

文献快读：“深度学习”理解人脸——机器可能比人厉害

“深度学习”理解人脸

——机器可能比人厉害

(临菲信息技术港公众号图文)

Deep Learning for Understanding Faces

Machines maybe just as good, or better, than humans

Rajeev Ranjan, Swami Sankaranarayanan, Ankan Bansal, Navaneeth Bodla, Jun-Cheng Chen, Vishal M. Patel, Carlos D. Castillo, and Rama Chellappa

该文为综述文献。讨论自动人脸识别系统所涉及的不同模块以及深度学习对每个模块的作用，最后给出至今依然存在的挑战性问题。

本文对于在人脸识别、安全、视觉监控和生物识别领域的工程技术人员和最终用户都有参考价值。

参考文献 112 篇。



我们能从人脸知道什么？

尽可能多地提取信息，例如：位置、姿势、性别、身份、年龄、情绪，等等。

- **三个基本模块**

人脸检测 (face detector)：在图像或视频中确定人脸的位置。

基准点检测 (fiducial point detector) : 定位重要的面部标志, 如: 眼睛中心、鼻尖、嘴角等。

特征描述 (feature descriptor) : 对人脸中提取的特性信息进行编码。

在给定面部的表示后, 通过度量获得它们之间的相似性得分, 然后与阈值比较, 从而判断两个人脸是否来自相同目标。

- **深度学习的引入**

从上世纪 90 年代初以来, 提出了许多人脸识别/验证算法, 但这些算法对条件和环境的依赖性很大, 例如, 当姿势、照明、分辨率、表情、年龄、背景等变化较大时, 算法的性能会严重下降。而且, 如果要从大量的视频中识别出对象, 算法必须快速而且鲁棒。

为应对上述挑战, 人们将深度学习应用于人脸识别, 提出了若干新的方法。其中, 深度卷积神经网络 DCNN (deep convolutional neural network) 被认为是有效的方法。

无约束图像人脸检测

DCNN 具有强大的深度特征提取能力。基于 DCNN 的人脸检测方法可分为两类: 基于区域 (region-based) 的方法和滑动窗口 (sliding-window) 方法。

- **基于区域的方法**

Faster R-CNN

目标检测的不断进步: 从 R-CNN、Fast R-CNN 到 Faster R-CNN。

基于 FasterR-CNN 的检测器, 使用共同 DCNN 产生提议区域 (proposals) 和分类人脸, 还可以同时回归每个面部提议的边界框坐标。

- **基于滑动窗的方法**

与基于区域的方法相比, 该方法的速度快, 并且可以仅使用以滑动窗口方式工作的卷积操作来实现。

- **单次检测器 (Single-shot detector , SSD)**

SSD 是基于滑动窗口的检测器, 但它不是以不同的比例创建图像金字塔, 而是利用 DCNN 中存在的内置金字塔结构。

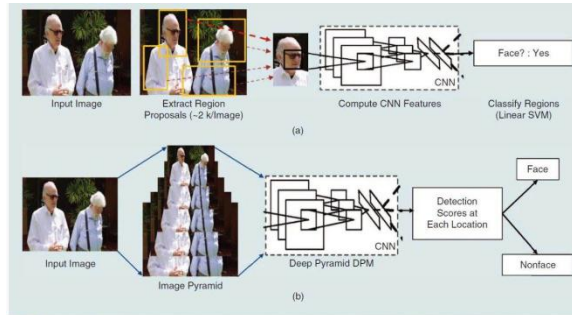


FIGURE 1. A typical DCNN-based face detector using the (a) region-proposal approach and (b) the sliding-window approach.

