

基于图像/视频的人脸和人体分析基础技术及其应用介绍

- 姓名：刘玉明
- 公司：云从科技
- Title：资深算法专家

百万AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

云从图像/视频技术的应用领域

云从科技作为国内AI四小龙之一，其图像/视频技术主要应用在金融、安防、交通、商业等领域，通过行业领先的人工智能、认知计算与大数据技术形成的整合解决方案。



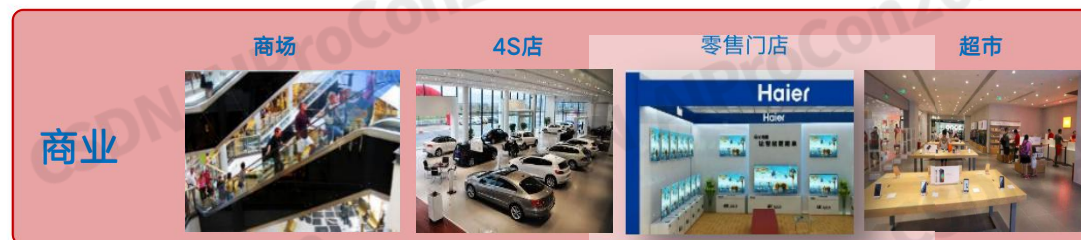
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

云从图像/视频技术的应用领域



百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

云从人脸和人体分析技术

人脸识别

人脸三维重建

动作识别

活体检测

人脸属性解析

行人属性解析

表情识别

人体ReID

客流统计

模型训练平台

模型发布引擎

底层平台：CWOS

百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

云从人脸和人体分析技术

人脸识别

人脸三维重建

动作识别

活体检测

人脸属性解析

行人属性解析

表情识别

人体ReID

客流统计

模型训练平台

模型发布引擎

底层平台：CWOS

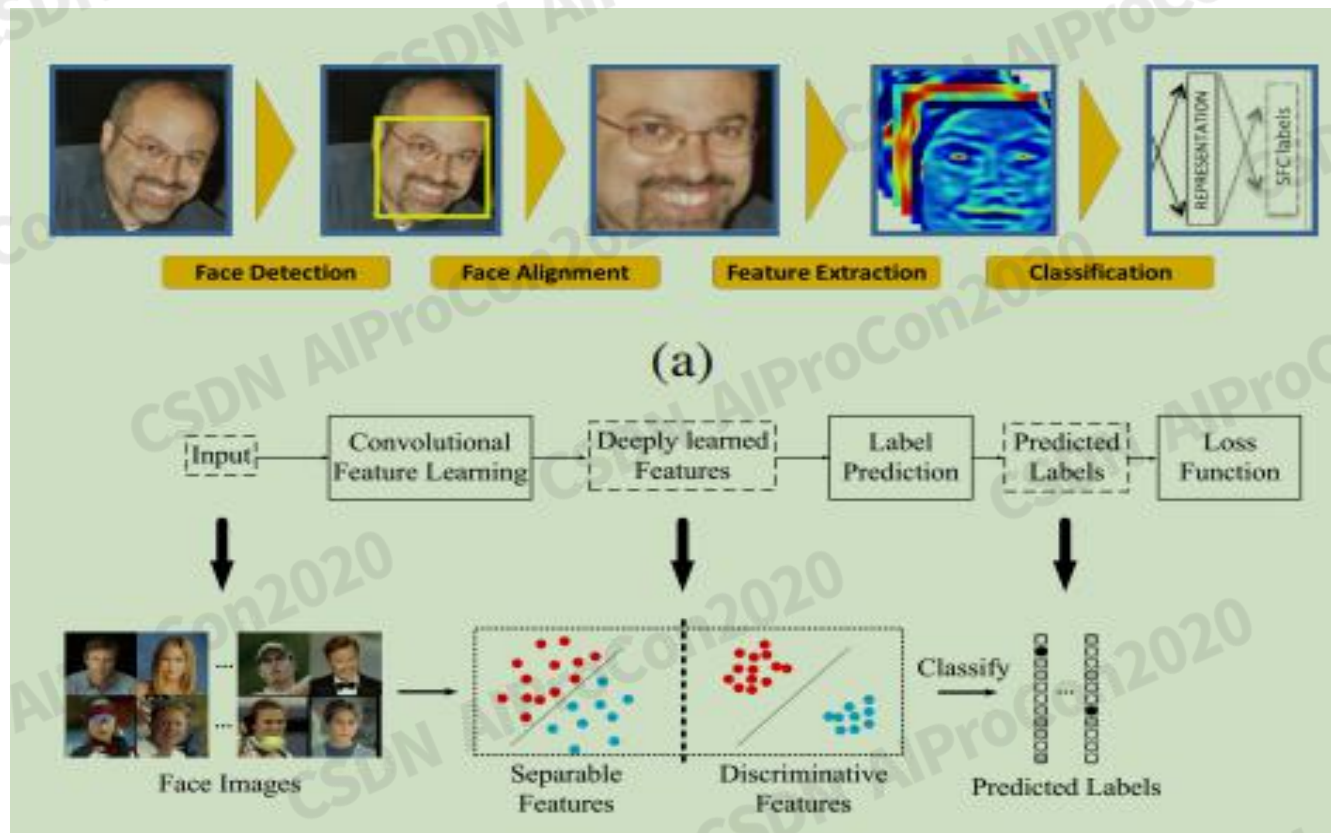
百万AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的一般流程



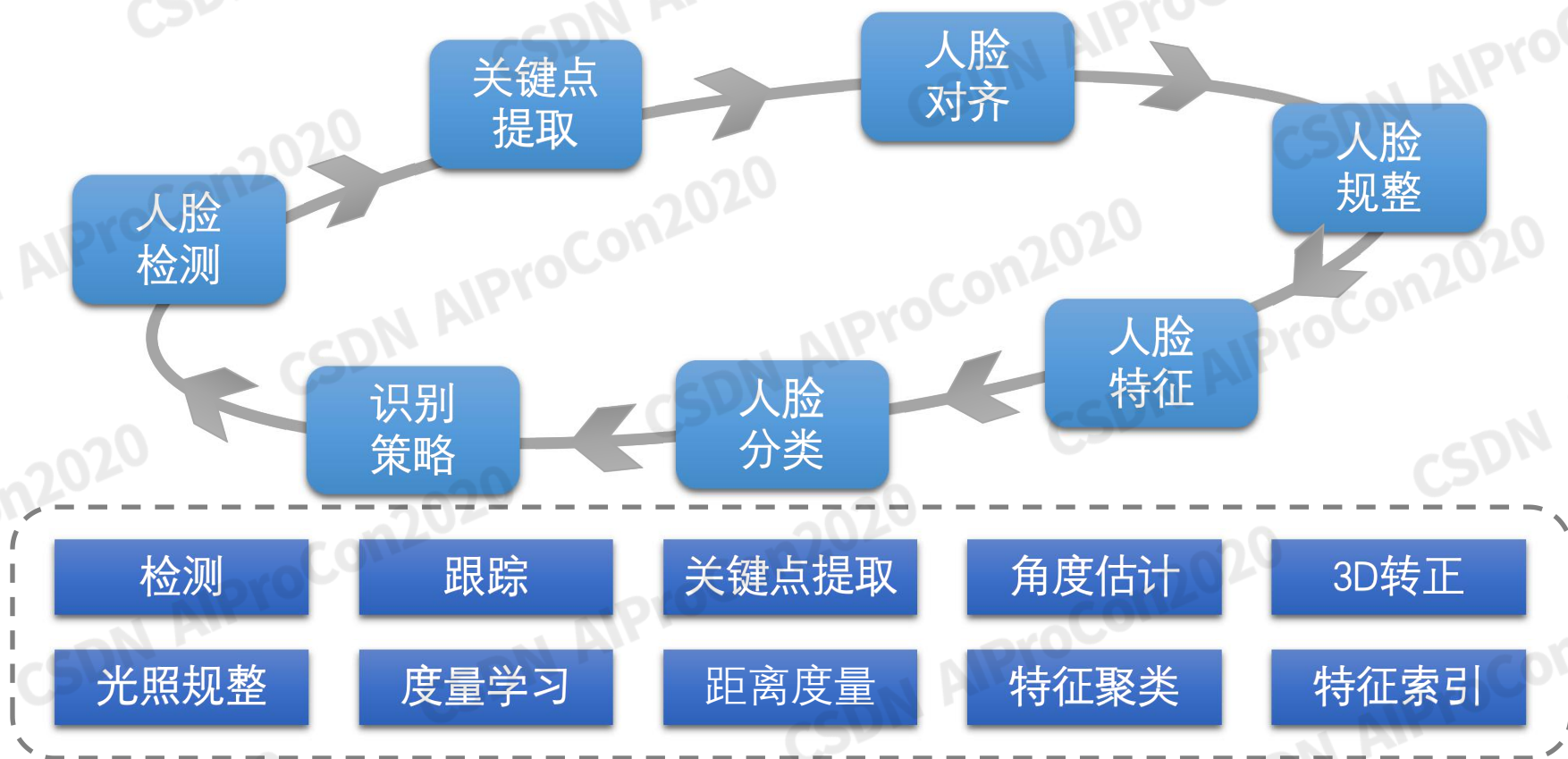
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的算法模块



