (\boldsymbol{\delta}_j^\ell)_{uv}.$$

$\mathbf{p}_i^{\ell-1}$是输入特征图中对应的图像块

误差 $\boldsymbol{\delta}_j^\ell = f'(\mathbf{u}_j^\ell) \circ \boxed{(\boldsymbol{\delta}_j^{\ell+1} * \tilde{\mathbf{k}}_{ij}^\ell)}$

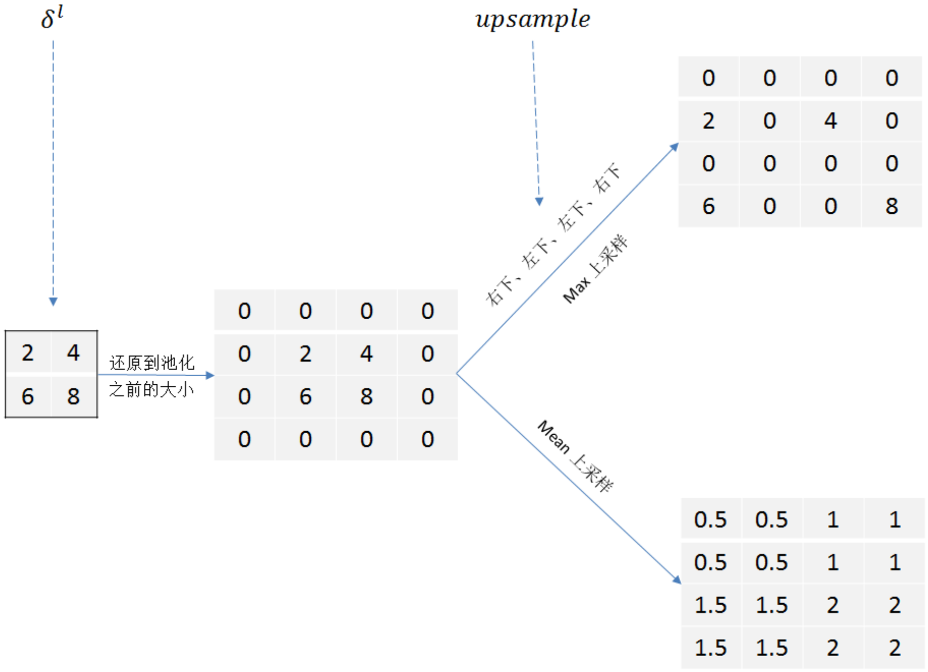
$\circ$ 表示元素级相乘


a b c
d e f
g h i
*
A B
C D
Aa+Bb+Cd+De
Ab+Bc+Ce+Df
Ad+Be+Cg+Dh
Ae+Bf+Ch+Di
rot180
$l$
$l+1$
D C
B A
*
$\delta_{11}^{(l+1)}$ $\delta_{12}^{(l+1)}$
$\delta_{21}^{(l+1)}$ $\delta_{22}^{(l+1)}$
$\delta_{11}^{(l+1)}A$
$\delta_{11}^{(l+1)}B+\delta_{12}^{(l+1)}A$
$\delta_{12}^{(l+1)}C$
$\delta_{11}^{(l+1)}C$
$\delta_{11}^{(l+1)}D+\delta_{12}^{(l+1)}C+\delta_{21}^{(l+1)}B+\delta_{22}^{(l+1)}A$
$\delta_{12}^{(l+1)}D+\delta_{22}^{(l+1)}B$
$\delta_{21}^{(l+1)}C$
$\delta_{21}^{(l+1)}D+\delta_{22}^{(l+1)}C$
$\delta_{22}^{(l+1)}D$

# 网络训练

深蓝学院
shenlanxueyuan.com

## 卷积神经网络的BP算法推导——池化层




**思考：**
池化层的BP
如何实现？

# 课程大纲

深蓝学院
shenlanxueyuan.com

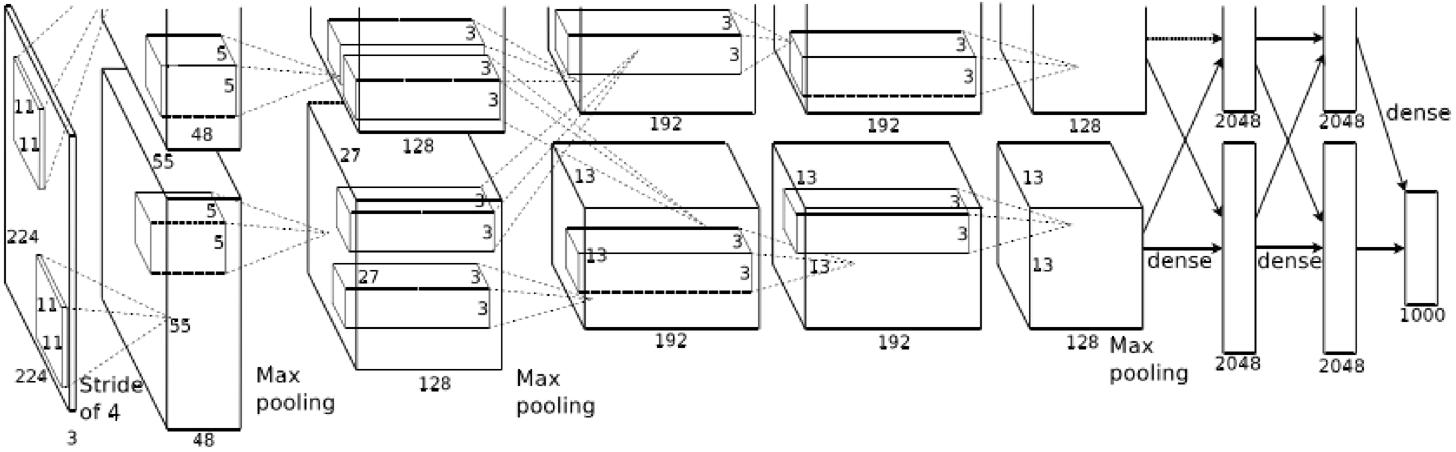
- ✓ 卷积神经网络的设置
- ✓ 卷积神经网络的训练
- ✓ 卷积神经网络的应用现状

