

颜值打分 Summary

1. 目标

构造一个颜值分类器，实现对输入一张人脸，给出该人脸的颜值分数（60-99，共40类）。

2. 数据准备

2.1 数据清洗

A	B	C	D	E
url	人工评价颜值	人工识别性别	人工推测年龄（数字）	备注
http://kg.qq.com/node/play?s=305483425_1503407598_462	78 男		30	面部装扮
http://kg.qq.com/node/play?s=119886200_1509067949_621	63 女		18	
http://kg.qq.com/node/play?s=9165381_1512991528_555	78 女		30	
http://kg.qq.com/node/play?s=305483425_1503409796_36	78 男		30	面部装扮
http://kg.qq.com/node/play?s=92132890_1514030481_900	84 女		25	
http://kg.qq.com/node/play?s=464319990_1512472221_807	75 女		12	儿童
http://kg.qq.com/node/play?s=190224242_1508840467_771	88 女		25	
http://kg.qq.com/node/play?s=117919764_1509330706_235	88 女		20	多人出镜
http://kg.qq.com/node/play?s=39881281_1504956989_992	65 女		45	
http://kg.qq.com/node/play?s=13337812_1506602746_733	78 女		18	
http://kg.qq.com/node/play?s=4890026_1512878619_274	80 女		23	人脸不全
http://kg.qq.com/node/play?s=497200813_1513865447_382	90 女		25	
http://kg.qq.com/node/play?s=13337812_1507103008_801	86 女		20	
http://kg.qq.com/node/play?s=3023976_1511528567_346	85 男		20	多人出镜、面部尺寸小
http://kg.qq.com/node/play?s=41699139_1506992770_274	-	-	-	已删除
http://kg.qq.com/node/play?s=133505494_1503373972_545	78 女		20	
http://kg.qq.com/node/play?s=90172436_1510455688_709	70 男		18	
http://kg.qq.com/node/play?s=391825896_1503914358_671	78 女		20	疑似非本人作品
http://kg.qq.com/node/play?s=37522577_1504782259_853	86 女		23	
http://kg.qq.com/node/play?s=183340674_1503313405_328	68 女		10	儿童
http://kg.qq.com/node/play?s=173751594_1505495217_134	87 女		23	
http://kg.qq.com/node/play?s=88066087_1502877485_618	90 女		23	
http://kg.qq.com/node/play?s=18297230_1504087500_964	87 男		23	吸烟
http://kg.qq.com/node/play?s=177521112_1504014156_553	-	-	-	未露脸
http://kg.qq.com/node/play?s=21293061_1502968388_329	-	-	-	已删除
http://kg.qq.com/node/play?s=41800374_1504181266_876	85 女		26	人脸不全
http://kg.qq.com/node/play?s=40411397_1508392362_328	87 女		24	
http://kg.qq.com/node/play?s=95624283_1513665070_238	75 女		20	
http://kg.qq.com/node/play?s=4960571_1509449761_856	78 女		18	

原始数据是全民K歌提供的每个视频url以及对应人工标注的groundTruth，其中有较多的噪声数据，包括人脸不全，未露脸，多人出镜等，我们先把这些情况的数据都予以剔除。

