

请输入您要查询的信息

- 1 从零开始学日语
- 2 从零开始学口语
- 3 网页制作学习

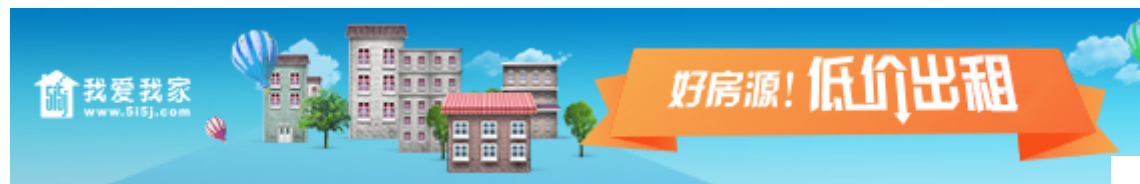


首页 脚本 编程 数据库 CMS 设计 动画 系统 网站 安全 制作 冲浪 组网 软件 手机 邮件 认

频道首页 >> IT博客 >> 研发管理 >> 文章内容

深度卷积网络CNN与图像语义分割

日期：2015/8/20 15:57:00 来源：本网整理



转载请注明出处: <http://xiahouzuoxin.github.io/notes/>

级别1：DL快速上手

级别2：从Caffe着手实践

级别3：读paper，网络Train起来

级别3：Demo跑起来

读一些源码玩玩

熟悉Caffe接口，写Demo这是硬功夫

分析各层Layer输出特征

级别4：何不自己搭个CNN玩玩

级别5：加速吧，GPU编程

关于语义分割的一些其它工作



网页制作学习



最新添加

CentOs7下GitHub入门复习

Git24--搭建Git服务器

OpenCV训练自己的人脸检测级连分类器并测试

python 实现识别手写 MNIST数字集的程序

Verilog基础知识5（ASIC设计中各个阶段的关键问
com.google.guava工具（1）

I2C详解

Verilog基础知识8（异步FIFO的FPGA实现）

Linux内存管理基础

说好的要笔耕不辍，这开始一边实习一边找工作，还摊上了自己的一点私事困扰，这几个月的东西都没来得及总结一下。这就来记录一下关于CNN、Caffe、Image Sematic Segmentation相关的工作，由于公司技术保密的问题，很多东西没办法和大家详说只能抱歉了。在5月份前，我也是一个DL和CNN的门外汉，自己试着看tutorials、papers、搭Caffe平台、测试CNN Net，现在至少也能改改Caffe源码（Add/Modify Layer）、基于Caffe写个Demo。这里希望把学习的过程分享给那些在门口徘徊的朋友。没法事无巨细，但希望能起到提点的作用！

下面有些链接若无效，请跳转原文阅读：<http://xiahouzuoxin.github.io/notes/html/深度卷积网络CNN与图像语义分割.html>

级别1：DL快速上手

UFLDL：<http://deeplearning.stanford.edu/tutorial/>

这是stanford Ng老师的教材，也刚好是以CNN为主，Ng老师教材的特色就是简洁明白。一遍看不懂多看两遍，直到烂熟于心，顺便把里面的Matlab Exercises完成了。

<http://deeplearning.net/tutorial/>

PRML作者给的python入门DL的tutorial，基于Theano Framework，有些偏向于RNN的东西。

一句简单的话描述：“深度学习就是多层的神经网络”。神经网络几十年前就有了，而且证明了“2层（1个隐层）的神经网络可以逼近任意的非线性映射”，意思就是说，只要我的参数能训练好，2层神经网络就能完成任意的分类问题（分类问题就是将不同类通过非线性映射划分到不同的子空间）。但2层神经网络存在的问题是：

如果要逼近非常非常复杂的非线性映射，网络的权值W就会很多，造成Train时候容易出现的问题就是Overfitting。所以大事化小，将复杂问题进行分割，用多层网络来逼近负责的非线性映射，这样每层的参数就少了。自然而然的网络就从2层变成了多层，浅网络(shallow)就变成了深网络(deep)。

