基于多任务多网络模型的人脸关键点定位

## 一、数据获取：

使用lfw和其他数据集图像和landmarks标签数据，一共约5万张人脸图像

## 二、训练数据集制作（HDF5）：

* **基于级联网络模型原理，主要分三批数据生成：**

1. **网络Level1生成（按照“全脸”，“眼睛和鼻子”以及“鼻子和嘴角”三种不同规格数据）**

提取人脸BBox图像，进行旋转，镜像等操作进行数据增强，归一化到39X39大小，依据landmarks数据，生成hdf5数据（input：预处理后的图像，labels：包含5个关键点坐标：左眼，右眼，鼻子，左嘴角和右嘴角）

（PS：1、5个关键点坐标为相对BBox坐标，而不是整图的绝对坐标；

2、旋转，镜像以及归一化后的图像，需要同时对landmarks数据同步映射更新）

**2、网络Level2生成（按照landmarks标注位置，在BBox的两组不同的sub BBox中，单独为每个关键点生成sub BBox规格数据）**

提取人脸BBox图像，，归一化到15X15大小，依据landmarks数据，生成hdf5数据（input：预处理后的图像，label：包含两组相同关键点的不同坐标：左眼或右眼或鼻子或左嘴角或右嘴角）

（PS：1、关键点坐标为相对sub BBox坐标，既不是整图的绝对坐标也不是BBox坐标；

2、归一化后的图像，需要同时对landmarks数据同步映射更新）

**3、网络Level3生成（同上， 只有sub BBox的尺寸不一致）**

* **编写prototxt模板文件, 为每个level训练数据，批量生成solver.prototxt, train.prototxt和deploy.prototxt文件：**

1. 网络模型数量计算：

Level1数据涉及到“全脸”，“眼睛和鼻子”以及“鼻子和嘴角”， 共需要训练3个CNN回归网络模型；

Level2数据涉及到5个关键点各两组，共需要训练10个CNN回归网络模型；

Level3数据涉及到5个关键点各两组，共需要训练10个CNN回归网络模型。

一共需要训练23个CNN回归网络模型

1. 网络结构：

4\*（convolution layer -> pooling layer -> relu layer）+ 2 \* (fc layer -> relu layer) + EuclideanLoss。

1. 编写python脚本generate\_batch\_prototxts.py生成69个prototxt文件

## 三、训练网络：

* **使用caffe的SGDSolver优化器和multiprocessing模块进行23个网络模型的并行训练：**

PS：

1. 实际训练过程中，有些模型可能不易收敛，需要微调学习率，批次大小等网络参数，进行修正；
2. 每个模型均训练迭代近100000次，网络均已收敛，生成23个不同iter\_100000.caffemodel模型参数。

## 四、测试网络：

* **采用级联网络结构进行测试（按照三个Level网络进行级联测试）：**

**测试效率：**

由于输入图像尺寸(39X39 -> 15X15 -> 15X15)很小，而且CNN回归网络只有6层，网络不深，在GPU测试环境下，可以实现每秒24帧的识别速率，基本可以实时进行人脸关键点的检测。  
**级联网络测试：**

1. 待测试图像利用Level1数据训练网络的前向传播**粗略**生成人脸的5个关键点位置；
2. 利用1中的5关键点位置，截取训练Level2数据时同样大小的尺寸图像进行前向传播生成人脸的5个**精确**关键点位置；
3. 利用2中的5个精确关键点位置，截取训练Level3数据时同样大小的尺寸图像进行前向传播对5个精确关键点位置进行**微调**。

**级联网络测试结构**