2.2 数据爬取

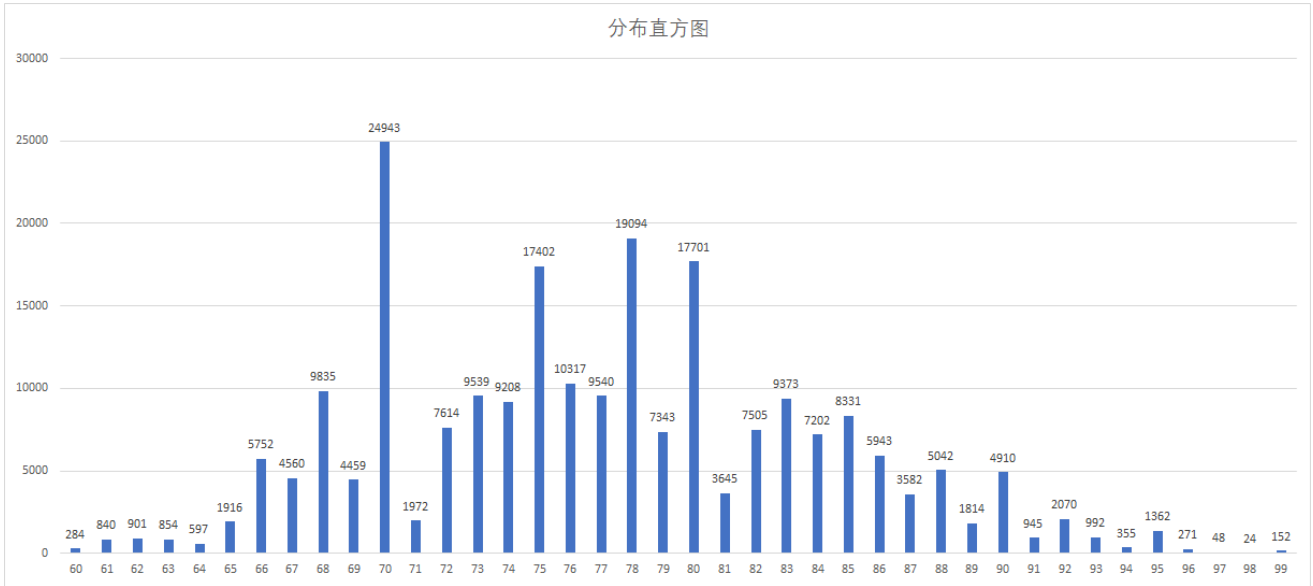
通过构造爬虫根据url爬取下载每个视频，共计25879个。

2.3 转化为图像

将下载后的视频每隔50帧取一张图像，每个视频只截取10张用作训练数据的准备。

2.4 统计数据分布

统计上一步生成的图像中，每个颜值分数（60-99）的对应的图像总数分布情况，画出直方图如下。



2.5 均衡数据

从上一步的直方图可以看出，数据的分布极为不均衡，大多数数据分布在颜值为70-90分之间，而两侧的数据就偏少，而且数量差距也较大，**70分的图像最多，有24943张，而98分的图像只有24张。**

为了降低数据的分布不均衡给训练带来的负面影响，我们做了如下的数据均衡，

- 先把所有图像按照5:1的比例划分为Train和Val集，避免重合。
- 对于每一个颜值分数，如果Train数据集中该分数图像**大于1000张**，则**只随机挑选出1000张**用作训练。
- 对于每一个颜值分数，如果Train数据集中该分数图像**小于1000张**，则**通过copy&paste凑齐1000张**图像用作训练。
- 对于每一个颜值分数，如果Val数据集中该分数图像**大于200张**，则**只随机挑选出200张**用作测试。
- 对于每一个颜值分数，如果Val数据集中该分数图像**小于200张**，则**通过copy&paste凑齐200张**图像用作测试。
- 最后保证用以训练的每个颜值分数的数据**各有1000张**，用以测试的每个颜值分数的数据**各有200张**。

