



人脸识别

从PCA到CNN

2015清华校友人工智能论坛
清华大学FIT楼多功能厅

报告人： 陈健生
清华大学电子工程系
2015-12-20

概述

- Face Recognition is easy, socially



- Face Recognition is difficult, technically

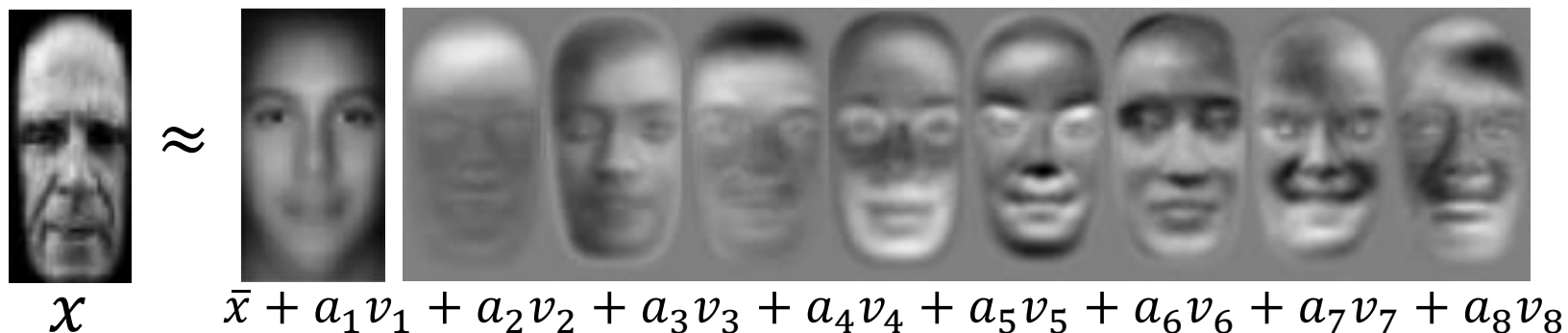


概述

- 人脸几何比例特征识别 (Kanade, 1973)
- 人脸的线性表示与降维 (Sirovich, 1987)
- **PCA特征用于人脸识别** (Turk&Pentland, 1991)
- **主动形状模型用于人脸配准** (Cootes etc, 1995)
- 其他线性表示用于人脸识别 (Belhumeur, 1997)
- **形变模型用于三维人脸重建** (Blanz&Vetter, 1999)
- **集联分类器用于人脸检测** (Viola&Jones, 2001)
- 非线性LBP特征用于人脸识别 (Ahonen, 2007)
- **CNN用于人脸检测、定位、识别 ...** (Since 2013)

PCA特征用于人脸识别

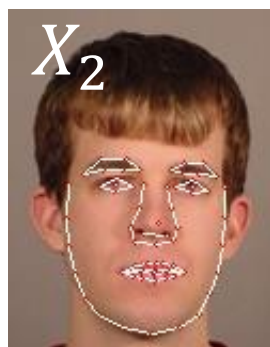
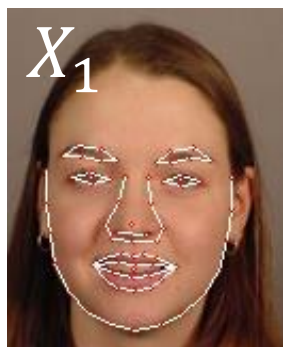
$$x \rightarrow \left(\underbrace{(x - \bar{x}) \cdot v_1}_{a_1}, \underbrace{(x - \bar{x}) \cdot v_2}_{a_2}, \dots, \underbrace{(x - \bar{x}) \cdot v_k}_{a_k} \right)$$



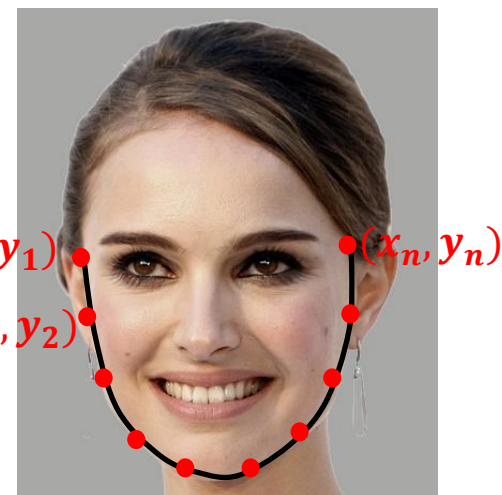
- 人脸图像视为 R^N 空间向量 x
- 人脸图像集合视为线性子空间
- 选取若干主方向 v_i 作为子空间的基
- 人脸图像在主方向上投影系数 a_i 为特征



主动形状模型用于人脸配准

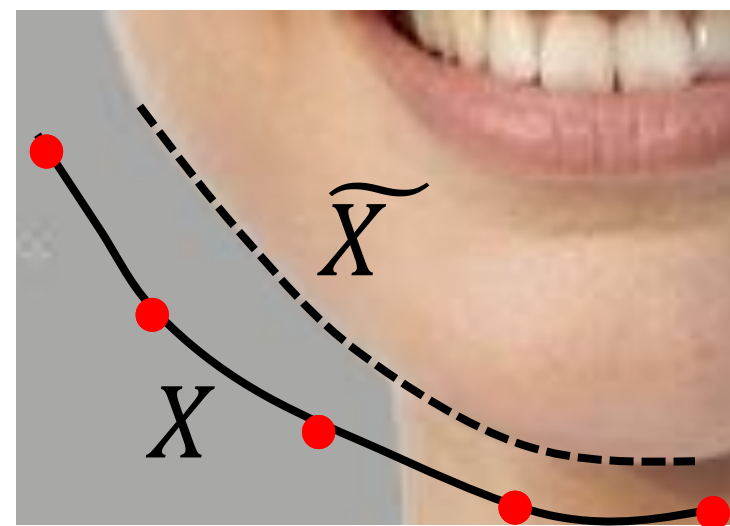


⋮

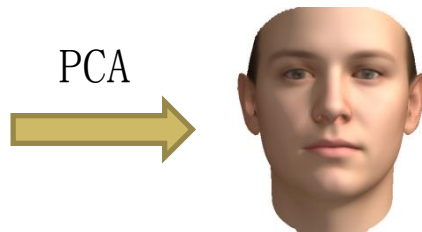
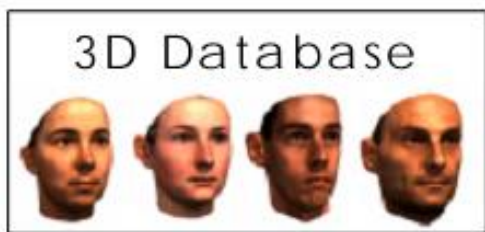