# 网络应用

深蓝学院
shenlanxueyuan.com

**按照卷积神经网络与模型中其他部分的关系分类：**

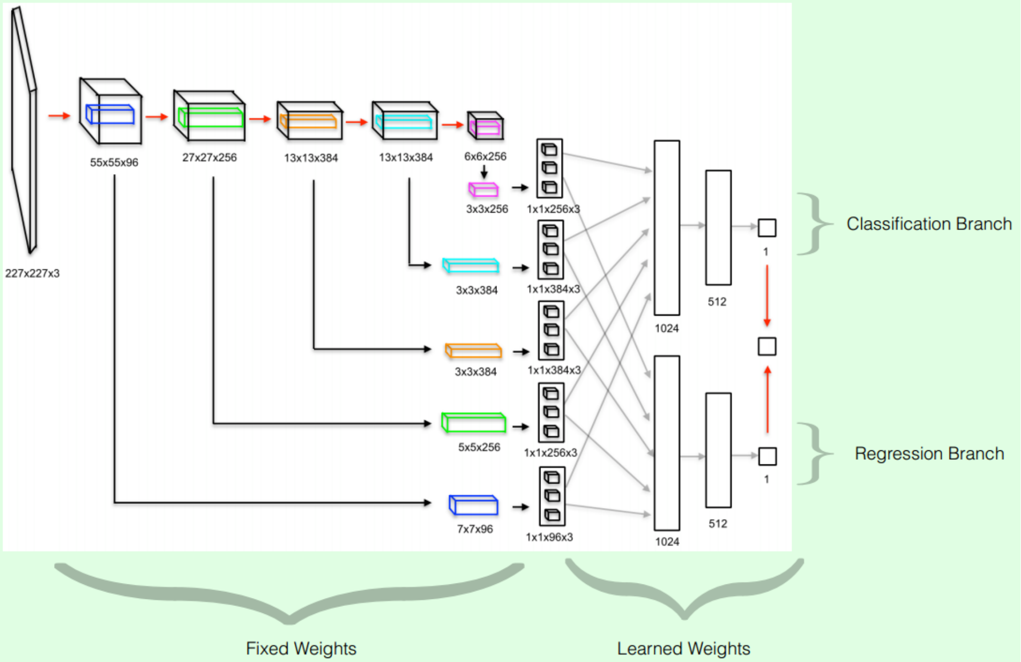
1. 与分类器（感知器）耦合，共同训练




A. Krizhevsk et al., ImageNet Classification with Deep Convolutional Neural Networks. NIPS 2012.

# 网络应用

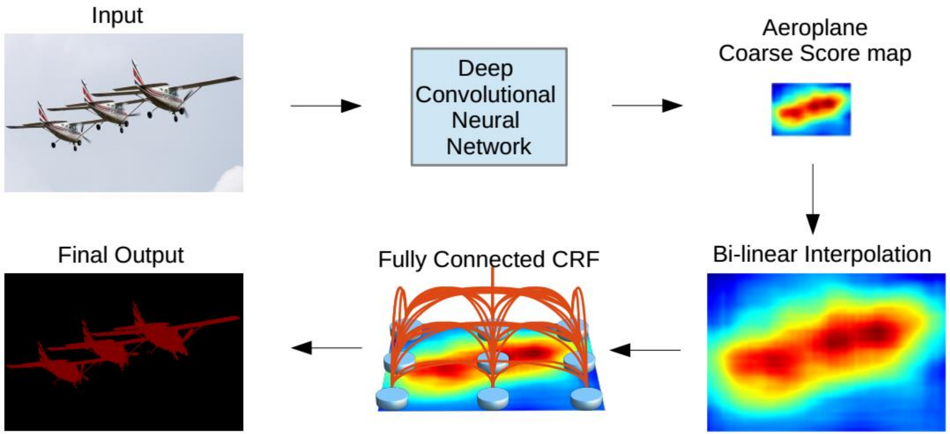
## 2. 与分类器相对独立，只采用网络学习的离线特征

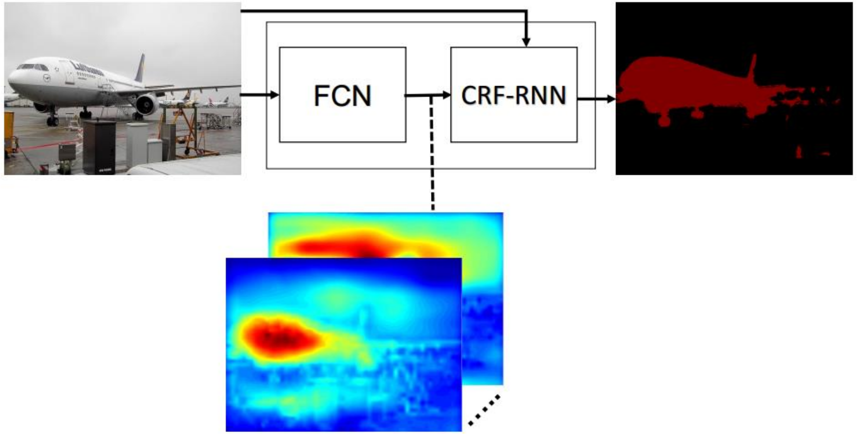



G. Bertasius et al., DeepEdge: A Multi-Scale Bifurcated Deep Network for Top-Down Contour Detection. CVPR 2015.

# 网络应用

## 3. 用网络来实现机器学习中的其他模型，以便于融入更大规模的网络中实现共同训练



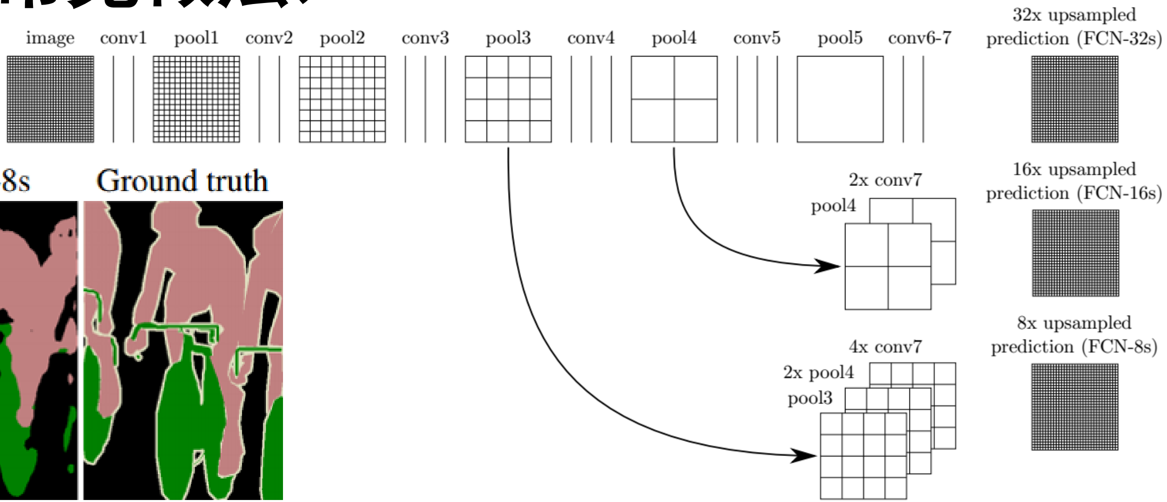

L. Chen, et al., Semantic image segmentation with deep convolutional nets and fully connected CRFs. ICLR 2015.

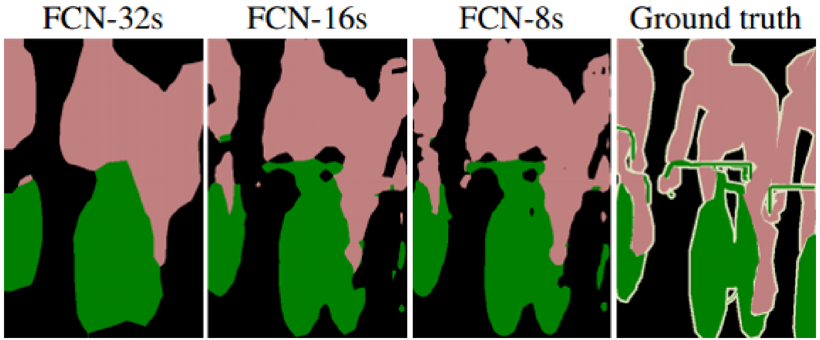
S. Zheng et al., Conditional random fields as recurrent neural networks. ICCV 2015.