2.6 生成LMDB

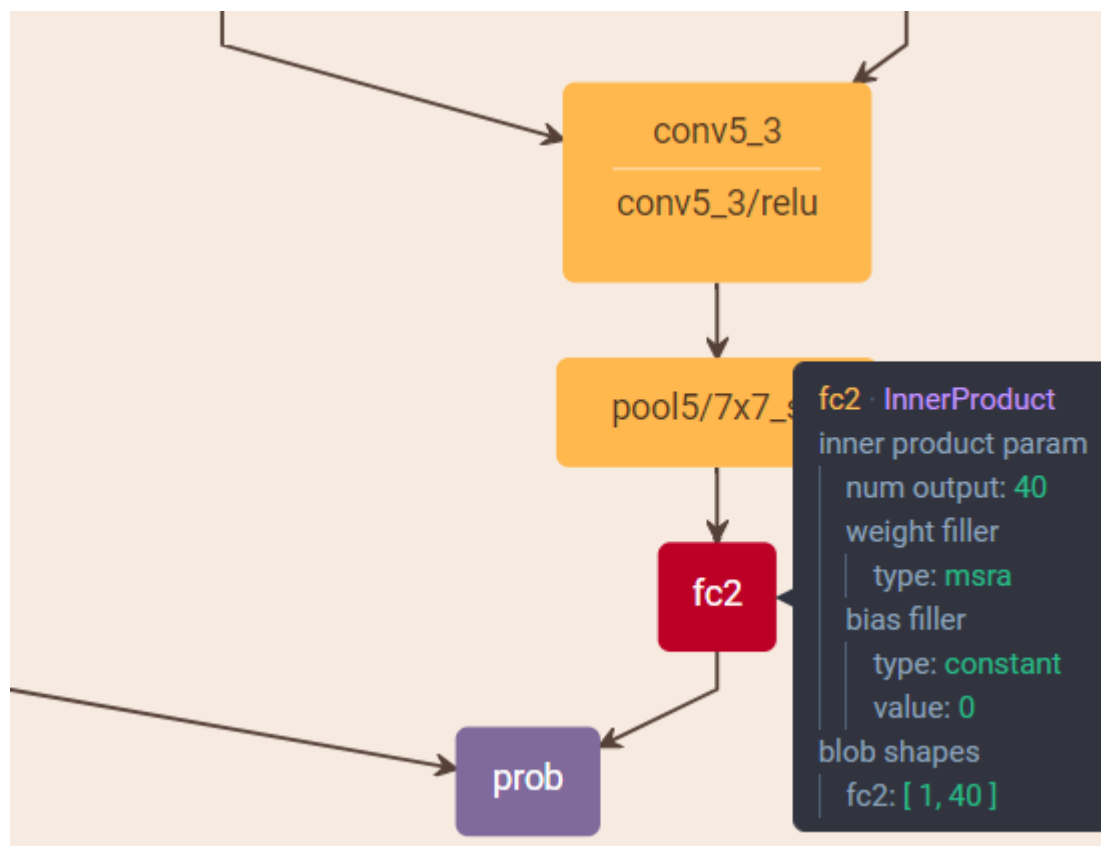
根据上一步的train.txt和val.txt生成lmdb用以在Caffe中读取数据。