但科研界的大牛们会这么傻吗，十几年前会想不到用多层网络来进行非线性映射？看看CNN最早的工作：<http://yann.lecun.com/exdb/publis/pdf/lecun-98.pdf> 那是98年的，Train了一个5层的CNN来进行MINIST数据集的数字图片分类。多层神经网络一直不火我觉得有这么两个原因：

神经网络中非线性的映射的极值优化问题本身是一个非凸问题，本身数学理论上的就对非凸优化问题缺少严格有效最优化方法的支撑。直到现在也依然对各层Layer的输出解释不清楚，但效果就是好，这还得归功于各种大神藏之捏之的各种Tricks

使用ISAPI过滤器开发来增强IIS的商品关联分析

Verilog基础知识7（FIFO深度计算）

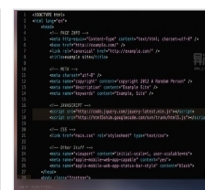
Maven学习总结（22）——Maven常session如何加入mysql库

go语言十大排序算法总结

猜你感兴趣



美工学习



编程学习入门



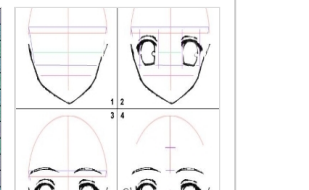
学习化妆



化妆学习



日语自学网



漫画学习

点击排行

深度卷积网络CNN与图像语义分割

SpringMVC中使用Interceptor拦截器

Permission denied (publickey) fatal:Could not read 解决 TortoiseGit 诡异的 Bad file number 问题

深度学习深度信念网络DBNs—简易详解

Deep Learning 系列（3）：CNN（卷积神经网络）

Xaml技术：浅谈Grid.ColumnDefinitions和

R语言与数据分析之八：时间序列--霍尔特指数平

Jenkins使用新经验

OpenWrt 上搭建代理服务器squid

数据与计算能力的问题。十来年前哪来随随便便就这么大的硬盘，哪里去找像ImageNet这样1000类的数据集。“大数据是燃料，GPU是引擎”，正是因为大数据的出现和GPU编程的出现带动了DL的进展，这些在10年前是做不来的。我在CPU与GPU上跑自己简化的Googlenet，GPU比CPU快10倍。

DL只是一个概念而已。对于做图像和视觉的就该一头扎到CNN(Convolutional Neural Network)，做自然语言的就该投入到RNN(Recurrent Neural Network)。我是做图像的。CNN的学习资料除了上面Ng的tutorial外，还有一个Stanford Li Fei-Fei教授的课程cs231：Convolutional Neural Networks for Visual Recognition，<http://cs231n.github.io/convolutional-networks/> 是Notes中一份关于CNN非常详细的资料。

级别2：从Caffe着手实践

先看看这个热身：贾扬清：希望Caffe成为深度学习领域的Hadoop，增加点学习的欲望，毕竟现在多少人靠着Hadoop那玩意儿挣着大钱。

接着请认准Caffe官方文档：<http://caffe.berkeleyvision.org/> 和Github源码：<https://github.com/BVLC/caffe>。毫不犹豫fork一份到自己的Github。然后就是照着INSTALL来Compile和Config Caffe了，值得注意的是，安装OpenCV的时候推荐使用源码安装。

先自己熟悉Caffe的架构，主要参考资料就是官网文档，我自己刚开始的时候也写了个小的ppt笔记：Diving into Caffe.pptx

接着就是实实在在地分析一个CNN，比如LeNet、AlexNet，自己在纸上画一画，下面那样



AlexNet

级别3：读paper，网络Train起来

当去搜索ICRL、CVPR、ICCV这些最前沿的计算机视觉、机器学习会议的时候，只要是涉及图像相关的深度学习实验，大都是基于Caffe来做的。所以，只要抓住1~2篇popular的paper深入，把论文中

Matlab学习-----GUI鼠标事件响应 (C
KVM上如何绑定虚拟机vcpu与物理C
TestNG的组测试和组中组测试
Erlang Rebar 使用指南之四：依赖管理
libyuv库的使用