- 点序列列表征形状 $X = [x_1, \dots, x_n, y_1, \dots, y_n]^T$
- 手工标注训练集 X_1, \dots, X_m , 训练统计模型
 $X = \text{均值}(X_i) + b \cdot \text{方差}(X_i)$

$$\begin{aligned} & \text{minimize } \|X(b) - \tilde{X}\|^2 \\ & \text{subject to } |b_i| \leq 3\sqrt{\lambda_i} \end{aligned}$$



形变模型用于三维人脸重建



$$\mathbf{s} = \bar{\mathbf{s}} + \sum_i \mathbf{s}_i \alpha_i = \bar{\mathbf{s}} + \mathbf{S} \alpha$$

$$\mathbf{t} = \bar{\mathbf{t}} + \sum_i \mathbf{t}_i \beta_i = \bar{\mathbf{t}} + \mathbf{T} \beta$$



$$S = (X1, Y1, Z1, \dots, Xn, Yn, Zn)^T$$

$$T = (R1, G1, B1, \dots, Rn, Gn, Bn)^T$$

分析：最小化欧式距离

$$E_I = \sum_{x,y} \| I_{input}(x, y) - I_{model}(x, y) \|^2$$



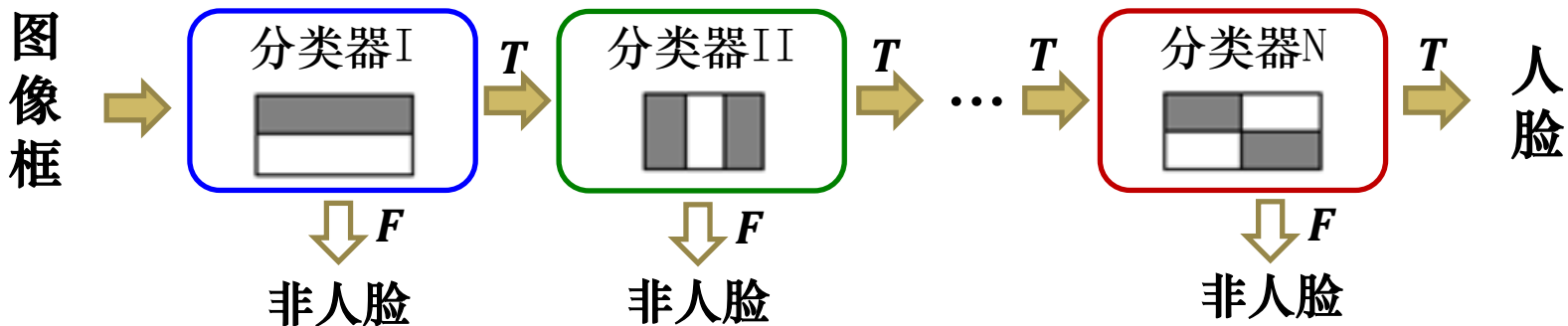
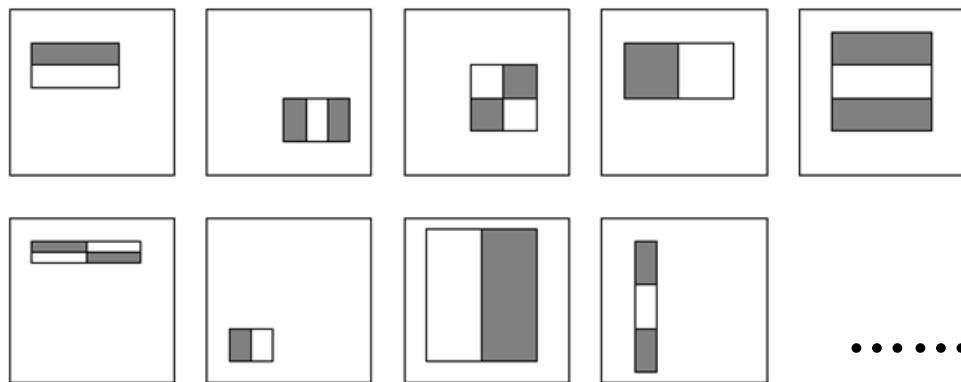
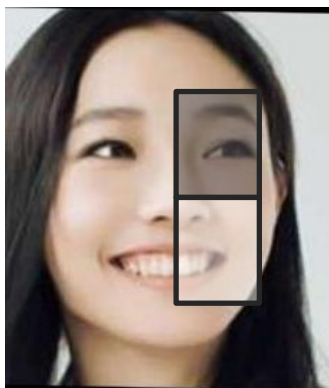
输入图像



合成图像



集联分类器用于人脸检测



- 弱分类器级联获得高分类精度
 - 积分图降低计算复杂度
 - 多尺度检测





Many more

我们的一些工作

$$\text{minimize } \sum_i \left(\sum_{j \in \omega_i} (\hat{I}(j) - \alpha_i I(j))^2 + \lambda (\alpha_i - R_i)^2 \right)$$

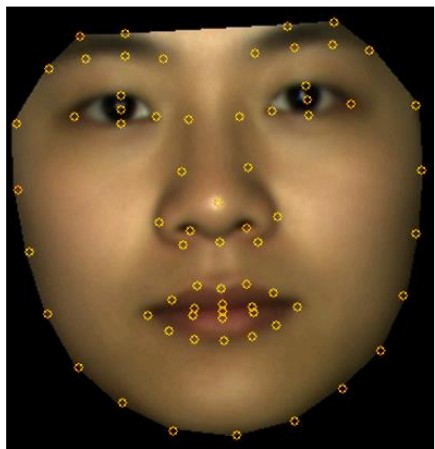
$$S \cdot \hat{I}_t^B = U \quad u_k = \lambda I_t^A(k) \sum_{i|k \in \omega_i} \frac{\tau_i}{\gamma_i + \lambda}$$

- 人脸图象光照转移
- 局部光照平缓假设
- 同时模拟投射阴影和附着阴影
- 有解析解, 转化为大型稀疏线性方程组

(Chen etc., 2010)

