

低分辨率人脸识别综述（1）：数据集

临菲信息技术港

人脸识别可以自动识别或验证一个人的身份，如今已经在智能手机上广泛应用。但人脸识别仍然存在挑战，包括在非受控环境下的低分辨率图像的识别和图像伪影下的识别。

低分辨率人脸识别（Low Resolution Face Recognition）是以低分辨率图像进行人脸人别，其像素 32×32 或以下（有人也称为极低分辨率）。低分辨率人脸图像一般来自非受控环境，除了低分辨率外，人脸识别技术还面临光照、比例、视点、部分遮挡和图像质量变化的挑战。

非受控场景下的人脸识别，有时也称为“野外人脸识别”（face recognition in the wild），近年来已经取得了显著的进展，其流行的基准数据库有：LFW、YTF 和 CelebA 等。如今，在这些数据集中，即使在实时限制条件下（每秒 30 帧以上），也有可能获得 96% 以上的识别准确率。

监控场景下的低分辨率人脸识别当前还是一个比较小的研究领域，其数据集有限，主要有 SCface、Point-and-Shoot、IJB-S、UCCSface、QMUL Survface 和 QMUL Tinyface，它们是无约束极低分辨率人脸识别的可用基准数据集。

本文首先介绍 3 种流行的非受控场景人脸识别数据集(LFW、YTF 和 CelebA)，然后介绍上述 6 种低分辨率人脸数据集。

LFW

LFW（Labeled Faces in the Wild）（野外标记人脸）是由美国马萨诸塞州立大学阿默斯特分校计算机视觉实验室整理完成的数据集[1,2]。

LFW 旨在研究无约束人脸识别问题。该数据集包含从网络收集的超过 13,000 张人脸图像。每张人脸都标有照片中人物的名字。照片中的 1680 人在数据集中有两张或更多张不同的照片。

现在有四组不同的 LFW 图像集，包括原始图像(LFW)和三种不同类型的“对齐”图像("aligned" images)。对齐的图像包括“漏斗图像（funneled images）”（ICCV 2007）、使用未公开的对齐方法的 LFW-a 和“深漏斗（deep funneled images）”图像（NIPS 2012）。其中，LFW-a 和深度漏斗图像对大多数人脸验证算法产生了优于原始图像和漏斗图像（ICCV 2007）的优越结果。

LFW 是世界上最具有影响力的人脸数据集之一。LFW 团队从 2007 年以来对 LFW 数据集和基准的生成及维护，有力地推动了在不受控环境中进行人脸识别的研究进展。到 2019 年被引用了 3700 多次，并被 Google 和 Facebook 等老牌公司用来测试他们的人脸识别准确率。

在首尔举行的 2019 年 IEEE 计算机视觉国际会议 (ICCV) 大会上, LFW 团队获得了马克·埃弗林厄姆奖 (Mark Everingham Award)。该奖的设立是为了纪念计算机视觉领域专家 Mark Everingham 和鼓励在计算机视觉领域做出突出贡献的研究人员或团队, 每年由 IEEE 计算机学会模式分析和机器智能技术委员会颁发。



