



QCon 全球软件开发大会
INTERNATIONAL SOFTWARE
DEVELOPMENT CONFERENCE

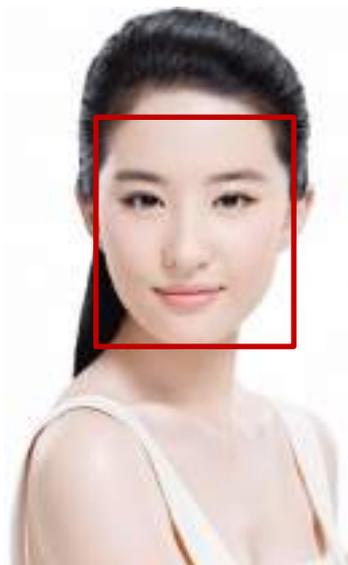
BEIJING 2018

苏宁无人店之人脸识别技术探讨

演讲者/王止观

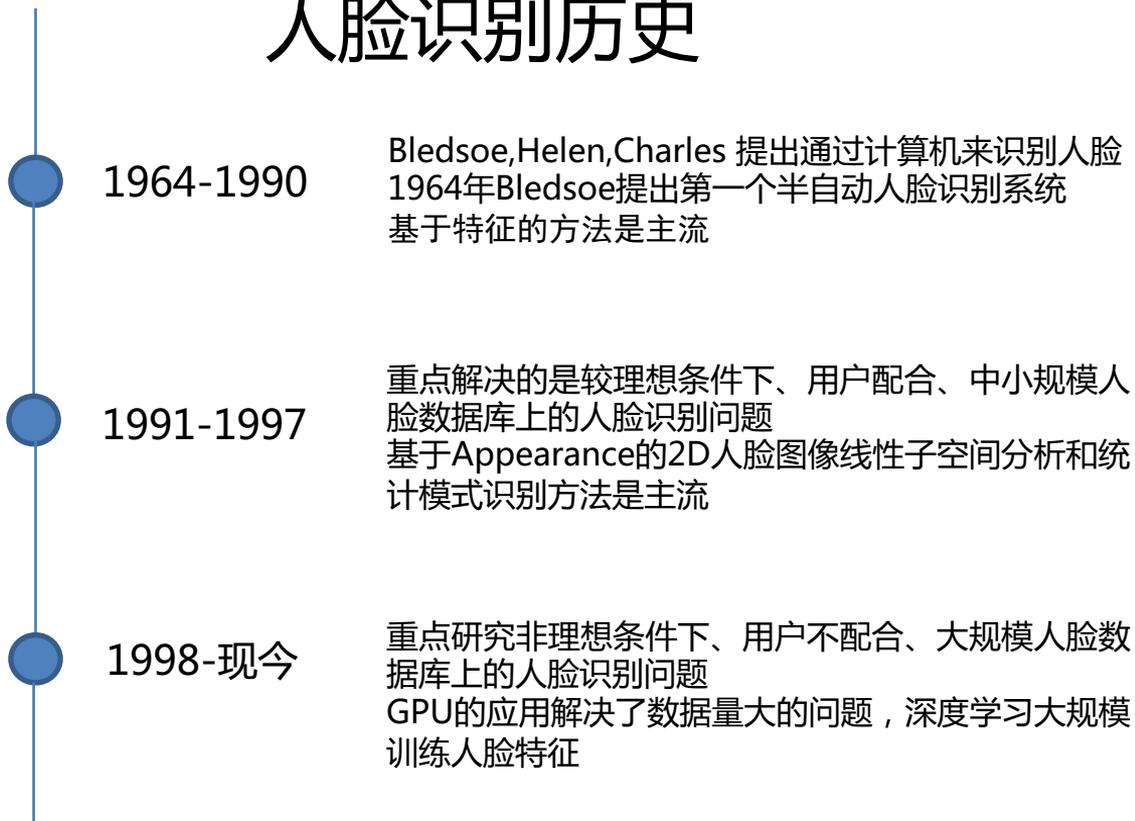
什么是人脸识别？

人脸识别，是基于人的脸部**特征信息**进行**身份识别**的一种生物识别技术。

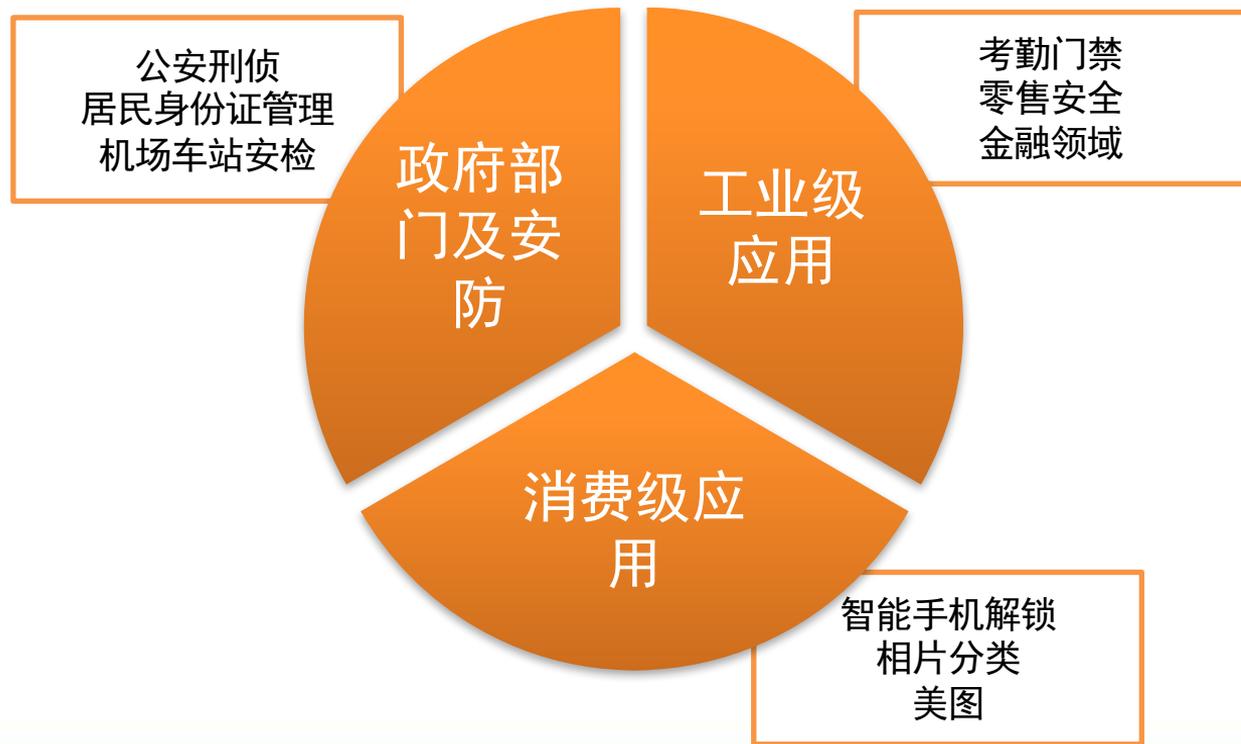


刘亦菲

人脸识别历史



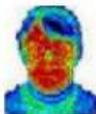
人脸识别的应用领域



人脸识别的优缺点



人脸



红外温谱

How

笔迹



语音



签名



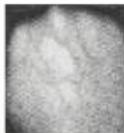
虹膜



视网膜



手形



手背脉管



指纹

生物特征识别：未来的身份验证方法！

• 人脸识别的优点

- 可以隐蔽操作，特别适用于安全问题、罪犯监控与抓逃应用
- 非接触式采集，没有侵犯性，容易接受
- 方便、快捷、强大的事后追踪能力
- 符合我们人类的识别习惯，可交互性强，无需专家评判

