

[首页](#)[论文下载](#)[杂志投稿](#)[论文发表百科](#)[FAQ](#)[论文参考书](#)[首页](#) > [论文发表百科](#) >[快捷服务](#)

基于cnn的人脸检测_tensorflow人脸识别_卷积神经网络人脸识别

论文发表网 发布时间：2016-08-18 00:03:27 来源：网络整理 浏览数：746



下载

完整？感觉不错，有学习店1-5元下载，方便，快速度文库、豆丁、道客巴巴均可快速方便下载，是您手。

论文发表

有论文找不到发表渠道？本站与省级、国家级、北大核心、南大核心CSSCI、CSCD、统计源科技核心等杂志社进行合作，为您提供稳妥的发表渠道。

论文代写

没时间没思路没资料不愿意自己写作？提交您的代写需求，由我们的专业硕博老师在24小时内为您私人定制，仅供您个人进行参考学习，完全无后顾之忧。

[摘要：1、 相关实际本篇专文首要讲授2013年CVPR的一篇行使深度进修做人脸特点面定位的典范

paper：《Deep Convolutional Network Cascade for Facial Point Detection》，论文的]

一、相关理论

本篇博文主要讲解2013年CVPR的一篇利用深度学习做人脸特征点定位的经典paper：《Deep Convolutional Network Cascade for Facial Point Detection》，论文的主页为：。网页提供了训练数据、测试demo，但是我却没有找到源码，所以只能自己写代码实现。这篇paper是利用深度学习搞人脸特征点定位的开山之作，想当年此算法曾经取得了state-of-art 的结果。后来face++发表的几篇paper的思想，都是在这篇paper的算法上进行扩展的。

如果之前没有学过类似DCNN的思想的话，那么会感觉相当难，至少我是这么觉得的。在之前学过的各种深度学习模型中，一般就只有一个CNN，可这篇paper是由十几个CNN组成的。我一看到文献中如下网络

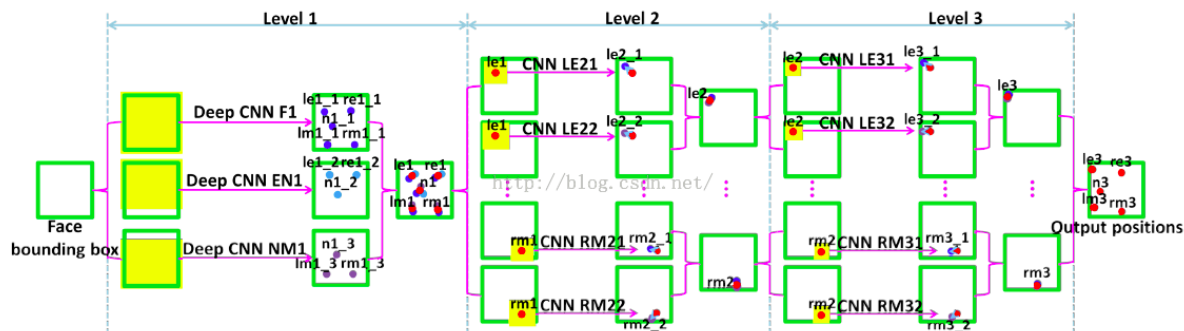
[论文下载](#)[专利申请](#)[论文发表](#)[教材挂名](#)[论文代写](#)[淘宝老店](#)

