

CSDN

博客 (http://blog.csdn.net?ref=toolbar)学院 (http://edu.csdn.net?ref=toolbar)

下载 (http://down.csdn.net?ref=toolbar)更多 ▾

Q

写博客

发布Chat (http://gitbook.cn/new/gitchat/activity?utm_source=csdnblog1)

登录 (https://passport.csdn.net/account/login?ref=toolbar)注册 (http://passport.csdn.net/account/mobileregister?ref=toolbar&action=mobileRegister)

👍

【人体姿态】Stacked Hourglass算法详解

原创

2016年05月17日 22:55:47

标签：深度学习 (http://so.csdn.net/so/search/s.do?q=深度学习&t=blog) /

人体姿态 (http://so.csdn.net/so/search/s.do?q=人体姿态&t=blog) /

deep-learn (http://so.csdn.net/so/search/s.do?q=deep-learn&t=blog) /

算法 (http://so.csdn.net/so/search/s.do?q=算法&t=blog) / 图像 (http://so.csdn.net/so/search/s.do?q=图像&t=blog)

📖 8192

Newell, Alejandro, Kaiyu Yang, and Jia Deng. “Stacked hourglass networks for human pose estimation.” arXiv preprint arXiv:1603.06937 (2016).

概述

本文使用全卷积网络解决人体姿态分析问题，截至2016年5月，在MPII姿态分析竞赛 (http://human-pose.mpi-inf.mpg.de/#results)中暂列榜首，PCKh（误差小于一半头高的样本比例）达到89.4%。与排名第二的CPM(Convolutiona Pose Machine)1方法相比，**思路更明晰，网络更简洁**。

作者给出了基于Torch的代码和模型 (http://www-personal.umich.edu/~alnewell/pose/)。单显卡，测试时间约130ms，使用cudnn4的训练时间约3天，比CPM方法有显著优势。

本篇博客结合源码，从无到有介绍Stacked Hourglass的搭建思路，之后介绍代价函数与训练过程，最后总结值得学习的思想。

模块化

本篇论文的源码体现了模块->子网络->完整网络的设计思想。

Residual模块

先来复习一下卷积层和pooling层的属性：

核尺寸（kernel）决定了特征的**尺度**；步长（stride）决定了降采样的**比例**；算子的通道数（channel）决定了输出数据的**层数/深度**



shenxiaolu1984 (http://bl...

+ 关注

(http://blog.csdn.net/shenxiaolu1984)

原创	粉丝	喜欢	未开通
66	501	0	(https://gite
			utm_sourc

他的最新文章

更多文章 (http://blog.csdn.net/shenxiaolu1984)

【优化】共轭函数(Conjugate Function)超简说明
(/shenxiaolu1984/article/details/78194053)

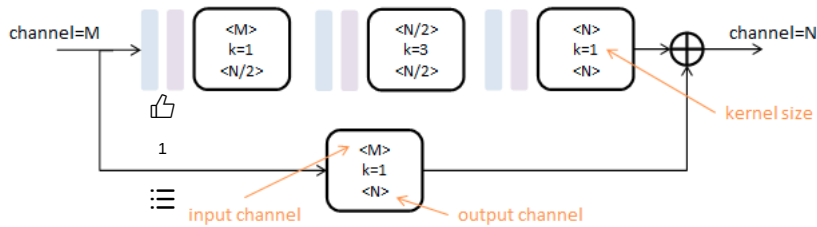
【优化】对偶上升法(Dual Ascent)超简说明
(/shenxiaolu1984/article/details/78175382)

【时间序列】时间序列分割聚类算法TICC
(/shenxiaolu1984/article/details/78134471)

【平价数据】SimGAN：活用合成数据和无监督数据
(/shenxiaolu1984/article/details/76160306)