的CNN在Caffe上复现了，就能找到一些感觉了。在这期间，下面一些论文是至少要读的：

LeNet-5: Gradient-Based Learning Applied to Document Recognition CNN首篇paper，虽然是1998年的文章，但依然值得仔细一读。

AlexNet: ImageNet Classification with Deep Convolutional Neural Networks 自我感觉是促进CNN的扛鼎之作，似乎很多所谓的Tricks在这篇文章中能找到，看这篇文章就是来学Tricks的。

Googlenet: Going Deeper with Convolutions ImageNet竞赛Number1，有效的Inception结构构建深层网络

VGGNet: Very Deep Convolutional Networks for Large-Scale Image Recognition ImageNet竞赛Number2，典型的卷积+Pooling方式构建深层网络，但是由于没有Googlenet中Inception的1x1的convolution用于减小网络厚度，世间上要比Googlenet慢一些。

具体到用CNN做Sematic Segmentation，利用到全卷积网络，对下面两篇进行了精读，并且都Caffe上复现过并用于分割任务，

FCNN: Fully Convolutional Networks for Semantic Segmentation

Deeplab: Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs

下面是几个月前我看过这两篇paper后做得ppt：

FCN for Sematic Segmentation.pptx

Semantic Image Segmentation With Deep Convolutional Nets and Fully Connected CRFs.ppt

级别4：Demo跑起来

读一些源码玩玩

caffe.proto

Convolution Layer

SoftmaxLossLayer

DataLayer



自己实现个IoULayer

熟悉Caffe接口，写Demo这是硬功夫

Caffe提供了好用的接口，包括matlab、C++、Python！由于特殊原因，我不能公开我C++和matlab的Demo源码以及其中的一些后处理技术，暂且只能给大家看一些分割的结果：



Sematic Segmentation结果



Sematic Segmentation结果

还有一个视频语义分割的结果，大家看看，热闹热闹就好，

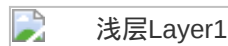


分析各层Layer输出特征

我一开始以为看看各层Layer的输出，能帮助我改进Net，可却发现错了，除了前几层还能看出点明亮或边缘信息外，网络后端Layer的输出压根就没办法理解。extract_featmat.cpp是我基于extract_features.cpp改的一个Caffe tool，放到tools目录下编译就好了，使用方法看help：

```
void print_help(void) { LOG(ERROR)<< "This program takes in a trained network and an input image, and then\n" " extract features of the input data produced by the net.\n" "Usage: extract_featmat [options]\n" " -model [pretrained_net_param]\n" " -proto [feature_extraction_proto_file]\n" " -img [rgb_image_name]\n" " -blobs [extract_feature_blob_name1[,name2,...]],reference .prototxt with " "\blob:\". [all] for all layers. \n" " -wr_prefix [output_feat_prefix1[,prefix2,...]], correspond to -blobs\n" " -wr_root [output_feat_dir], optional, default \"./\", make sure it exist\n" " -gpu [gpu_id],optional,if not specified,default CPU\n"; }
```

下面图是一些Layer的输出blob，从结果可以看出，前面的layer还能看到一些边缘信息，后面的layer就完全看不出原图像相关的信息了，



浅层Layer1

浅层Layer1



浅层Layer2

浅层Layer2



浅层Layer3

浅层Layer3



深层Layer

深层Layer

级别5：何不自己搭个CNN玩玩

虽然还是个新手，关于搭建CNN，还在慢慢在找感觉。我觉得从两方面：

利用已有的网络，使劲浑身解数找它们的缺点，改进它们

熟读Googlenet和VGGnet那两篇paper，两者的CNN结构如下：



GoogleNet



GoogleNet



VGGNet

VGG不是Weight Filter不是非常厚么，卷积操作复杂度就高。而Googlenet通过Inception中1x1的Convolution刚好是为了减少Weight Filter的厚度，我最近一段时间在尝试做的事就是将VGG中的Layer用Googlenet中的Inception的方式去替代，希望至少在时间上有所改进。