精品期刊

福建茶叶：北大核心，适合高职二本三本院校的难以发表核心而又要评副高级高级职称的老师，100%包发表，专业基本不限。

现代电子技术、计算机仿真：北大核心、CSCD，适合XX大学或2本以上院校的硕士讲师以上作者，100%包发表，限理工专业。

结构图片：



作为菜鸟的我，就已经晕了，然后在看一下文献的一些公式，我彻底没了耐性。因为我看外文paper，



直接看图、看公式，可这篇paper的图，一上来就把我吓到了。之前学过反卷积网络，FCN、R-CNN、OverFeat、Siamese Network、NIN等等各种网络，感觉花个一两天的时间，总可以把算是学了这篇paper之后，感觉难度完全不是一个等级的，因为初次接触这种DCNN，花了好几周才把源码实现出来，过程相当痛苦，好吧，还是不罗嗦了，这篇博文将仔细讲解paper算法的实现与不一样的CNN模型，也就是传说中的DCNN（这篇paper，后来人们又把这种多层次的CNN模型称为DCNN）。接着我将先简单介绍，文章的主要思想。

1、RCNN回顾

paper的总思想是一种从粗到精的定位思想，如果你之前已经对R-CNN有所了解，我们就回忆一下现在的一些物体检测的大体算法思路，因为我觉得RCNN也是一种从粗定位，到精识别的过程，跟本篇paper的思想很相似，下面是RCNN的流程图(RCNN文献《Region-based Convolutional Networks for Accurate Object Detection and Segmentation》)：

商业经济研究：北大核心，适合各层次大学院校的经济管理大类专业的老师，通过率较高。

职教论坛、教学与管理：北大核心，适合职业院校或大学的作者，要求职业教育或基础教育相关内容。

中国药房、重庆医学：北大核心，适合地市级三甲医院的医生。

经济研究导刊：2016年中国科技核心期刊目录（社会科学卷）收录期刊，包发，经济管理专业，作者不限。

论文发表相关期刊

社科期刊：发表与社会科学专业相关的论文

经管期刊：发表与经济管理专业相关的论文

科技期刊：发表与科学技术工程科技专业相关的论文

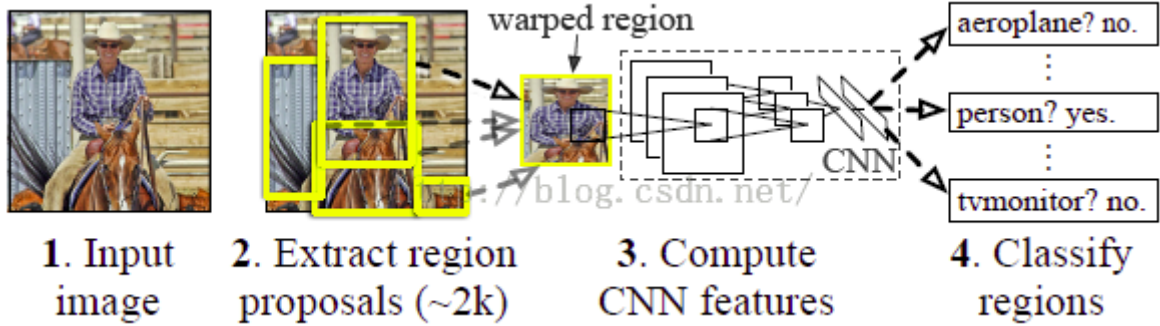
教育期刊：发表与教育教学专业相关的论文

医学期刊：发表与医学医疗医药专业相关的论文

文艺期刊：发表与文学艺术专业相关的论文

法律期刊：发表与法律法制专业相关的论文

省级期刊：省级管理机构主办或管理的期刊，评初级、中级可用



首先输入一张图片，然后RCNN的思路是：

(1)采用传统的方法，先把一张图片中各个可能的物体的bounding box分割出来



各个可能的裁剪出来的物体检测框，输入cnn，进行特征提取，最后进行物体分类识别

算法的总体思路很简单。其实这也是一种从粗到精的思想，就是我们对输入的一整张图片先进行定位出物体的bounding box，也就是物体的大体的位置。然后进行裁剪(bouding box)，这样我们的图片范围就变小了，外界背景因素的干扰就减小了，然后我们在把裁剪后的图片，作为CNN的输入，这样有利于提高精度。总之，假设你要用CNN做人脸识别，那么你不应该把一整张图片，包括风景、背景、身体、头发这些背景因素都扔进CNN中，这样的精度会比较小。你应该是先用人脸检测器，把人脸部位的图片裁剪出来，送入CNN中，精度才会较高。OK，可能你觉得，我这边讲的话题有点废话，其实《Deep Convolutional Network Cascade for Facial Point Detection》这篇paper的思想就是这样，很简单的一个思想。