在线课程

本文使用的初级模块称为Residual Module，得名于其中的旁路相加结构（在这篇论文中2称为residual learning）



第一行为卷积层，由三个核尺度不同的卷积层（白色）串联而成，间插有Batch Normalization（浅蓝）和ReLU（浅紫）；

第二行为跳级路，只包含一个核尺度为1的卷积层；如果跳级路的输入输出通道数相同，则这一路为单位映射。

所有卷积层的步长为1，padding为1，不改变数据尺寸，只对数据深度（channel）进行变更。

Residual Module由两个参数控制：输入深度M和输出深度N。可以对任意尺寸图像操作。

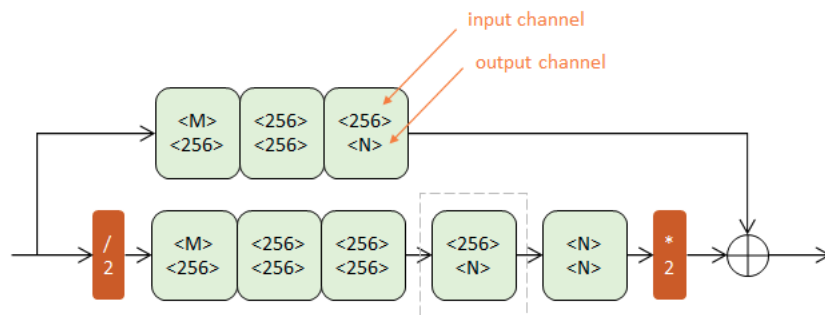
其实许多网络已经隐含了模块化的思想，例如AlexNet中重复出现的conv+relu+pool模式。

作用：Residual模块提取了较高层次的特征（卷积路），同时保留了原有层次的信息（跳级路）。不改变数据尺寸，只改变数据深度。可以把它看做一个保尺寸的高级“卷积”层。

Hourglass子网络

Hourglass是本文的核心部件，由Residual模块组成。根据阶数不同，有不同的复杂程度。

一阶Hourglass



上下两个半路都包含若干Residual模块（浅绿），逐步提取更深层次特征。但**上半路**在原尺度进行，**下半路**经历了先降采样（红色/2）再升采样（红色*2）的过程。
降采样使用max pooling，升采样使用最近邻插值。

另一种进行升采样的方法是反卷积层 (Deconv)，可以参看这篇解决分割问题的Fully Convolutional 论文 (<http://blog.csdn.net/shenxiaolu1984/article/details/51348149>)。



深度学习部署系统构建
([http://edu.csdn.net/course/detail/578?](http://edu.csdn.net/course/detail/578?utm_source=blog9)
([http://edu.csdn.net/huiyi](http://edu.csdn.net/huiyi/course/detail/578?utm_source=blog9)
course/detail/578?
utm_source=blog9)



搜索引擎翻译技术分享
(http://edu.csdn.net/huiyi
utm_source=blog9)
Course/detail/590?
utm_source=blog9)

热门文章

【目标检测】Faster RCNN算法详解
(shenxiaolu1984/article/details/51152614)
📖 69357

【目标检测】Fast RCNN算法详解
(shenxiaolu1984/article/details/51036677)
📖 51794

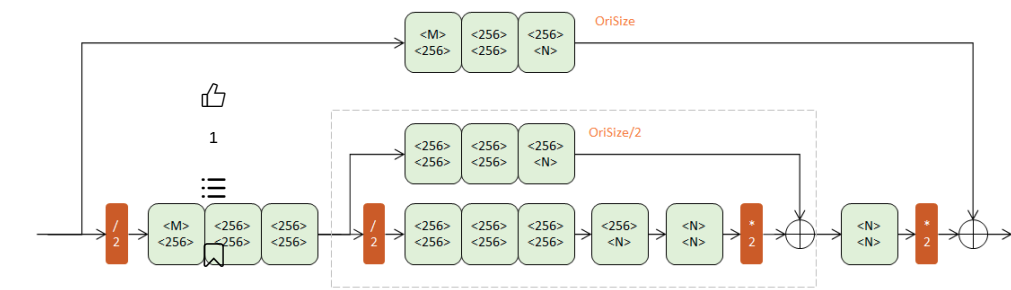
【目标检测】RCNN算法详解
(shenxiaolu1984/article/details/51066975)
📖 36792

【目标跟踪】KCF高速跟踪详解
(shenxiaolu1984/article/details/50905283)
📖 29117

【深度学习】生成对抗网络Generative Adversarial Nets
(/shenxiaolu1984/article/details/52215983)
📖 28315

二阶Hourglass

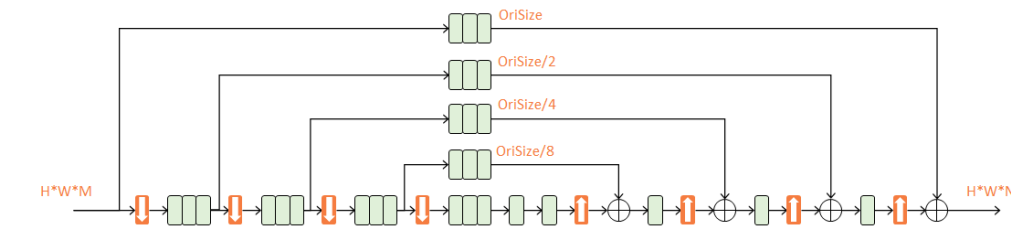
把一阶模块的灰框内部分替换成一个一阶Hourglass（输入通道256，输出通道N），得到二阶Hourglass：