百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的难点

影响人脸图像表现的因素

人与摄像设备的位置关系（距离角度等）

光照环境条件

摄像设备

图像存储质量

年龄变化

意外损伤



饰物（眼镜帽子等）

化妆、整容

精神状态

健康状况

面部毛发(头发, 胡须)

- 成像条件复杂性
- 不同个体相似性
- 同一个体易变性

百万
AI

2020
ProCon
万人开发者大会

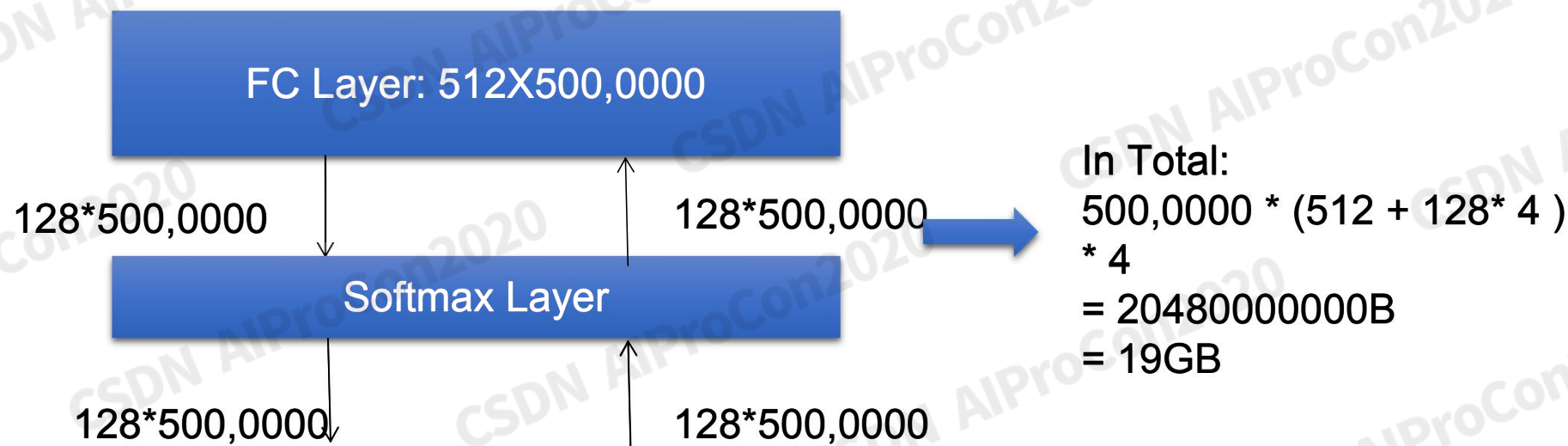
7月3日-4日

CSDN

人脸识别的难点

- 海量数据训练

- academia scale: ~4-5百万 images, ~1-10万 identities
- industry scale: >2000万 images, >500万 identities



7月3日-4日

CSDN

大规模数据采集



大规模数据采集阵列

- 可移动、可拆卸
- 多角度、毫秒级同步
- 现场实景（大厅、广场等）拍摄



联合实验室

- 银行、公安、民航、政通
- 实际应用场景
- 亿级数据量

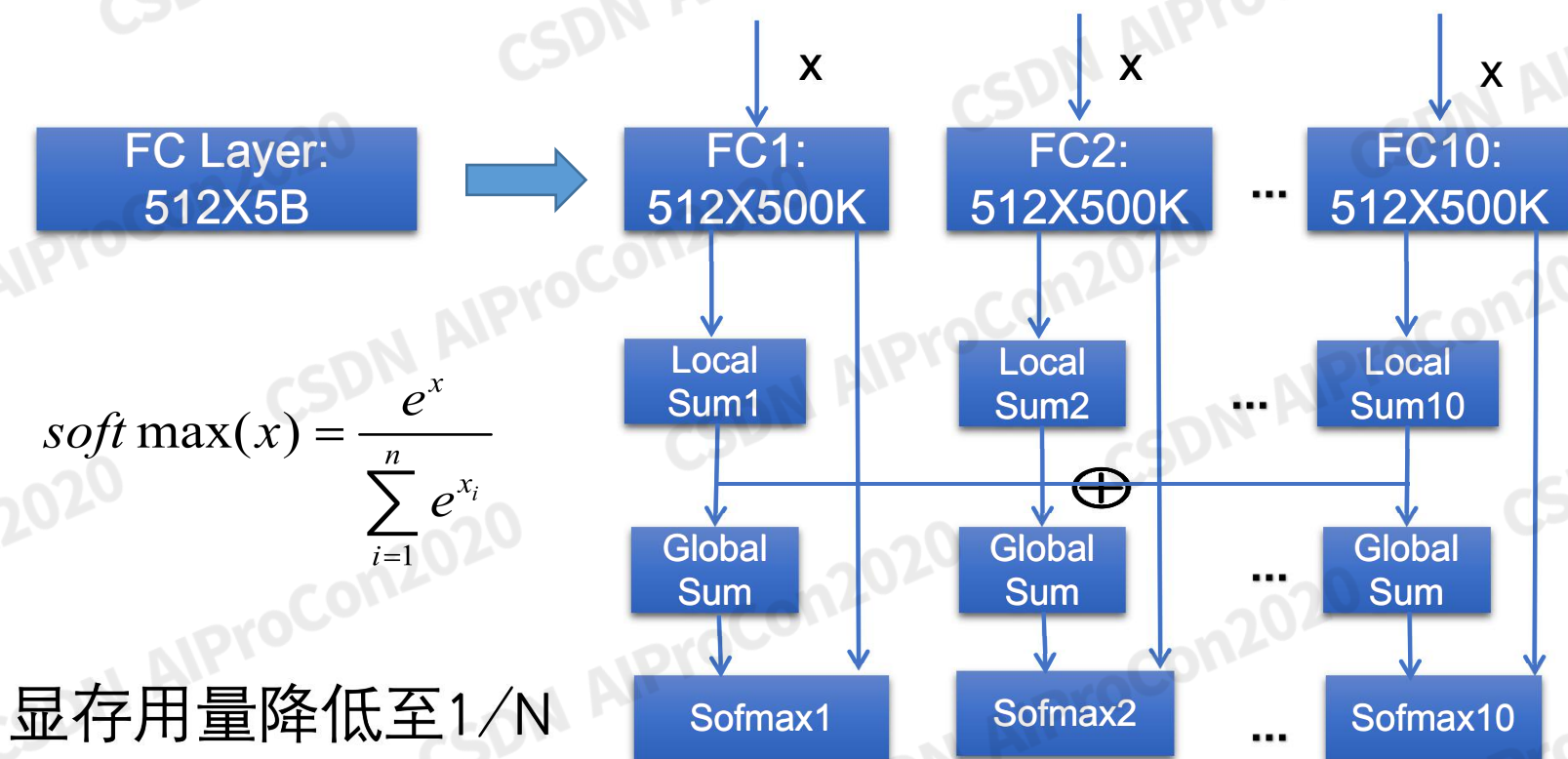
百万
学AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