2.7 计算Mean

根据上一步生成的train_lmdb计算训练数据的均值，在读入图像后都减去该均值之后再送入网络进行分类。

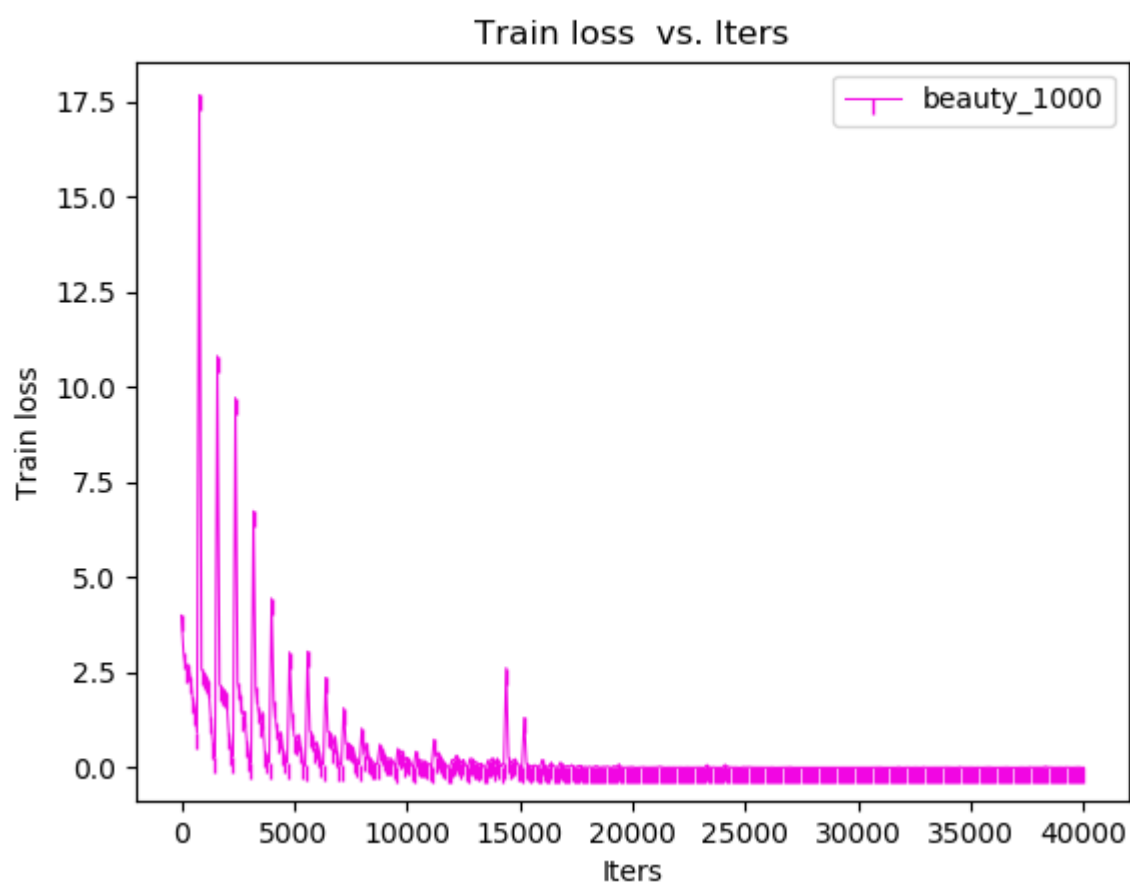
3. 网络结构

由于是一个40类的分类任务，这次选用网络结构为ResNet-50，实现过程使用Caffe的深度学习框架，使用舒汉之前在中国人脸训练得到的model进行Finetune。