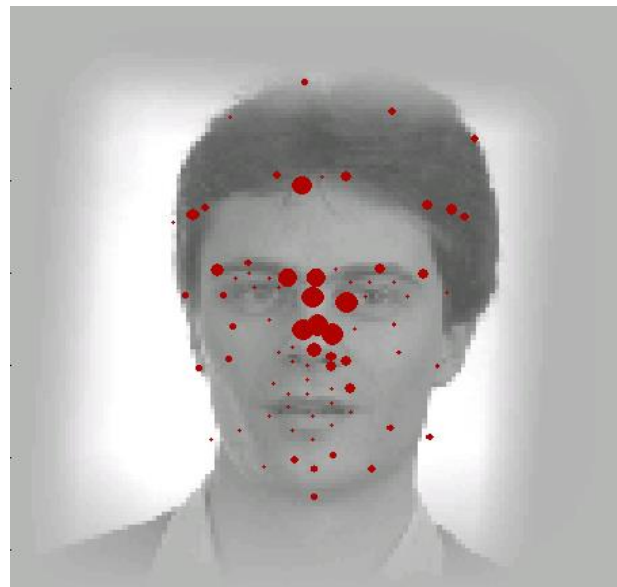


我们的一些工作



minimize $\sigma_G / (\mu_G - \Psi(I^T * \omega, r) / r)$

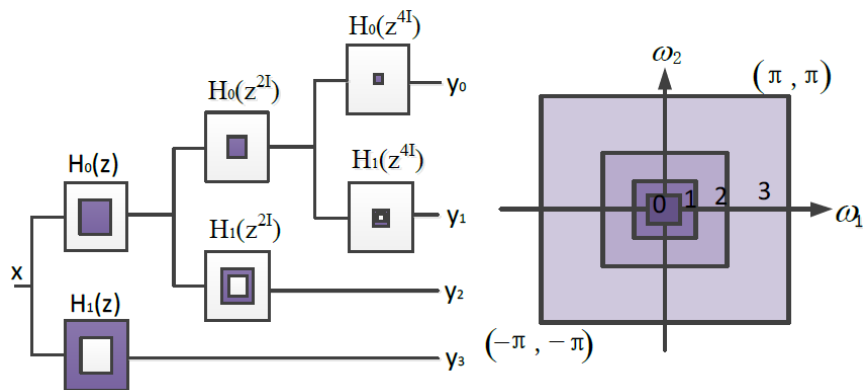
subject to $\sum_{i=1}^N \omega_i = 1$



- 不同人脸区域对于识别率的贡献
- 分数层融合, 最大化识别率指标
- 转化为拟凸优化问题, 多项式时间求解
- 有效提升多种识别算法识别率

(Gong etc., 2011)

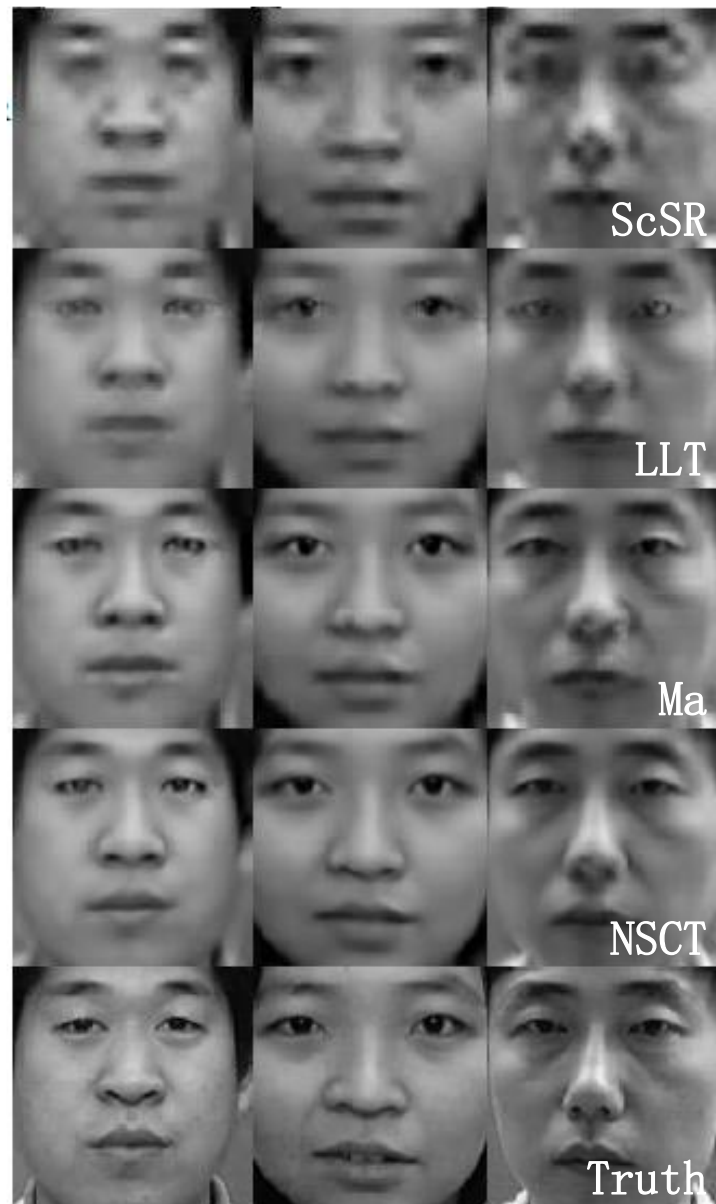
我们的一些工作



Non-Subsampled Contourlet Transform

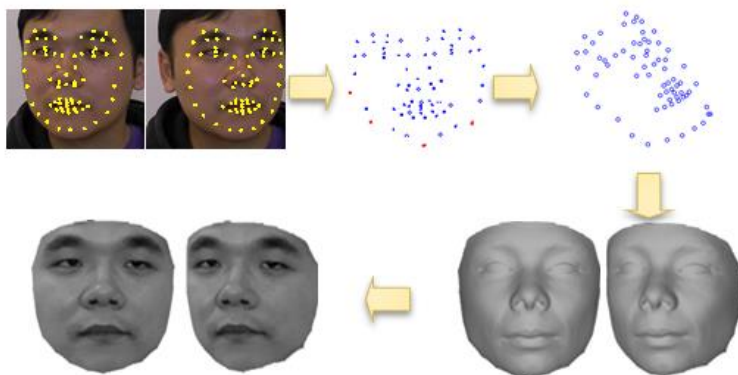
- 在多尺度NSCT域采用嵌入式学习
- 对图像不同频段分离学习
- 更好的频域信息保持特性

(Liu et al., 2012)