Labeled Faces in the Wild



LFW 团队获得 2019 Everingham 奖

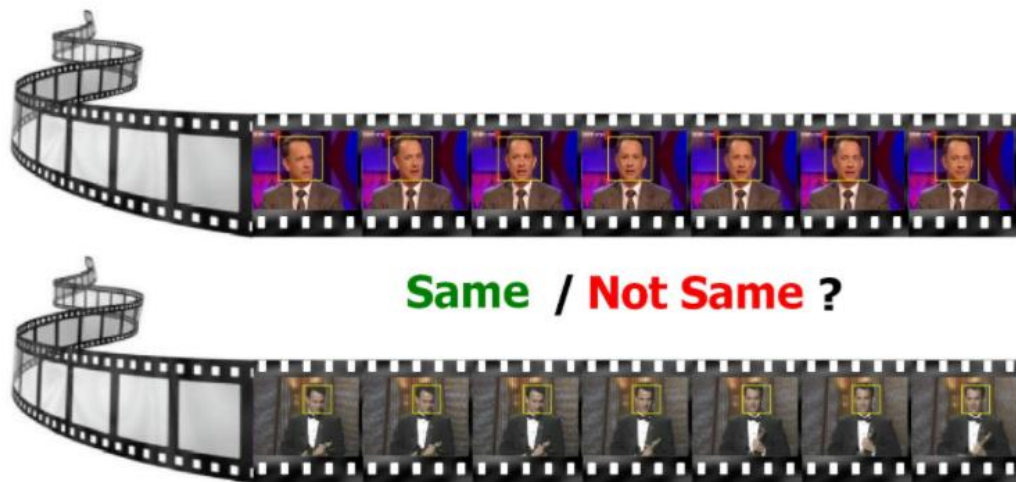
YTF

YTF (YouTube Faces DB) 数据集是从 YouTube 下载视频组成的人脸视频数据集, 由以色列特拉维夫大学于 2011 年发布[3,4]。

YTF 是一个用于研究视频中无约束人脸识别问题的人脸视频数据集。该数据集包含 1595 个不同人的 3425 个视频, 所有视频均从 YouTube 下载。最短剪辑时长为 48 帧, 最长剪辑为 6070 帧, 剪辑的平均长度为 181.3 帧。

在设计视频数据集和基准时, 遵循了“野外标记人脸”LFW 图像集的示例[1], 以生成大量视频以及表明每个视频中出现的人的身份的标签。

此外, 还发布了基准测试, 旨在衡量视频配对技术在这些视频上的性能, 并使用完善的描述符方法为这些视频中出现的人脸提供描述符编码。



YouTube Faces

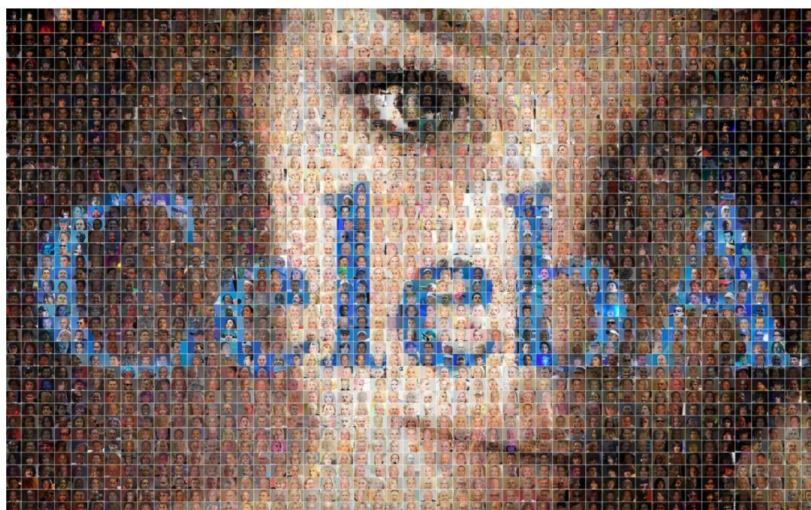
CelebA

CelebFaces Attributes Dataset (CelebA) 是一个大规模的人脸属性数据集，由香港中文大学发布[5,6]。

CelebA 该数据集拥有超过 200K 的名人图像，每个图像有 40 个属性注释。该数据集中的图像涵盖了大的姿势变化和背景杂波。

CelebA 多样性大、数量多、注释丰富，包括 10177 个身份，202599 张人脸图像，以及 5 个地标 (landmark) 位置，每张图像 40 个二元属性注释。

该数据集可用作人脸属性识别、人脸检测、地标（或面部部分）定位以及人脸编辑和合成。



CelebA



CelebA 样本示例

SCface

SCface- Surveillance Cameras Face Database, 是一个静态人脸图像数据集, 由克罗地亚萨格勒布大学 (University of Zagreb) 电气工程与计算机学院于 2009 年发布[7,8]。

