

# 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	清理数据
标注的数据规模	500w 张图片
标注工具	需求方不提供工具
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
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

- 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有一个图片
  - 因为训练时使用的是图片中部 128\*128 像素的部分，在这种情况下，如果脸部

特别大以至于有部分器官，比如下巴，被截取到图片外面，是要删除的；如果脸部特别小，截取的是人的上半身，也是要删除的。

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
- 戴墨镜和口罩的人脸，如果对于年龄和性别判断影响不大，也是要保留的。因为这批数据很重要

## 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，需要删除。戴墨镜和帽子的在不影响年龄和性别判断的前提下留下

这方面的标准如下：

- 墨镜情况下不可再遮住口鼻
- 帽子情况下不可遮住口鼻，眼睛在阴影中可以大体分辨或者眼睛不被阴影遮挡
- 正常情况下，与上同理，眼睛遮住了不可遮口鼻，口鼻遮住不可遮住眼睛。遮住一个问题不大。
- 给一个通式的话，就是，眼镜和口鼻(嘴巴鼻子算一项)不可同时遮住。
- 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

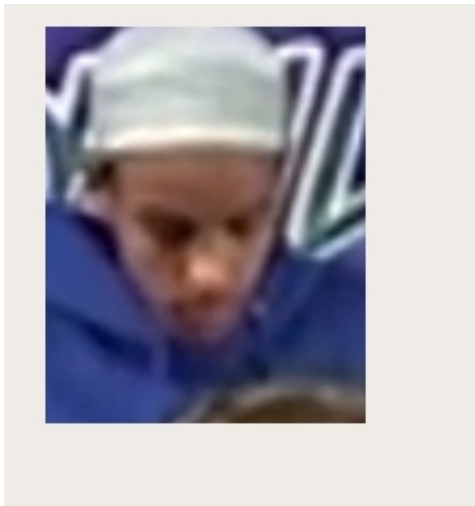
- 标注精度 很高
- 质量要求 高

## 标注细节

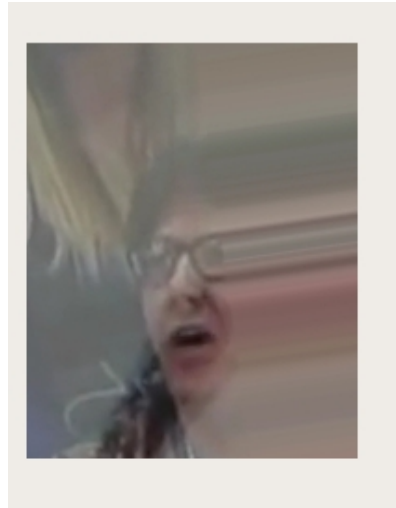
标注流程如下。

标注流程(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：

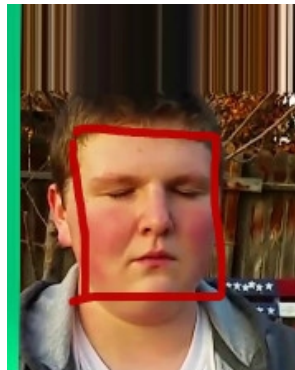
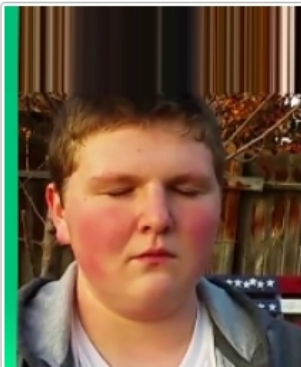


极度模糊，像素极低



因为帧截取而导致的模糊

但是下面的情况可以保留:



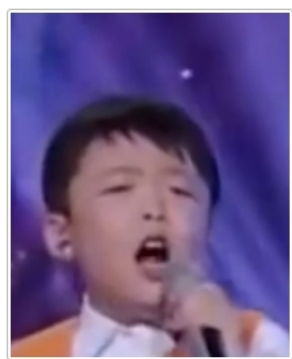
因为我们最后使用的只是红框框出的部分，所以如果头顶下巴、脸颊外沿两侧有变形的话，图片是可以保留的。

轻微模糊的图片可以保留：  
(比如下面两张图)



