

IBM PowerAI 人工智能 12 小时编程马拉松大赛-我是大侦探

--第十组优秀作品简介(吴沐, 陈子俊)

1 团队背景

首先感谢 IBM 和 CSDN 对本次比赛的大力支持, 比赛的成功与你们密不可分。在本次比赛中我们获得了二等奖(“柯南将”), 优秀的作品不开团队成员的努力, 在这个活动中我们收获颇丰, 接触到了目前的流行的, 前沿的技术, 更重要的是结识了优秀的人才。



吴沐:北京航空航天大学,基础数学博士。在团队中合作开发中,搭建 Cafffe 开发环境。对机器学习理论和算法有很深入的研究。提出了提升向量树多分类算法和基于随机森林的密度估计方法。目前主要从事人脸识别算法的研发,包括人脸检测,人脸标志点检测,人脸跟踪和人脸辨识。



陈子俊：中国科学院文献情报中心，长期保存研发，27岁，计算机科学硕士。在团队中合作开发，Tensorflow 搭建测试。有非常丰富的项目经验，在多个方面都有兴趣，有好奇心。主要从事 javaweb 开发，新框架新技术的探索。目前在在数字资源保存，分布式大数据，非结构化数据挖掘方面的工作。

2 项目详情

2.1 灵感来源

定睛：【基本解释】眼睛盯住一处,形容视线集中。《东周列国志》第一回：“宣王定睛看时，乃上大夫杜伯，下大夫左儒。” 定睛是日常生活中常用到的一个词。我们觉得这个词里面体现了人类视觉认知系统一个基本原理就是“coarse-to-fine”。这次比赛我们采用的MTCNN 算法来检测人脸和辨识人脸遮挡，其级联的思想从某种程度上体现了“coarse-to-fine”这个基本原理。

2.2 研发背景

在 csdn 的官方网站上面看到了这次赛的介绍，瞬间热血沸腾。在准备这次比赛的过程中，我们仔细地分析本次比赛的要点和难点。本次比赛的任务有两个：首先要从图片中检测出人脸的所处的位置，其次判断出这个人的脸部是否有遮挡。所以这次比赛是一个关于物体

检测和物体分类的比赛。比赛的难点有：首先，不同于日常的研发，我们需要在 12 个小时内完成开发环境搭建，数据处理，算法训练和测试结果提交；其次由于待检测人脸面部有严重的遮挡，所以无法直接使用开源的人脸检测器，需要定制人脸检测；最后比赛还要看算法的计算性能，所以网络不能太复杂，否则运行时间太长，计算性能太差。在仔细地分析了前面提到的问题后，我们决定使用级联 MTCNN 作为我们的 region proposal 网络。考虑到计算性能，加上比赛现场发现，待检测人脸其实比较容易检测出来，所以我们只使用了两级 MTCNN 作为我们的 region proposal 网络 然后在连接一个 5 层的简单的 CNN 训练 BBox regression 和 Classification。

2.3 实现思路/算法实现/框架的设计

通过使用人脸识别引擎，对图片进行识别和标记，使用比赛提供的 50% 的标签文件量对当前使用 cnn 网络进行训练，使用训练好的模型对验证图片进行识别，最终将人脸部分和遮挡部分标记出来：

- 1). 使用两个级联 cnn，获得人脸矩形框的 proposal
- 2). 将前一阶段获得的候选区域 crop 成 80×80 的图片，同时在该区域附近随机 crop 10 倍的数据
- 3). 计算 crop 出的数据与 ground truth 的 iou, 将所有 $iou > 0.2$ 的区域作为正样本，这些正样本进一步分为 face(label=1)和 mask(label=2)两类，而 $iou < 0.2$ 的区域作为负样本(label=0)
- 4). 修改一个已有人脸识别器，达到同时分类和 bbox regression
- 5). 将已有的人脸识别器的 confidence 与后来训练出模型的 confidence 的乘积作为最终的 confidence 输出（假设两个识别器是独立训练出来的，那么他们判断一个矩形框是哪

一类的概率，可以认为是两个识别器概率的乘积。)