两个层次的下半路组成了一条两次降采样，再两次升采样的过程。两个层次的下半路则分别在原始尺寸（OriSize）和1/2原始尺寸，辅助升采样。

四阶Hourglass

本文使用的是四阶Hourglass：



每次降采样之前，分出上半路保留原尺度信息；
每次升采样之后，和上一个尺度的数据相加；
两次降采样之间，使用三个Residual模块提取特征；
两次相加之间，使用一个Residual模块提取特征。

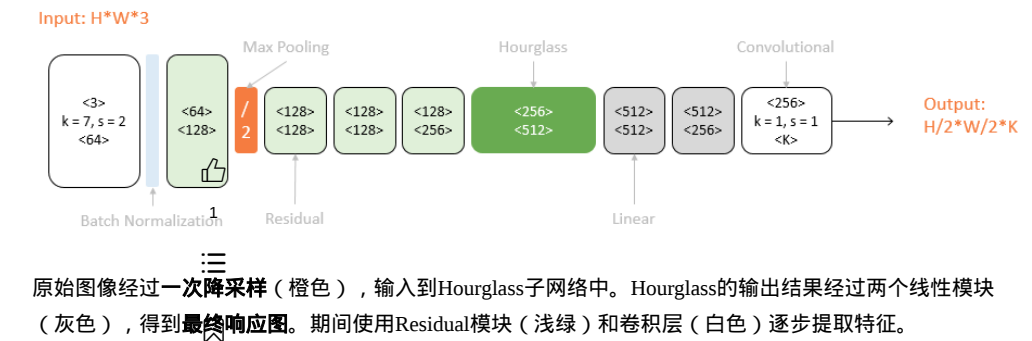
由于考虑了各个尺度的特征，本文不需要像CPM (<http://blog.csdn.net/shenxiaolu1984/article/details/51094959>)方法一样独立地在图像金字塔上多次运行，速度更快。

作用：n阶Hourglass子网络提取了从原始尺度到 $\frac{1}{2^n}$ 尺度的特征。不改变数据尺寸，只改变数据深度。

完整网络结构

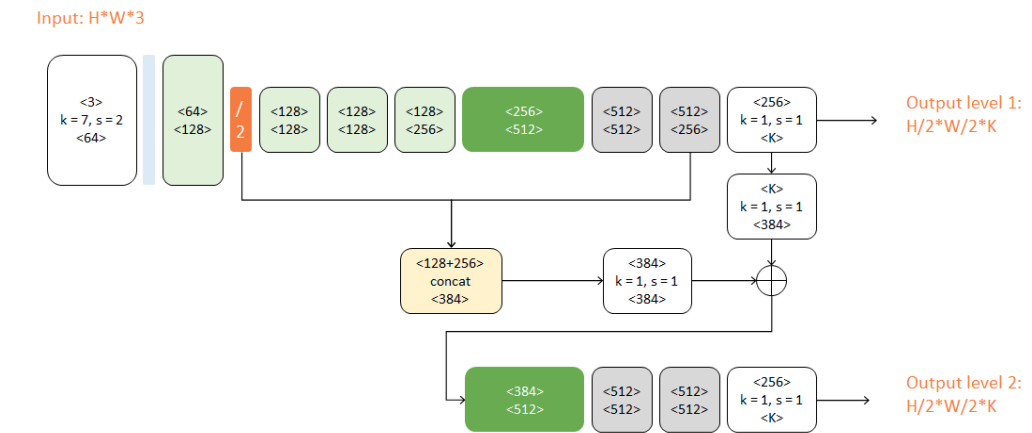
一级网络

以一个Hourglass（深绿色）为中心，可以从彩色图像预测**K个人体部件**的响应图：



二级网络

本文使用的完整网络包含两个Hourglass：



代价函数与训练

对于 的输入图像，每一个hourglass级都会生成一个 的响应图。对于每个响应图，都比较其与真值的误差作为代价。这种做法和CPM (<http://blog.csdn.net/shenxiaolu1984/article/details/51094959>)方法类似，都体现了**中继监督**(intermediate supervision)的思想。