高清的图片（可以看清皮肤纹路）必须保留，如下图：

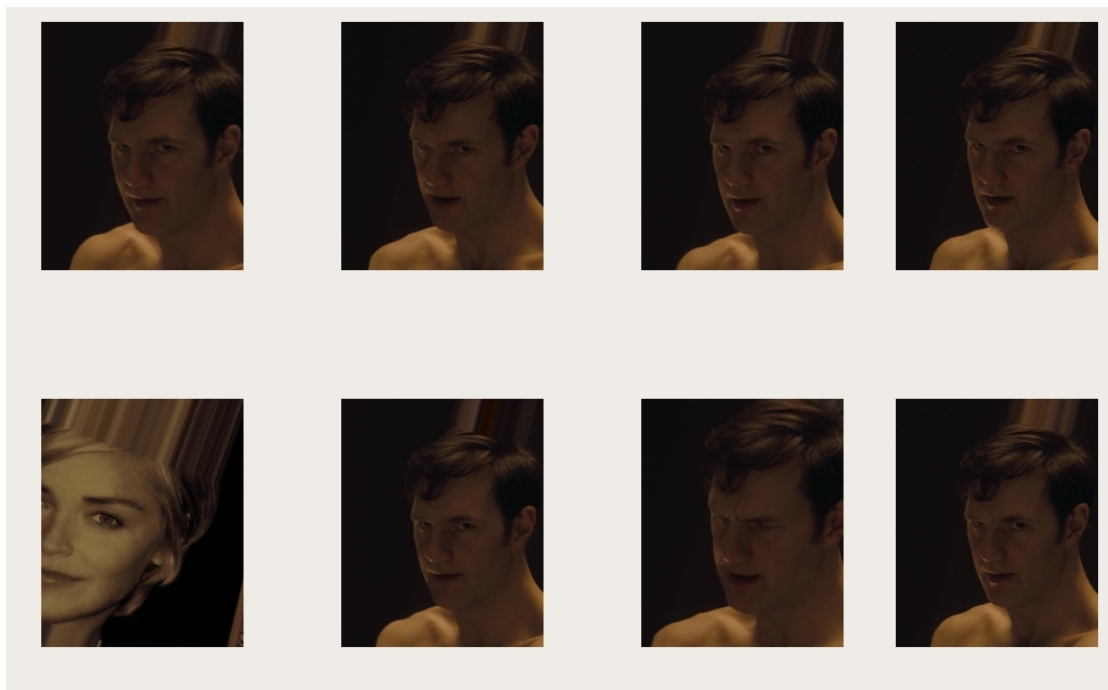




682 inn

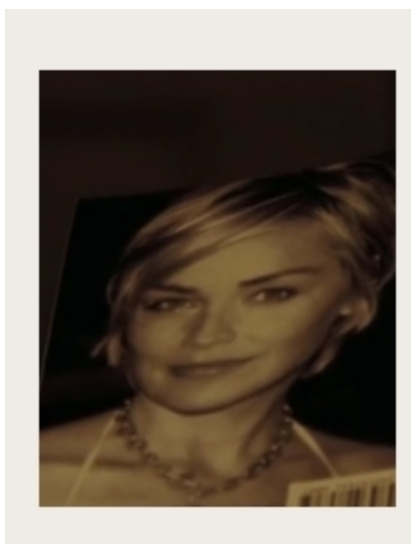
2.如果一个文件夹内有两个人的图片，删除掉其中一个人（图片数量较少的那个）的图片。

这个小文件夹中有两个人脸，应该删除女人脸

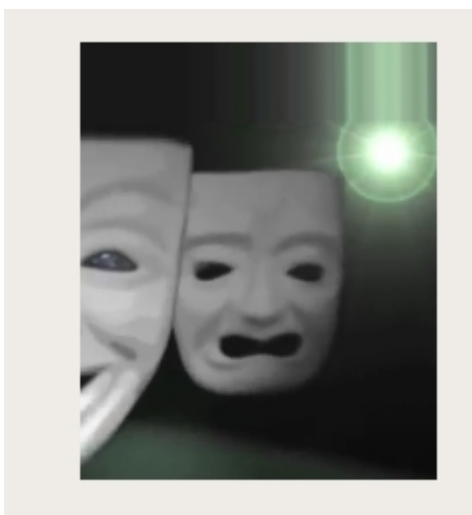


但是应该注意的是，为了追求样本的多样化，如果该文件中有两个人的图片 **A** 和 **B**，其中一个人 **A** 在前面的文件夹中出现较多次，则无论这个人谁的数目多谁的数目少，都请删除样本量已经足够多的那个人 **A** 的脸。

