注1：一个文件夹约2万样本，以下代码均以文件夹为单位批量处理。

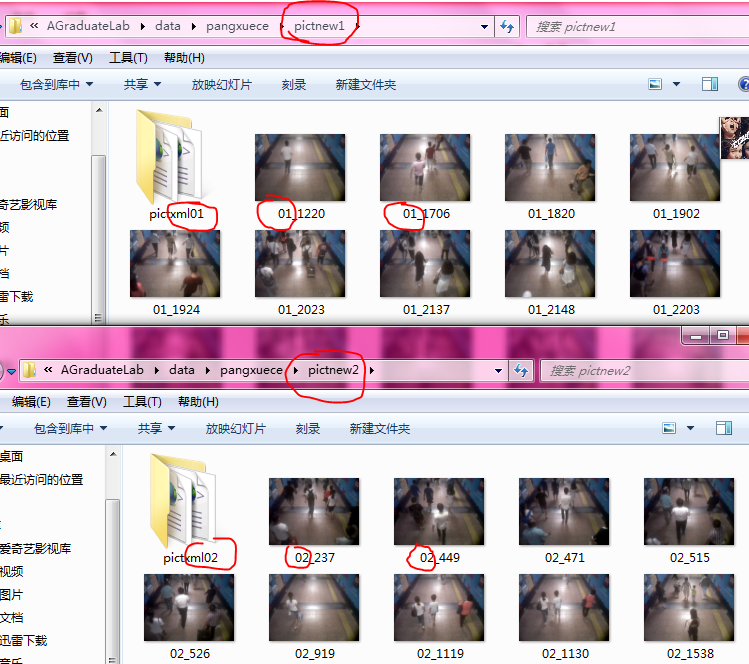
注2：解压田老师实验室给的样本时，如：解压yangben1后，将 其命名为pict1，再解压yangben2…依次进行。………..

（因为他们给的样本只是压缩文件名不同，解压后的文件名都是pict，若不先将解压的文件夹重命名，第二个压缩文件解压时会覆盖第一个已解压的文件。）

注3：我们需要60000个样本，大概需要12000+张图，一个文件夹的2万图片中约能取出2000+可用图片，所以总共需要6个文件夹。如**注2**，已有2个文件夹（可找田凯新同学拷贝，我已初步处理，可直接标标签），仍需处理4个文件夹的图片（解压后按pict3, pict4…命名）。

1. 先筛选2万图片。（以单位文件夹为单位进行批量处理）

每隔9张图片取一张图（如0-10张图，取0和10两张），取完后重命名，将其移到新文件夹，按pictnew3，pictnew4…依次命名。

Eg: pictnew2====>包含pictxml02文件夹（存xml文件，文件名与所标图片一致），02\_XXXX.jpg，trainval\_2.txt**三种**文件

所用代码：rename.py（pict1中只能有图片类的文件，不能有xml和txt）

需修改行：10行，11行按代码注释修改路径；19行，按文件夹修改03，04…等。

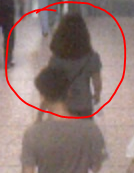
1. 将处理好的pictnew1——pictnew6等文件夹内的图片进行标签。（田凯新同学负责安排）

侧身------------sidehead（人的姿态是侧身）



半身------------halfhead（图片边缘只有拍到一半的人身）

双人----------------- Doublehead（两个头，平行紧挨/前后紧邻等情况）

头肩---------------------backhead（比较正的头肩）



标记工具会自动将xml标签文件名存为和图片名一致。

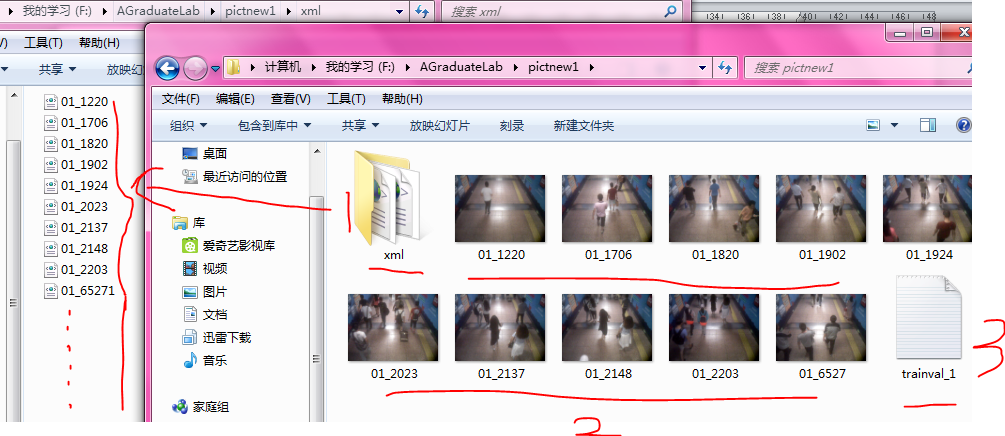
1. 修正标记后的xml文件。（注意查验xml文件夹数量与图片数量一致）

使用代码：changexml\_etree.py

需修改行：11行pictnew1-6的文件夹路径下的xml文件路径

15行的参照11行路径做相应修改，且trainval\_1.txt，对应修改为trainval\_3.txt，trainval\_4.txt……trainval\_6.txt依次类推。

处理好的文件夹应为以下模式：



1. 核验标框（解析xml中标框的坐标信息，对图片打框，并生成打框后的图片副本）。

使用代码：bbox.py

修改代码：16，17，18行路径依照注释修改即可。



其中20行路径为windows自带字体路径，系统版本不同可能需要做出相应修改，如报错，请百度（自己系统版本+SIMLI.TTF字体路径）。

最终将在bbox1文件夹下生成打框图片副本。（字体清晰度与像素有关）



如实在看不清框下文字，可修改字体大小，上例为22（注意修改过大可能超出图片边界）。