2、paper的思想

国家级期刊：国家级管理机构主办或管理的期刊，评初级、中级可用

北大核心期刊：北京大学中文核心期刊目录收录的期刊，评副高、高级可用

南大核心CSCSI期刊：南京大学中文社会科学索引目录收录的期刊，评高级可用

统计源科技核心期刊：为统计源科技核心目录所收录，分为社科类和自然科学类，评中级、副高可用

CSCD期刊：中国自然科学索引数据库所收录的期刊，评副高、高级可用

EI/SCI期刊：EI工程索引、SCI自然科学索引所收录的国际期刊，效力等同或高于北大核心\CSCSI期刊

其他增值服务

专利申请：发明专利、实用新型专利、商标注册等，适用于公司或个人

教材主编副主编挂名：教师出版教材挂名主编副主编评职称有效，正规出版社，个人出书镀金



本篇文章的主题，特征点定位。以上面的图片为例，假设我们要定位出上面的5个人脸特征点，从精定位的过程如下（下面我先用最简单的理解方式，讲解粗到精的定位过程，先不根据paper的算法进行讲解，等明白了粗定精的定位过程，我们再结合paper的算法进行细讲）：

(1)、level1。首先我们要做的第一步，把人脸图片裁剪出来，而不是直接把上面的一整张图片，扔到CNN中，因为上面图片包含的范围太大了，我们需要使得输入CNN的图片范围越小越好，比如裁剪成下面的图片，然后输入CNN中（我们只需要保证要定位的5个特征点包含在里面就可以了）：

- Given example images $(x_1, y_1), \dots, (x_n, y_n)$ where $y_i = 0, 1$ for negative and positive examples respectively.
- Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0, 1$ respectively, where m and l are the number of negatives and positives respectively.
- For $t = 1, \dots, T$:

1. Normalize the weights,

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$$

so that w_t is a probability distribution.

2. For each feature, j , train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to w_t , $\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$.
3. Choose the classifier, h_t , with the lowest error ϵ_t .
4. Update the weights:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$



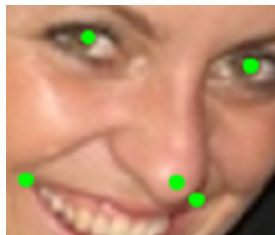


图 2

因此网络的第一层次CNN模型，我们的目标就是要定位出包含这5个特征点的最小包围盒，缩小搜索范围（paper没有这一层，它所采用的算法是直接采用人脸检测器，但是如果采用cnn，先定位出5个特征点的bounding box 精度会比直接采用人脸检测器，定位精度来的高，这个可以从后面face++发表的paper:



《Fine Facial Landmark Localization with Coarse-to-fine Convolutional Network Cascade》中看到，
会在另外一篇博文中讲解，这边只是为了方便理解）

1、CNN模型的输入：原始图片（图1）

2、CNN模型的输出：包含五个特征点的bonding box，预测出bounding box后，把它裁剪出来，得

(2)、level2。接着我们采用CNN，粗定位出这五个特征点，如下图所示（换了一张示例图片，将就一下）：



图 3

上面示例图中，蓝色的点是正确的点；然后红色的点是，我们采用本层次网络CNN模型，预测定位出来的特征点。可以看到红色的点和正确的蓝色点，之间还是有很大的误差，这就是所谓的粗定位，只能大体的搜索到各个点的位置，精度还有待提高。这一层次，又称之为网络特征点的初始定位层，很粗糙的一个定位。然后根据我们cnn的粗定位点，也就是红色的点，作为中心，进行裁剪出一个小的矩形区域，进一步缩小搜索范围：

