

人脸年龄及性别标注需求文档

标注内容

标注目的	为了训练 Attribute model 中 pose 导致的年龄和性别不准确问题
标注数据集描述	大批量从视频中截取的人脸数据，人脸有各种朝向
待标注目标	1.清理数据 2.标注人脸的年龄和性别
标注的数据规模	500w 张图片
标注工具	需求方提供工具
交付数据格式	仍然保留原来的数据格式，只需要删除不好的图片，并且修改每个小文件夹下的 resutl.txt 文本就好，其他的后处理我自己来做
后续是否还有相同任务	应该不会再有了
重要性	比较重要
紧急程度：	高

标注要求

标注范围

1. 清理过程中需要删除的数据包括：
- 删除所有非真人脸（卡通、3D 模型、雕塑、镜中倒影等）
 - 删除像素太低或者光线问题而不清晰的人脸，即给人第一印象上感觉从这样模糊的照片里面无法分辨出年龄
 - 删除错误截取的图片，比如图片中是人的上半身，而不是脸部占据图片的大部分面积；或者是纯背景纯景色的图片
 - 侧脸程度大于 90° 的图片应该删除
 - 删除在同一张图片中出现两个人脸的图片，在这种情况下，往往是主人脸的后面会出现另外一个人的五官
 - 删除掉脸部被严重遮挡的图片，比如帽子遮住了双眼，这种情况下只通过脸的下半部分较难判断年龄；再比如话筒遮住了嘴巴（遮住下巴的一部分的那种图片可以保留）；在比如用手遮住脸的较大部分，这些都需要删除。戴墨镜的可以留下
 - 如果一个小文件夹内有两个人的图片，删除掉其中一个人（图片数量较少的那个）的图片，即一个小文件夹中只能够有一个图片

2. 清理过程中需要保留的数据包括：

- 所有正常人脸的任何可以看清基本面貌的朝向，包括正向，小于 90°仰面，小于 90°低头，小于 90°侧面。这些各种脸部朝向都需要保留，用来提高脸部处于各种方向时对于年龄和性别判断的准确度

需要特殊处理的情况

- 待标注物被遮挡时，如果被遮挡范围较大，比如帽子遮住了双眼，这种情况下只通过脸的下半部分较难判断年龄；再比如话筒遮住了嘴巴（遮住下巴的一部分的那种图片可以保留）；在比如用手遮住脸的较大部分，这些都需要删除。戴墨镜的可以留下
- 待标注物特别模糊时，像素太低或者光线问题而不清晰的人脸，即给人第一印象上感觉从这样模糊的照片里面无法分辨出年龄

标注精度说明

- 标注精度 很高
- 质量要求 高

标注细节

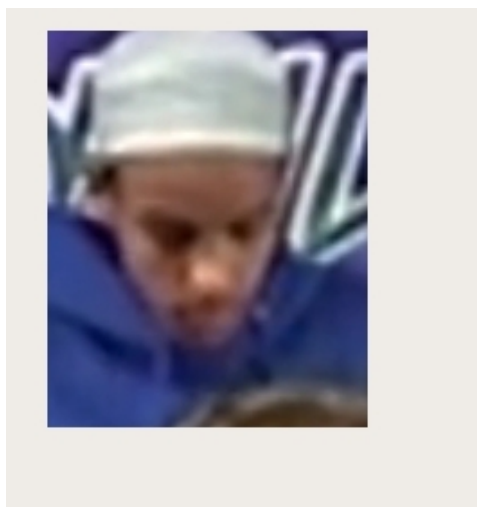
整个标注过程分为两个过程，无标注工具的清理过程，使用标注工具的 label 过程。标注流程如下。