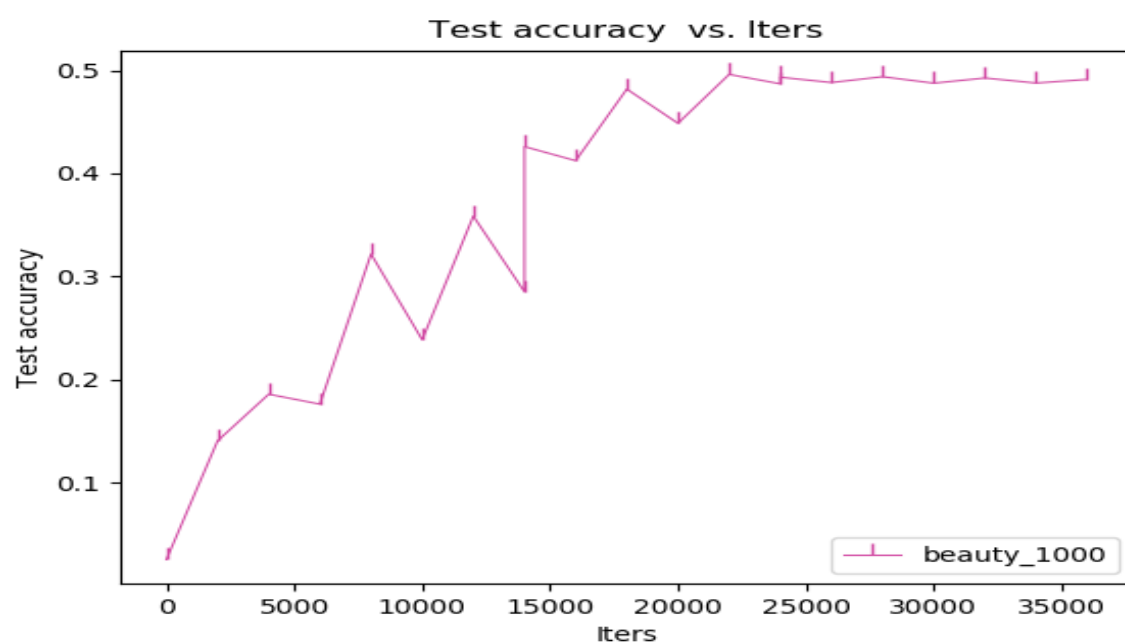


4. 训练结果

4.1 TrainLoss-Iters图



4.2 TestAccuracy-Iters图

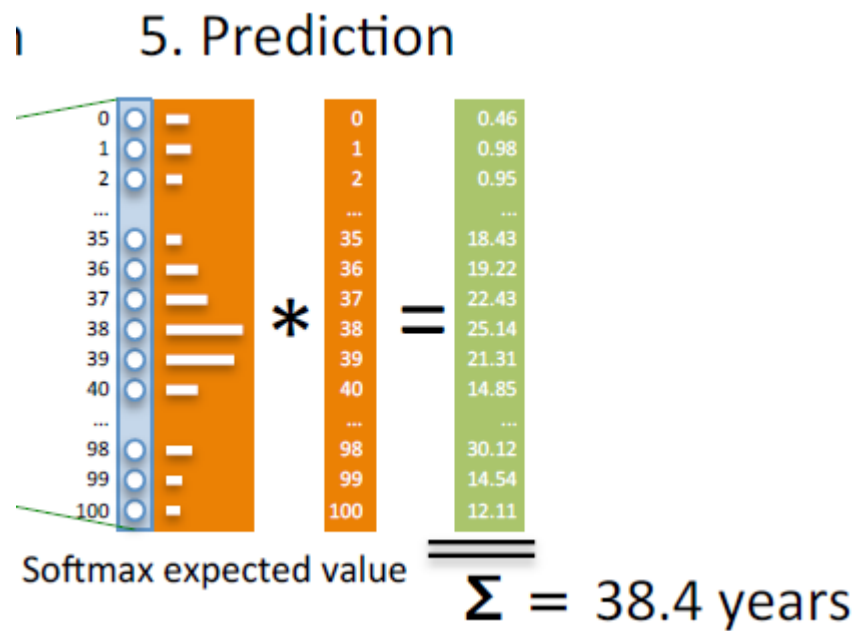


4.3 训练与测试时间

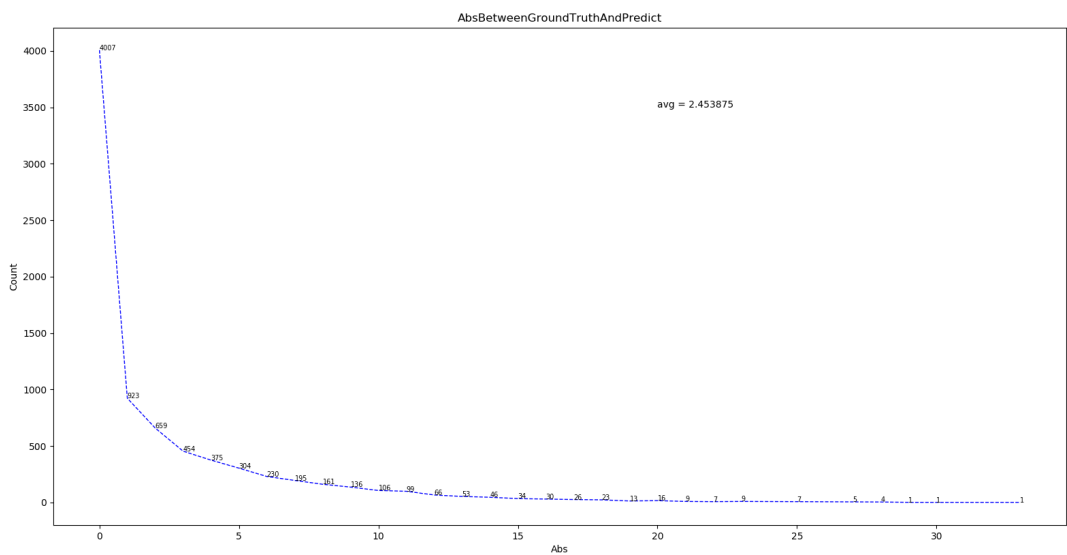
阶段	数据量/张	时间
训练	40000	14.6 hour
测试	8000	5 min