# 网络应用

**按照特征图的利用方式分类：**

1. 只用最后一层（常见做法）
2. 部分层共同使用

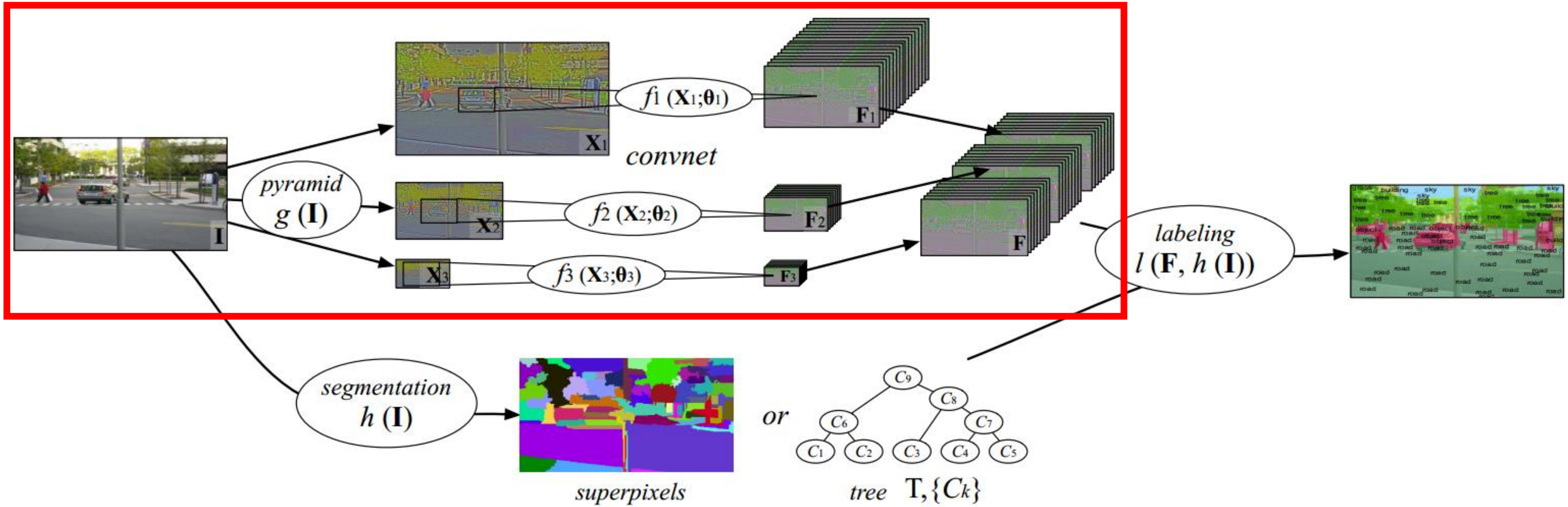



J. Long et al., Fully convolutional networks for semantic segmentation. CVPR 2015.

# 网络应用

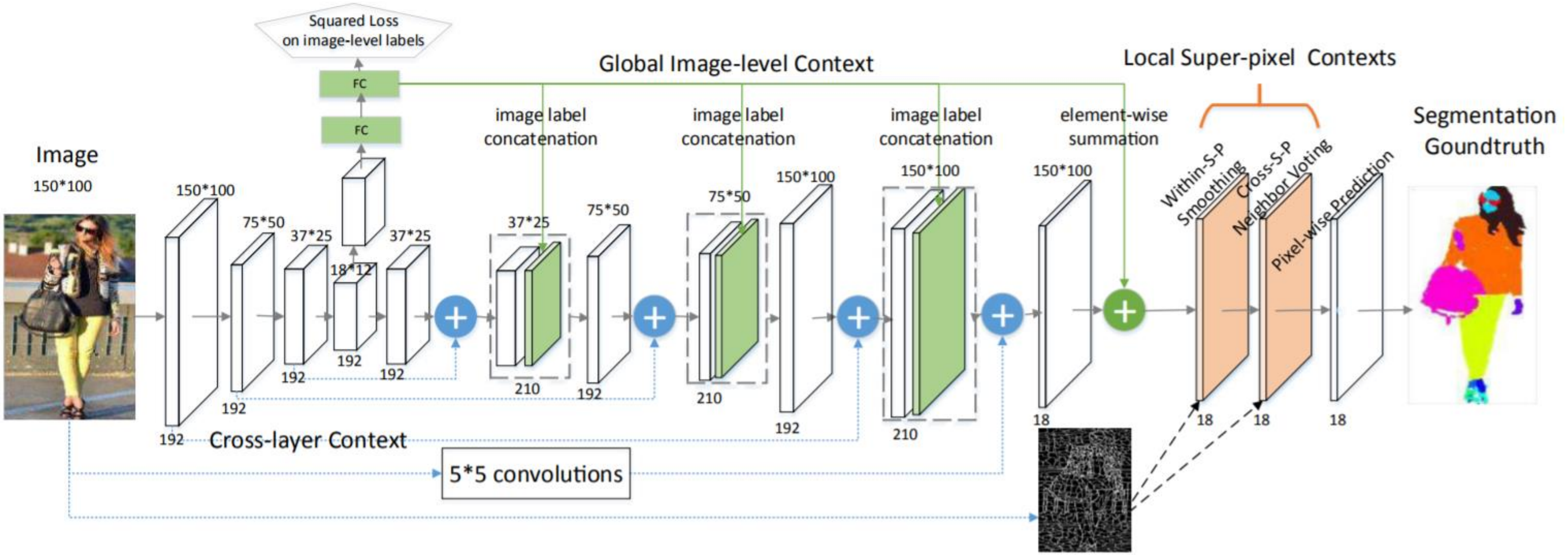
## 3. 所有层共同使用




C. Farabet et al., Learning Hierarchical Features for Scene Labeling. PAMI 2013.

# 网络应用

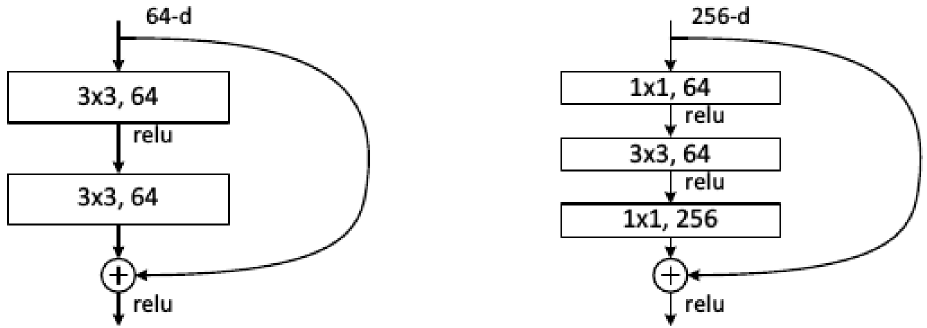
## 4. 部分层进行融合




X. Liang et al., Human Parsing with Contextualized Convolutional Neural Network. ICCV 2015.

# 网络应用

## 5. 各个层全连接



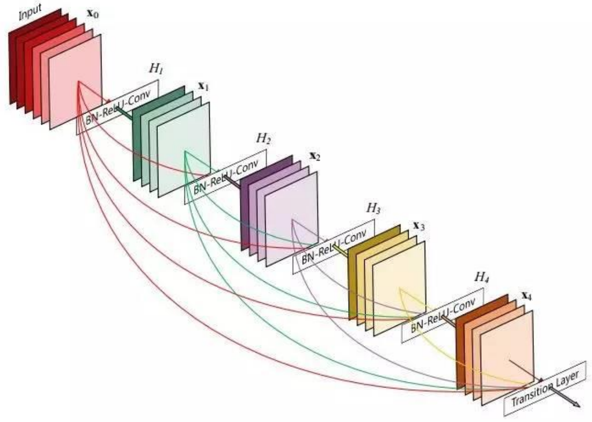
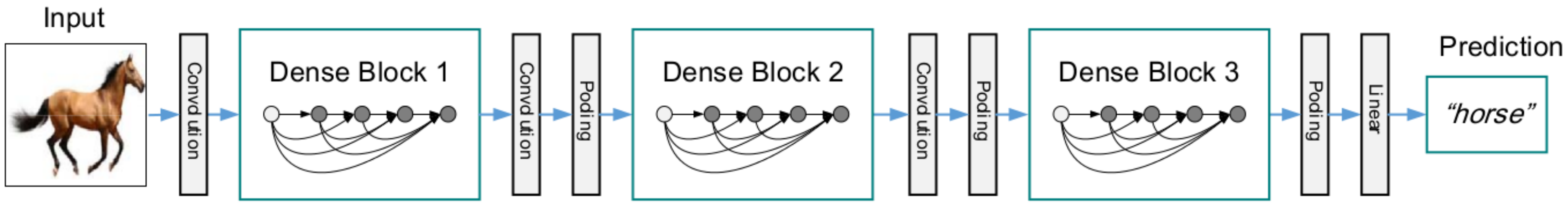




Figure 1. A 5-layer dense block with a growth rate of $k = 4$. Each layer takes all preceding feature-maps as input.