标注流程：

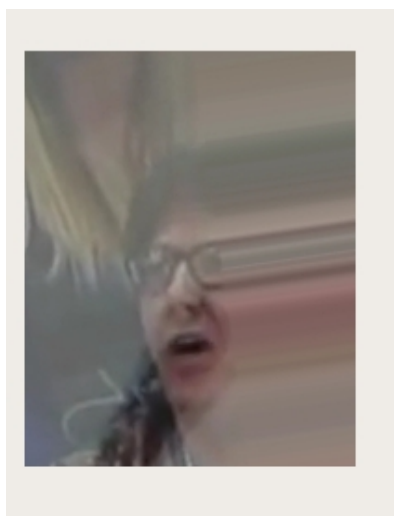
一、文件夹内图片清理：

1.如果整个文件夹内的图片都不清晰，即像素较差或者光线极暗或者极亮，给人第一印象上感觉从这样模糊的照片里面无法分辨出年龄，则删除掉整个文件夹；如果一个文件中只有个别图片有这种情况，则删除掉那张图片

比如下面的情况：

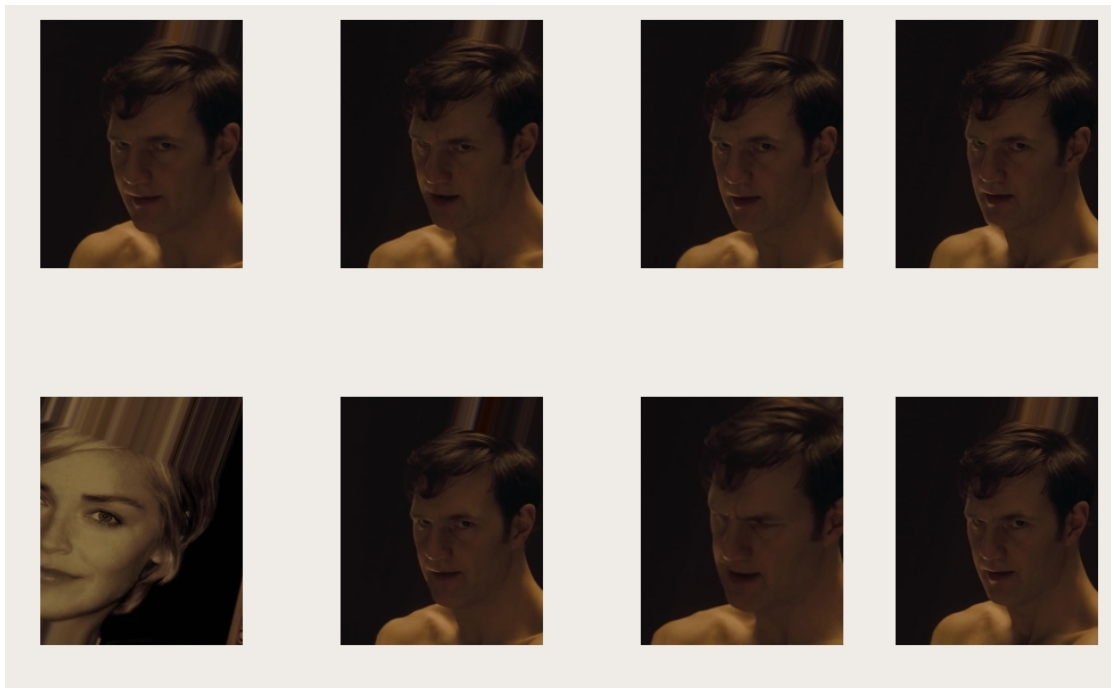


极度模糊，像素极低



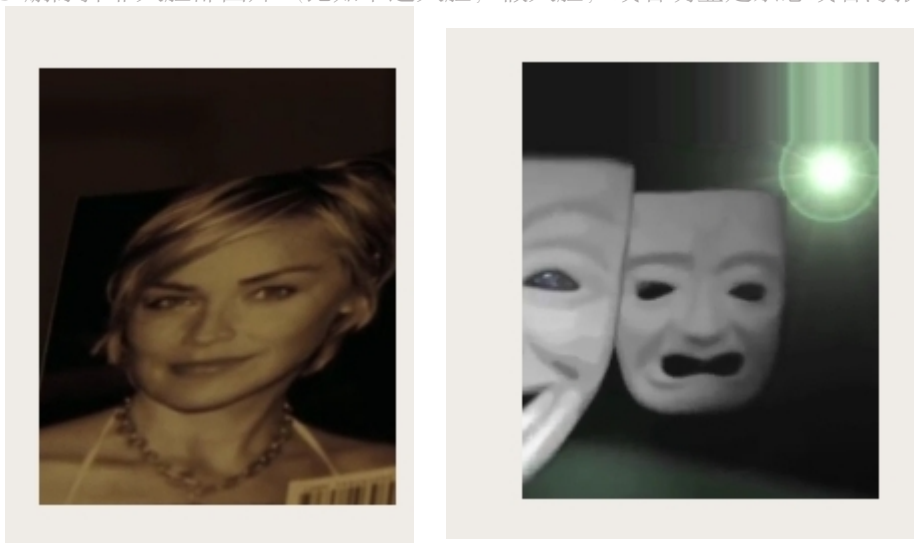