SCface 图像是在不受控制的室内环境中使用五个不同质量的视频监控摄像机拍摄。该数据集包含 130 名受试者的 4160 张静态图像 (可见光和红外光谱)。来自不同质量相机的图像模拟真实世界的情况, 并支持鲁棒人脸识别算法测试, 强调不同的执法和监视用例场景。

该数据集的特点可概括如下:

使用了不同质量和分辨率的相机;

图像是在不受控制的照明条件下拍摄;

图像是从不同距离拍摄;

监控图像中的头部姿势是商业监控系统的典型特征, 即摄像头放置在对象头部略上方, 这使得识别要求更高: 此外, 在监控摄像头记录期间, 这些人并没有看向一个固定点;

数据集包含 9 个适合头部姿势建模和/或估计的不同姿势图像;

数据集包含 130 个对象的图像, 足以排除纯巧合获得的性能结果 (纯巧合识别的机会小于 0.7%);

识别和验证场景都是可能的。



Scface



为一个对象设置一个图像集的示例



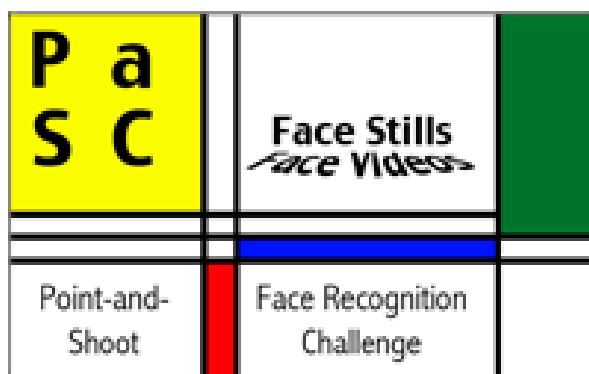
不同姿势图像的示例

PaSC

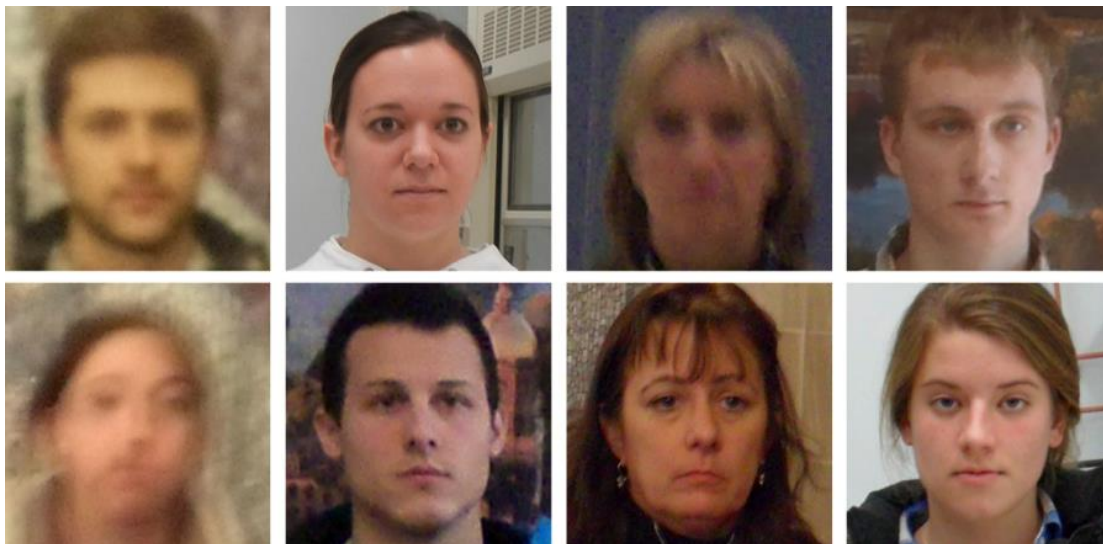
Point and Shoot Face Recognition Challenge (PaSC)的发布时间是 2013 年，由美国国家技术标准局 (NIST) 发布[9,10]。