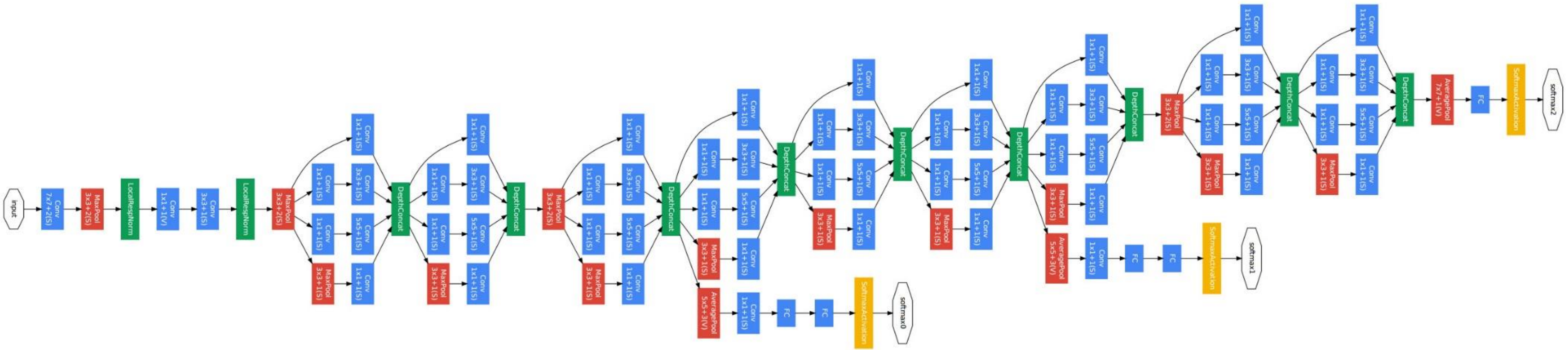



G. Huang et al., Densely Connected Convolutional Networks. CVPR 2017.

# 网络应用

**按照利用卷积神经网络解决的问题分类：**

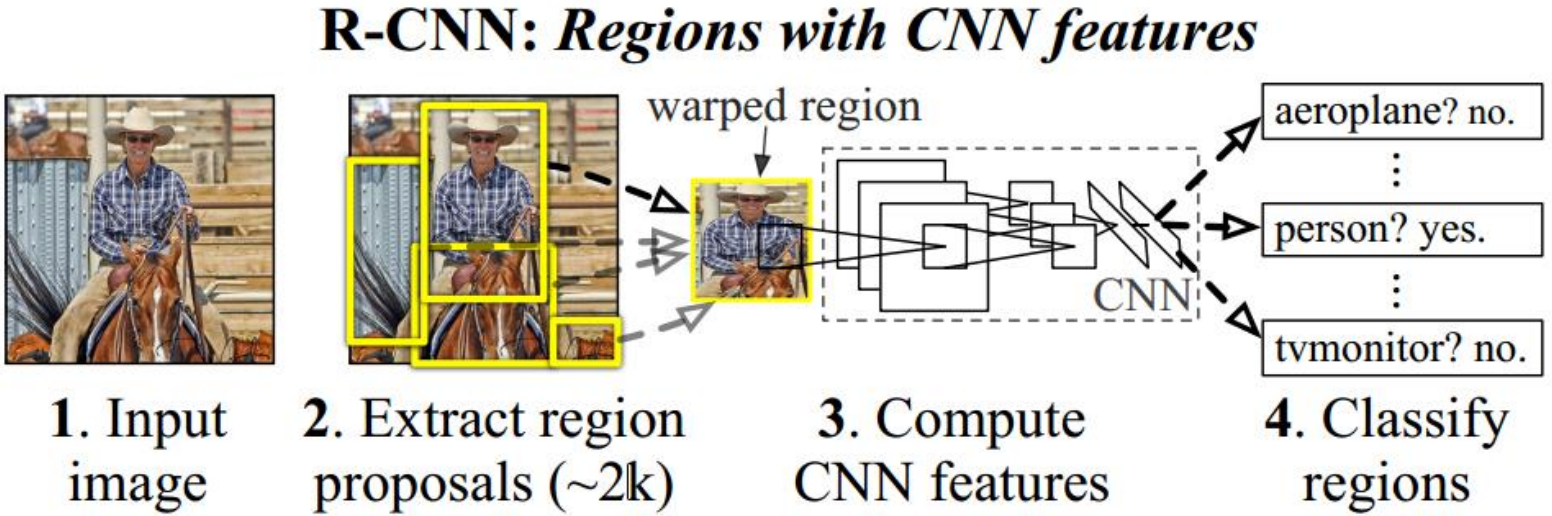
1. 图像分类与目标识别



C. Szegedy et al., Going deeper with convolutions. CVPR 2015.

# 网络应用

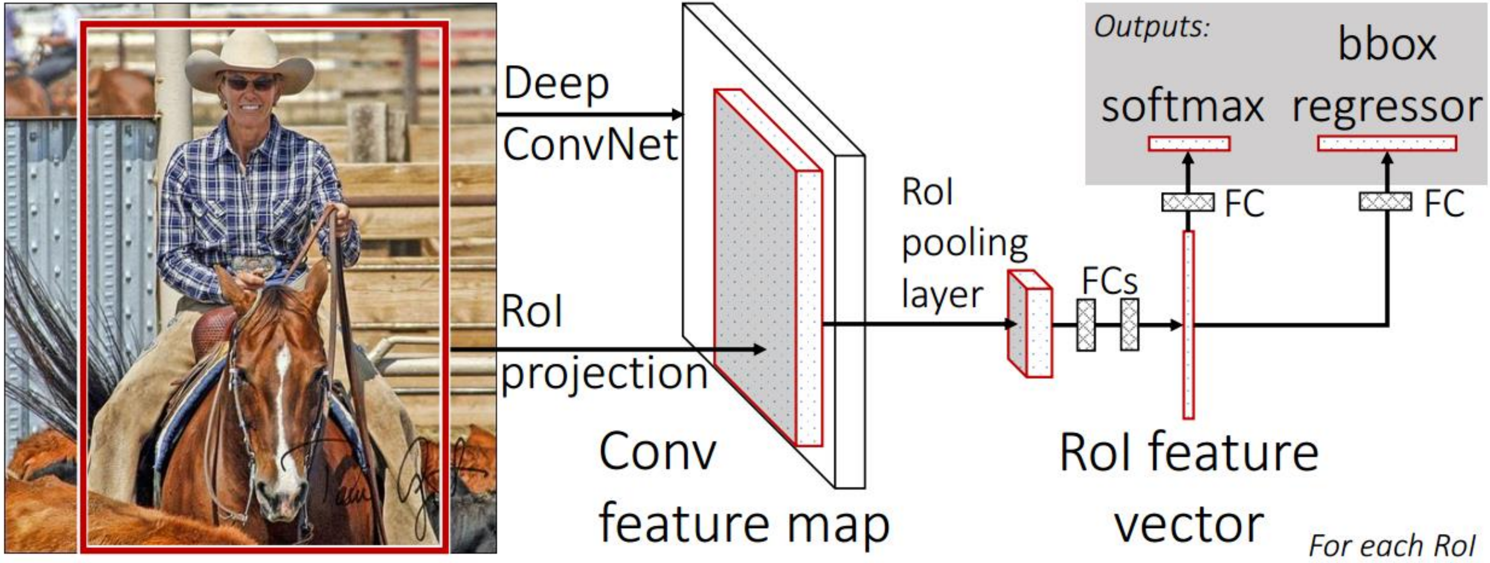
## 2. 目标检测(1)