因为帧截取而导致的模糊，同时也是镜中的反光

2.如果一个文件夹内有两个人的图片，删除掉其中一个人（图片数量较少的那个）的图片



这个小文件夹中有两个人脸，应该删除女人脸

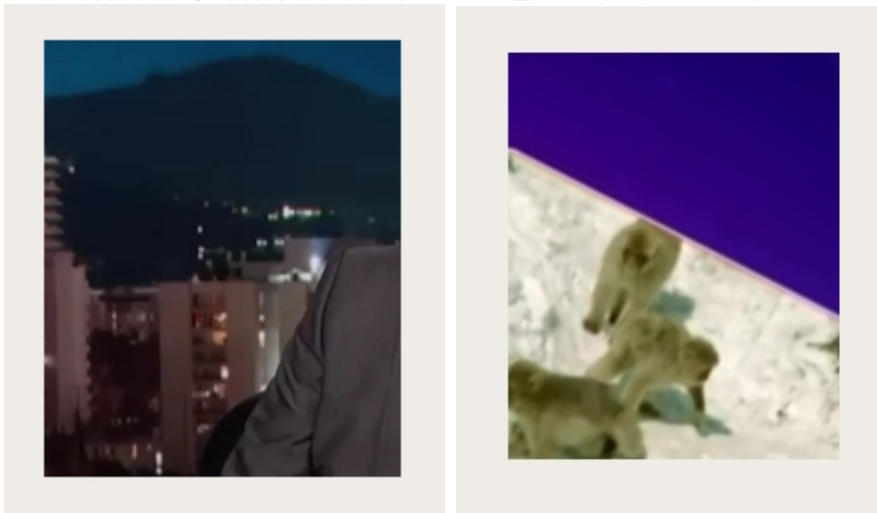
3.删除掉非人脸部图片（比如卡通人脸，假人脸，或者明显是杂志或者海报中的人脸）



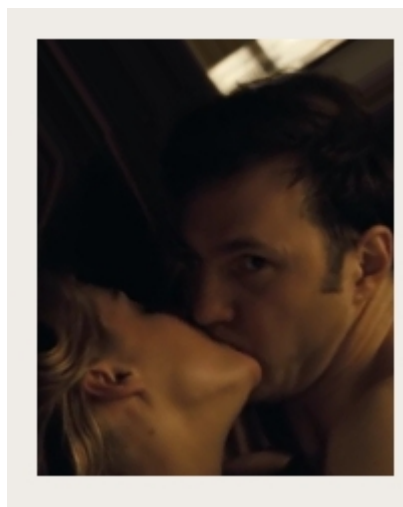
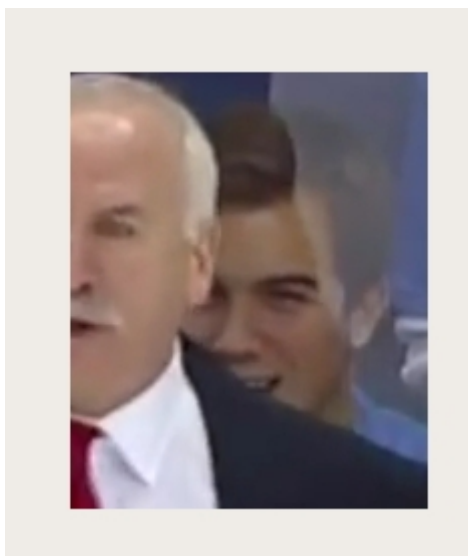
杂志中的人脸，扭曲变形