$16 \times 16 \rightarrow 128 \times 128$

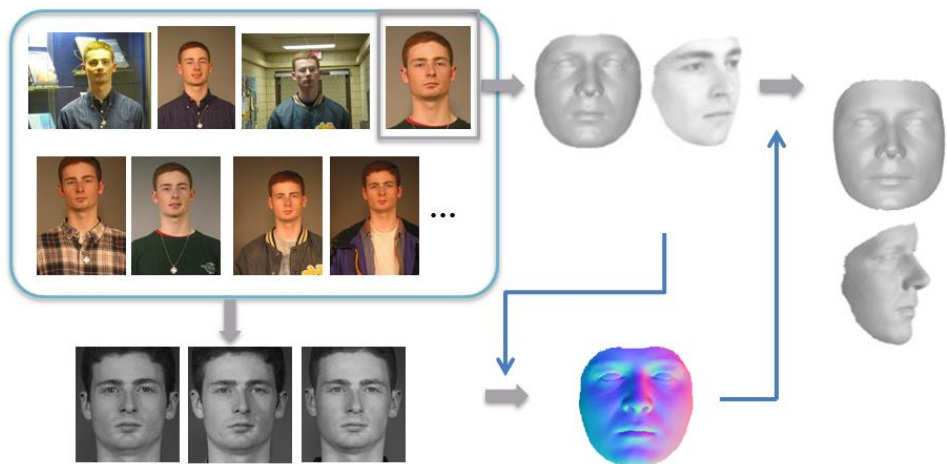
我们的一些工作



(Yang et al., 2013)



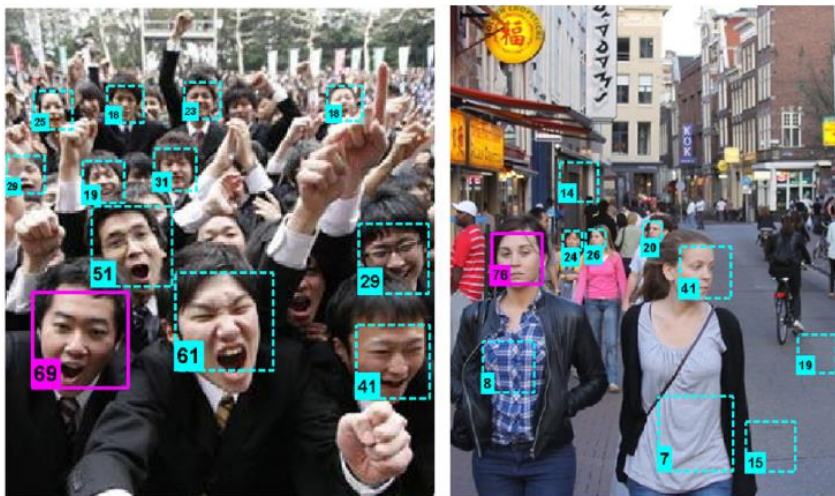
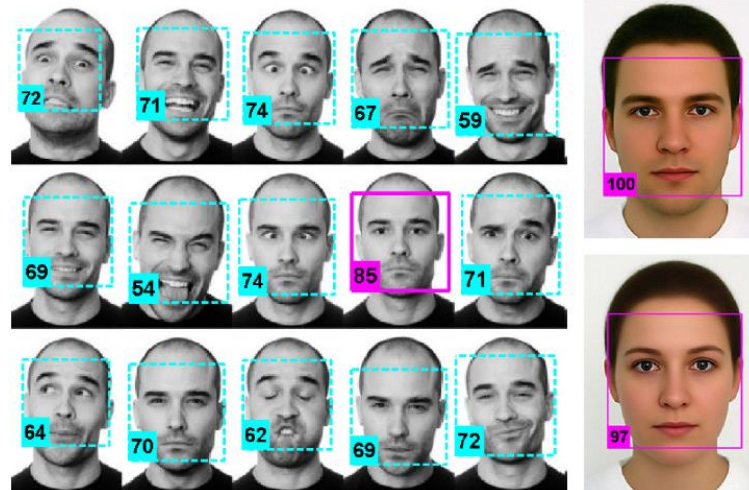
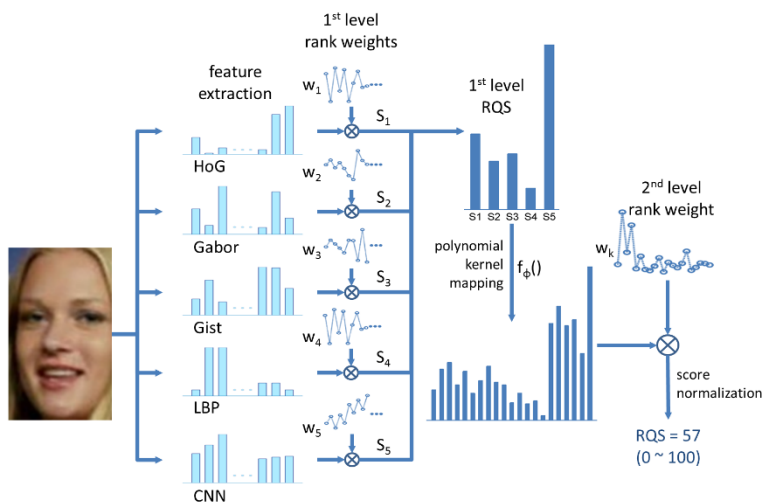
(Chen et al., 2014)



(Yang et al., 2014)

- 基于人脸对称性的三维重建
- 基于SFM的跟踪鲁棒三维重建
- 多源图像的精细化三维重建

我们的一些工作



- 人脸图象质量判别框架
- 排序学习扩展训练样本空间
- 多特征融合提高鲁棒性

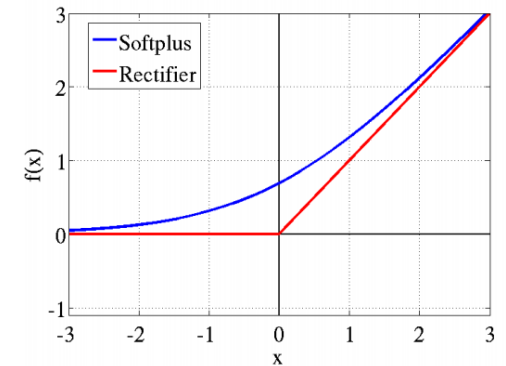
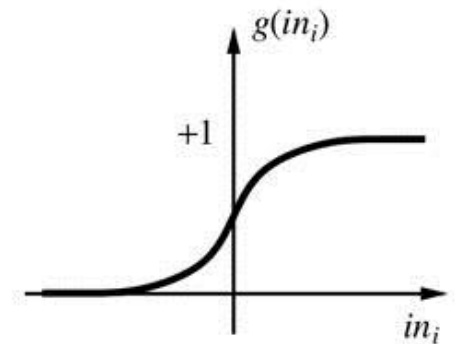
(Chen et al., 2015)