5. 测试结果

这里我们根据颜值问题的特殊性：**颜值分数相近的人脸主观看来差别不会很大**（比如88分和89分的颜值打分其实看起来较为接近），所以我们借鉴一篇年龄预测的论文（[Link](#)），将测试网络结果的输出做了一些修改，原来是挑选出预测概率最大的作为predict分数，修改为把**预测概率与颜值分数（60-99）**做一个**加权平均**，把计算的结果作为predict分数，加权平均的示意图如下。



根据上述的加权平均得分，我们对Val集的所有图片(40*200=8000张)进行预测，再将预测拿到的结果与它的 groundTruth 计算MAE（mean absolute error）值，结果表明**超过一半的人脸颜值预测值与其groundTruth相等**，**84%的误差在5分以内**，总体MAE=2.453875。



6. 结果分析可视化

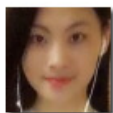
6.1 Good Case

6.1.1 相差0分的Samples



gt:80_

pre:80.0031572106_
19975664_151074...



gt:80_

pre:80.0198932943_
85245745_149036...



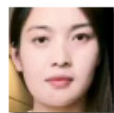
gt:80_

pre:80.4077844341_
85245745_150478...



gt:81_

pre:80.9999992079_
45854781_150564...



gt:81_

pre:81.0000060423_
8757616_1481440...



gt:83_

pre:83.2966555432_
231435449_14880...



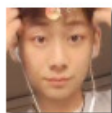
gt:83_

pre:83.3368760357_
19949930_146335...



gt:84_

pre:83.8164445431_
35714516_151384...



gt:84_

pre:83.9990341461_
23225693_146788...



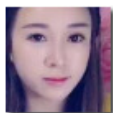
gt:85_

pre:84.9966855214_
231901188_15027...



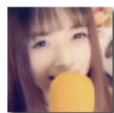
gt:87_

pre:87.000092659_
463466287_15123...



gt:87_

pre:87.0064502562_
216436207_15092...



gt:87_

pre:87.2290840662_
3621179_1493043...



gt:88_

pre:87.9285128484_
24593265_143090...



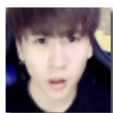
gt:88_

pre:87.9975538217_
127837326_15111...



gt:89_

pre:88.9999998801_
5320291_1467638...



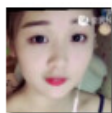
gt:90_

pre:89.9896368264_
105160120_14957...



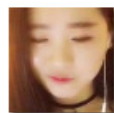
gt:90_

pre:90.0723456048_
1640608_1440339...



gt:90_

pre:90.1687328959_
73461102_150458...



gt:92_

pre:91.5489737684_
29432347_148777...

6.1.2 相差1分的Samples



gt:75_
pre:74.4293354279_
313094923_14812...



gt:75_
pre:76.2138833202_
122638273_14983...



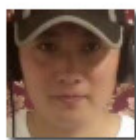
gt:76_
pre:74.9888833875_
193582322_15154...



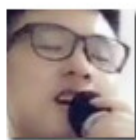
gt:76_
pre:77.0719935543_
32301659_149670...



gt:78_
pre:78.9222642466_
23692438_
1514286707_290_



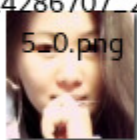
gt:78_
pre:79.0329772383_
182332662_15009...