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）



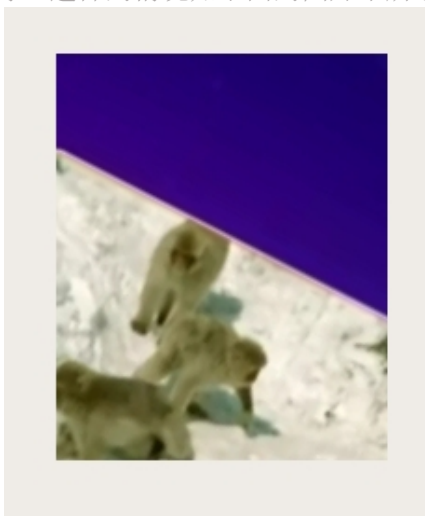
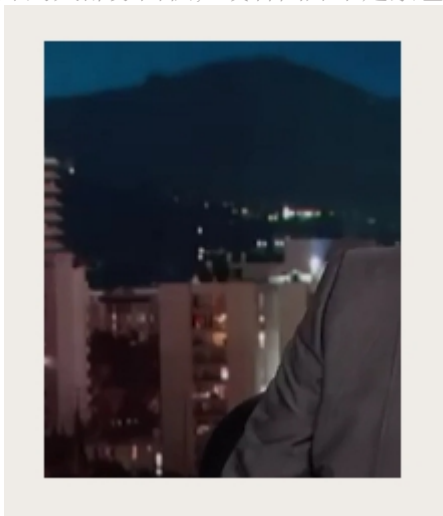
者面具



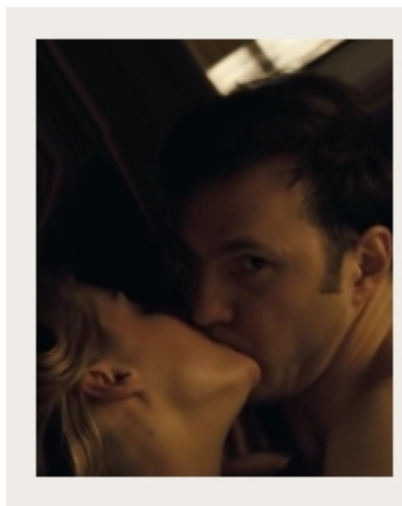
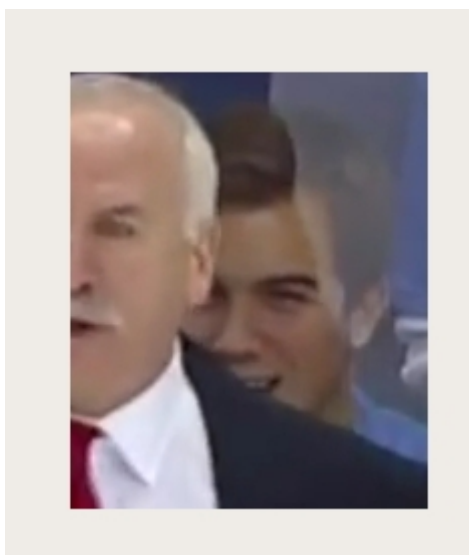
杂志中的人脸，扭曲变形

卡通人脸或

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，两个人脸部占据图片的大小大致向等

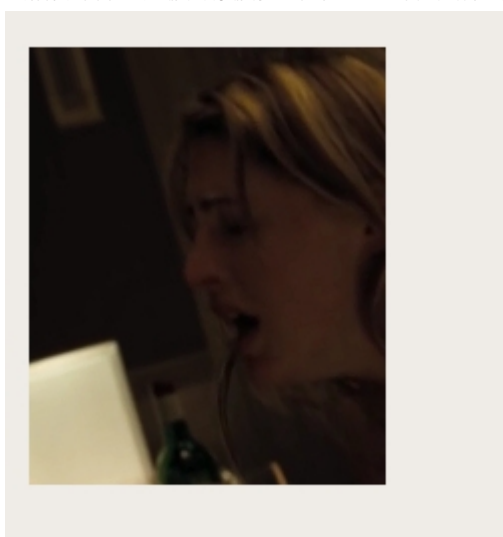


如果图片中，主要角色的人脸在图片正中大概 128\*128 的位置，而且没有被次要角色的脸部器官遮挡，即次要角色只出现在图片的边缘位置的话，则不要删除，可以保留。  
比如下面这样：

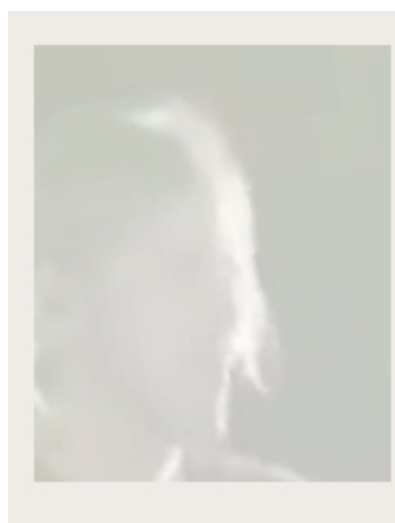


这种是可以保留的

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

但是对于黑人来说，他们本来就很黑，所以看起来像是光线暗一样，在那种情况下，只要

能够分辨清楚这个黑人的年龄和性别，图片就可以保留，比如下图：



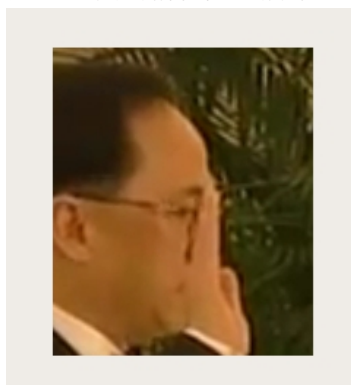
7.删除掉脸部被严重遮挡的图片，但是戴墨镜的在可以判断年龄和性别的前提下可以留下。