R. Girshick et al., Rich feature hierarchies for accurate object detection and semantic segmentation. CVPR 2014.

# 网络应用

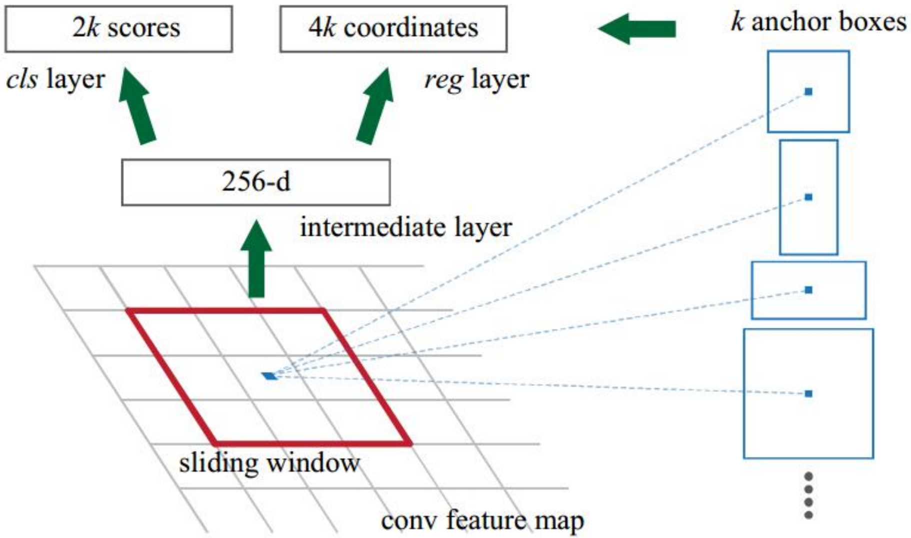
## 2. 目标检测(2)