一点总结

- Many **smart** ideas ...
- 在不同的问题层面取得了**较好的效果**
- 针对不同的子问题需构建**不同的模型框架**
- 利用**多种**机器学习工具 (PCA, LDA, SVM, ...)
- 将问题近似到**线性空间**中加以解决
- 采用**handcrafted**非线性特征 (LBP, HoG, SIFT, ...)
- 训练数据的扩充对于性能的提升渐趋**不明显**
- Fail cases 对人而言往往**较难理解**

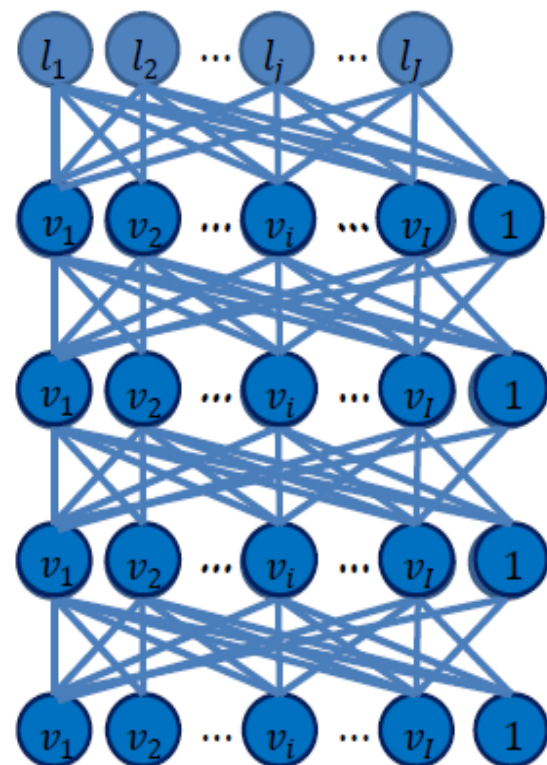
深度学习

- 利用简单而统一的形式实现非线性描述能力



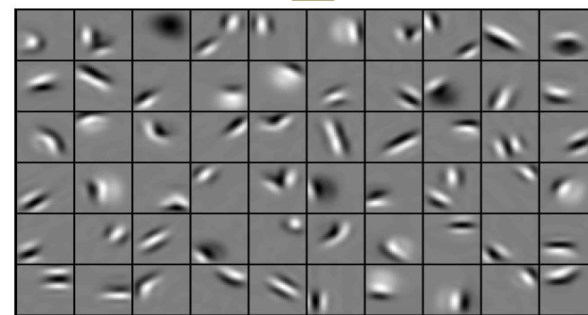
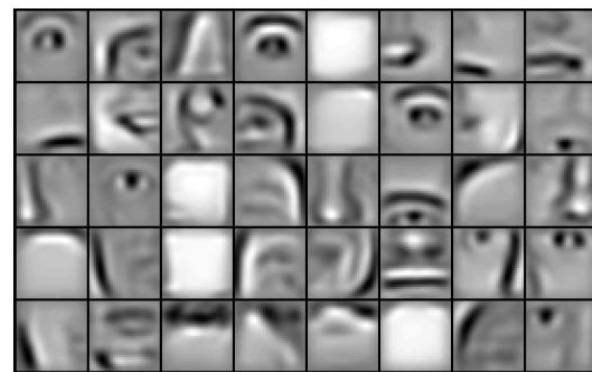
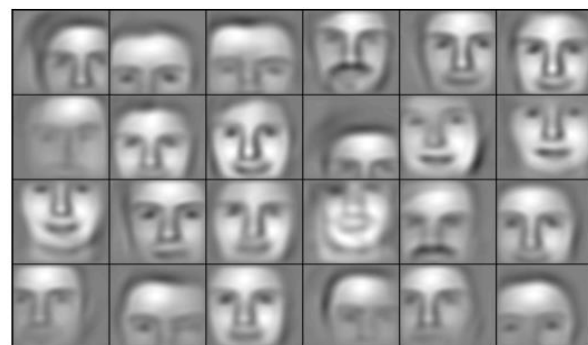
深度学习

- 利用简单而统一的形式实现非线性描述能力
- 深层的结构基于较少的节点构建更复杂的模型空间



深度学习

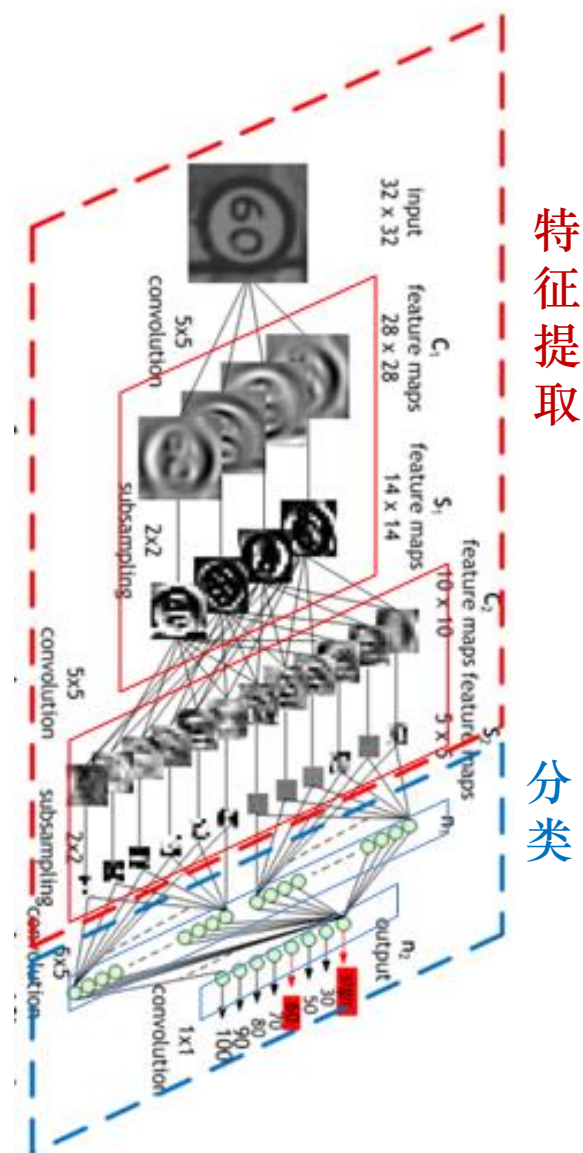
- 利用简单而统一的形式实现非线性描述能力
- 深层的结构基于较少的节点构建更复杂的模型空间
- 替代handcraf自动学习出具有鉴别能力的多层次特征



(Lee etc., 2009)