鼻子

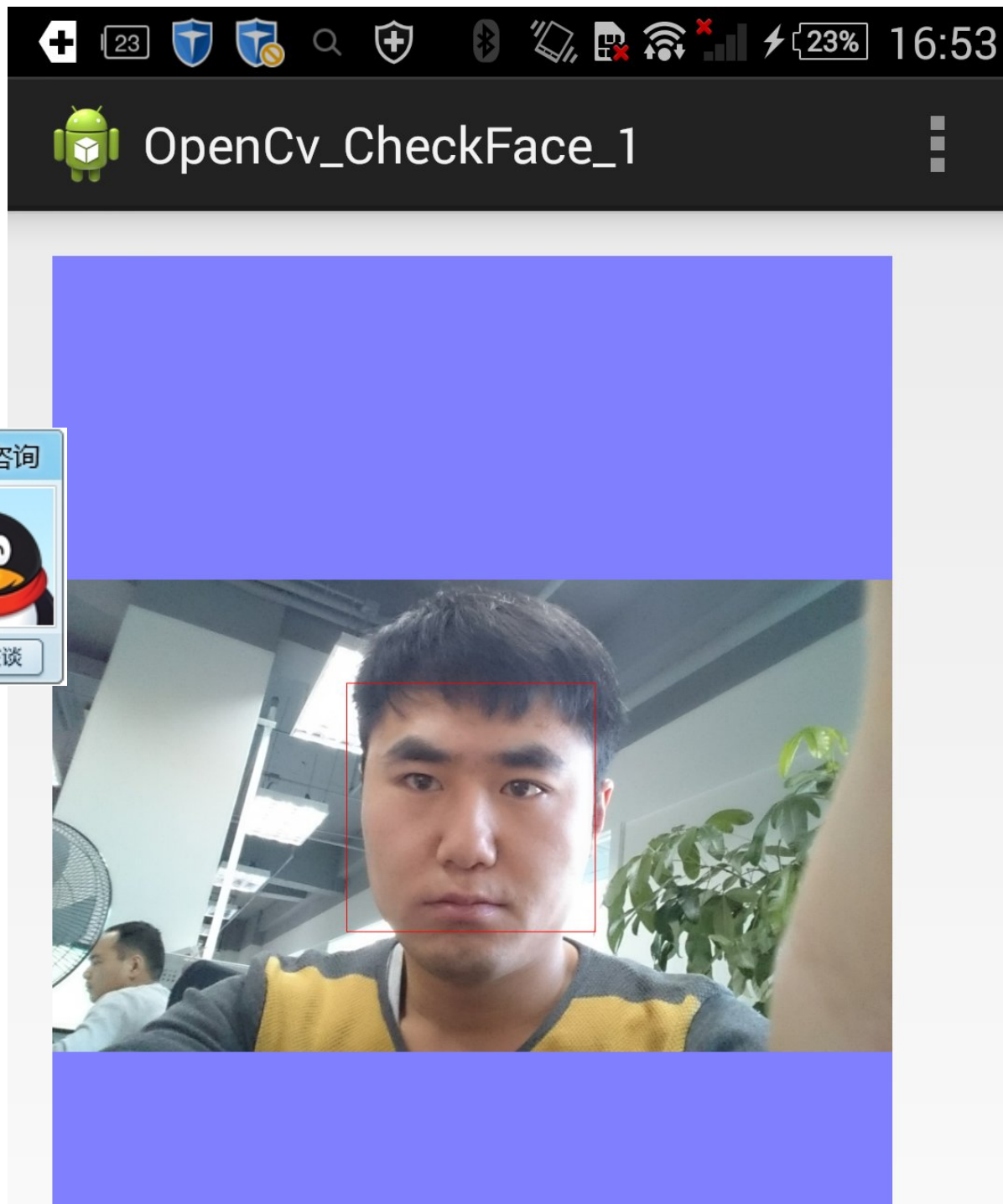
左右嘴角



眼睛

CNN模型输入：包含五个特征点的bounding box图片（也就是图 2）

CNN模型输出：预测出五个特征点的初始位置，得到图3的红色特征点位置，预测出来以后，进行裁剪，把各个特征点的一个小区域范围中的图片裁剪出来，得到图4





level3。这一层由称之为精定位。

因此接着我们就要分别设计5个CNN模型，用于分别输入上面的5个特征点所对应的图片区域了，然后用于分别定位，找到蓝色正确点的坐标。通过图4的裁剪，我们的搜索的范围一下子小了很多，就只有小小的一个范围而已了。这边需要注意，各个部位的CNN模型参数是不共享的，也就是各自独立工作，5个CNN用于分别定位5个点。每个CNN的输出是两个神经元（因为一个CNN，只定位1个特征点，一个特征点，包含了(x, y)两维）。声明：这一层次的网络，文献不仅仅包含了5个CNN，它是用了10个CNN，每个特征点有两个CNN训练预测，然后进行平均，我们这里可以先忽略这一点，影响不大。

本层次CNN模型输入：各个特征点，对应裁剪出来的图片区域，如图4

本层次CNN模型输出：各个特征点的精定位位置。

OK，到了这里，基本讲完了，从粗定位到精定位的思想了，如果看不懂，就得结合文献，反复的读了，因为只有懂了这个思想，才能进行下一步。中场休息，改天再继续写.....

算法总体思路：

1、利用人脸检测器，裁剪出人脸图片，作为CNN模型的输入。

DCNN是一个由多个CNN组层的网络模型，.....

一、level1

在网络的第一层次上：由三个卷积神经网络组成，这三个卷积神经网络分别命名为：F1（网络的输入为一整张人脸图片），EN1（输入图片包含了眼睛和鼻子）、NM1(包含了鼻子和嘴巴区域)。这三个卷积网络区别在于输入的图片区域不同。，如下图所示：