• 人脸识别的不足

- 不同人脸的相似性小
- 识别性能受外界条件的影响大

人脸识别的挑战

影响人脸图像表观的因素

人与摄像设备的位置关系（距离角度等）

光照环境条件

饰物（眼镜帽子等）

摄像设备

化妆、整容

图像存储质量

精神状态

年龄变化

健康状况

意外损伤

面部毛发(头发，胡须)



人脸识别分类

人脸验证(verification)



1:1 身份验证模式本质上是计算机对当前人脸与人像数据库进行快速人脸比对并得出是否匹配的过程，可以简单理解为证明你就是你。

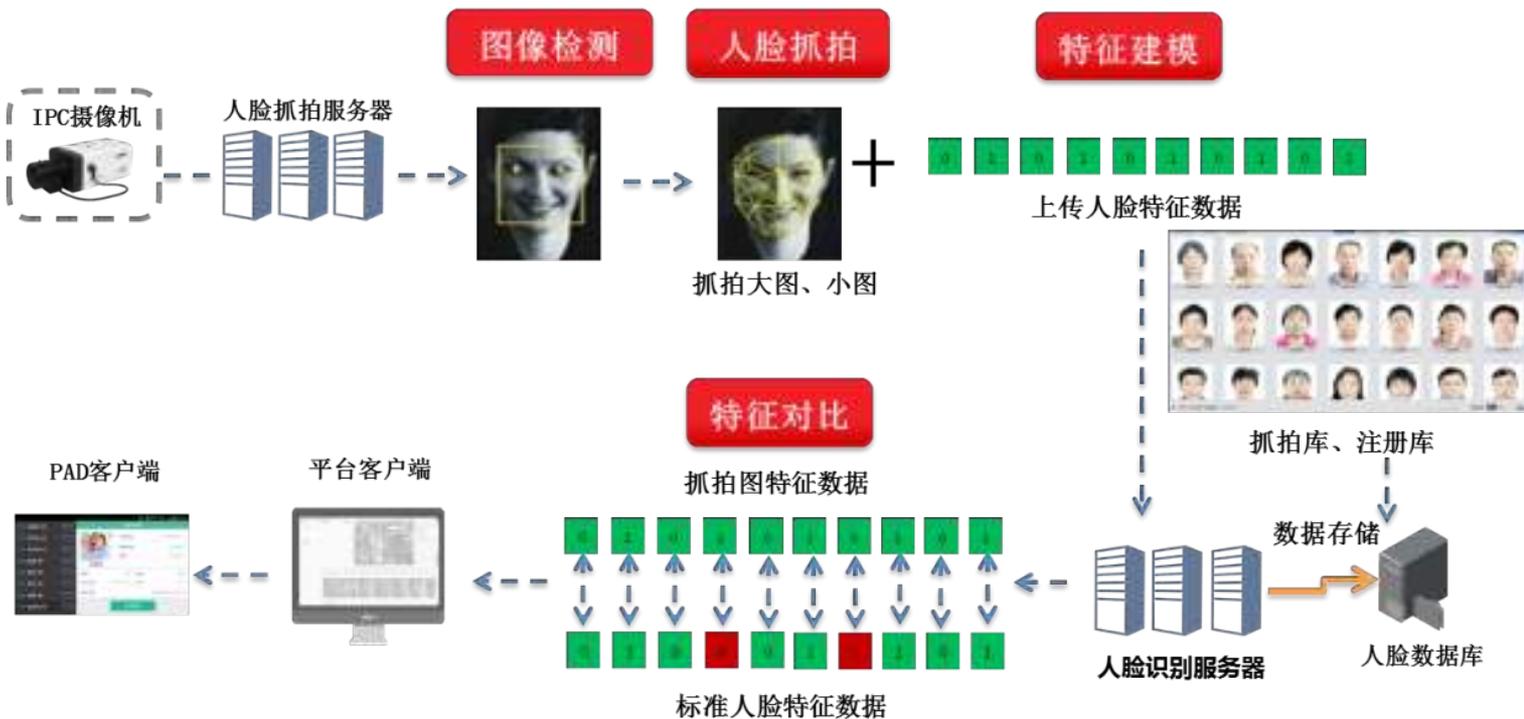
“刷脸”登机、验票、支付都属于1:1的人证核验！

人脸识别(identification)



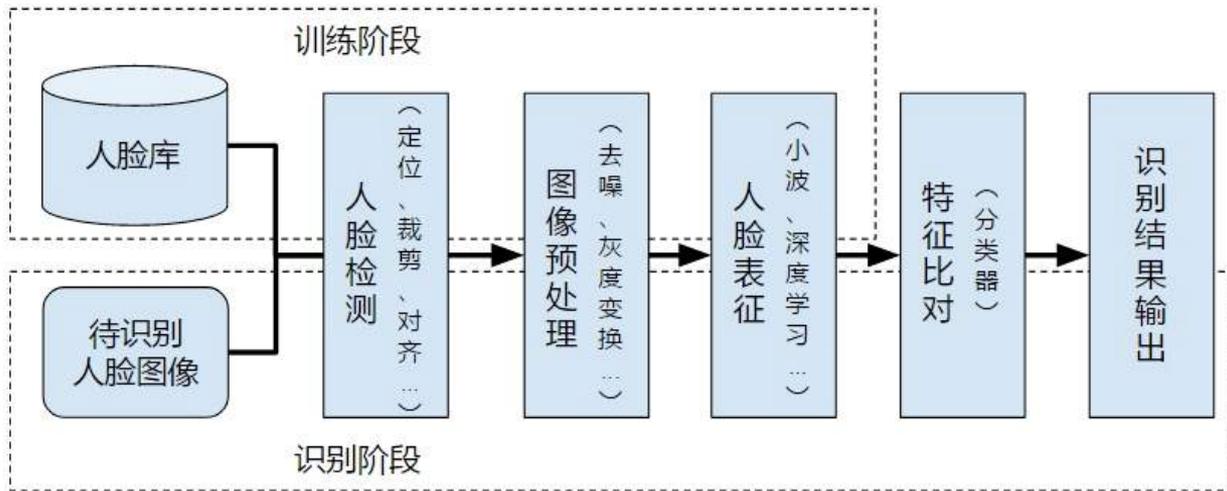
1:N人脸识别，是1:N进行图像匹配比对的过程，从一群人中识别你是谁，一般用于开发“人脸识别门禁系统”和“VIP人脸识别系统”。