卡通人脸或者面具

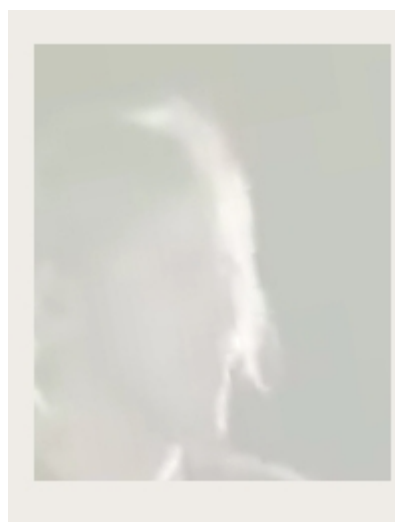
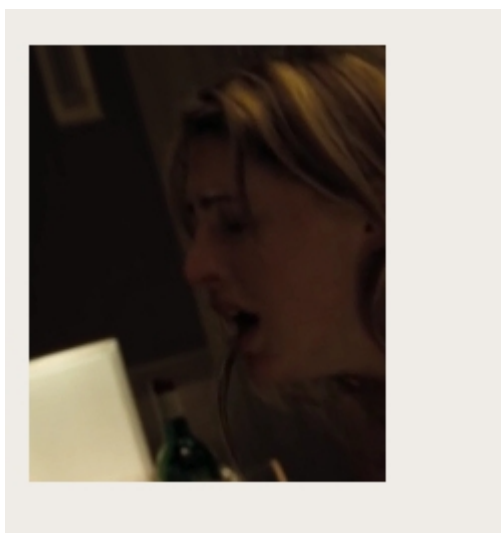
4.删除掉错误截取的图片或者误检的图片，比如图片中是人的上半身，而不是脸部占据图片的大部分面积；或者图片中是景色等。这样的情况如下面的图片中所示



5.删除掉一张图片中出现两个人脸的情况，在这种情况下，往往是主人脸的后面会出现另外一个人的五官



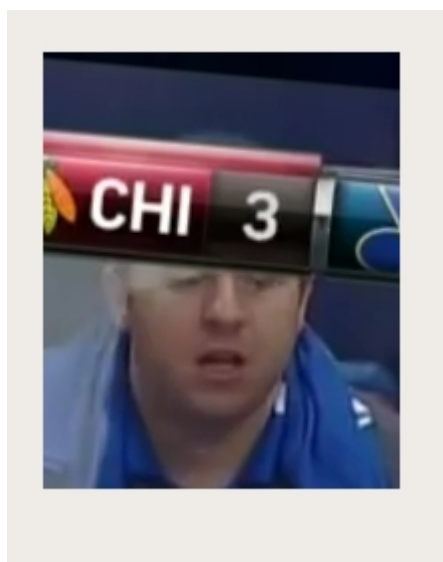
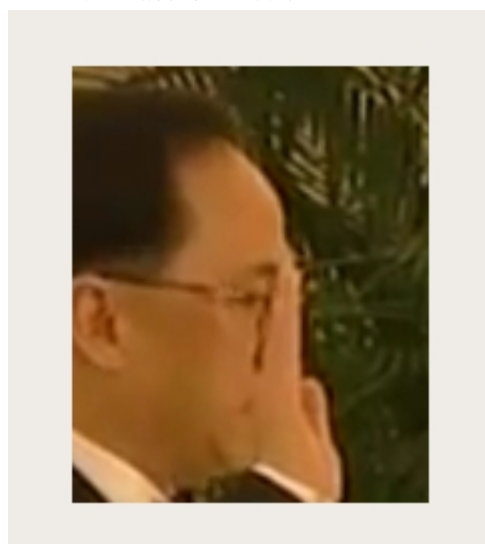
6.删除掉光线极暗/极亮导致让人看不清的图片，