大规模训练-FC层并行



$$\text{soft max}(x) = \frac{e^x}{\sum_{i=1}^n e^{x_i}}$$

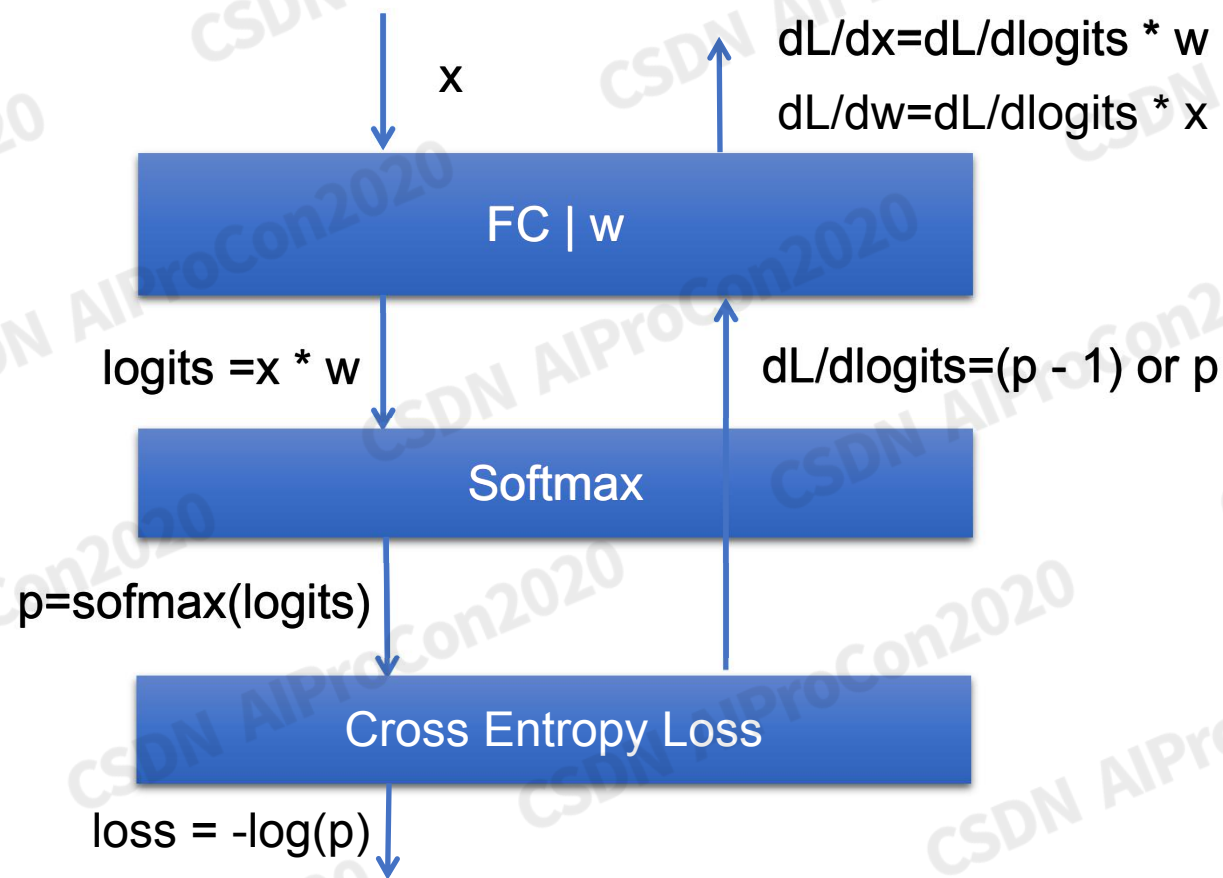
显存用量降低至1/N



7月3日-4日

CSDN

大规模训练-显存复用



➤ 复用方式:

- 得到 p 后, logits不再有用
- 得到 $dL/d\text{logits}$ 后, p 不再有用
- $dL/d\text{logits}$ 显存复用

➤ 收益:

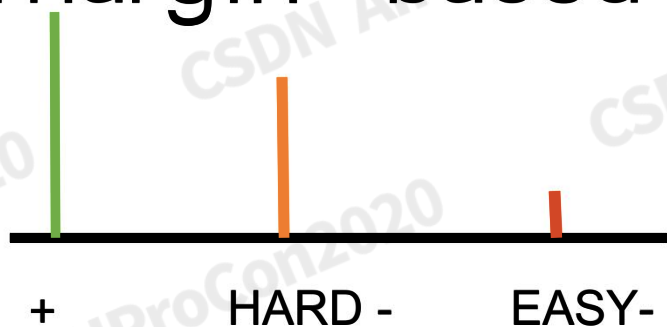
- 用量进一步减少至约 $1/3$



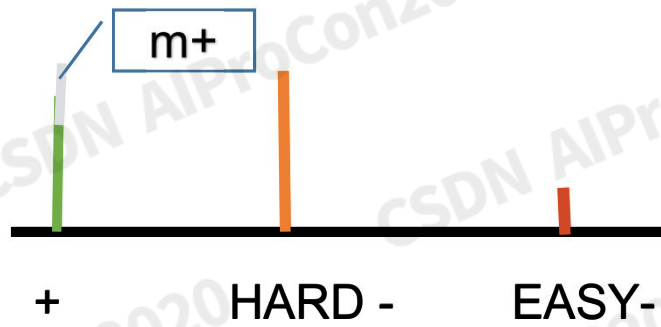
7月3日-4日

CSDN

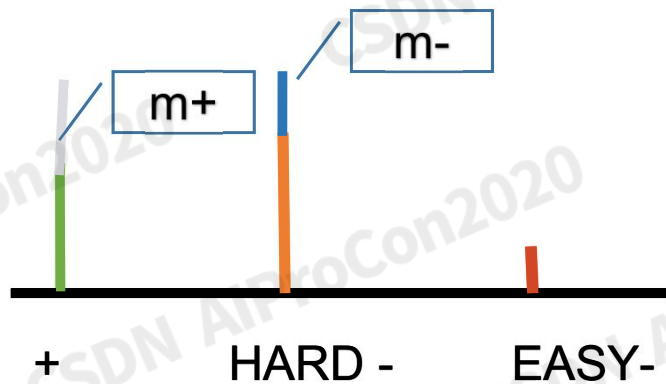
margin-based softmax



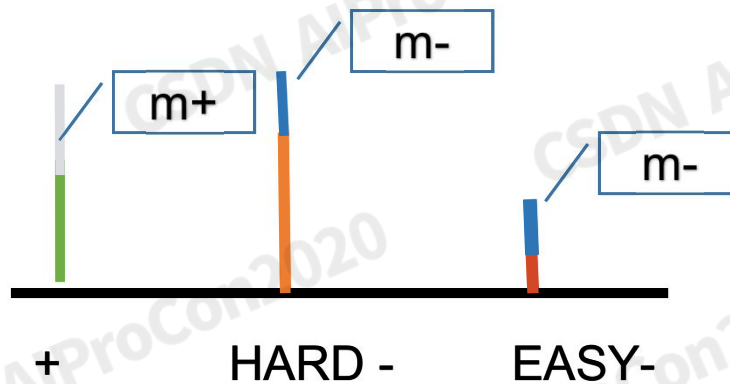
Original softmax



Arc softmax(state-of-the-art)



Support vector softmax



Ours



7月3日-4日

CSDN

人脸识别的应用



is

- 认证合一
- 人脸解锁
- 人脸确认

人脸1: 1



who

- 人脸考勤
- 人脸布控
- 人脸检索
- VIP识别

人脸1: N

百万
AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的应用-智慧金融



- 刷脸门禁
- 刷脸闸机
- 刷脸取款

百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的应用-智慧金融



刷脸支付



刷脸登录



二次身份核验

百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的应用-智慧机场



场景一：一张脸走遍机场，全流程、自服务、新体验



百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的应用-智慧门店

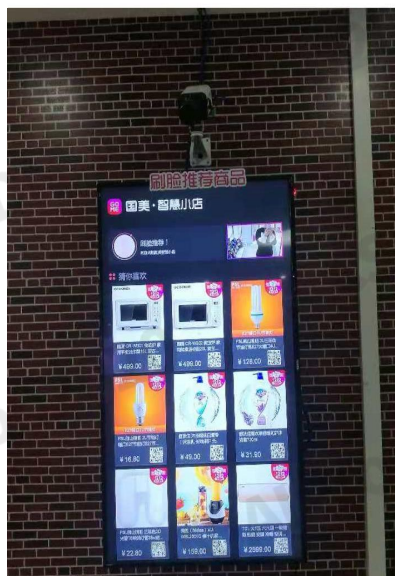
感应人员属性、身份
(人脸、ReID、语音)



后台分类推荐
(推荐算法)



前端呈现交互
(视觉语音AR)



百万
学AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

人脸识别的应用-智慧门店



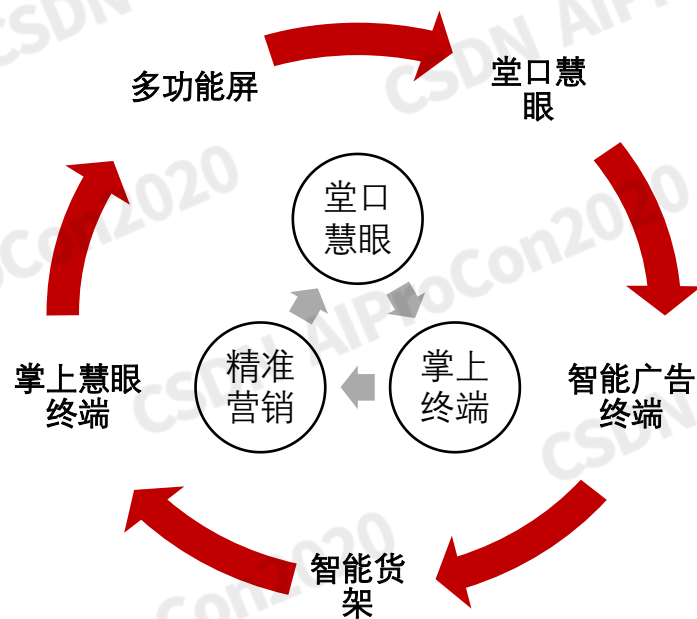
堂口慧眼：通过人脸识别技术针对进出客户进行人脸识别、熟客分析识别、客流统计（人数、人次、性别、年龄进行统计）