人脸识别系统



人脸识别步骤

- 人脸检测与对齐校准
- 人脸特征提取与特征模型建立
- 人脸特征匹配
- 人脸识别结果



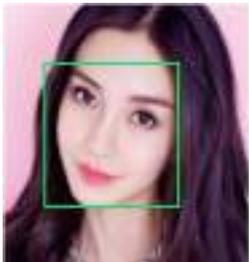
人脸检测与对齐校准



人脸检测与对齐矫正

- 人脸识别第一步
- 人脸**关键点**检测
- 根据参考点，**仿射变换**对齐
- 人脸**水平**，**居中**
- 人脸特征模型输入

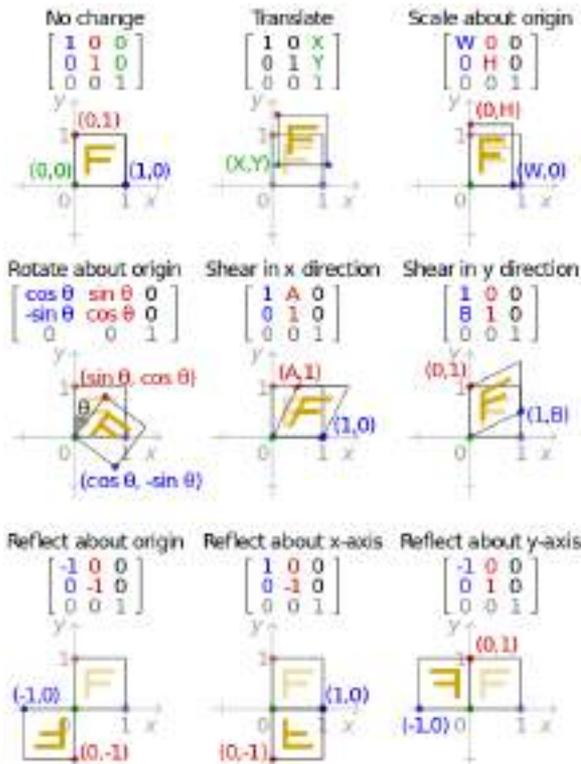
人脸检测



人脸关键点检测



人脸对齐



人脸检测

基于规则/知识方法

人脸模式的变化满足一定的规律，所以可以归纳描述人脸特征的规则，如灰度分布、比例关系、纹理信息等。

基于模板的方法

固定模板法

可变形模板法

基于不变特征的方法，如彩色信息

人脸的肤色在彩色空间中的分布相对比较集中，所以可用来检测和跟踪人脸

基于外观学习的方法---目前的主流方法

将人脸检测视为从非人脸样本中识别人脸样本的模式识别问题，通过对人脸样本集和非人脸样本集的学习产生分类器。

mtcnn人脸检测结果

Fddb人脸检测准确率：95.04%



人脸特征提取与特征模型建立



人脸特征提取

传统算法

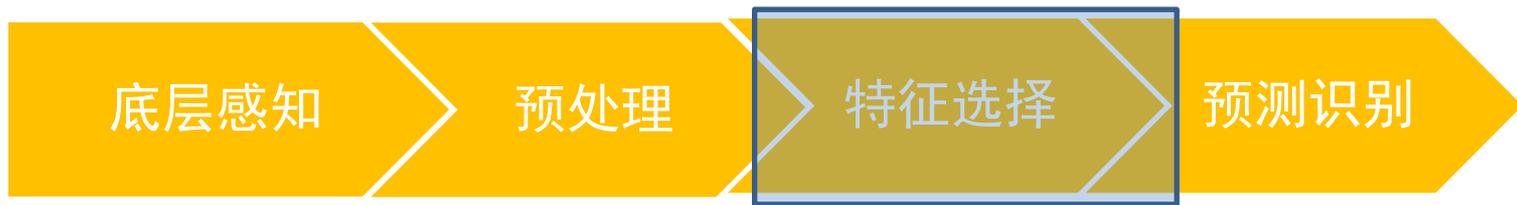
1. 基于**面部器官**的特征提取
2. 基于**模板**的特征提取
3. 基于**代数**方法的特征提取
4. 基于**弹性匹配**法的特征提取