可改进的空间：由于时间原因，region proposal network 没有重新训练，很可惜。

由于比赛的标签是 bbox 和 label 绑在一起的，所以之前准备的一个分类的 semi-supervised network 没有用上很可惜。

2.4 Caffe, Tensorflow 技术的应用的经验

Caffe 是一个深度学习框架，他是由 Berkeley Vision and Learning Center (BVLC) 和社区的开发者共同完成的，Yangqing Jia 在 UC Berkeley 攻读博士学位期间完成的。

Caffe 具有一些非常鲜明的特色：1. 极具表现力的架构推动了应用的产品研发。模块和优化配置使用配置的方式完成而不是代码 (Hard-Coding)，使用一个简单的标志位就可以在 CPU 运算和 GPU 运算之间自由切换。2. 代码的易扩展性对代码编写起到了推动的作用。在 Caffe 发展的第一个年头就有超过了 1000 个开发者贡献代码，对 Caffe 的发展起到的至关重要的作用。开源世界的发展离不开这些热爱钻研的程序员。3. 高效性是科研和工业领域非常推崇的。Caffe 可以使用一个 Nvidia K40 gpu 在一天之内处理完 60M 的数据。有理由相信 Caffe 在同领域中是最快最便捷的实现。4. 社区化。Caffe 已经在高效科研领域，初创公司和大规模的工程应用中被广泛使用，计算机视觉，语音，多媒体等不同方向都有他的身影。作为一个成熟的深度学习框架，Caffe 的优点不言自明。

Tensorflow 是谷歌行研发的新一代人工智能学习系统，其命名来源于本身的运行原理。Tensorflow 是将复杂的数据结构传输至人工智能神经网络中进行分析 and 处理过程的系统。Tensorflow 是又大名鼎鼎的谷歌研发的新一代人工智能学习系统，是一个适用范围非常广泛的开源软件包。开发者可以使用这套系统建立，训练神经网络，用于检测，揭秘模型，关联，模拟人类进行学习和理解世界。目前它同时被谷歌应用于科研和生产，意在替代原有的闭源

软件如 DistBelief。Tensorflow 最初由 Google brain 团推研发，用于 Google 内部，Tensorflow 可被用于语音识别或图像识别等多项机器深度学习领域，它可在各种设备上运行。作为完全开源的软件，任何人都可以用。Tensorflow 是一个库平台的系统可运行于 linux，windows，mac 系统，安装简单易得。开发语言可以使用 python。根据其硬件支持方面又可以分为 CPU 版本和 GPU 版本，后者可以借助于 Navida 强大的 cuda 核心，充分发挥 gpu 的计算性能，在海量计算中为使用争取效率和时间上的巨大利益。目前 Tensorflow 笔者认为处于成熟上升期，相对于 Cafffe 这个“老兵”Tensorflow 还是有一些路要走的。作为一个年轻，活力十足的框架，它的未来必将是一片宽广。

3 在项目开发中印象最深的事情

1. 比赛当天，我们都准时到达了赛场。赛场坐落于东三环，交通便利，场地环境整洁宽敞，软硬件条件都很给力，着实感受到了主办方对此次比赛的重视和强大的支持，这里要给主办方 IBM 和 CSDN 一个大大的赞！
2. 感谢 IBM 和 CSDN 组织与支持，我们在短时间内组成了一个队伍，比赛的前一天，我们在 QQ 上面经过简单的自我介绍之后，直入主题，研究题目，思考解决方案，这个是理工科工作人员的优点-直率，高效。
3. 比赛当天，在大家还在研究方案的时候，我们已经着手进行一些实践了，有了思路就做，有了想法就尝试，在临近比赛尾声，我们终于递交了一个版本的成果，接着我们就着手思考更好的改进，由于我们的 IOU 方面并不是特别理想。在接下来的短短几分钟内，有做出了第二版和第三版，分别对准确度和效率进行了进一步的提升，越是在紧张的时刻我们越是可以做出更好的产品来。团队成员同心协力，直奔主题，互相谦让，为了同一个目标努力，这是团队取得成果的先决条件和良好结果的有力推手。

4 对项目的一句话口号

追求无止境！

5 团队照片/作品照片



-- 2017.3.21