从头搭建一个CNN用于解决实际问题。一个词：搭积木。

先搭一个简单的，比如说就3层：卷积-Pooling-卷积-Pooling-卷积-Pooling，先把这个简单的网络训练以来，效果不好没关系，我们接着往上加，直到满意为止。但是这里面有一个finetune的技巧，那就是用浅层的网络训练weight结果去初始化或finetune深层网络。这也是为什么不直接一开始就搭建深层网络的原因，级别1里就说过，深度网络的Train是个非凸问题，是个至今难解决的大问题，网络初始化对其收敛结果影响很大，finetune就这样作为Deep Network中一项重要的tricks而存在了。finetune除了由浅至深逐级初始化帮助收敛外，还有一个作用：将自己的网络在一个非常非常大的数据集上(现在最大的ImageNet)进行Train，这个Train的结果再拿去作为实际要解决的问题中用于初始化，这样能增加网络的泛化能力。然而，当Net遇到问题时，如何去改进？这是个大问题，不说现在没太多经验，有也是盲人摸象的感觉，暂且搁下不提，待后期多做实验分析整理，希望能有所收获。

就我现在的水平，最多也就修炼到这一级了。这一级还要多花功夫，读paper，多思考，继续练。。。

级别6：加速吧，GPU编程

呃，这个实在还没开始着手去做，但迟早是要做的，说了“大数据是燃料，GPU是引擎”的，怎么能不懂引擎呢.....

关于语义分割的一些其它工作

CRF：CRF在图像分割中是最常见的refine后处理手段。在CNN中目标是做成end-2-end的CRF，实习这段时间也做过不少这部分的工作，Oxford有篇CRF-RNN的paper，将denseCRF重新解释成RNN来进行end-2-end的Training



分享到

结合grabcut交互式分割，或者SLIC超像素分割等方法进行边缘精细化的处理

不闲扯淡了.....

注：由于“实习上班+实验室+论文+刷leetcode+私事”占用时间的关系，好不容易抽出一个上午+一个晚上整理了一下，暂时想到这么多，算列个提纲吧，文章中不少具体细节有机会再补充。

版权声明：本文为博主原创文章，未经博主允许不得转载。

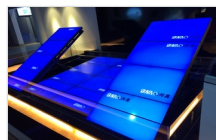
本文转载自xiahouzuoxin博客,版权归xiahouzuoxin所有

分享到： QQ空间 新浪微博 腾讯微博 人人网 微信

猜你喜欢：



化妆学习



液晶拼接



漫画学习



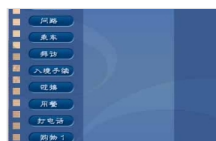
中式装修



网页制作学习



荨麻疹好了



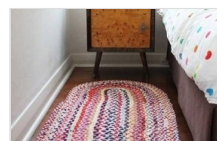
从零开始学日语



零基础学习英语



电动老爷车



手工地毯



从零开始学口语



美工学习



编程学习入门



发酵床养鹅



今日肉牛价

本文相关：

Git flow的分支模型与及常用命令简介

使用ssh公钥实现面密码登录：RSA公钥

需求分析的关键方法

Git版本控制

Atitit. 软件设计 模式 变量 方法 命名最佳实践 vp820 attilax总结命名表大全

Windows环境下msysgit安装git flow



递归降序遍历目录层次结构，并按文件类型计数

我的openwrt开发相关文章

unix网络编程几种模型比较

我的openwrt学习笔记（四）：OpenWrt源代码下载

为你推荐

研发管理

项目管理 软件

工程管理软件

卷积

比基尼除毛

销售管理

三室两厅装修样板间

语义

学网

分割

项目管理软件

从零开始学日语

研发管理

百度一下

上一条：算法学习之排序算法（五）（快速排序）

下一条：Git flow的分支模型与及常用命令简介



免责声明 - 关于我们 - 联系我们 - 广告联系 - 友情链接 - 帮助中心 - 频道导航

Copyright © 2017 www.zgxue.com All Rights Reserved