深度学习算法

1. 深度**卷积神经网络**架构
2. 模型训练**自我学习**特征

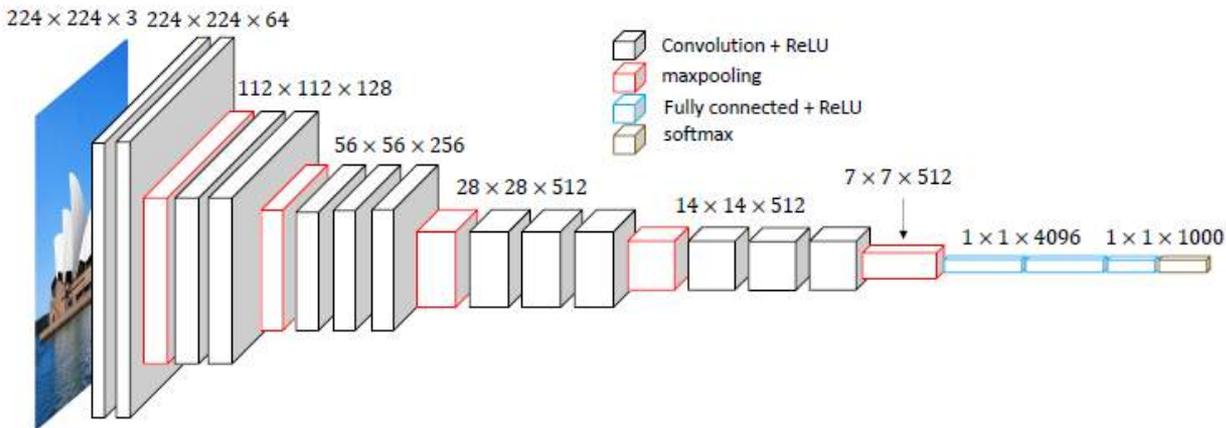


深度学习的优势



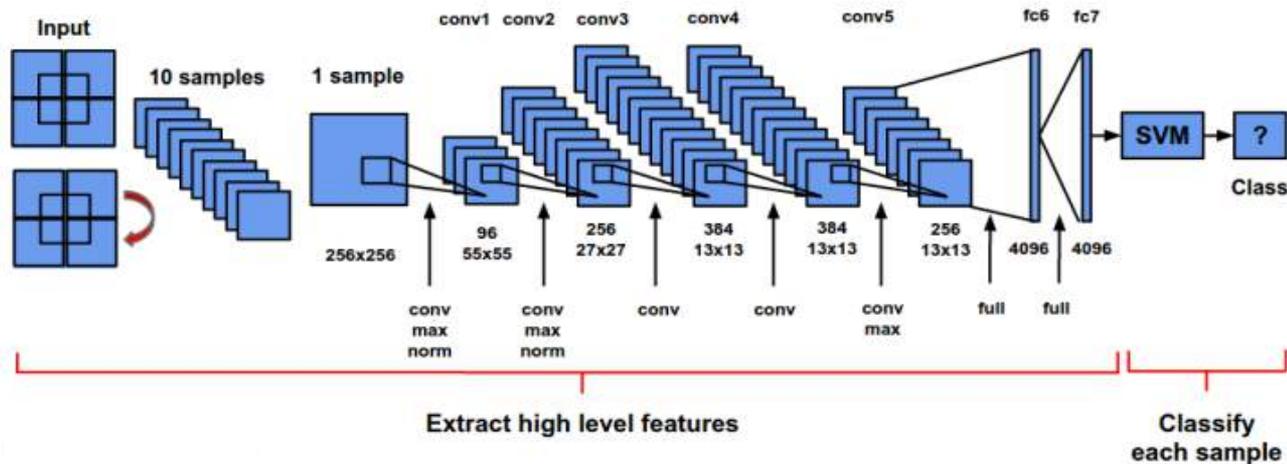
- 良好的**特征表达**，对最终算法的准确性起了非常关键的作用
- 识别系统的计算和测试主要集中在**特征提取**部分
- **特征的样式**目前一般都是人工设计的，靠人工提取特征
- 深度学习的优势：
 - 1.强调了数据的**抽象**
 - 2.强调了特征的**自动学习**
 - 3.自主学习的特征更为**可靠**

人脸特征模型之深度神经网络架构



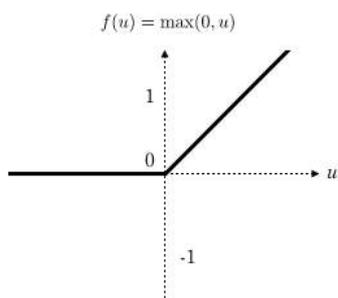
- 卷积层 (Convolutional layer)
- 激活函数(ReLu)
- 池化层 (Pooling layer)
- 全连结层 (fully connected layer)
- 损失函数层 (loss function)

AlexNet

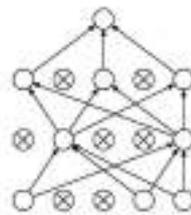
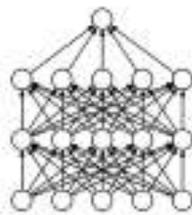


- 2012年 Alex Krizhevsky等人提出，相比11年，精度提升近10个百分点
- 历史性突破
- 8层：5卷积层+3层全连接层
- 卷积核大小：11*11，5*5，3*3