深度学习

- 利用简单而统一的形式实现非线性描述能力
- 深层的结构基于较少的节点构建更复杂的模型空间
- 替代handcraft自动学习出具有鉴别能力的多层次特征
- 便捷的端到端学习框架构建



深度学习

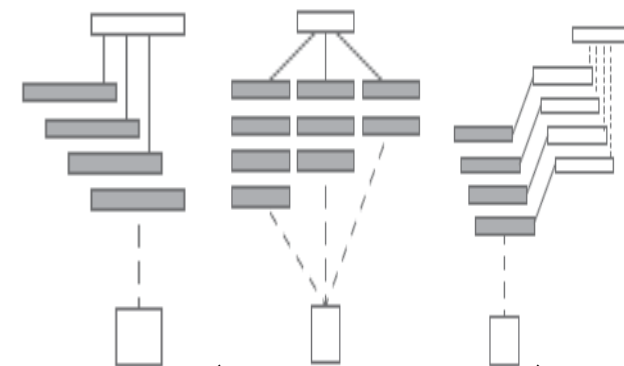
- 利用简单而统一的形式实现非线性描述能力
- 深层的结构基于较少的节点构建更复杂的模型空间
- 替代handcraf自动学习出具有鉴别能力的多层次特征
- 便捷的端到端学习框架构建
- 灵活的模型架构可以适应问题的变化以及数据的快速增加



ImageNet



MegaFace



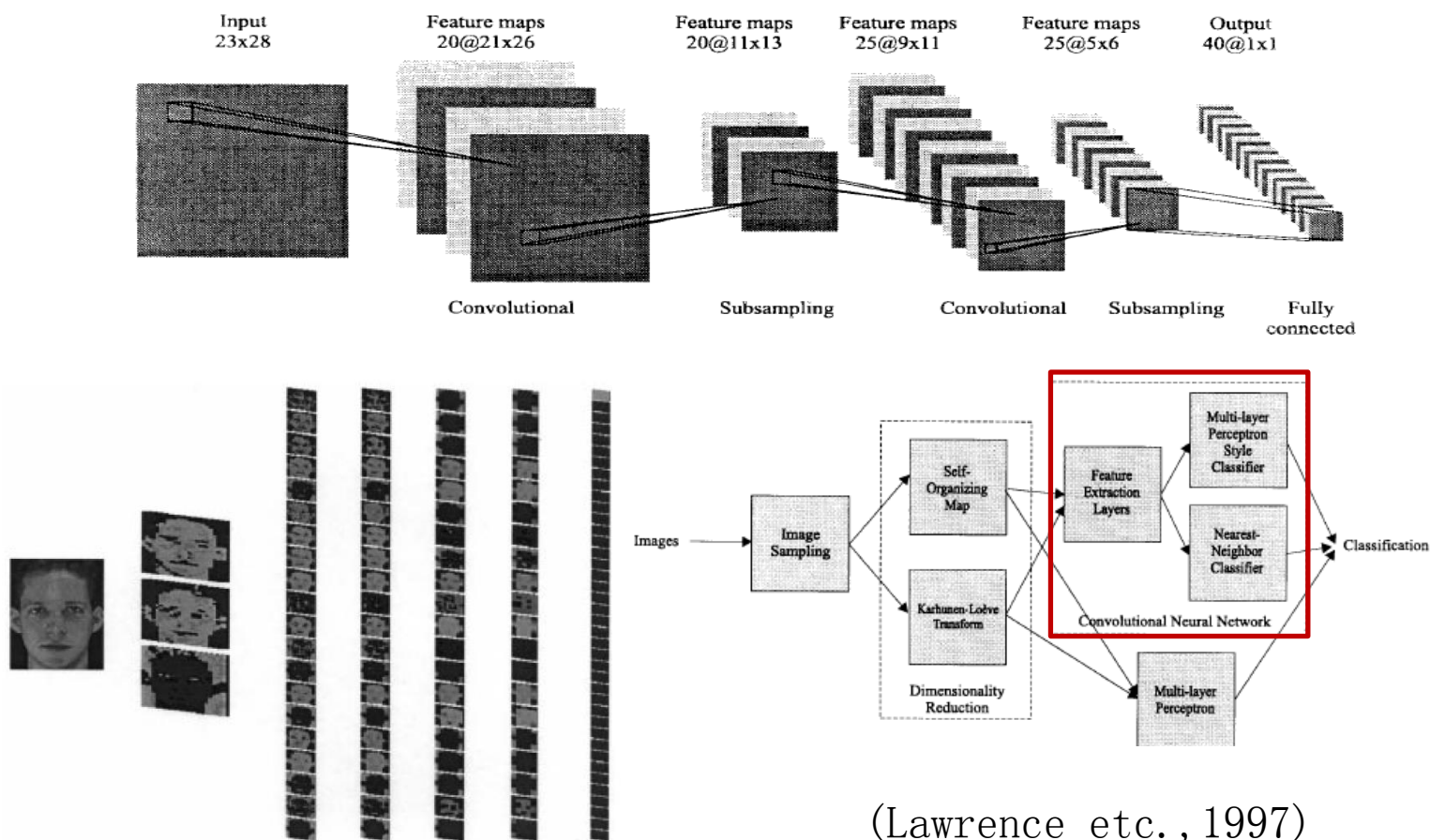
(Xie etc., 2015)

深度学习

- 数据的增加往往带来性能的提升
- 数据的数量往往比数据的质量还重要
- 多种训练方式可以更加充分地利用数据
 - 图像的反转、随机切割、几何扰动等
 - 基于图像对的学习 $N \rightarrow N^2$
 - 三元组 (Triplet) 学习 $N \rightarrow N^3$
- 方便地实现多尺度特征及多模型的融合
- 多种形式的损失函数灵活使用

深度学习与人脸识别

- 2013年起逐步成为主流, 但很早即有相关研究发表

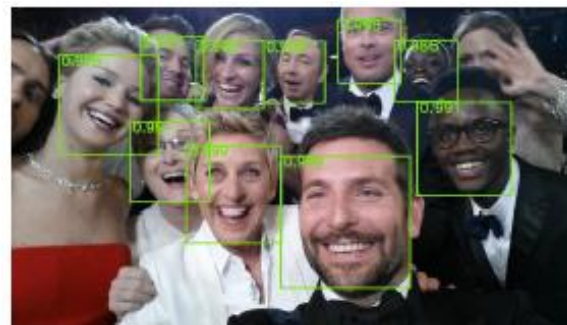
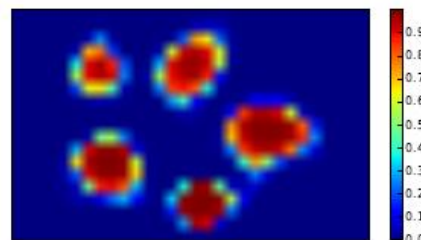