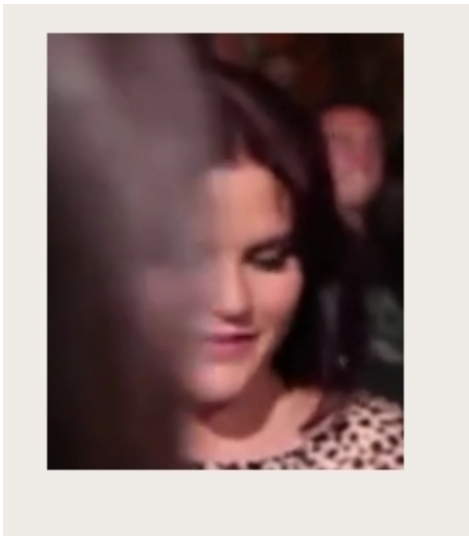


光线极暗

光线极亮

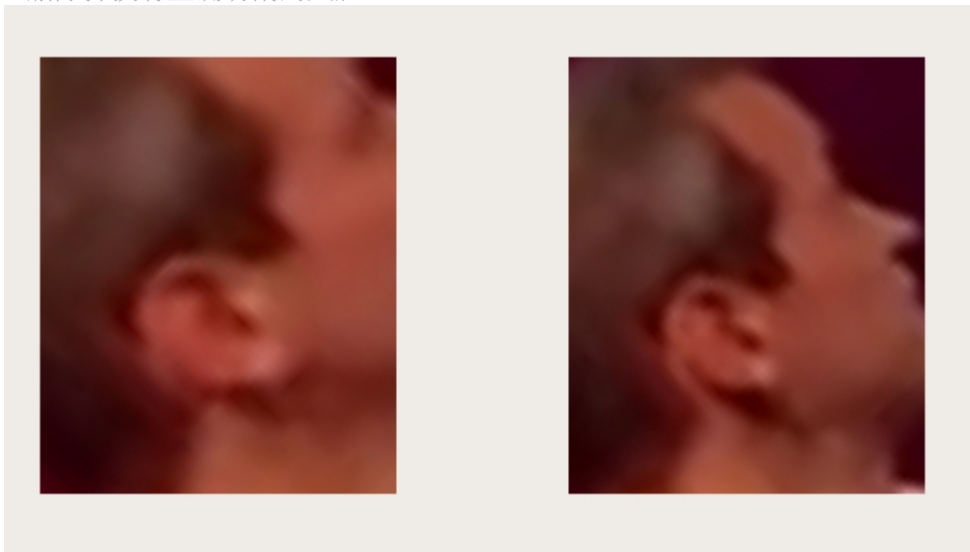
7.删除掉脸部被严重遮挡的图片，比如帽子遮住了双眼，这种情况下只通过脸的下半部分较难判断年龄；再比如话筒遮住了嘴巴（遮住下巴的一部分的那种图片可以保留）；在比如用手遮住脸的较大部分，这些都需要删除。戴墨镜的可以留下
严重遮挡的情况如下所示：