智能广告机：属性分析，感知客户年龄、性别、心情等属性标签并根据客户标签推荐商品

智能货架：顾客驻留时间、人货对应数据采集。

智慧终端：将VIP识别、熟客识别信息推送至营销人员前端，给出具体营销方案及话术。

多功能辅屏：结账过程中，将顾客人脸与购物信息进行关联，实现交易信息脸谱化。补全会员信息。



百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

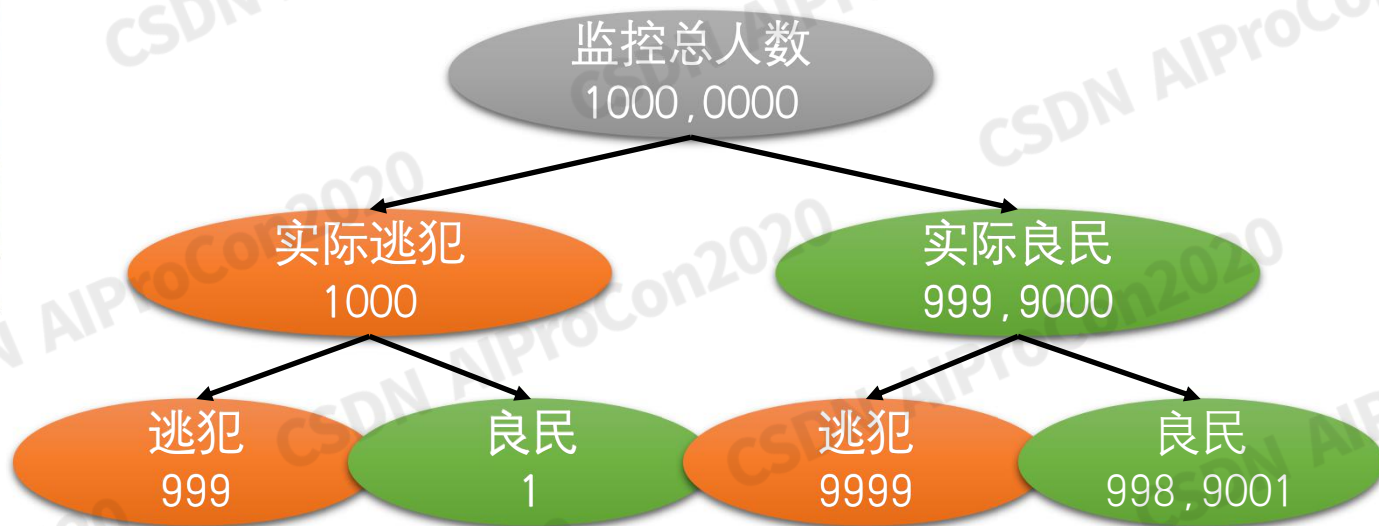
极高准确率的意义



人脸识别抓逃犯



说好的99.9%呢?



警察:

- 平均出警10次还抓不到一个, You out!



技术员:

- 逃犯比例太低咧
- 迭代模型, 再加3个9



百万学AI

2020 ProCon
万人开发者大会

7月3日-4日

CSDN

云从人脸和人体分析技术

人脸识别

人脸三维重建

动作识别

活体检测

人脸属性解析

行人属性解析

表情识别

人体ReID

客流统计

模型训练平台

模型发布引擎

底层平台：CWOS

百万AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

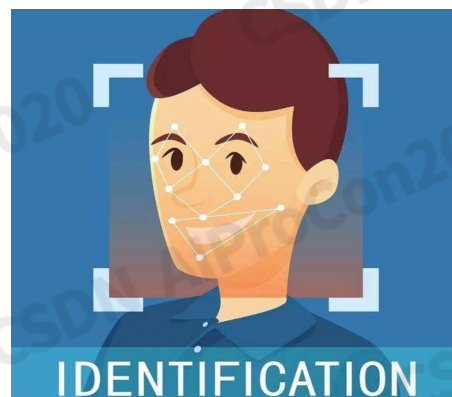
活体检测方式

配合式



用户体验不佳，验证时间长

非配合式



易受到使用场景和
摄像头硬件性能的影响

百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

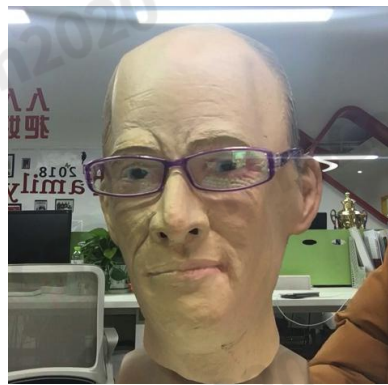
典型的Hack攻击



面具



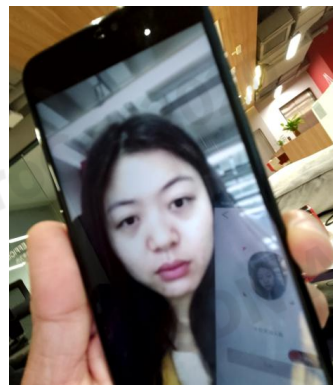
3D生成



头模



照片



视频

百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

活体检测技术

动作活体

要求用户做出随机给出的转头、眨眼、张嘴等面部动作，算法判断是否跟要求的动作一致

TOF活体

基于TOF模组，获取可见光图像、近红外图像和深度图像，算法区分真人和道具

红外活体

基于红外双目摄像头获取可见光和红外图像，算法通过两种图像进行学习，区分真人和攻击道具

3D结构光活体

基于结构光摄像头，获取可见光图像、近红外图像和深度图像，算法区分真人和道具



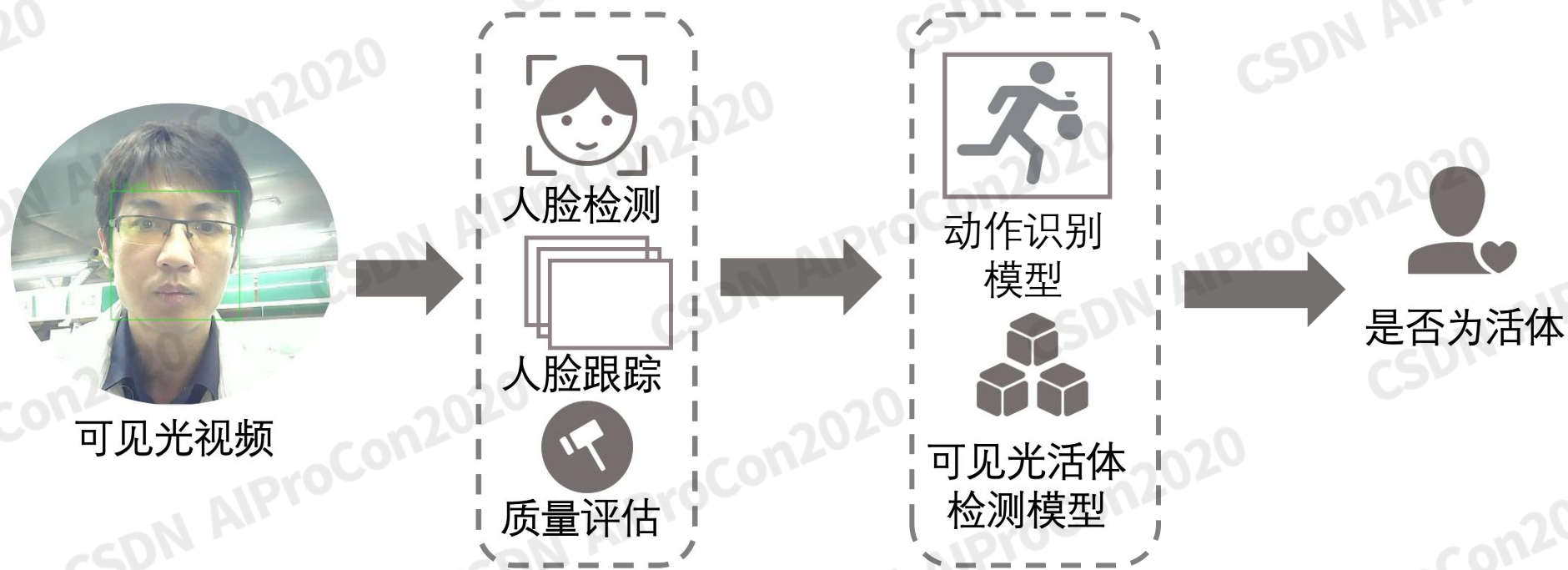
百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