- 需要删除。戴墨镜和帽子的在不影响年龄和性别判断的前提下留下

这方面的标准如下：

- 墨镜情况下不可再遮住口鼻
- 帽子情况下不可遮住口鼻，眼睛在阴影中可以大体分辨或者眼睛不被阴影遮挡
- 正常情况下，与上同理，眼睛遮住了不可遮口鼻，口鼻遮住不可遮住眼睛。遮住一个问题不大。
- 给一个通式的话，就是，眼镜和口鼻(嘴巴鼻子算一项)不可同时遮住。

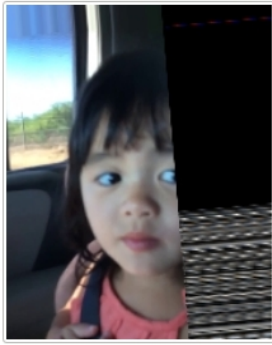
严重遮挡的情况如下所示：



这张图里只露出了一个眼睛，不可以

但下面的图片可以保留，因为在两眼和口鼻中，只遮了一只眼睛的一部分，不影响。





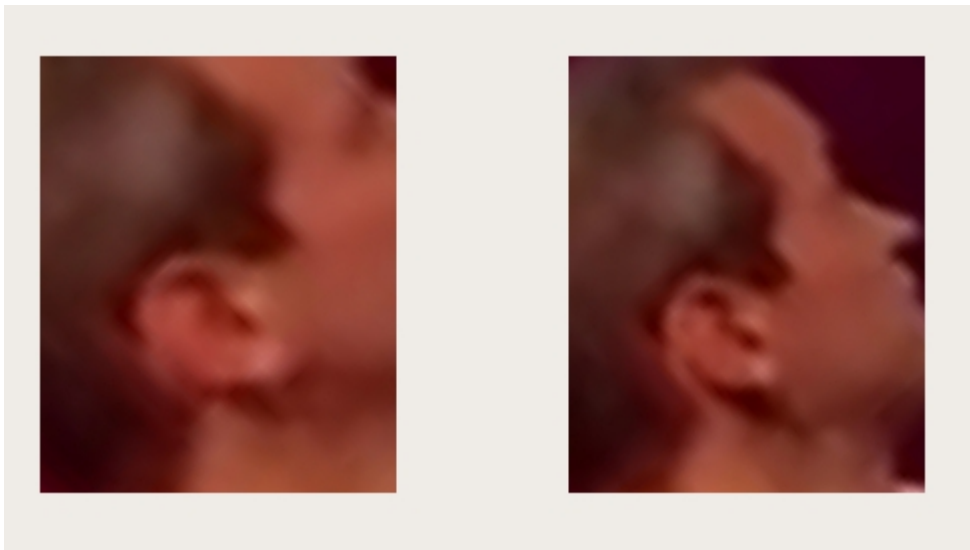
10.jpg

脸部上方有文字或者条幅等，只要不涉及遮住左眼、右眼、口鼻这三者中的两者，就可以保留。

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：

因为训练时使用的是图片中部  $128*128$  像素的部分，在这种情况下，如果脸部特别大以至于有部分器官，比如下巴，被截取到图片外面，是要删除的，如下图：



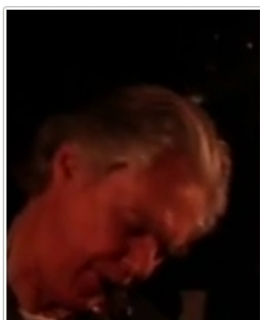
特别小，截取的是人的上半身，也是要删除的。比如下图



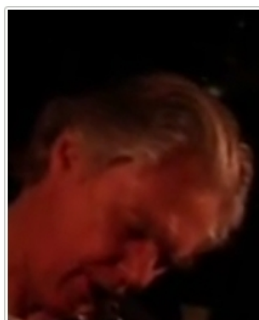
同理，因为是要的正中的方框，歪斜的图片也应该删除：

如果脸部





878.jpg  
14.9 kB  
12月 12 2016



879.jpg  
15.7 kB  
12月 12 2016

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。

## 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

---