(Lawrence etc., 1997)

深度学习与人脸识别

近年来几乎成功地应用于人脸识别的各个子问题！

- 多姿态鲁棒人脸检测

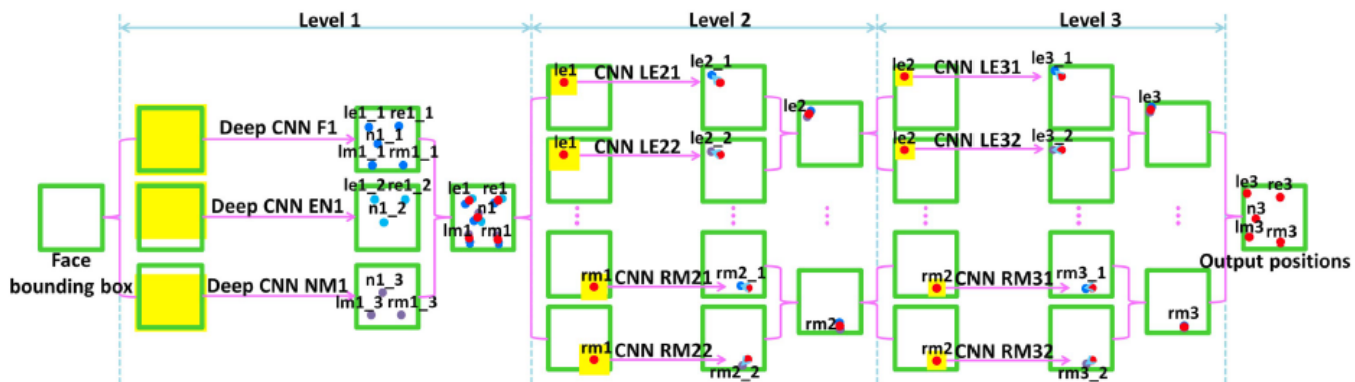


(Farfadi et al., 2015)

深度学习与人脸识别

近年来几乎成功地应用于人脸识别的各个子问题！

- 多姿态鲁棒人脸检测
- 人脸特征点定位与配准

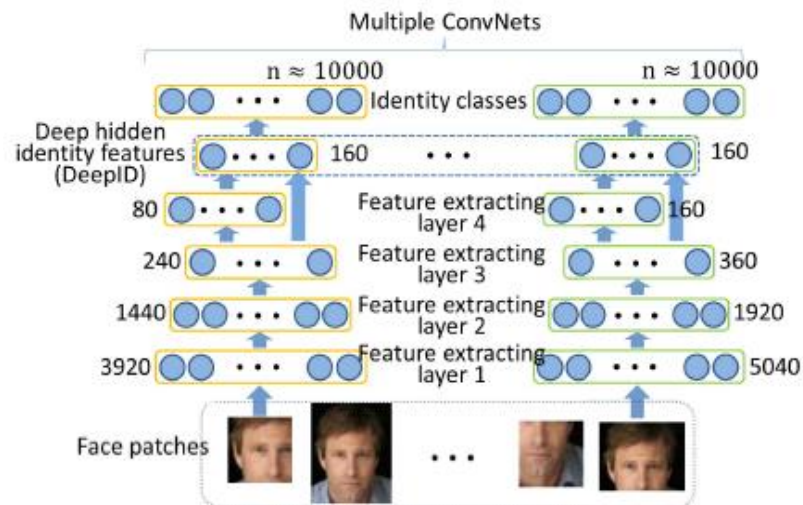


(Sun et al., 2015)

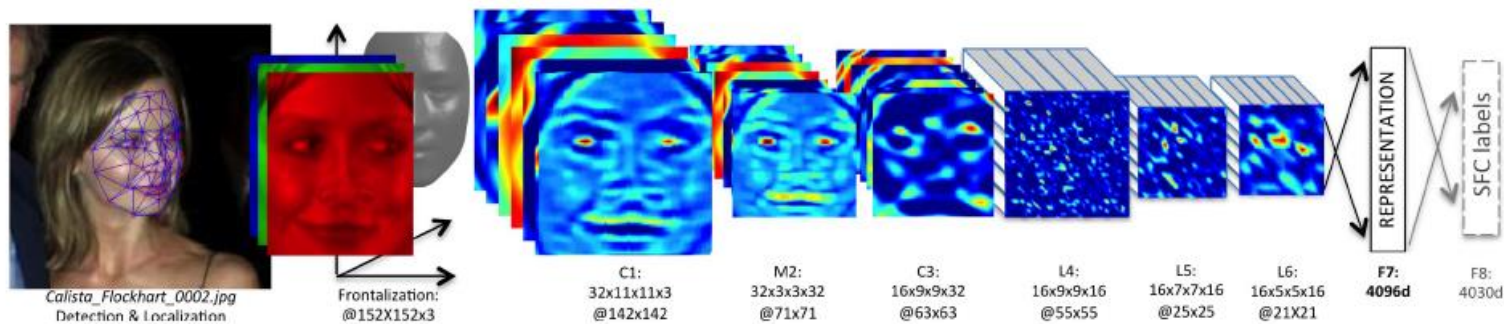
深度学习与人脸识别

近年来几乎成功地应用于人脸识别的各个子问题！

- 多姿态鲁棒人脸检测
- 人脸特征点定位与配准
- 人脸特征提取与识别



(Sun et al., 2014)



(Taigman et al., 2014)

深度学习与人脸识别

近年来几乎成功地应用于人脸识别的各个子问题！

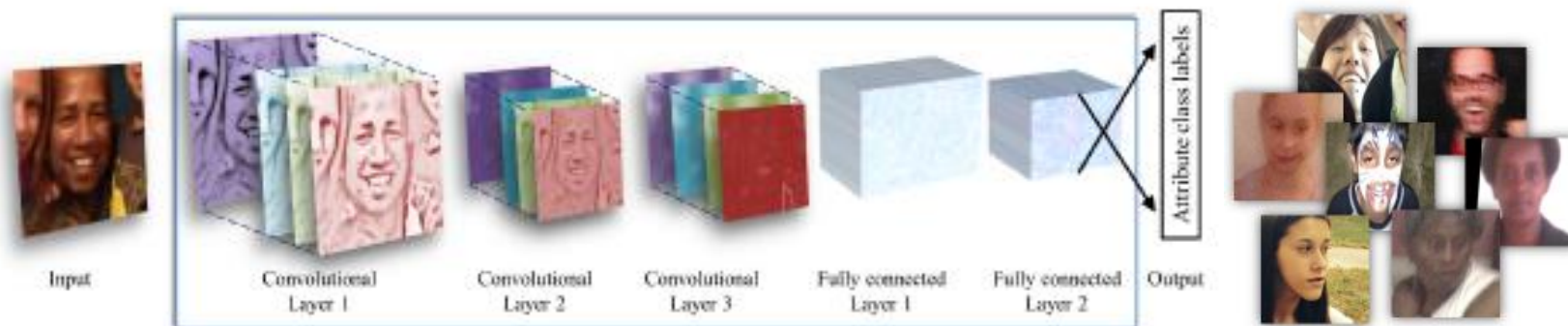
- 多姿态鲁棒人脸检测
- 人脸特征点定位与配准
- 人脸特征提取与识别
- 性别判断与年龄估计