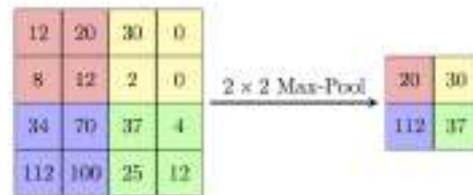
AlexNet



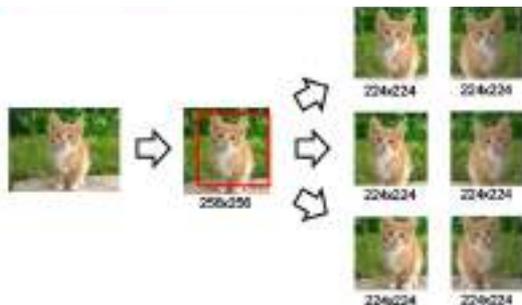
Relu代替Sigmoid



Dropout防止过拟合



Max-pooling 代替 average-pooling

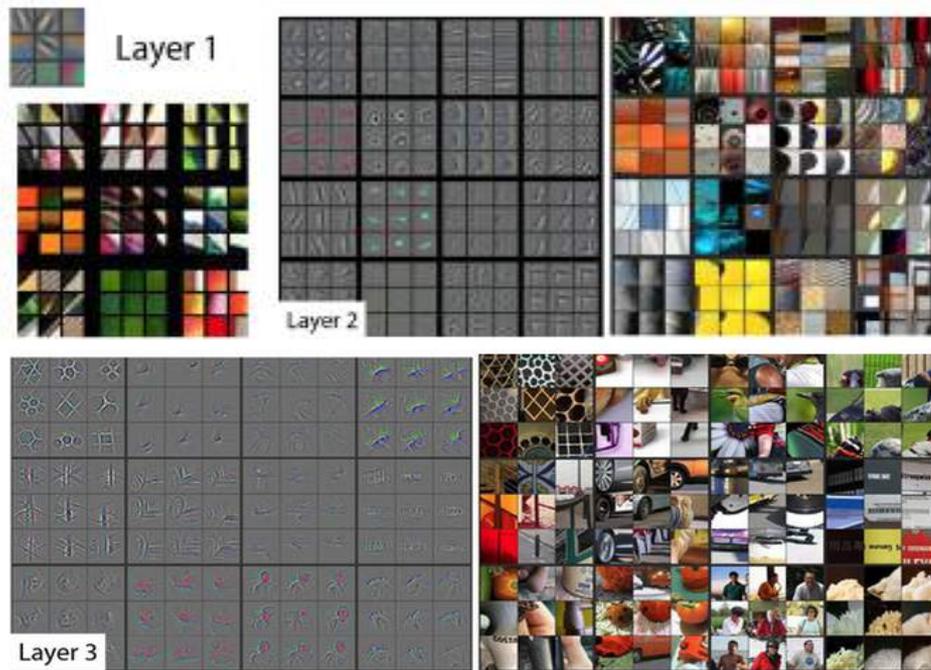


数据增强扩充训练集



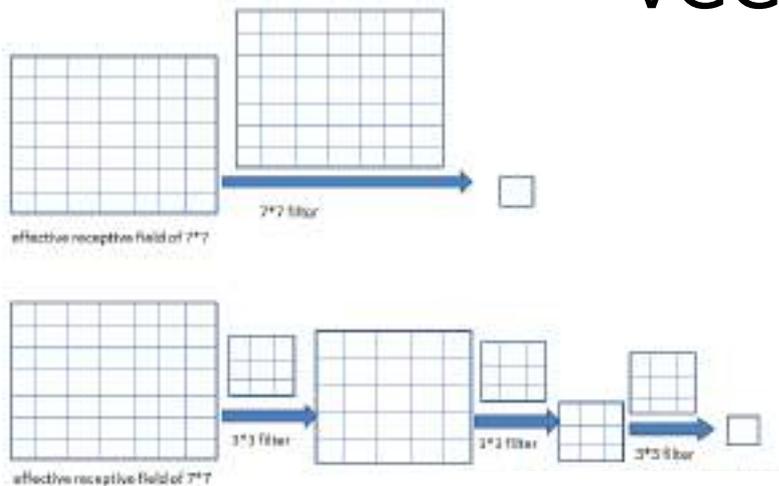
引入GPU，多GPU训练模型，提高速度

神经网络特征图



- 随着网络加深，特征表征能力越强

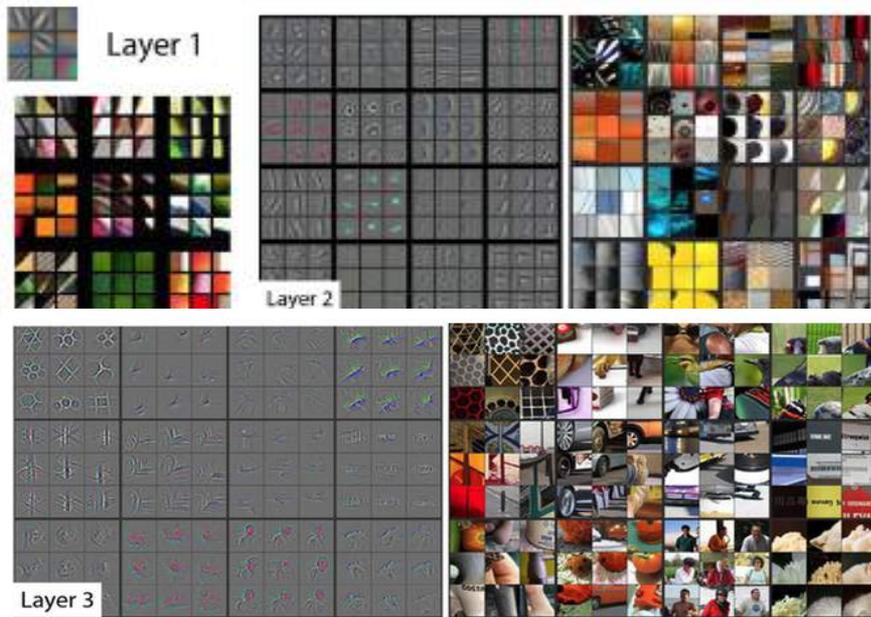
VGGNet



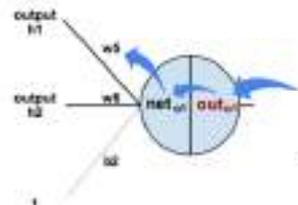
- 2014年亚军，牛津大学提出
- 它的特点也是连续conv多，计算量**巨大**
- 选用3*3卷积核的优点：
 - 使网络结构**更深**，学习到的特征**更多**，结果更具有**判别性** (discriminative)
 - 三层3*3的卷积核比一层7*7的卷积核所需要的参数**更少**

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

如何训练更深层的网络？



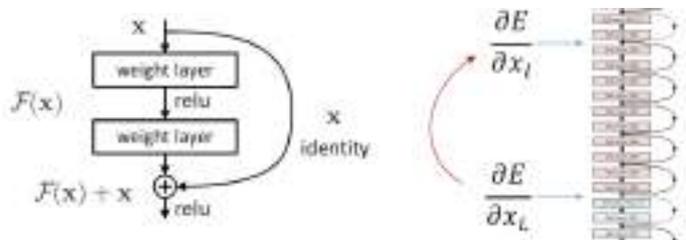
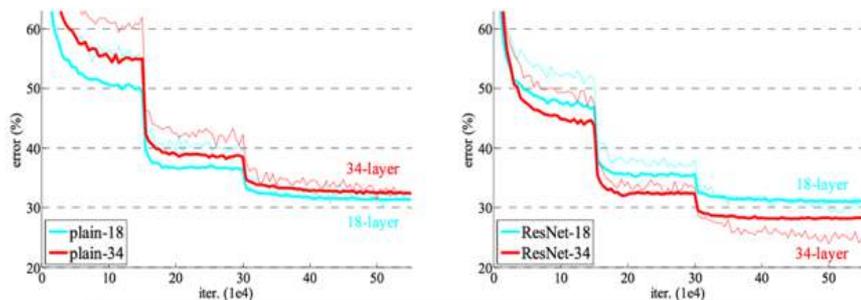
- 随着网络加深，特征表征能力越强
- 梯度计算



$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial w_5}$$

- 简单叠加层数会产生梯度消失

残差神经网络ResNet



- 进一步缓解梯度消失问题
- 梯度由“乘”变“加”

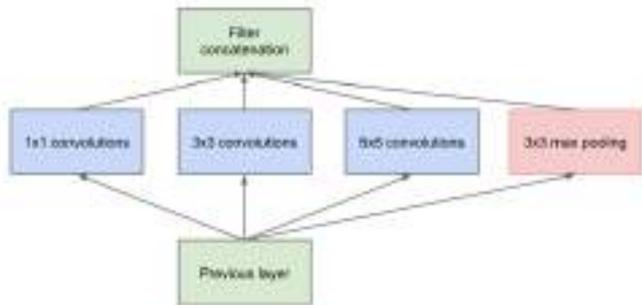
$$x_L = x_l + \sum_{l=i}^{L-1} F(x_l)$$

$$\frac{\partial E}{\partial x_l} = \frac{\partial E}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial E}{\partial x_L} \left(1 + \frac{\partial}{\partial x_l} \sum_{l=1}^{L-1} F(x_l) \right)$$