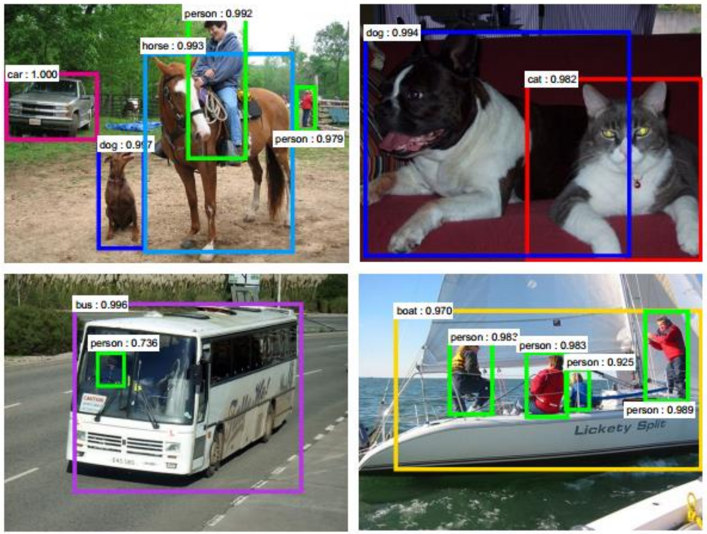



R. Girshick, Fast R-CNN. ICCV 2015.

# 网络应用

## 2. 目标检测(3)

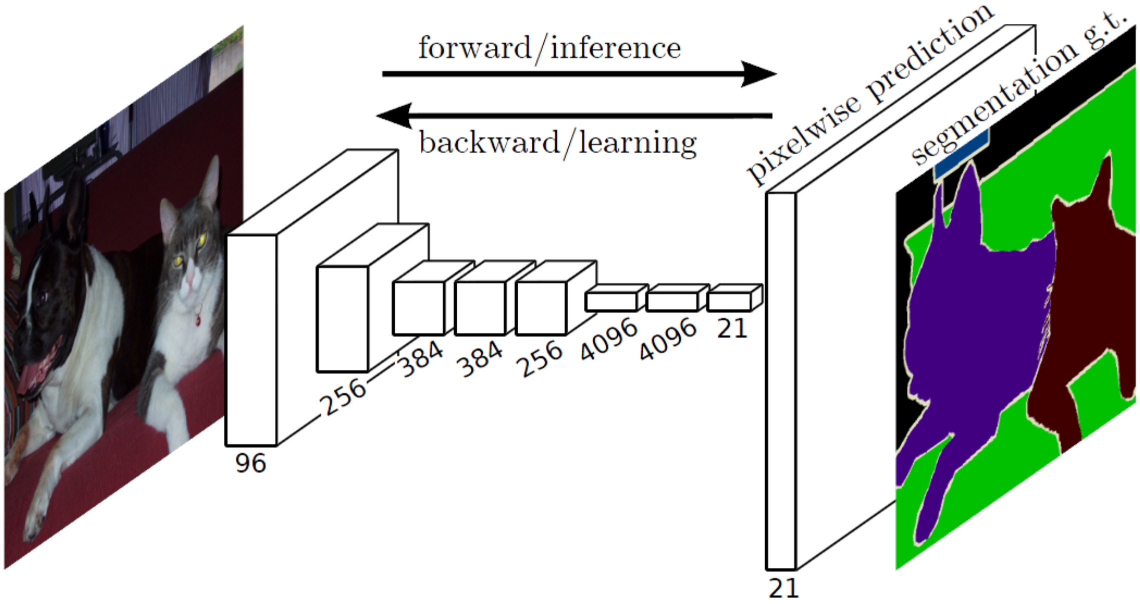



S. Ren et al., Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. NIPS 2015.

# 网络应用

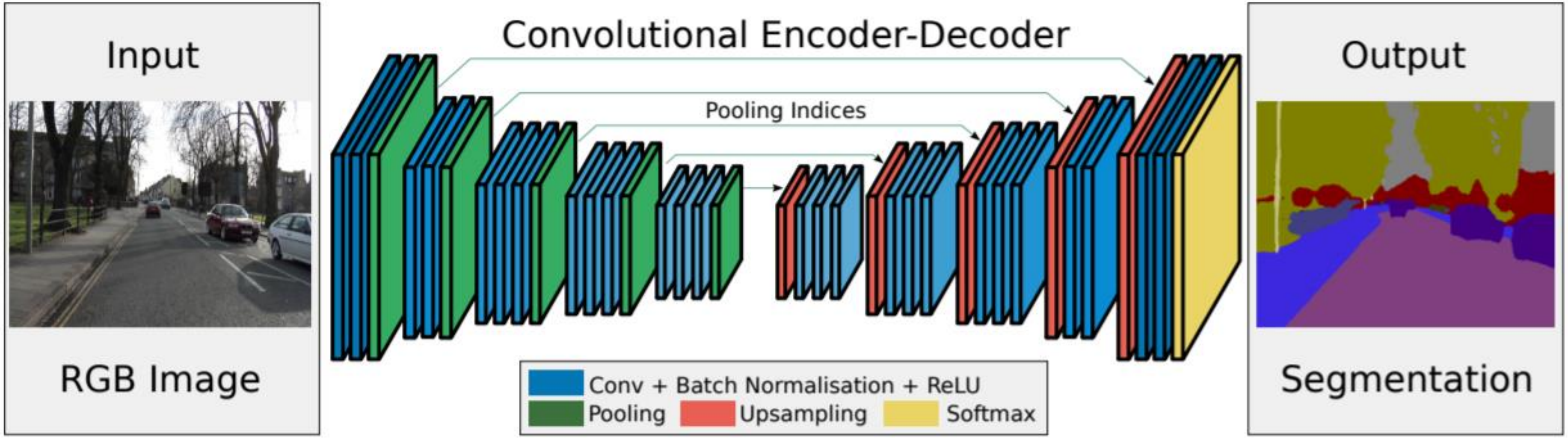
## 3. 语义分割




J. Long et al., Fully convolutional networks for semantic segmentation. CVPR 2015.

# 网络应用

## 4. 场景解析

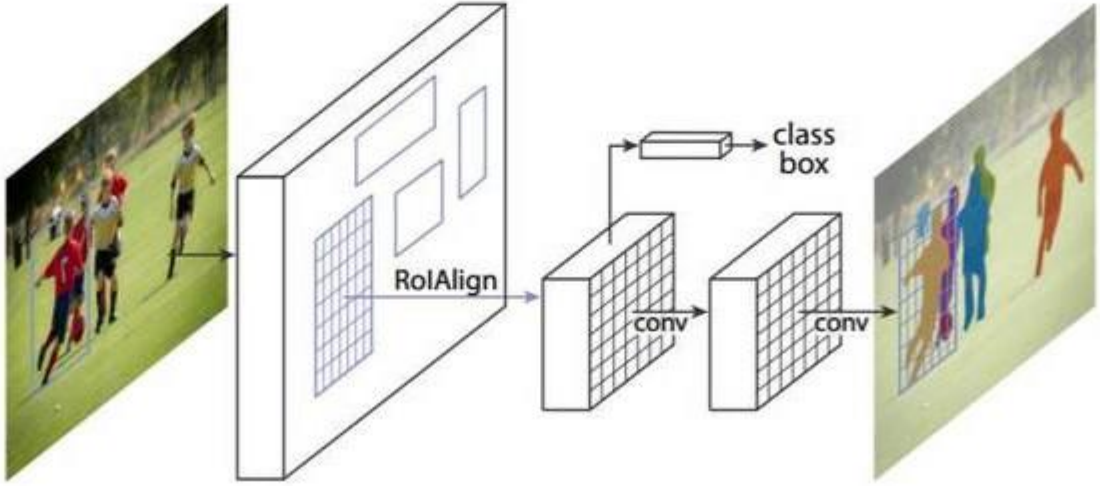



V. Badrinarayanan et al., SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, arXiv 2016.

# 网络应用

## 5. 样例级分割Mask R-CNN

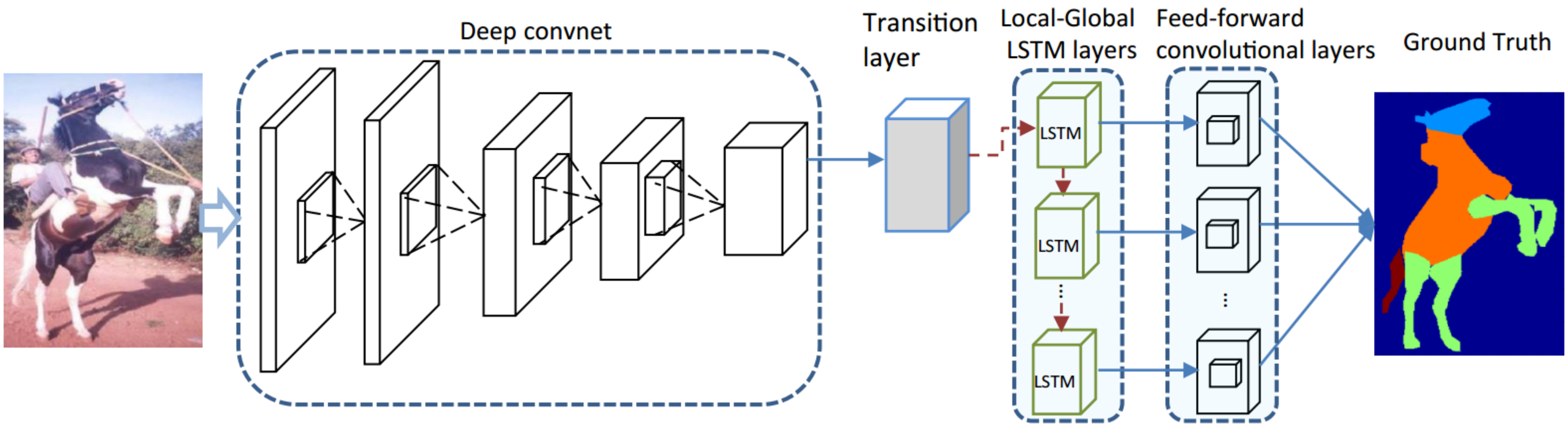
## （与faster R-CNN结构类似）




K. He et al., Mask R-CNN. CVPR 2017.

# 网络应用

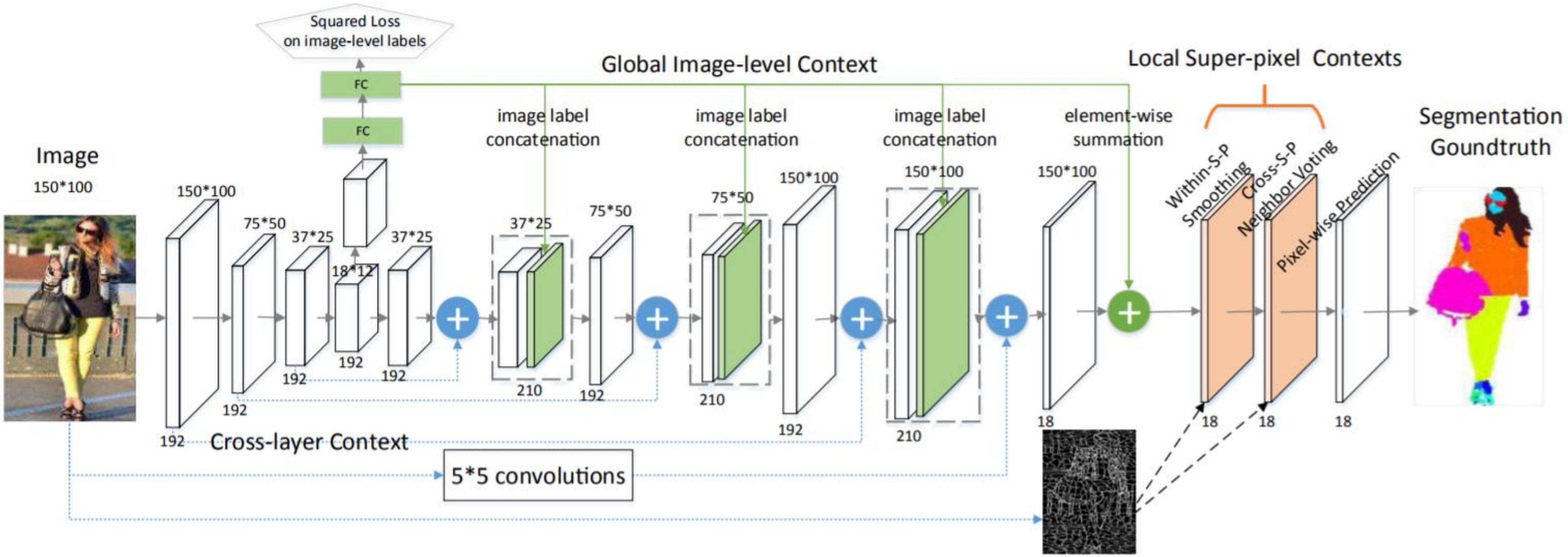
## 6. 目标/人体解析(1)