在源码中，整个网络的输出结果包含每个级别的响应图，但在测试中只使用最后一级结果。这是因为torch的代价函数只能绑定在输出数据上。

使用cudnn4，在单个TitanX GPU(12G显存)上训练MPII (<http://human-pose.mpi-inf.mpg.de/>)数据，本文方法需要3天时间。

总结

本论文中值得学习的思想如下：

- 使用**模块**进行网络设计
- 先降采样，再升采样的**全卷积**结构
- **跳级**结构辅助升采样
- **中继监督**训练 1



1. Wei, Shih-En, et al. "Convolutional Pose Machines." CVPR, 2016 ↩

2. He, Kaiming, et al. "Deep Residual Learning for Image Recognition." arXiv preprint arXiv:1512.03385 (2015). ↩

3. Wei, Shih-En, et al. "Convolutional Pose Machines." CVPR, 2016 ↩

版权声明：本文为博主原创文章，未经博主允许不得转载。



dp_BUPT (/dp_BUPT) 2017-09-01 14:34 3楼

(/dp_BUPT)你好。中间层加的监督loss，CMP的groud truth每个Stack里的是不一样的，只是用的loss一样。而stacked-hourglass里groud truth都是一样的。想问问，这些中间层的loss都是放到最后一起反向传播计算吗？还是各层单独固定训练？

回复

sanqinsheng0236 (/sanqinsheng0236) 2017-08-17 21:57 2楼

(/sanqinsheng0236)里得出是三路数据融合的时候使用了串接（concat）的方式？因为我太不懂Lua，看了pytorch的复现好像没有看到cat,看论文里的解释也是三路直接加起来，所以希望得到指导，谢谢！

回复

rimless (/rimless) 2016-12-10 00:41 1楼

(/rimless)你写的太好了！感觉中继监督真的是很有用的技术。。

回复

相关文章推荐



阅读小结：Stacked Hourglass Networks for Human Pose Estimation (/layumi1993/article/details/52459385)

arXiv: <https://arxiv.org/pdf/1603.06937v2.pdf> github: <https://github.com/anewell/pose-hg-train> W...

Layumi1993 (<http://blog.csdn.net/Layumi1993>) 2016-09-07 14:34 2328

【人体姿态】Convolutional Pose Machines (/shenxiaolu1984/article/details/51094959)

2016年CVPR深度学习估计人体姿态的Convolutional Pose Machines算法

 shenxiaolu1984 (<http://blog.csdn.net/shenxiaolu1984>) 2016-04-29 18:33  12494



程序员同事工作3年，却不如我这个应届生

今年毕业找工作，工资比工作3年的前辈还要高，真的让我意识到，选择改变命运，如果不是当时....



1

(http://www.baidu.com/cb.php?c=lgF_pyfqHmsrH61rH60IZ0qnfK9ujYzP1D4PW630Aw-5Hc3rHnYnHb0TAq15HLPWRznjb0T1dBuHm3rHnzP10vPvP9PWbv0AwY5HDdnj63n1mYPWm0lgF_5y9YIZ0IQzq-uZR8mLPbUB48ugfEIAqspynEmybk5LNYUNq1ULNzmvRqmhkEu1Ds0ZFb5HDk0AFV5H00TZcqn0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1civrHnz0AqLL)



【行人识别】Deep Transfer Learning for Person Re-identification

(/shenxiaolu1984/article/details/53607268)

解决行人识别中的Re-Identification问题：判断两次出现的人是否是同一个人。在Market 1501竞赛中名列榜首。

 shenxiaolu1984 (<http://blog.csdn.net/shenxiaolu1984>) 2016-12-22 16:02  2516

TensorFlow中cnn-cifar10样例代码详解 (/diligent_321/article/details/53130913)



TensorFlow是一个支持分布式的深度学习框架，在Google的推动下，它正在变得越来越普及。我最近学了TensorFlow教程上的一个例子，即采用CNN对cifar10数据集进行分类。在看源代码...

 diligent_321 (http://blog.csdn.net/diligent_321) 2016-11-11 17:44  9952

Stacked Hourglass Networks for human pose estimation

(/human_recognition/article/details/51707934)

摘要 这篇文章提出了一个新颖的ConvNet架构，应用于人体姿态估计。作者认为重复使用 bottom-up，top-down能够提升网络性能。作者将这个网络命名为“stacked hourglass...”