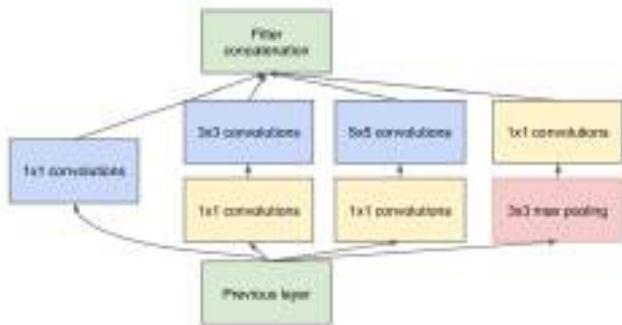
- 加快收敛速度
- 提高准确率



Inception-ResNet



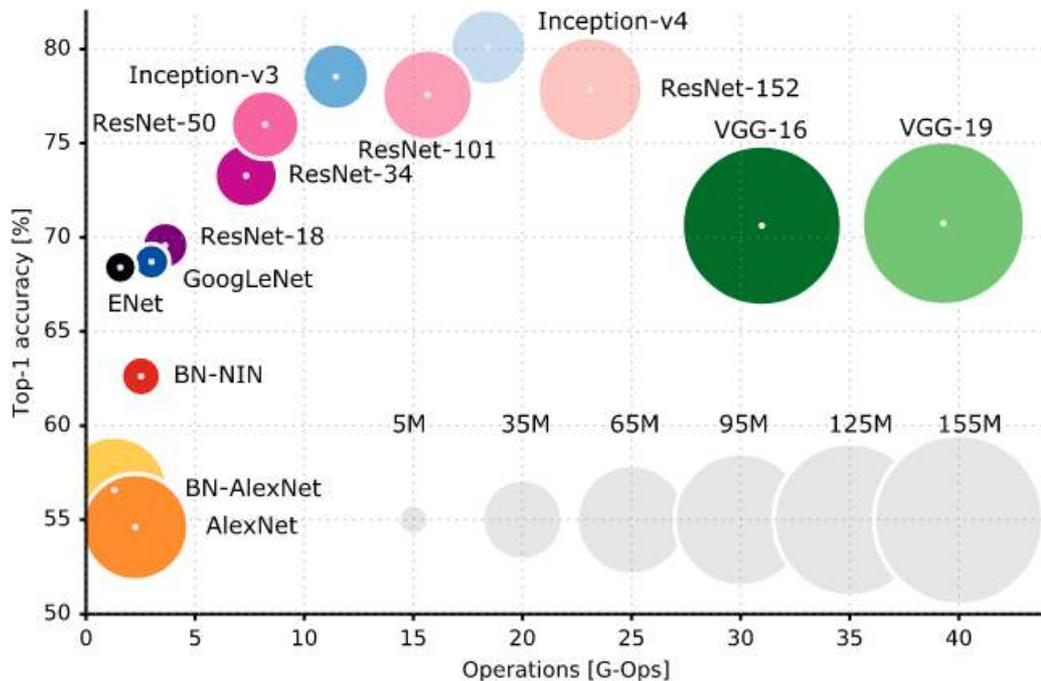
(a) Inception module, naive version



(b) Inception module with dimensionality reduction

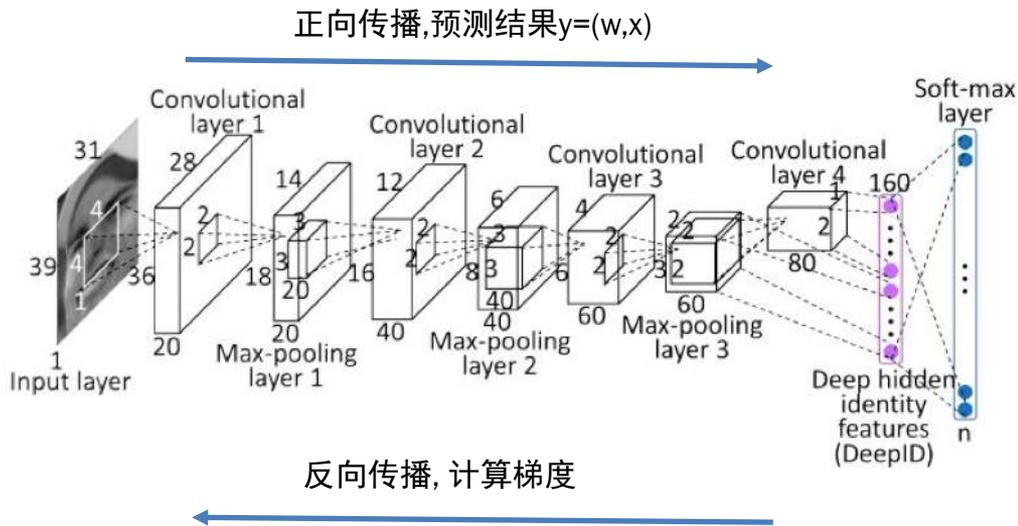
- Inception是GoogleNet中的重要结构
- 高效表达**特征**
- 在同一层中使用**不同大小**的卷积核 3×3 , 5×5
- 提升**不同尺度**图像的检测能力
- 1×1 卷积核**降维**
- Inception结构加入到ResNet中将**多尺度**的特征融合

不同架构对比

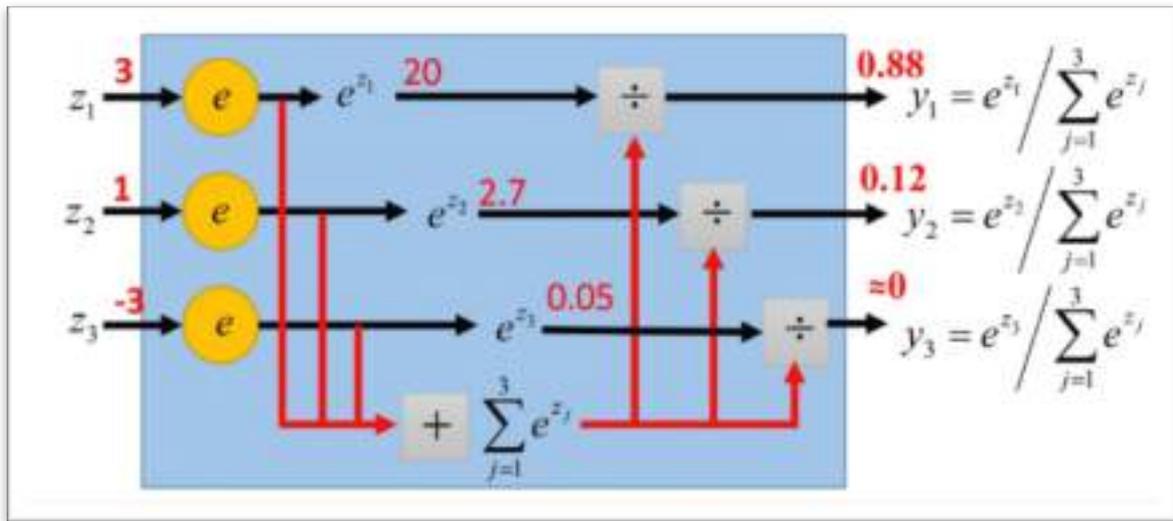


损失函数

- 表示**预测结果**和**真实结果**之间的差异，
- 是需要优化的**目标函数**
- 指导训练：
 - 输入:原始图像，label
 - 输出:神经网络的每层权重值 w
 - 初始化**权重值** w
 - 通过神经网络计算出**预测**结果
 - 计算出**损失函数**
 - 对 w 求导得到梯度，更新 w
 - **梯度下降**的方法得出损失函数最小时的 w