X. Liang et al., Semantic Object Parsing with Local-Global Long Short-Term Memory. CVPR 2016.

# 网络应用

## 6. 目标/人体解析(2)

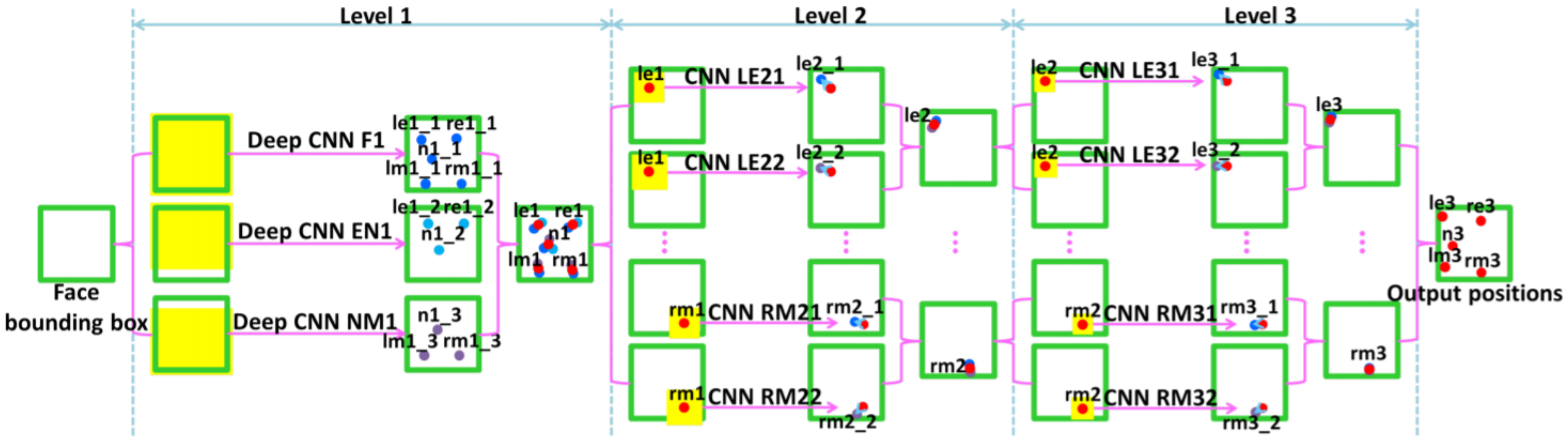



X. Liang et al., Human Parsing with Contextualized Convolutional Neural Network. ICCV 2015.

# 网络应用



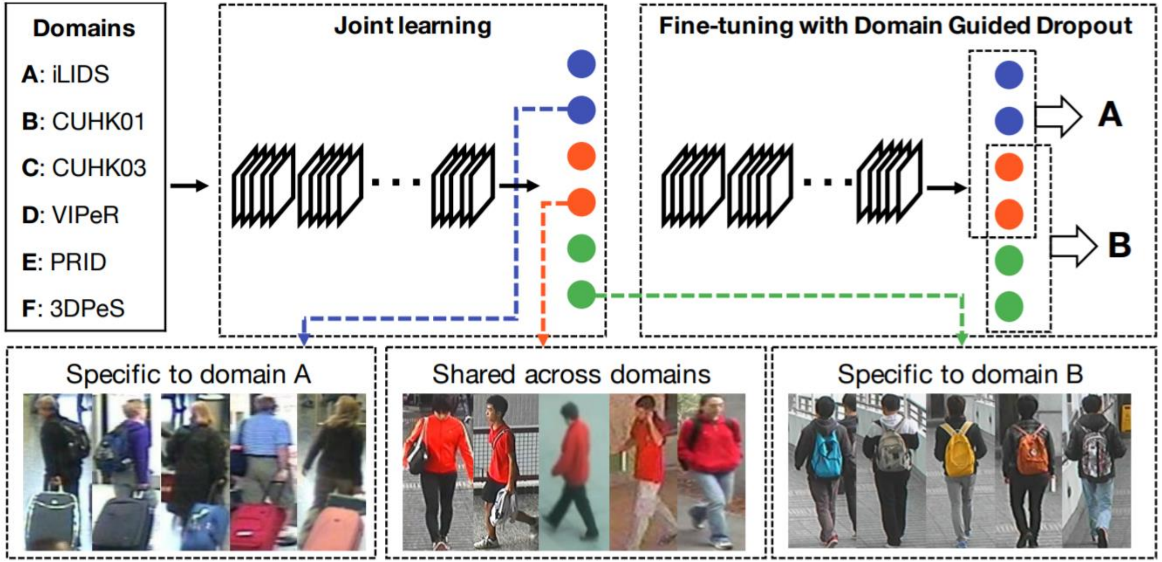
## 7. 人脸关键点检测




Y. Sun et al., Deep Convolutional Network Cascade for Facial Point Detection. CVPR 2013.

# 网络应用

## 8. 行人再识别（人脸识别）

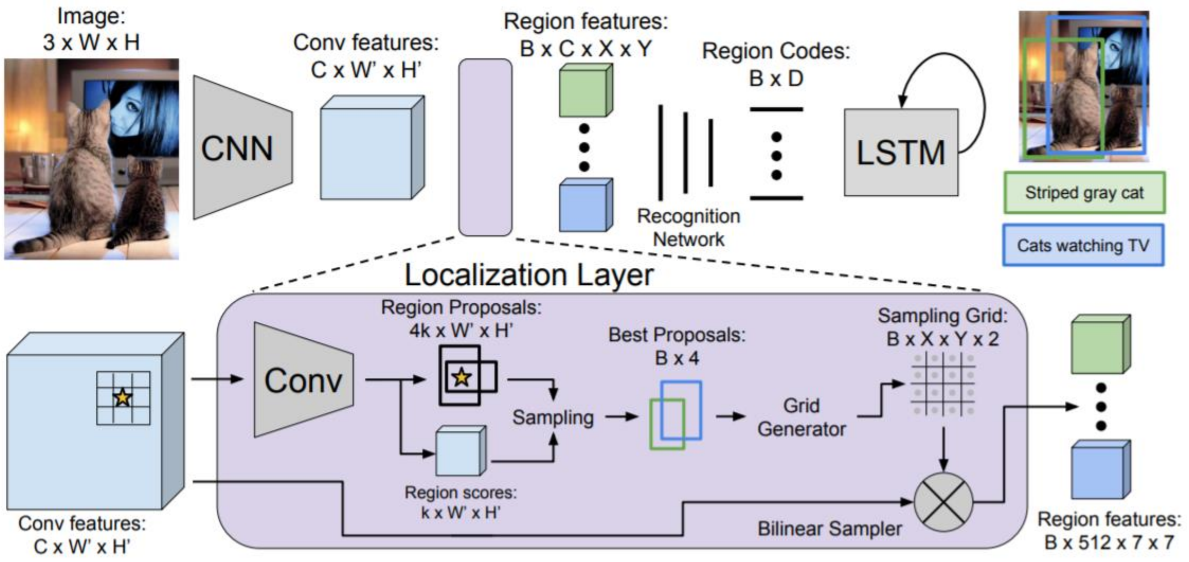



人脸识别？
人脸比对？

T. Xiao et al., Learning Deep Feature Representations with Domain Guided Dropout for Person Re-identification. CVPR 2016.