 u013068978 (<http://blog.csdn.net/u013068978>) 2016-06-18 23:55  2085

【深度学习】聚焦机制DRAM(Deep Recurrent Attention Model)算法详解



(/shenxiaolu1984/article/details/51518578)

Visual Attention基础，Multiple object recognition with visual attention算法解读。

 shenxiaolu1984 (<http://blog.csdn.net/shenxiaolu1984>) 2016-06-28 22:14  4897

较详细的单例实现模式地址 (/loudspeaker1987/article/details/6524050)

<http://wenku.baidu.com/view/2cddb3ea81c758f5f61f670a.html> 还有一种就是通过枚举：public enum Elvis{ INSTANC...

 loudspeaker1987 (<http://blog.csdn.net/loudspeaker1987>) 2011-06-03 09:17  46

【图像分割】Fully Convolutional Networks for Semantic Segmentation



(/shenxiaolu1984/article/details/51348149)

全卷积网络用于图像语义分割，2015 CVPR Oral，算法详解。

 shenxiaolu1984 (<http://blog.csdn.net/shenxiaolu1984>) 2016-05-11 15:33  6115



《Towards Viewpoint Invariant 3D Human Pose Estimation》--深度图领域人体姿态估计的CNN算法 (/zhangboshen/article/details/70833062)

《Towards Viewpoint Invariant 3D Human Pose Estimation》-深度图领域人体姿态估计的CNN算法 这篇文章是ECCV 2016的一篇3D人体姿态估计的文...

 zhangboshen (<http://blog.csdn.net/zhangboshen>) 2017-04-26 21:56  940



论文实践学习 - Stacked Hourglass Networks for Human Pose Estimation (/zziahgf/article/details/72763590)

Stacked Hourglass Networks for Human Pose Estimation - Demo CodeStacked Hourglass Networks for Human...

 oJiMoDeYe12345 (<http://blog.csdn.net/oJiMoDeYe12345>) 2017-05-26 09:48  1145



论文阅读理解 - Stacked Hourglass Networks for Human Pose Estimation (/zziahgf/article/details/72732220)

Stacked Hourglass Networks for Human Pose Estimationkeywords 人体姿态估计 Human Pose Estimation 给定单张RGB图像, ...

 oJiMoDeYe12345 (<http://blog.csdn.net/oJiMoDeYe12345>) 2017-05-25 17:48  1177

基于410c开发板和摄像头的人体姿态估计（预研分析） (/andymfc/article/details/52537218)


最近自己想用410c来做一个运动相关的DIY，用410c开发板搭载摄像头来实现对自己的姿态的跟踪，用到健身运动中，完成对自己所作的动作的定位和标准评判，实时提醒自己所作健身动作的标准程度，前期进行了一...

 ANDYMFC (<http://blog.csdn.net/ANDYMFC>) 2016-09-14 13:22  1384



http://download.csdn.net/detail/qq_15297293/7335599

人体手姿态捕捉 (http://download.csdn.net/detail/qq_15297293/7335599)



2014-06-17 2.16MB

[下载](#)



<http://download.csdn.net/detail/u014734471/9986309>

基于Kinet+openni的人体骨架提取及姿态识别 (<http://download.csdn.net/detail/u014734471/9986309>)

UFLDL——Exercise: Stacked Autoencoders栈式自编码算法 (/danieljianfeng/article/details/41926365)

2017-09-19 22:42 19.93MB [下载](#)

实验要求可以参考deeplearning的tutorial， Exercise: Implement deep networks for digit classification 。 本实验仍然是...

 Daniel_djf (http://blog.csdn.net/Daniel_djf) 2014-12-14 17:08  2192



<http://download.csdn.net/detail/qianxu050/1565652>

基于支持向量机的多种人体姿态识别. (<http://download.csdn.net/detail/qianxu050/1565652>)



2009-08-11 17:17 635KB

[下载](#)



http://download.csdn.net/detail/qq_26487681/8492137

姿态测量算法研究 (http://download.csdn.net/detail/qq_26487681/8492137)

DCM姿态估计算法分析 (/whyscience/article/details/55218134)

2015-03-11 17:00 104KB [下载](#)

DCM姿态估计算法分析飞控算法 以下内容主要是阅读Direction Cosine Matrix IMU: Theory (by William Premerlani and Paul Bizard) ...

 whyscience (<http://blog.csdn.net/whyscience>) 2017-02-15 20:22  696