损失函数Softmax



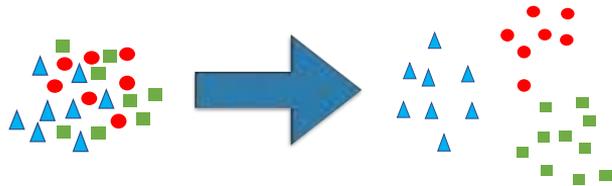
更清晰的描述了预测模型与理想模型的距离

交叉熵：
$$L_s = -\sum_{i=1}^n \log \frac{e^{w_{yi}^T x_i + b_{yi}}}{\sum_{j=1}^n e^{w_j^T x_j + b_j}}$$

特点：不同种类分得比较开

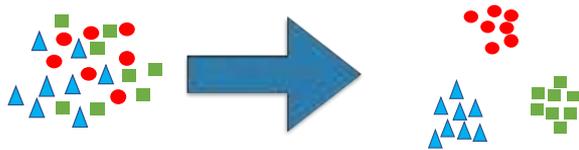
加入center loss的Softmax

- Center loss计算步骤：
 - 1:每个类中心位置 c_{y_i}
 - 2:类元素离所属class中心位置 c_{y_i} 的距离
 - 3:求和得到center loss



- Softmax+center loss

$$L = -\sum_{i=1}^n \log \frac{e^{\mathbf{w}_{y_i}^T \mathbf{x}_i + b_{y_i}}}{\sum_{j=1}^n e^{\mathbf{w}_j^T \mathbf{x}_i + b_j}} + \frac{\lambda}{2} \sum_{i=1}^n \left\| \mathbf{x}_i - \mathbf{c}_{y_i} \right\|_2^2$$



特点：不同种类分的很开，并考虑了同类距离

Angular Softmax

$$L = -\sum_{i=1}^n \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}}$$

$$L = -\sum_{i=1}^n \log \frac{e^{\|W_{y_i}^T\| \|x_i\| \cos(\theta_{y_i, i}) + b_{y_i}}}{\sum_{j=1}^n e^{\|W_j^T\| \|x_i\| \cos(\theta_{j, i}) + b_j}}$$

$$L_{\text{modified}} = -\sum_{i=1}^n \log \frac{e^{\|x_i\| \cos(\theta_{y_i, i})}}{\sum_{j=1}^n e^{\|x_i\| \cos(\theta_{j, i})}}$$

$$L_{\text{ang}} = -\sum_{i=1}^n \log \frac{e^{\|x_i\| \cos(m\theta_{y_i, i})}}{e^{\|x_i\| \cos(m\theta_{y_i, i})} + \sum_{j \neq y_i} e^{\|x_i\| \cos(\theta_{j, i})}}$$

二分化特例：

- 1: 归一化权重向量 w_1, w_2 , bias=0
- 2: 计算特征向量和权重向量的夹角 θ_1, θ_2
- 3: 通过夹角判断该特征归属

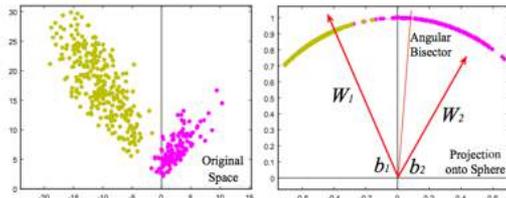
Modified Softmax 决策边界:

$$\|x\|(\cos \theta_1 - \cos \theta_2) = 0$$

Angular Softmax 决策边界:

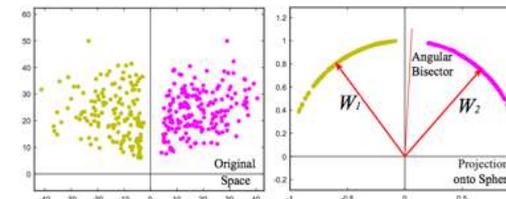
$$\|x\|(\cos m\theta_1 - \cos \theta_2) = 0 \text{ for class 1}$$

$$\|x\|(\cos \theta_1 - \cos m\theta_2) = 0 \text{ for class 2}$$



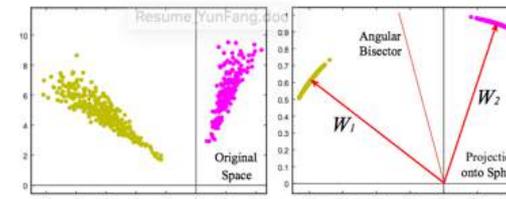
(a) Original Softmax Loss

(b) Original Softmax Loss



(c) Modified Softmax Loss

(d) Modified Softmax Loss



(e) A-Softmax Loss

(f) A-Softmax Loss

人脸特征匹配



特征匹配

- 相似度计算

L2 norm 欧式距离

$$sim = d_{XY} = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$$

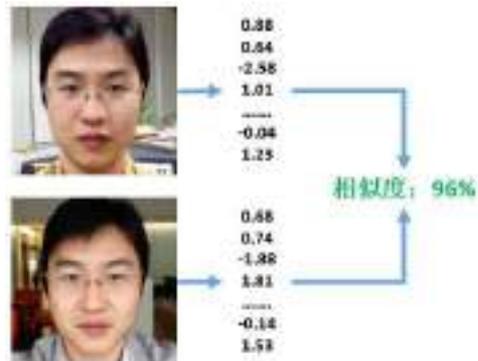
Cosine 相似度

$$sim = \cos(\theta) = \frac{X \cdot Y}{\|X\| \|Y\|} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \times \sqrt{\sum_{i=1}^n y_i^2}}$$

- 相似度排序

L2 norm 欧氏距离 从小到大

Cosine 相似度 从大到小



人脸识别结果



训练集预清理

- 训练集：MS-Celeb-1M 1049 W图片，9W人，存在大量错误的图片
- 预清理(clean list)

错误类型	清理方式
分类错误，有对应类别	分到对应类别
分类错误，无对应类别	扔掉
没有人脸	扔掉
人脸被大面积遮挡	扔掉

- 过滤器(filter)

测试集 (lfw)