# 网络应用

## 9. 图像描述

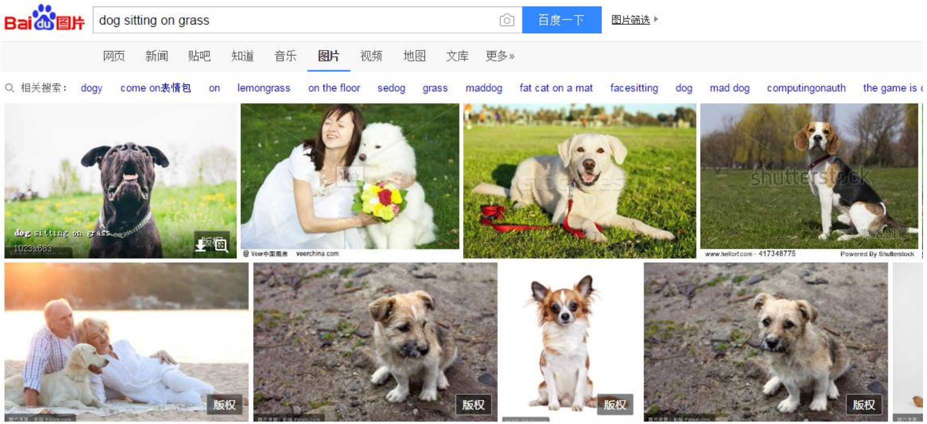




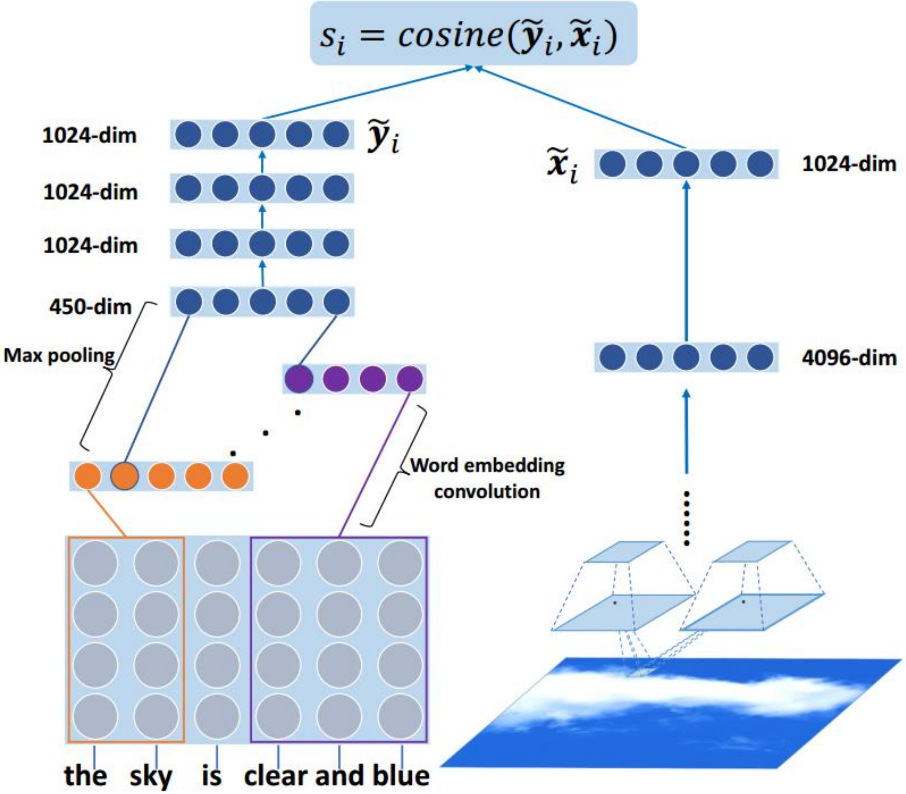
J. Johnson et al., DenseCap: Fully Convolutional Localization Networks for Dense Captioning. arXiv 2015.

# 网络应用

## 10. 跨模态检索

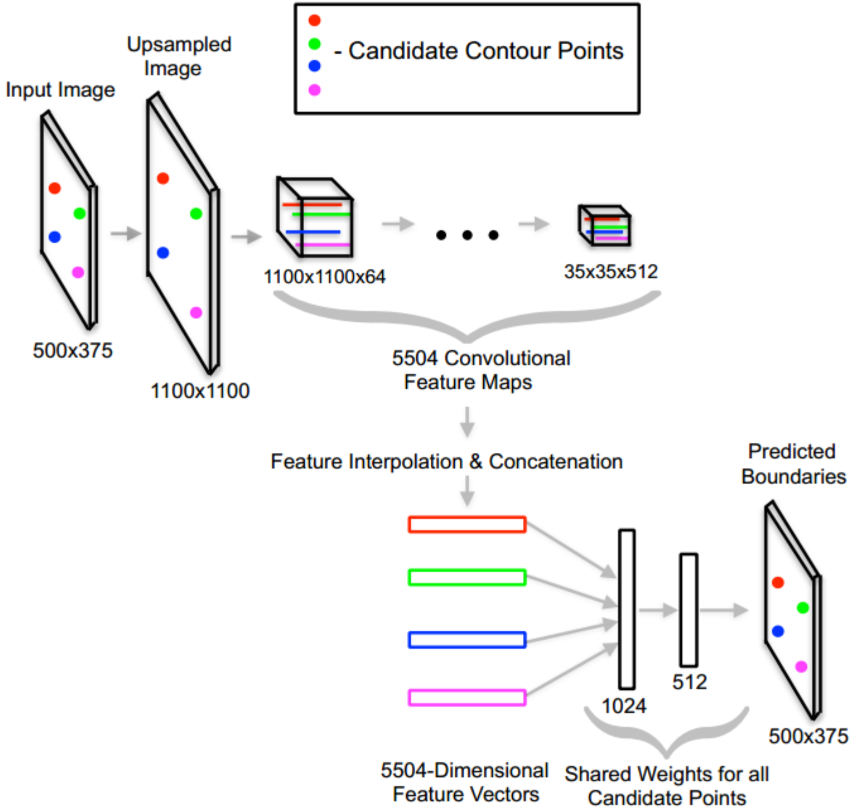



Y. He et al., Cross-Modal Retrieval via Deep and Bidirectional Representation Learning. TMM 2016.



# 网络应用

## 11. 边缘检测：用网络学到的高层语义特征指导低层视觉问题




G. Bertasius et al., High-for-Low and Low-for-High: Efficient Boundary Detection from Deep Object Features and its Applications to High-Level Vision. ICCV 2015.

# 网络应用

深蓝学院
shenlanxueyuan.com

## 12. 其他

- ✓ 图像生成、图像风格化、
- ✓ 显著性检测、图像质量评估、
- ✓ 图像检索、
- ✓ 医学图像分割、
- ✓ 视频目标跟踪、视频分割……

# 网络应用

深蓝学院
shenlanxueyuan.com

## 香港中文大学多媒体实验室：

人体部件：人脸关键点检测、人脸识别、人脸解析……

人体：姿态估计、人体解析、行人检测、行人再识别……

人群：人群分割、人群计数、人群行为分析与预测……

http://www.ee.cuhk.edu.hk/~xgwang/

# 参考资料


1. J. Bouvrie, Notes on Convolutional Neural Networks. Technical report 2006.

2. Y. LeCun et al., Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE 1998.

3. A. Krizhevsk et al., ImageNet Classification with Deep Convolutional Neural Networks. NIPS 2012.

4. Y. Bengio et al, Deep learning. https://github.com/HFTrader/DeepLearningBook .

5. 卷积神经网络相关的众多论文（CVPR、ICCV、ECCV、T-PAMI、arXiv等）。

# 在线问答



Q&A

# 感谢各位聆听！

Thanks for Listening