廉价的“傻瓜式”相机与社交网络相结合，使得普通民众也可使用人脸识别技术。他们想要抓拍照片、拍摄视频、上传，并让他们的朋友、家人和熟人或多或少地被自动识别。尽管问题很简单，但是在这种情况下人脸识别是很困难的。这是 PaSC 创建的动因。

PaSC 包括静止图像和视频。静止照片有 9376 张，共 293 人。根据实验设计，图像采集时，系统地改变传感器、位置、姿态和摄像机距离。PaSC 还包括 2802 段 265 人的简单动作视频，这 265 人静态图像 293 人的子集。



PaSC



从 PaSC 中的静止图像中提取的裁剪人脸图像。这些图像展示了傻瓜式成像、照明、运动模糊和对焦不佳时出现的一些复杂情况。

IJB-S

美国国家技术标准局 (NIST) 在 2015 年召开的 CVPR 上发布了 IJB-A (IARPA Janus Benchmark-A) 图像集，后来又分别发布了 IJB-B 和 IJB-C。2018 年发布了 IJB-S (IARPA Janus Benchmark – Surveillance) [11]。

IJB-S 数据集包含 202 个主体身份的静态图像和视频。数据收集于 2017 年 11 月的两个周末，地点是美国国防部的一个训练设施。之所以选择这个地点，是因为它可以用来模拟执法和国家安全的视频监控收集用例。

IJB-S 有 350 个监控视频，总共 30 个小时，5656 个注册图像和 202 个注册视频。平均而言，每名受试者都出现在 12 段监控录像中。每个图像或帧中的所有可见面都有由人类专家注释的相应边界框坐标。在 IJB-S 数据集的整理过程中，总共产生了超过 1000 万个手工注释。

IJB-S 很具挑战性，是严重低质量的图像和视频。



IJB-S Panasonic WV-SW395 frames



IJB-S Speco O4P30X frames

IJB-S 示例



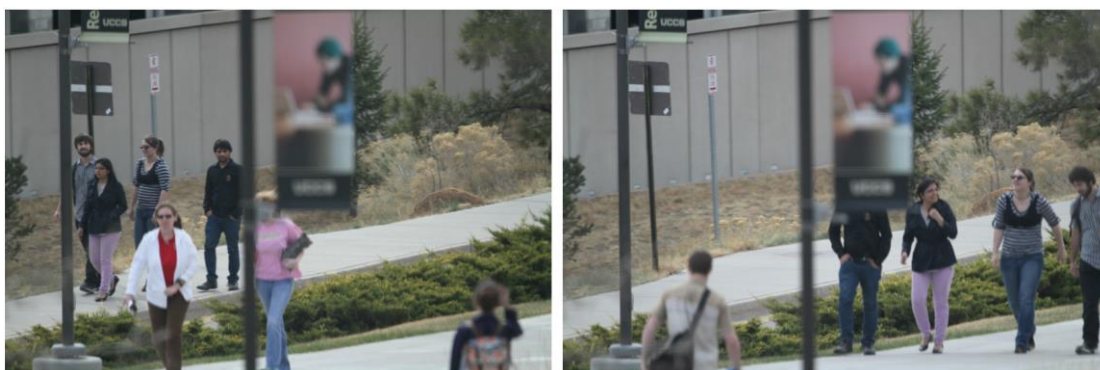
IJB-S 测试数据集的示例图像

UCCS

UCCS 数据集由美国科罗拉多大学科罗拉多斯普林斯分校 (University of Colorado Colorado Springs) 发布于 2013 年[12-14]。