动作活体检测流程



百万AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

红外双目的技术原理

红外
双目
相机



屏幕
红外
图像



主要技术原理：

- 红外光谱反射率受介质影响原理
- 双目深度信息测量原理
- 深度学习提取红外人脸特征进行分类
- 视频图像人脸纹理、边缘、角点检测

技术特点：

- 红外双目天然防止了视频、平面介质等攻击手段
- 体验方便，无需动作
- 受光线影响小，场景适用广

百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

红外双目活体检测流程



可见光图像



红外光图像



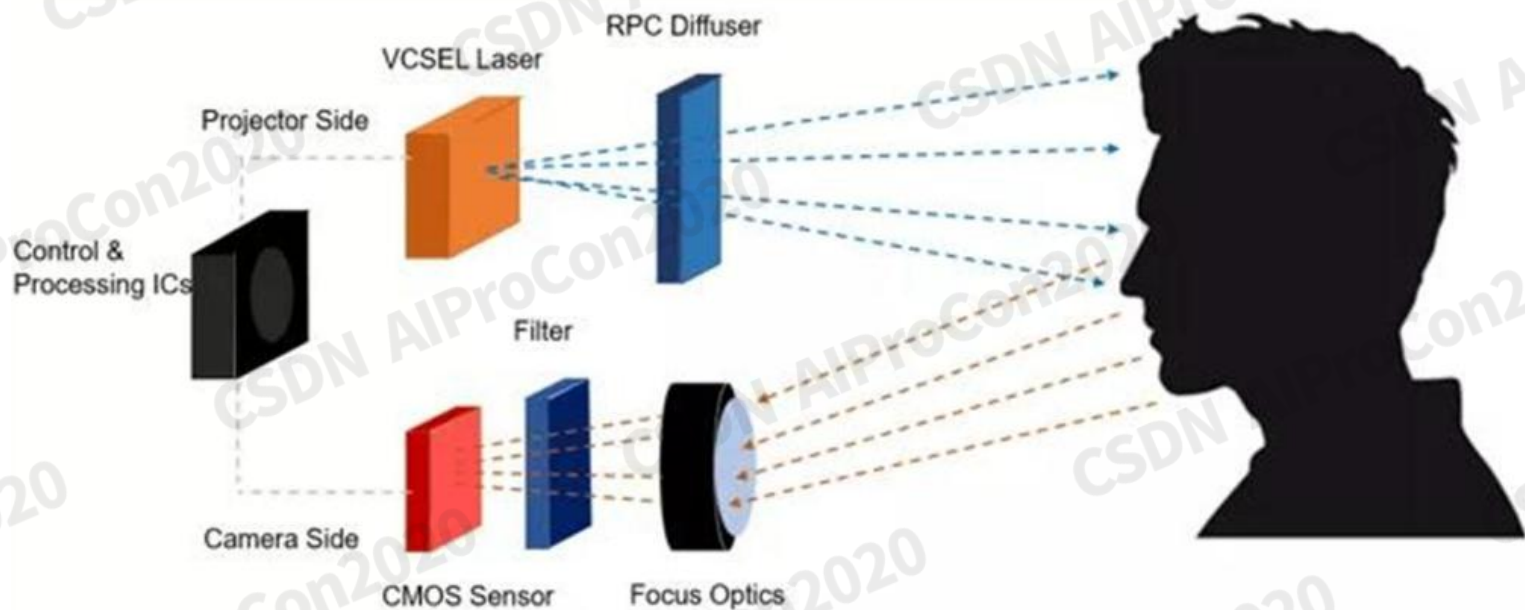
百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

TOF技术原理



TOF模组包含发射模块、接收模块。发射模块向场景发射红外光，经过物体反射后，接收模块接收到反射的红外光计算深度信息。根据激光往返的时间长短和其固定的飞行速度，计算出物体表面上的点与TOF相机之间距离。

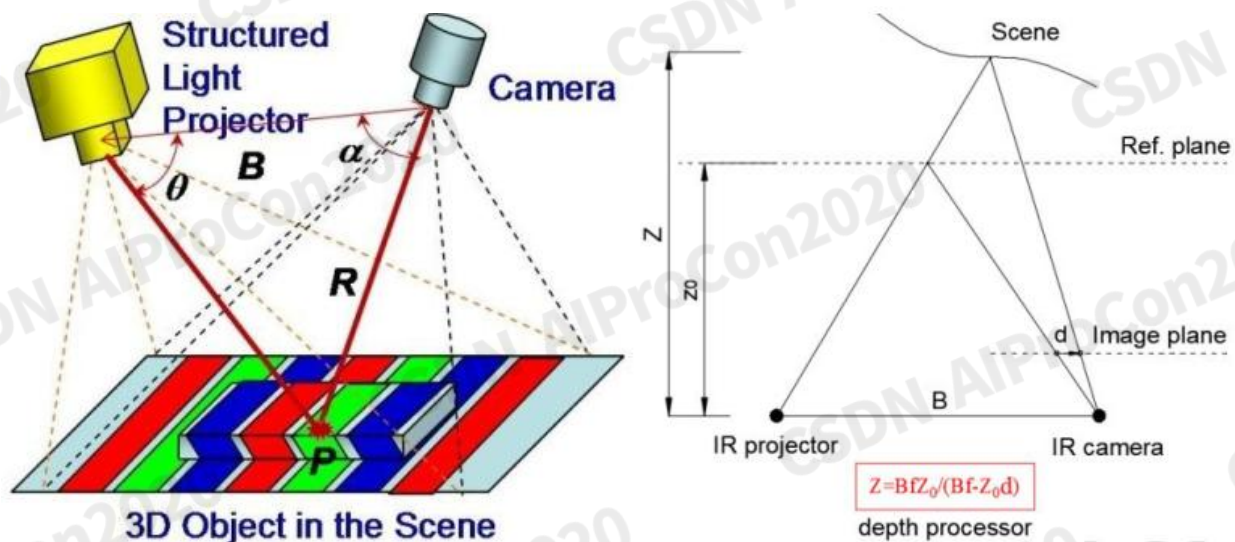
AI

2020
ProCon
万人开发者大会

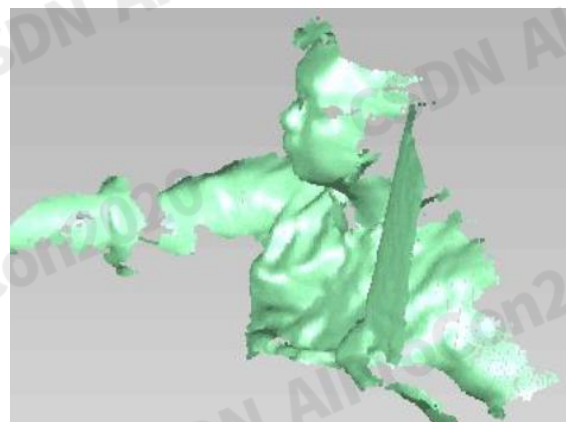
7月3日-4日

CSDN

3D结构光原理



结构光摄像头包含结构光图案发射器，能够向场景投射出固定的结构光图案（散斑图案），并通过红外相机捕捉目标的红外结构光图案。深度计算处理器接受红外结构光图像后执行深度计算算法并输出目标场景的深度图像。



AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

TOF VS 3D结构光

特性	3D结构光	时差测距技术
基本原理	散斑结构光	飞行时间
激光光源	散斑	均匀面光源
分辨率	中	低
速度	中	快
抗光照	低	中
模块大小	大	大
硬件成本	高	高
工作距离	0.2m~1.2m	0.4m~5m