Method	Accuracy
Best from [10]	77.8 ± 1.3
Best from [23]	79.3 ± 0.0
Proposed using single crop	85.9 ± 1.4
Proposed using over-sample	86.8 ± 1.4

性别判别

Method	Exact	1-off
Best from [10]	45.1 ± 2.6	79.5 ± 1.4
Proposed using single crop	49.5 ± 4.4	84.6 ± 1.7
Proposed using over-sample	50.7 ± 5.1	84.7 ± 2.2

年龄估计



(Levi et al., 2015)

深度学习与人脸识别

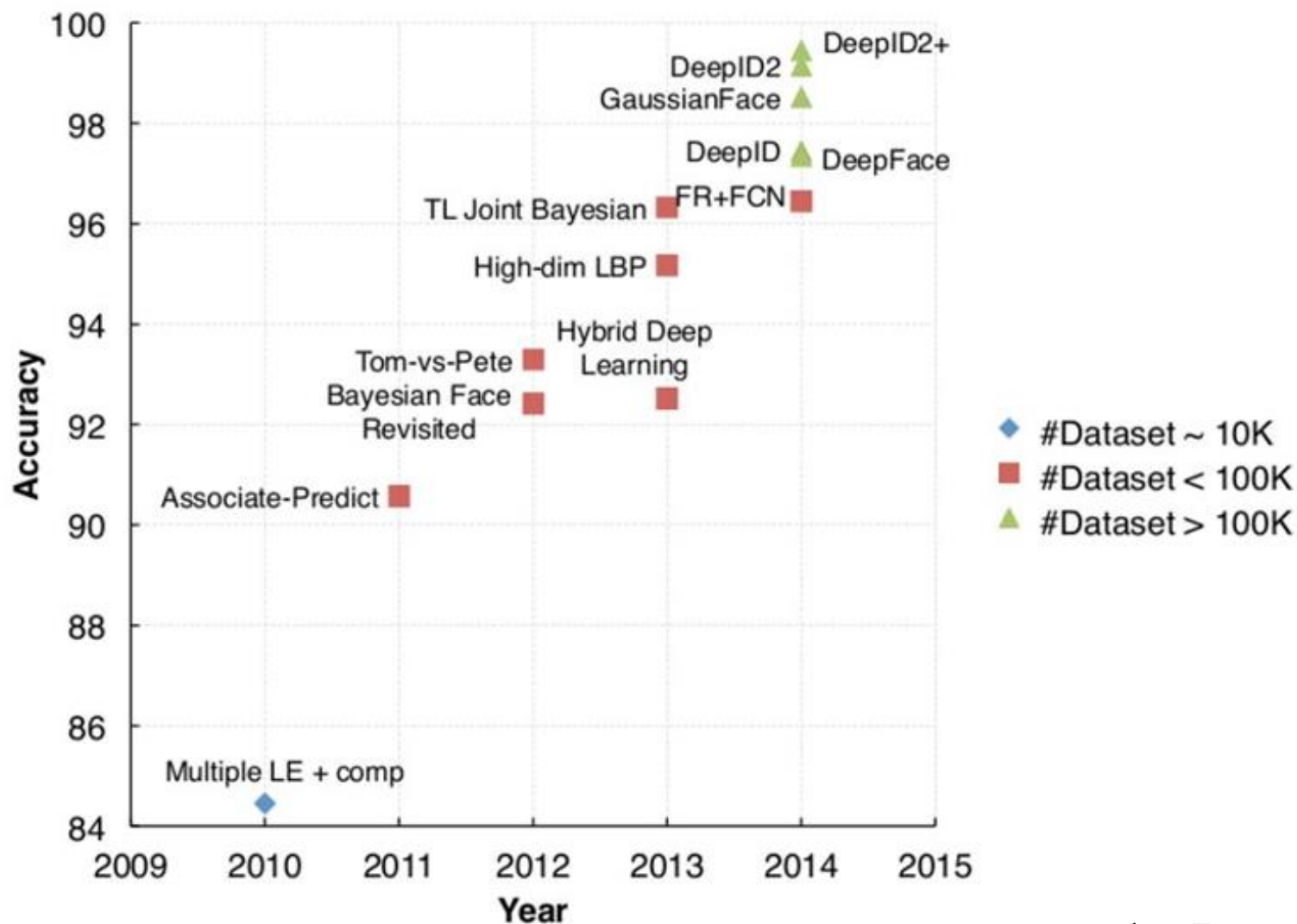
训练集规模

Model	Identities	Images
DeepID	10, 000+	202, 000+
WebFace	10, 000	986, 000+
DeepFace	4, 000+	120, 000, 000+
FaceNet	8, 000, 000+	200, 000, 000+

网络结构规模

Model	Layers	Accuracy
DeepID	7	97.45% \pm 0.26%
DeepFace	8	97.35% \pm 0.25%
Face++	10	99.50% \pm 0.36%
FaceNet	22	99.60% \pm 0.09%

深度学习与人脸识别



深度学习与人脸识别

计算复杂度

训练:	<i>PCA</i> hours	<i>CNN</i> weeks
识别:	<i>PCA</i> 10^7 乘加 10^7 参数 < 80%	<i>CNN</i> 2.5×10^7 乘加 1.3×10^6 参数 > 95%

- * 输入人脸图像像素数= 10000
- * *PCA*降维维度= 1000
- * *CNN*为7层网络以及常用的参数配置

Why CNN ?

- One ‘stupid’ idea ...
- 在不同问题层面以可比复杂度取得了**明显更好**的效果
- 针对不同的子问题采用**统一**的模型框架
- 几乎只用**一种**数学方法 (BP + some tricks)
- 在**高度非线性**空间中解决问题
- 特征通过训练**自动获得**
- 训练数据的扩充对于性能的提升**非常明显**
- Fail cases 对人而言似乎**更易理解**



Thank
you