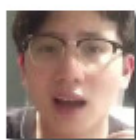
gt:79_
pre:78.2620929017_
190123017_15069...



gt:80_
pre:78.964433658_
100830366_15025...



gt:81_
pre:80.3312199614_
6823373_1491717...



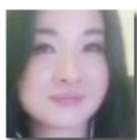
gt:81_
pre:82.0125539174_
129283064_15087...



gt:82_
pre:80.9999771702_
11126248_151272...



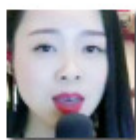
gt:82_
pre:81.4033444658_
7467603_1462022...



gt:84_
pre:84.7535722979_
57602782_150350...



gt:84_
pre:84.9018347658_
7003718_1496281...



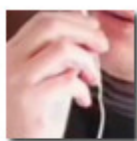
gt:84_
pre:85.494779313_
85125103_151452...



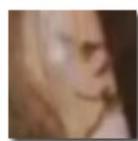
gt:85_
pre:83.9222110304_
93208830_149801...

6.2 Bad Case

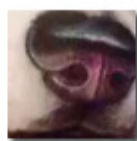
6.2.1 非人脸或者人脸不全



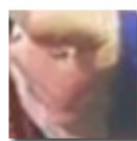
gt:63_
pre:92.537938184_
359538686_
1514343086_2598_
8_2.png



gt:69_
pre:87.8263764953_
11532513_
1513208202_4235_
8_2.png



gt:91_
pre:68.1724407199_
21935182_
1473602877_529_
6_0.png



gt:95_
pre:62.1056755682_
258224448_
1503283565_618_
9_1.png

6.2.2 GroundTruth错误



gt:63_
pre:85.5994177359_
175893759_
1513440686_1966_
8_0.png



gt:67_
pre:89.7030772377_
293531831_
1505721637_2832_
3_0.png



gt:70_
pre:90.6328943385_
312464051_
1510407408_218_
3_0.png

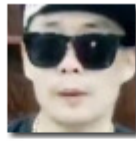


gt:88_
pre:65.3986611538_
10645215_
1508754343_1519_
10_0.png

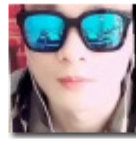
6.2.3 墨镜



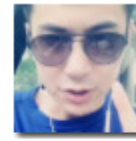
gt:68_
pre:88.4438104935_
142191975_
1492423783_4920_
3_1.png



gt:70_
pre:94.6957499535_
127455645_
1514119163_4195_
8_1.png



gt:85_
pre:62.3080577985_
58368183_
1509262952_2529_
10_0.png



gt:95_
pre:65.8092487778_
357794045_
1497433594_1917_
1_0.png

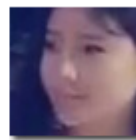
6.2.4 侧脸



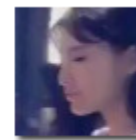
gt:60_
pre:86.9512199788_
232614049_
1493176716_2294_
1_1.png



gt:95_
pre:73.1478342992_
258224448_
1503283565_618_
6_0.png



gt:96_
pre:68.3378394412_
258224448_
1493042414_713_
8_0.png

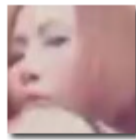


gt:96_
pre:70.9334399742_
258224448_
1493042414_713_
9_0.png

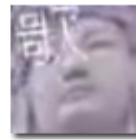
6.2.5 模糊



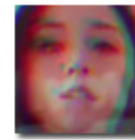
gt:61_
pre:82.6331224177_
418255460_
1504527882_35_9_
0.png



gt:61_
pre:83.0827494388_
23388778_
1509934037_2007_
8_0.png



gt:66_
pre:88.6584669524_
44745992_
1477045692_870_
5_0.png



gt:93_
pre:66.4937243527_
57767244_
1512886609_597_
6_0.png