百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

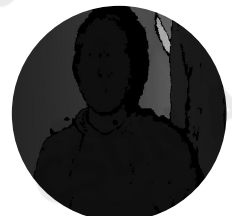
TOF & 3D结构光活体检测流程



可见光图像



红外光图像



深度图像

图像对齐



是否为活体

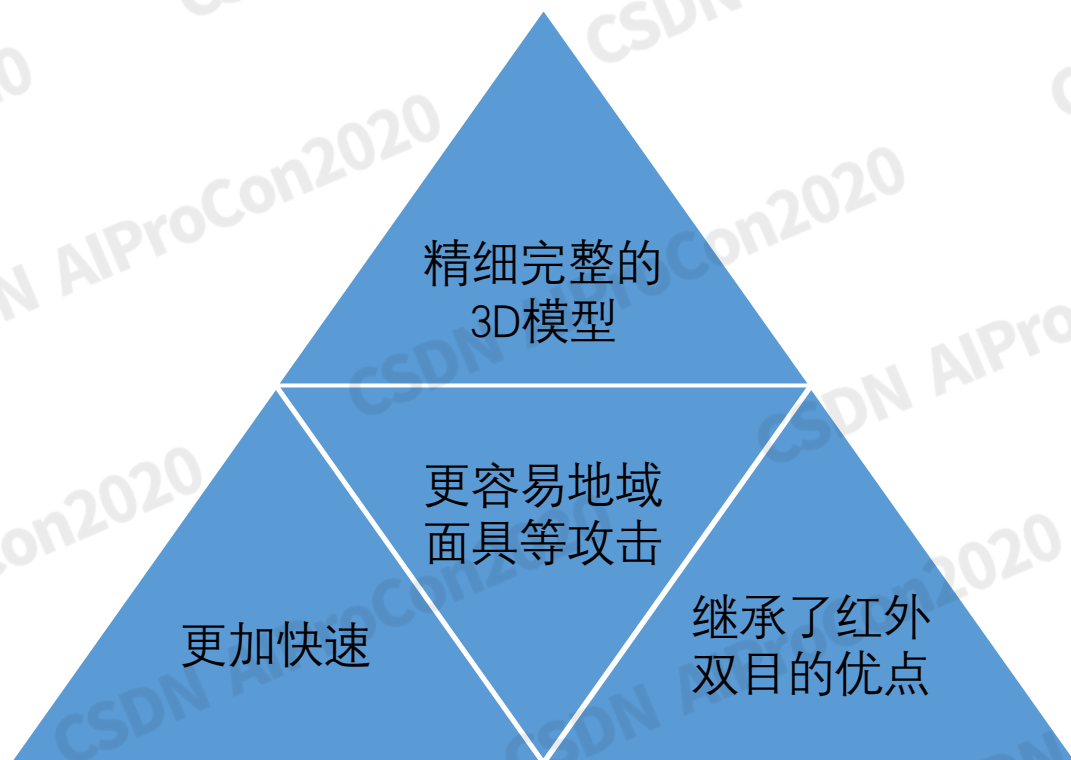
百万AI

2020 ProCon 万人开发者大会

7月3日-4日

CSDN

TOF/结构光对比红外双目的优势



- TOF、结构光和近红外都是依靠红外光
- TOF和结构光能够获取到深度信息，二近红外只能获取到二维平面信息
- TOF和结构光尤其是对面具等具有三维结构的攻击手段优势明显

百万AI

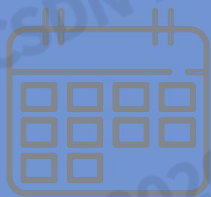
2020
ProCon
万人开发者大会

7月3日-4日

CSDN

活体检测应用举例

金融场景



ATM



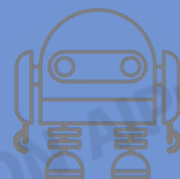
STM超级柜员机



自助发卡机



智能pos



厅堂机器人



App远程身份核验



柜面双录



刷脸支付PAD



收银机



自助购物机

身份验证



考勤



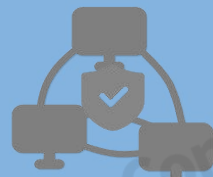
门禁/闸机



人脸识别



智慧鉴身



联网人证合一



7月3日-4日

CSDN

活体模组



比邻星红外双目
活体安全模组

CW-FI3105A

适用于各种自助机具



比邻星红外双目
活体安全模组

CW-MN2051B

适用于POS机、
PAD、双屏收银机



北极星结构光
活体安全模组

CW-LC2308-A

POS机、PAD、收银机、
自助机具



大角星ToF活体
安全模组

CW-LC1505-B

POS机、PAD、收银
机、自助机具

百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

云从人脸和人体分析技术

人脸识别

人脸三维重建

动作识别

活体检测

人脸属性解析

行人属性解析

表情识别

人体ReID

客流统计

模型训练平台

模型发布引擎

底层平台：CWOS

百万AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

人体ReID的定义

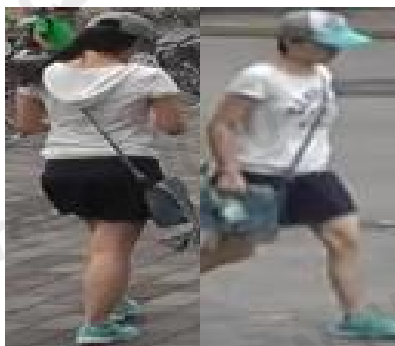
跨镜追踪 (Person Re-Identification, 简称 ReID) 技术是现在计算机视觉研究的热门方向, 主要解决跨摄像头跨场景下行人的识别与检索。该技术可以作为人脸识别技术的重要补充, 可以对无法获取清晰拍摄人脸的行人进行跨摄像头连续跟踪, 增强数据的时空连续性。



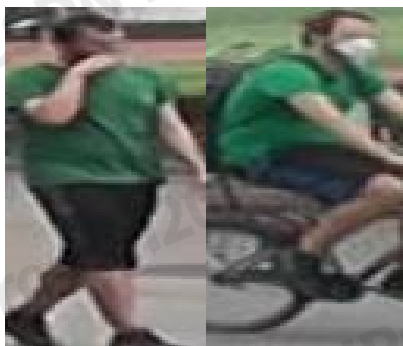
7月3日-4日

CSDN

人体ReID技术难点



姿态



无正脸照



配饰



遮挡

更多因素：

- 相机拍摄角度
- 图片模糊
- 室内室外环境变化

百万AI

2020
AI ProCon
万人开发者大会

7月3日-4日

CSDN

数据集和评价指标

	Market1501	DukeMTMC-reID	CUHK03
拍摄地点	清华大学	Duke大学	香港中文大学
图片数量	32217	36441	13164
行人数量	1501	1812	1467
摄像头	6	8	10

评价指标:

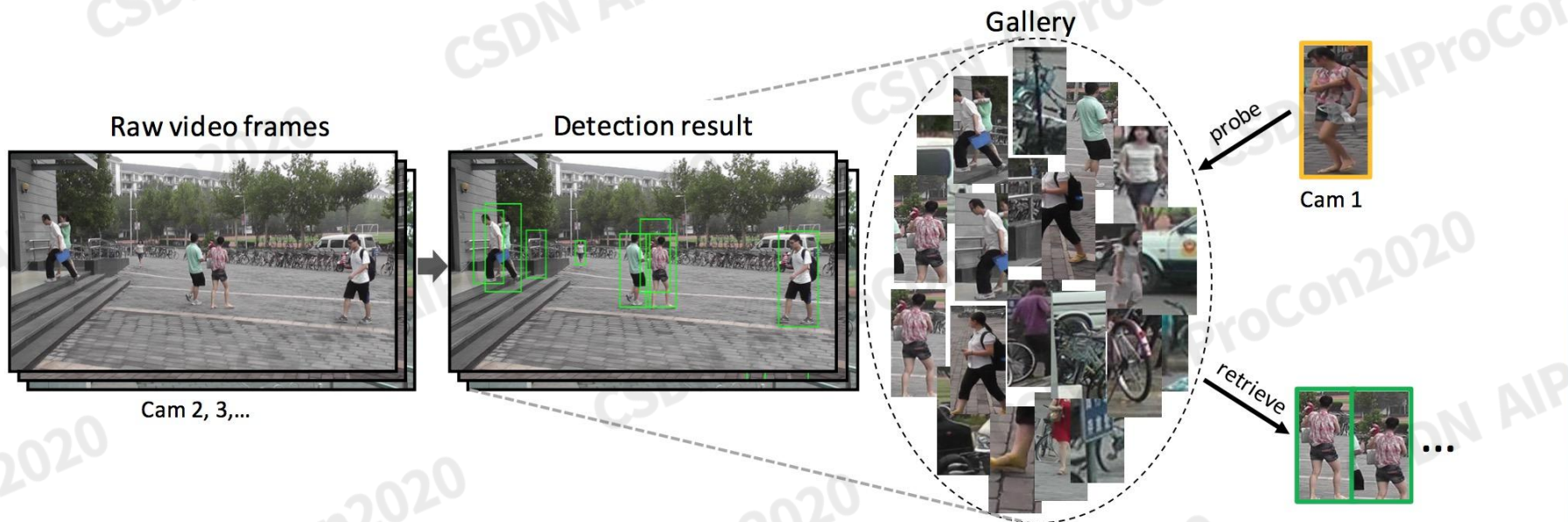
- Rank1: 首位命中率 (Rank-1 Accuracy)
- mAP: 平均精度均值 (Mean Average Precision, mAP)