8.文件夹中脸部的所有朝向，比如仰头，低头，侧脸，都应该留下，不应该删除，除非脸部有较大异物遮挡。但是如果图片中只有后脑勺，还是请删掉呀～

9.删除掉没有正确切割的人脸：



10.如果有问题，请即使把相应图片发到小群里面，我们告诉你要不要删除～

文件夹内常用按键：

Enter 进入文件夹

Delete 删除文件/文件夹

Back 返回上一层文件夹

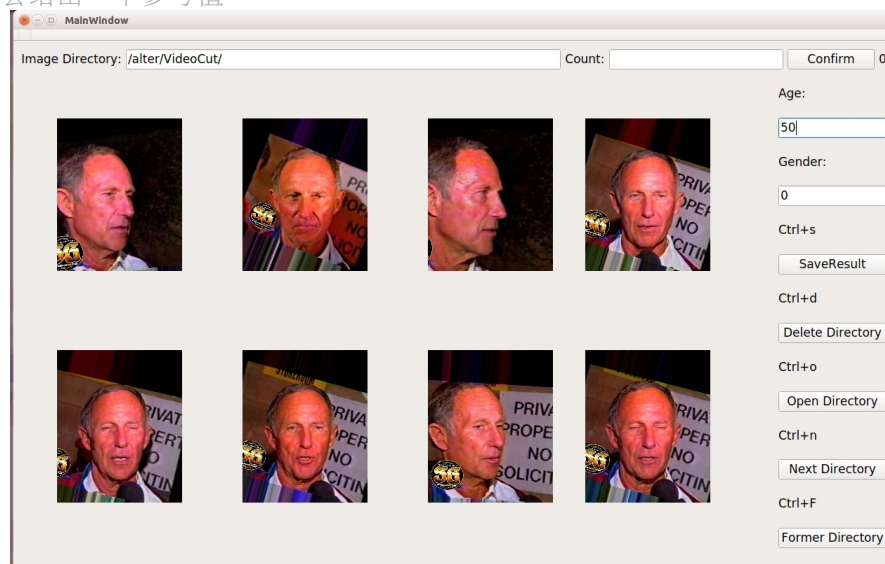
Ctrl+鼠标滚轮向上滚动 文件夹内图片放大

Ctrl+鼠标滚轮向下滚动 文件夹内图片缩小

Ctrl+鼠标框选或点击 选中文件夹内多个图片。

二、使用标注工具进行年龄性别标注

标注要求：根据自己的第一眼判断来更改年龄和性别，在年龄框 **Age** 和性别框 **Gender** 中会给出一个参考值。



性别中：0 表示男性，1 表示女性
性别和年龄都需要调整

标注流程：

>环境设置：

1.将邮件附件中的 Qt 压缩包解压到主目录下，比如/home/sensetime。

2.之后执行下面的命令：

export LD_LIBRARY_PATH=/home/sensetime/Qt/5.7/gcc_64/lib/:\${LD_LIBRARY_PATH}

其中/home/sensetime/应该替换为 Qt 解压的直接路径

3.运行附件中的 AppRun

./AppRun

>开始标注前：

1. 在 Image Directory 输入需要标注的大文件夹的文件地址，比如当前发送过来的数据集大文件是 CutVideo 时，假设放在路径 /alter/ 下，则在文件夹地址中输入/home/sensetime/CutVideo/(最后一个‘/’必须有)

2. 在 Count 中输入上次标注的位置，比如上次标注到第 455 个文件夹，则这次开始标注时输入 456

3.点击 confrim，或者在键盘输入 Shift+Enter 来载入开始标注的文件夹

>开始标注时：

1.光标会停在 Age（年龄）行，修改完年龄后按 Tab 键跳转到 Gender（性别）行修改性别。如果觉得建议年龄和性别可以，也可以不修改。

2.按键盘 Ctrl+n 键或者点击 Next Directory 按键跳转到下一个文件夹，修改后的年龄和性别会自动保存,也可以点击 Save Result 按键或者按键盘 Ctrl+S 来保存结果。

3.如果需要回退到上一行，则可以按键盘 Ctrl+f 键或者点击 Former Directory 按键跳转到上一个文件夹，再按下 ctrl+n 即可跳转到下一个文件夹并且内容自动保存

4.如果在标注过程中遇到清理过程中清理不完善的数据，可以点击 Open Directory 或者按 Ctrl+O 打开当前文件夹进行清理。也可以点击 Delete Directory 或者按 Ctrl+D 直接删除掉整个文件夹。

其他注意事项

文档更新记录

日期	版本号	更改人	更新内容
12/25/2016	V1.0	王孜晗	创建标注文档

标注须知