层次网络

构：输入为整个人脸框，输出为我们所要预测的五个特征点。F1的网络结构如下：

一张人脸图片大小为39*39，然后通过卷积神经网络，输出一个10维的特征向量（5个特征点）。

NM1的网络结构和F1相同，不过输入图片的大小不同。

那么三个网络是怎么连接在一起的？我们通过平均的方法，把重复预测的特征点进行位置平均。

网络结构的设计：

Paper分析了三个影响精度的重要因素：

(1)对于leve11来说，我们需要保证网络的深度，主要是考虑到特征提取的难度，因为leve11输入的图片区域比较大.....原理解释请参考paper

(2)

第一个层次的网络用于预测粗定位，接着第二、三个网络用于精定位。

本文来自论文发表网，转载请注明本文网址：

<http://www.lwqklm.com/lunwen/wiki/article-73932-1.html>

↓ 我想下载这篇论文的完整版本

✍ 我想代写一篇原创定制的论文

📄 我想获得渠道将论文发表出去



你好，需要在杂志上 **快速发表** 文章吗？ **全国唯一官方代理商**



信息与电脑



防灾减灾工程学报



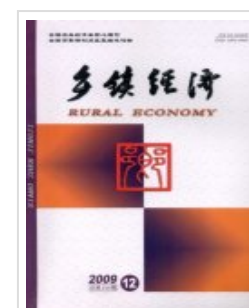
江苏警官学院学报



心理技术与应用



湖南交通科技



乡镇经济

[网站地图](#) | [RSS订阅](#) | [关于我们](#) | [免责声明](#)

Copyright © 论文发表网 www.lwqklm.com All Rights Reserved 鲁ICP备16012746



扫一扫，论文发表更轻松

稍后再说

现在咨询

论文下载

论文发表

论文定制

专利申请

出书挂名