➤ lfw公共测试集：最流行的人脸测试集

- 数据量小
- 识别是不是同一个人
- 6000 对人脸 3000+ 3000-

同一人



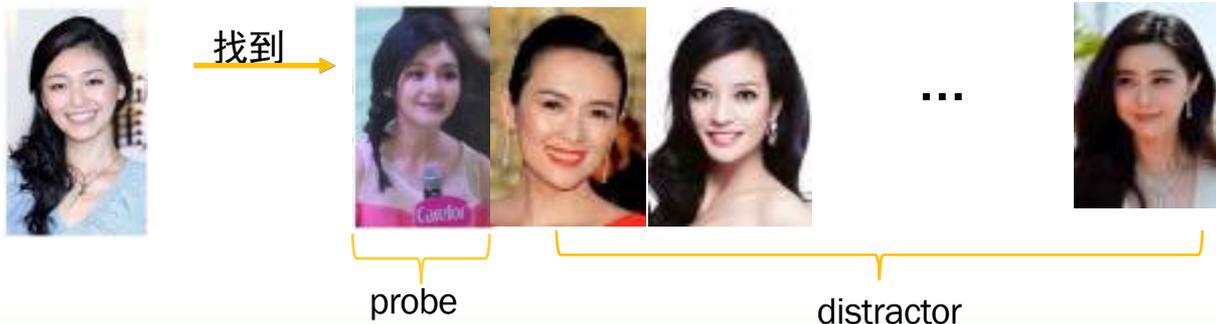
不同人



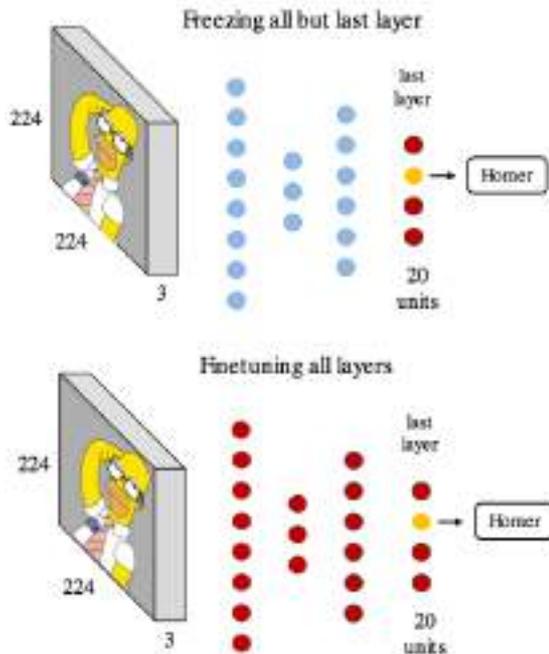
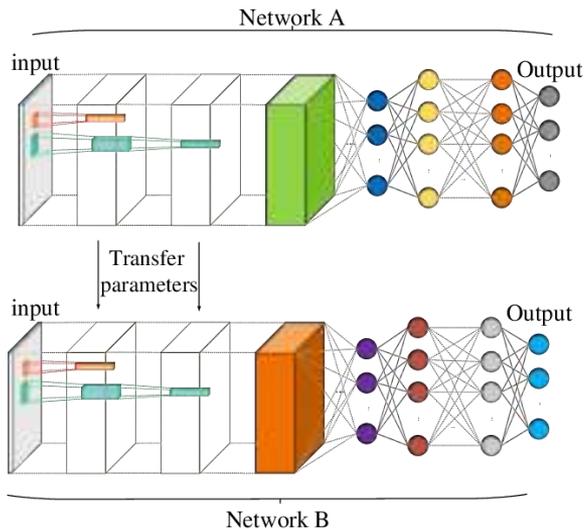
测试集 (Megaface)

➤ MegaFace : 近两年的人脸测试数据集

- 数据量大
- 在distractor图片中识别probe图片
- probe : facescrub数据集中选取80个名人脸 (3530张人脸图片)
- distractor: 大小 10, 100, 1000, 10000, 100000, 1000000(百万)



参数微调



- 参数微调
 - MS-Celeb-1M 包含各个种族
 - 我们主要针对亚洲人脸
 - 特征相似，但仍有区别
 - 将预训练的模型进行微调
 - 亚洲人脸特征

- 如何微调
 - 微调分类器参数，freeze 其他架构中的权重
 - 微调整个架构中的参数

实验结果 (LFW)

公司	lfw测试结果
北京飞搜科技	99.40%(2016)
云图睿视	99.43%(2016)
Face++	99.50%(2016)
商汤	99.55%(2016)
颜鉴	99.64%(2016)
百度	99.77%(2016)
腾讯	99.80%(2017)
平安科技	99.83%(2017)
苏宁	99.83%(现在)

Identification Rate vs. Distractors Size

Algorithm	Date Submitted	Set 1	Set 2	Set 3	Data Set Size
Orion Star Technology (Intern)	3/21/2018	88.330%			Large
IBUG, Deepnight	2/6/2018	88.067%	86.038%	85.050%	Large
SiningUS AILab	3/21/2018	86.8318947%	86.2618947%	85.2618947%	Large
IntelVisual	2/11/2018	83.120%	83.532%	83.136%	Large
ULBee - Face Team	3/21/2018	82.120%			Large
Vocord - deepVc V2	84/21/2017	81.768%	81.711%	81.704%	Large
MTDP_ITC	12/21/2017	87.889%	85.877%	87.184%	Large
TUPUTECH	12/23/2017	88.550%	86.957%	86.819%	Large
Video++	1/5/2018	85.74%	85.757%	85.706%	Large
THU CV-6 Lab	10/13/2017	84.521%	84.513%	84.514%	Large
YeeceerAILab_FaceCNN_v1	8/21/2017	84.291%	84.265%	84.257%	Large
Orion Star Technology (in cloud)	3/21/2018	83.889%			Large
Yeeceer Lab (Tencent Best-Image)	84/08/2017	83.22%	83.267%	83.268%	Large
OrionAI Tech	12/15/2017	83.184%	83.177%	83.144%	Large
FaceTag V1	12/18/2017	82.411%	82.376%	82.364%	Large
Argus_v1	2/5/2018	81.437%	81.624%	81.638%	Large
King Sun	84/05/2017	81.336%	81.266%	81.264%	Large
DeepDress V2	3/25/2017	81.289%	81.268%	81.268%	Large
IBUG (Reported by Author)	84/26/2017	80.277%			Small
Vocord-deepVc V2	12/1/2018	80.252%	80.155%	80.241%	Large
Facestar Lab	12/18/2017	79.426%	79.367%	79.356%	Large

- MegaFace 挑战赛I
- Distractor 大小为
一百万
- rank1的准确率
- 排名第三

参数微调结果

- 训练集（爬网数据60万）
 - 选择28182个人(类)
 - 条件：人脸照片>5张
- 测试集（门店数据）
 - 6000对 3000+ 3000-

模型	准确率
一类单模型	94.85%
参数微调模型	98.76%

总结

- 人脸识别背景介绍
- Mtcnn人脸检测
- 尝试了不同深度学习架构和不同损失函数进行模型训练
- 对微软训练数据进行了清理，进一步提高了模型预测准确率
- 人脸特征模型在公共数据集上得到不错的结果
- 爬网数据进行参数微调，在门店数据集进行测试



关注QCon微信公众号，
获得更多干货！

Thanks!



主办方 **Geekbang** & **InfoQ**
极客邦科技

Q&A

邮箱：doris.wang@ussuning.com

苏宁招聘

部门： 020研发中心数据洞察产品部

职位： 产品经理，数据分析师

联系人： 杜海峰经理

邮箱： 14070348@cnsuning.com