UCCS 展示了可操作的人脸识别场景的特性。数据集的第 1 版包含来自 308 个人的 6337 张图像。使用装有 Sigma 800mm F5.6 EX APO DG HSM 镜头的佳能 7D 相机获取图像。摄像机放置在办公室内，聚焦在离办公室 100m 处的室外人行道上，产生 18 兆像素的场景图像。图像以 100ms 的间隔拍摄，产生大约 10 张不同焦点的照片，每个特定间隔有多个视图和表情。第二天同一个人以相同的时间间隔出现在镜头前的几率很高，例如上学的学生。这样就有一个人在多天的多个序列。图像包含晴天和下雪天等各种天气条件，以

及包含各种遮挡，如太阳镜、冬帽、毛皮夹、树枝、电线杆等。



UCCS 数据集的示例图像



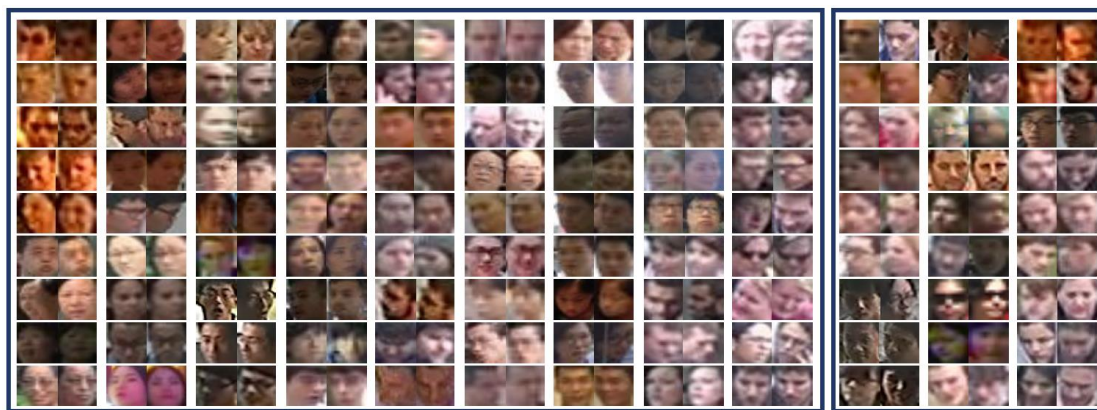
不同的姿势和模糊度

QMUL-Survface

QMUL-Survface 由伦敦大学玛丽女王学院 (QMUL) 2018 年发布[15,16]，是以开发低分辨率监视人脸图像构建的。

该数据集包含 463507 张人脸图像，15573 个不同身份，这些图像是在跨越广阔空间和时间的真实世界不合作监视场景中捕获的。这些低分辨率的人脸图像是原生的，而不是从原始图像低采样人工合成。由于在开放空间和不受限制的时间内出现任意大量的非目标人员（干扰者），因此在开放式设置中人脸识别通常更困难，这对于监视人员搜索场景来说是典型的。

QMUL-Survface 发布者使用该数据集评估了人脸识别性能。除了低分辨率问题，还有其它不受控制的协变量和噪声，例如照明变化、表情、遮挡、背景杂波和压缩伪影。所有这些因素都会在不同程度上导致推理的不确定性（见下图）。