7月3日-4日

CSDN

人体ReID的基本流程



RawFrame

Detect

ReID

常用数据集

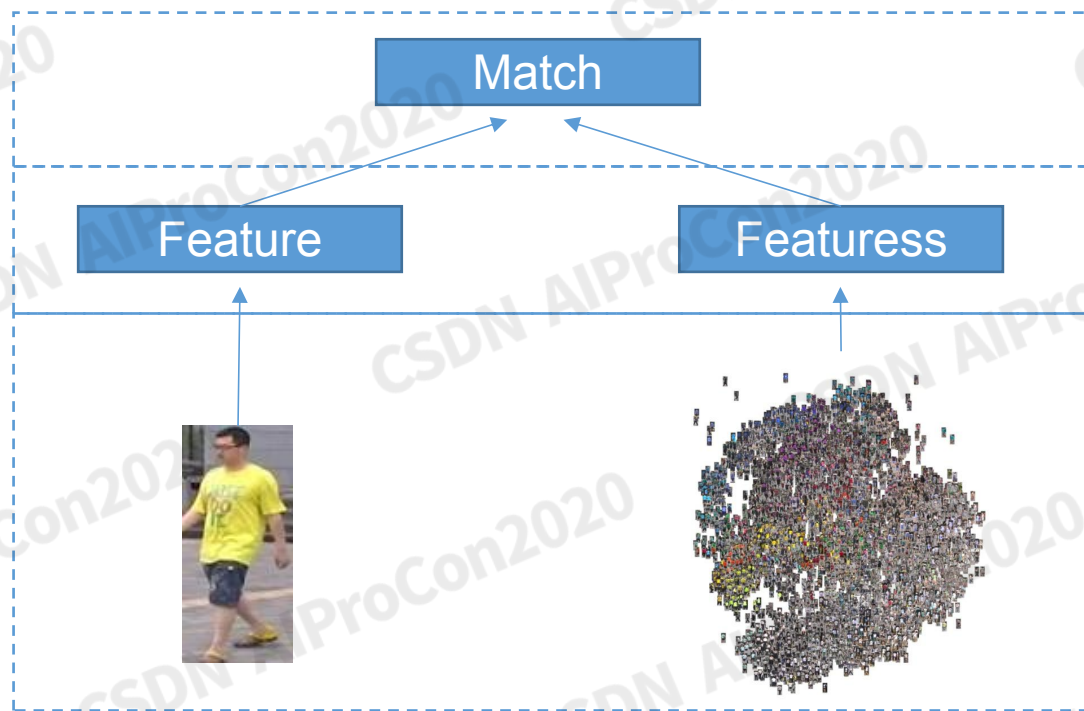
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

实现人体ReID的一般步骤



先抽取特征再进行比对

1. 检索图经过网络抽取图片特征 (Feature)
2. 底库里的所有图片全部抽取图片特征 (Feature)
3. 将检索图与地库图的特征计算距离 (例如欧式距离)
4. 根据计算距离进行排序, 排序越靠前表示是相似率越高



7月3日-4日

CSDN

度量学习

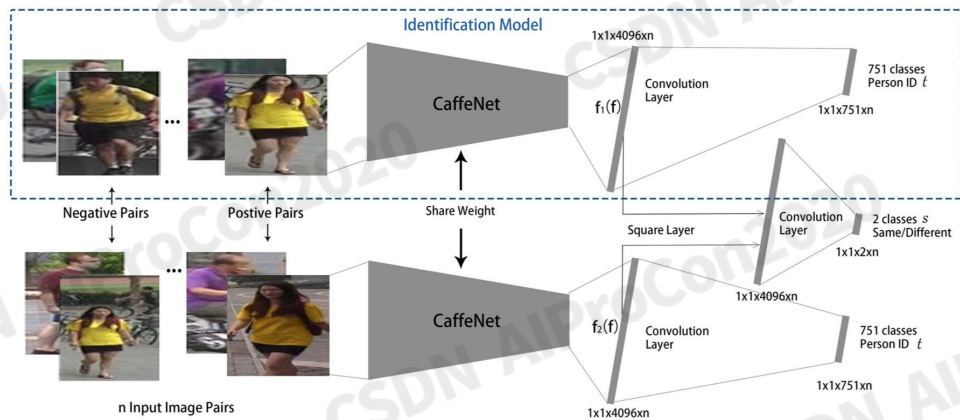


Fig. 3. The proposed model structure. Given n pairs of images of size 227×227 , two identical CaffeNet models are used as the non-linear embedding functions and output 4,096-dim embeddings f_1, f_2 . Then, f_1, f_2 are used to predict the identity t of the two input images, respectively, and also predict the verification label s jointly. We introduce a non-parametric layer called Square Layer to compare high level features f_1, f_2 . Finally, the softmax loss is applied on the three objectives.

Method	Single Query	
	rank-1	mAP
BoW + KISSME [3]	44.42	20.76
SL [31]	51.90	26.35
Multiregion CNN [32]	45.58	26.11
DADM [33]	39.4	19.6
CAN [34]	48.24	24.43
DNS [35]	55.43	29.87
Fisher Network [36]	48.15	29.94
S-LSTM [37]	-	-
Gate Reid [6]	65.88	39.55
CaffeNet-Basel. [26]	50.89	26.79
Ours(CaffeNet)	62.14	39.61
VGG16-Basel. [27]	65.02	38.27
Ours(VGG16)	70.16	47.45
ResNet-50-Basel. [28]	73.69	51.48
Ours(ResNet-50)	79.51	59.87

➤ SoftmaxLoss + ContrastiveLoss

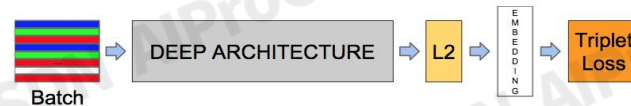


Figure 2. **Model structure.** Our network consists of a batch input layer and a deep CNN followed by L_2 normalization, which results in the face embedding. This is followed by the triplet loss during training.

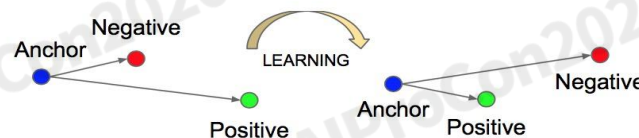


Figure 3. The **Triplet Loss** minimizes the distance between an *anchor* and a *positive*, both of which have the same identity, and maximizes the distance between the *anchor* and a *negative* of a different identity.

Market-1501 SQ		
mAP	rank-1	rank-5
69.14	84.92	94.21
60.71	81.38	92.34
58.06	78.50	91.18

➤ TripletLoss



7月3日-4日

CSDN

局部特征学习

- 方法一:
- 根据人体关键点取局部patch

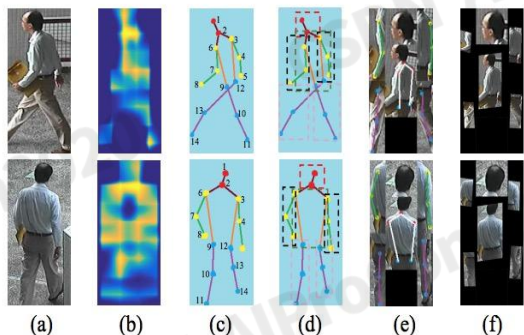


Figure 1. Illustration of part extraction and pose normalization in our Feature Embedding sub-Net (FEN). Response maps of 14 body joints (b) are first generated from the original image in (a). 14 body joints in (c) and 6 body parts in (d) can hence be inferred. The part regions are firstly rotated and resized in (e), then normalized by Pose Transform Network in (f).

dataset	CUHK03		Market1501	
	rank1	rank1	mAP	rank1
method				
Global Only	79.83	71.89	52.84	76.22
Part Only	53.73	47.29	31.74	55.67
Global+Part	85.07	76.33	62.20	81.74
Global+Part+FEN	87.15	77.57	62.58	83.05
Global+Part+FWN	86.41	77.62	62.58	82.69
PDC	88.70	78.29	63.41	84.14

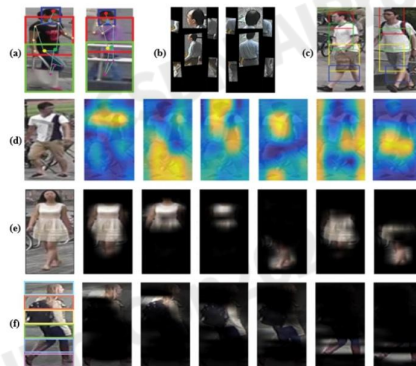


Figure 1. Partition strategies of several deep part models in person retrieval. (a) to (e): Partitioned parts by GLAD [31], PDC [27], DPL [35], Hydra-plus [22] and PAR [37], respectively, which are cropped from the corresponding papers. (f): Our method employs a uniform partition and then refines each stripe. Both PAR [37] and our method conduct “soft” partition, but our method differs significantly from [37], as detailed in Section 2.