区域建议和滑动窗检测

找到重要的面部关键点和头部朝向

- 基于模型
 - 例如 AAM, ASM, and CLM 等。
- 基于级联回归

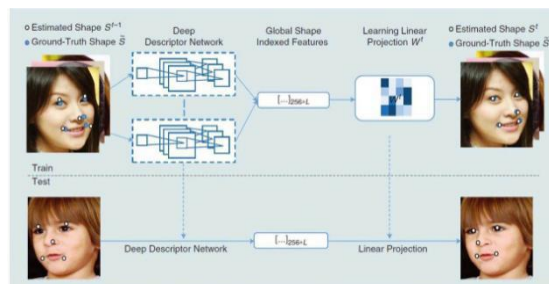


FIGURE 4. A model diagram of regression-based keypoint estimation method using DCNNs. (Figure used with permission from [11].)

关键点检测

人脸识别与验证

- 基于深度学习的人脸鲁棒特征学习
- 人脸判别度量学习
- 实现
- 训练数据集
- 性能

LFW 数据集

IJB-A 基准数据集

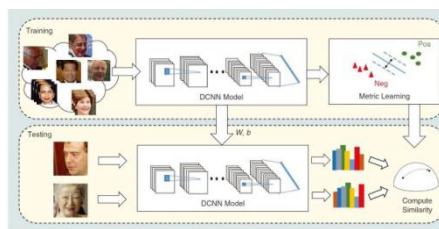


FIGURE 5. A general pipeline for training and testing a face verification system using DCNNs.

DCNN 人脸验证测试

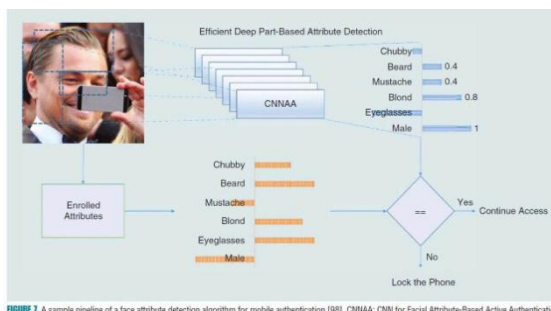
面部属性

识别面部属性，如 性别、表情、年龄、肤色等。这些属性对于类似图像检索、情感检测和移动安全等应用非常有用。

最近 DCNN 被用于属性分类，展示了令人印象深刻的结果。面向深度属性的姿态对齐网络 PANDA (Pose Aligned Networks for DeepAttributes)，基于部件模型与深度学习相结合，用以训练姿态归一化的 DCNN，达到了先进的技术性能。

针对年龄和性别，使用多种数据集进行了研究。

Samangouei 等人 2016 年设计了适用于移动设备的高效 DCNN 体系结构。



移动认证人脸属性检测

人脸分析的多任务学习

在多任务框架中较早使用 DCNN 的是 HyperFace (2016 年) 和 TCDCN (2014 年)。

HyperFace 本质上是同时实现人脸检测、标志定位、头部姿态估计和性别分类的算法。

TCDCN 算法使用基于 MTL 的 CNN 进行面部标志检测、连同离散头部偏航估计、性别识别、微笑预测和眼镜检测等。

最近提出的 All-in-One Face (2017 年)，是单一的 DCNN 模型，同时完成人脸检测、脸部标志定位、人脸识别、三维头部姿态估计、微笑检测、面部年龄估计和性别分类等。

开放性问题

- **人脸检测**

依然是光照、面部表情、视点和遮挡等带来的面部外观的变化，以及模糊和低分辨率等其他因素的挑战。

- **基准检测**

面对极端姿态、低光照、小和模糊的人脸图像带来的挑战，大规模标注和无约束的数据集，将使得人脸对齐系统更加稳健。

- **人脸识别/验证**

由于图形卡的存储量限制带来的问题，例如用大规模数据集那样的在线训练方法在端到端的使用。另一个具有挑战性的问题，是在深度网络中加入全动态视频处理，以实现基于视频的人脸分析。

文献： [Deep Learning for Understanding Faces Machines may be just as good, or better, than humans. IEEE Signal Processing Magazine | January 2018 |](#)

(本文原载：微信公众号“临菲信息技术港”)