QMUL SurvFace 示例，左：配对；右：未配对



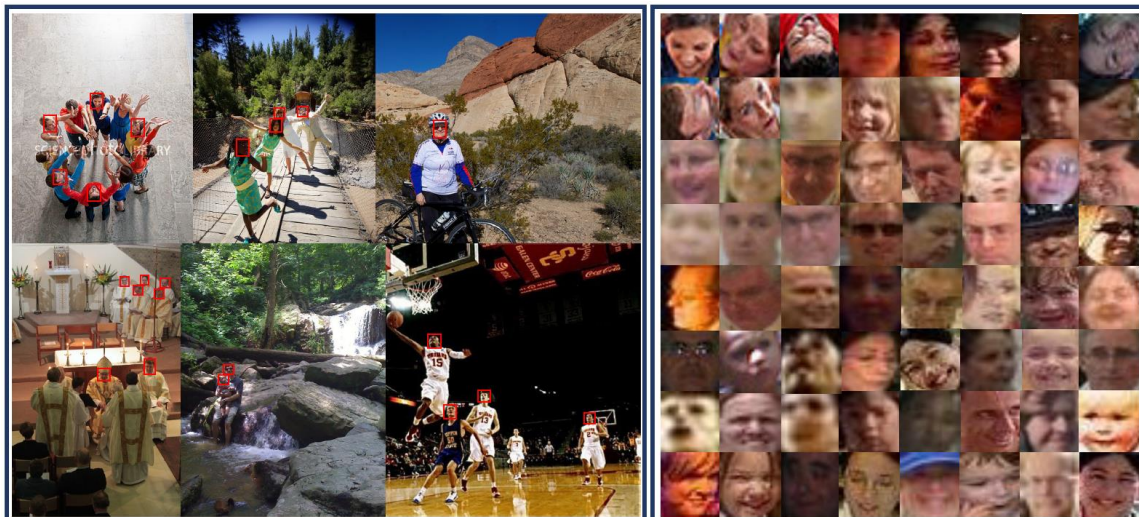
原生监控图像人脸检测 左：自动检测的人面；右：失败案例

Tinyface

Tinyface 也是由伦敦大学玛丽女王学院（QMUL）2018 年发布[17，18]。

TinyFace 数据集有 169403 张原生低分辨率人脸图像（平均 20×16 像素），来自 5139 个标记的身份，用于 1:N 识别测试。

TinyFace 中的所有低分辨率人脸都是从公共网络数据中收集的，这些数据是在姿势、光照、遮挡和背景不受控制的条件下采集的。



TinyFace 图像自动检测示例

参考文献

- [1] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.
- [2] <http://vis-www.cs.umass.edu/lfw/index.html>
- [3] L. Wolf, T. Hassner, and I. Maoz. Face recognition in unconstrained videos with matched background similarity. In CVPR 2011, pages 529–534, June 2011.
- [4] <http://www.cs.tau.ac.il/~wolf/ytfaces/results.html>
- [5] Z. Liu, P. Luo, X. Wang, and X. Tang. Deep learning face attributes in the wild. In Proceedings of International Conference on Computer Vision (ICCV), December 2015.
- [6] <http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>
- [7] M. Grgic, K. Delac, and S. Grgic. Siface - surveillance cameras face database. Multimedia Tools Appl., 51:863–879, 02 2011.
- [8] <https://www.siface.org/>
- [9] J. R. Beveridge, et al., The challenge of face recognition from digital point-and-shoot cameras. In 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pages 1–8, Sep. 2013.
- [10] <https://www.nist.gov/programs-projects/point-and-shoot-face-recognition-challenge-pasc>
- [11] N. D. Kalka, B. Maze, J. A. Duncan, K. OrConnor, S. Elliott, K. Hebert, J. Bryan, and A. K. Jain. Ijb-s: larpa janus

surveillance video benchmark. 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), pages 1–9, 2018.

[12] A. Sapkota and T. E. Boulton. Large scale unconstrained open set face database. In 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pages 1–8, Sep. 2013.

[13] <https://vast.uccs.edu/UCCSfaces/>

[14] <https://vast.uccs.edu/Opensetface/>

[15] Z. Cheng, X. Zhu, and S. Gong. Surveillance face recognition challenge. arXiv preprint arXiv:1804.09691, 2018.

[16] <https://qmul-survface.github.io/>

[17] Z. Cheng, X. Zhu, and S. Gong. Low-Resolution Face Recognition. arXiv preprint arXiv:1811.08965, pages 1–16, nov 2018.

[18] <https://qmul-tinyface.github.io/>



临菲信息技术港



临菲信息技术港公众号



临菲学堂



临菲少年