Models	Feature	dim	Market-1501			
			R-1	R-5	R-10	mAP
IDE	pool5	2048	85.3	94.0	96.3	68.5
IDE	FC	256	83.8	93.1	95.8	67.7
Variant 1	\mathcal{G}	12288	86.7	95.2	96.5	69.4
Variant 1	\mathcal{H}	1536	85.6	94.3	96.3	68.3
Variant 2	\mathcal{G}	12288	91.2	96.6	97.7	75.0
Variant 2	\mathcal{H}	1536	91.0	96.6	97.6	75.3
PCB	\mathcal{G}	12288	92.3	97.2	98.2	77.4
PCB	\mathcal{H}	1536	92.4	97.0	97.9	77.3
PCB+RPP	\mathcal{G}	12288	93.8	97.5	98.5	81.6
PCB+RPP	\mathcal{H}	1536	93.1	97.4	98.3	81.0

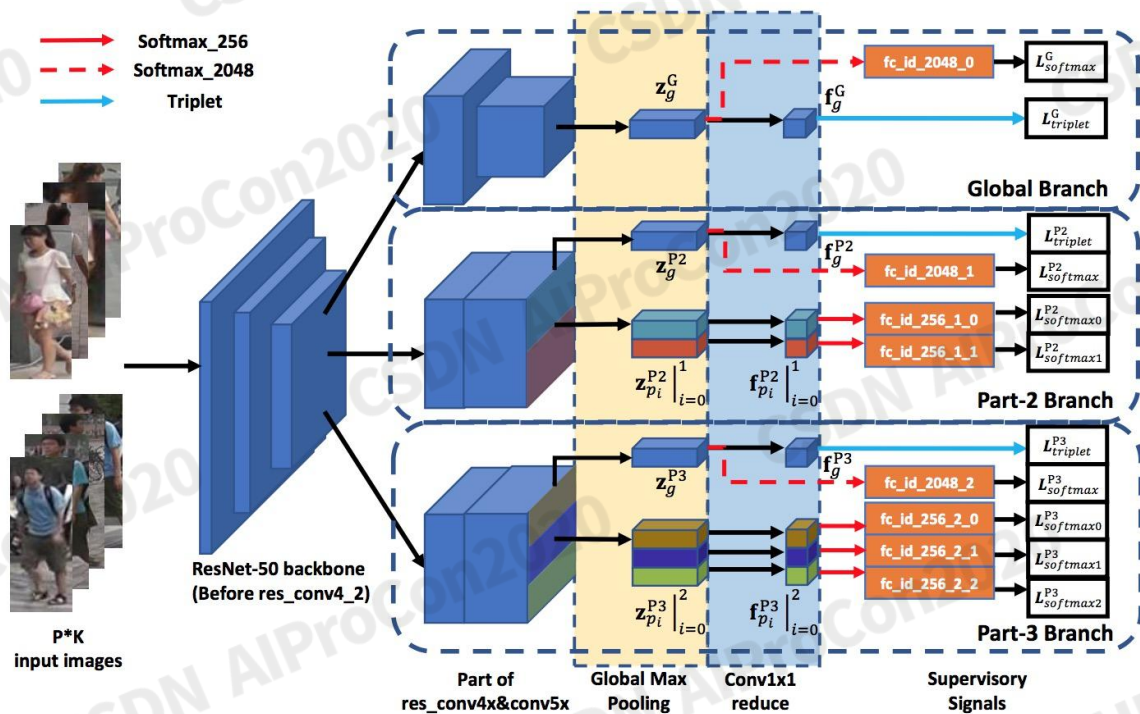
- 方法二:
- 横向切分为多个patch, 获取局部特征
 - RPP加attention权重



7月3日-4日

CSDN

云从多粒度网路 (MGN)



Loss: 8 Softmax + 3 Triplet

- Global branch
 - softmax & triplet
- N-Part Global features
 - softmax & triplet
- N-Part Local features
 - softmax

Learning Discriminative Features with Multiple Granularities For Person Re-Identification, CVPR 2018



7月3日-4日

CSDN

云从多粒度网路 (MGN)

Methods	Single Query		Multiple Query	
	Rank-1	mAP	Rank-1	mAP
MSCAN[18]	80.3	57.5	86.8	66.7
DLPA[40]	81.0	63.4	-	-
SVDNet[31]	82.3	62.1	-	-
PDC[30]	84.1	63.4	-	-
TriNet[14]	84.9	69.1	90.5	76.4
JLML[20]	85.1	65.5	89.7	74.5
DML[38]	87.7	68.8	91.7	77.1
DPFL[6]	88.6	72.6	92.2	80.4
HA-CNN[21]	91.2	75.7	93.8	82.8
GP-reid[2]	92.2	81.2	94.7	87.3
PCB [32]	92.3	77.4	-	-
Deep-Person[3]	92.3	79.6	94.5	85.1
Aligned-ReID[37]	92.6	82.3	-	-
PCB+RPP[32]	93.8	81.6	-	-
MGN(Ours)	95.7	86.9	96.9	90.7
TriNet(RK)[14]	86.7	81.1	91.8	87.2
GP-reid(RK)[2]	92.2	90.0	94.2	91.2
Aligned-ReID(RK)[37]	94.0	91.2	-	-
MGN(Ours, RK)	96.6	94.2	97.1	95.9

Table 2. Comparison of results on Market-1501 with Single Query setting (SQ) and Multiple Query setting (MQ). "RK" refers to implementing re-ranking operation.

Methods	Rank-1	mAP
PAN[43]	71.6	51.5
FMN[9]	74.5	56.9
SVDNet[31]	76.7	56.8
PSE[27]	79.8	62.0
HA-CNN[21]	80.5	63.8
Deep-Person[3]	80.9	64.8
PCB[32]	83.3	69.2
GP-reid[2]	85.2	72.8
MGN(Ours)	88.7	78.4

Table 3. Comparison of results on DukeMTMC-reID.

Methods	Labeled		Detected	
	Rank-1	mAP	Rank-1	mAP
BOW+XQDA[41]	7.9	7.3	6.4	6.4
LOMO+XQDA[22]	14.8	13.6	12.8	11.5
IDE[42]	22.2	21.0	21.3	19.7
PAN[43]	36.9	35.0	36.3	34.0
SVDNet[31]	40.9	37.8	41.5	37.3
HA-CNN[21]	44.4	41.0	41.7	38.6
MLFN[4]	54.7	49.2	52.8	47.8
PCB[32]	-	-	61.3	54.2
PCB+RPP[32]	-	-	63.7	57.5
MGN(Ours)	68.0	67.4	66.8	66.0

Table 4. Comparison of results on CUHK03 with evaluation protocols in [45].

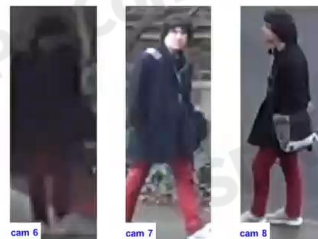
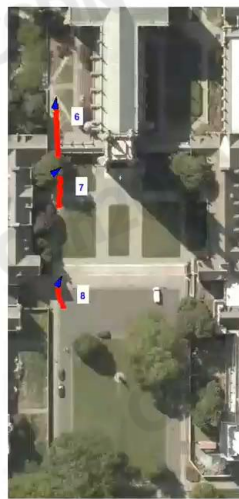
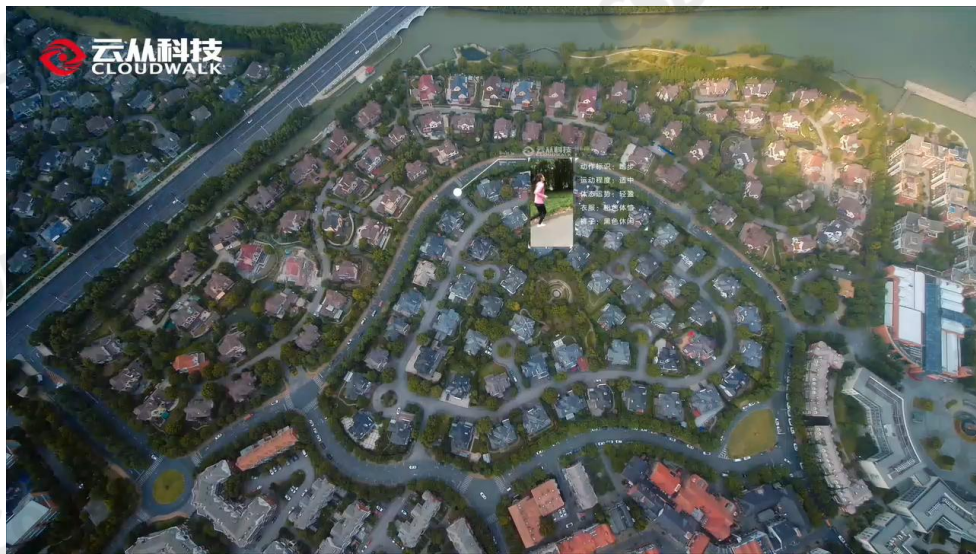
三个主要的
数据集
都达到
SOTA



7月3日-4日

CSDN

应用举例-安防



根据日常轨迹挖掘小偷、逃犯等可疑人员

百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

应用举例-安防



已知嫌疑人图片，搜索历史监控视频获得更多嫌疑人图片

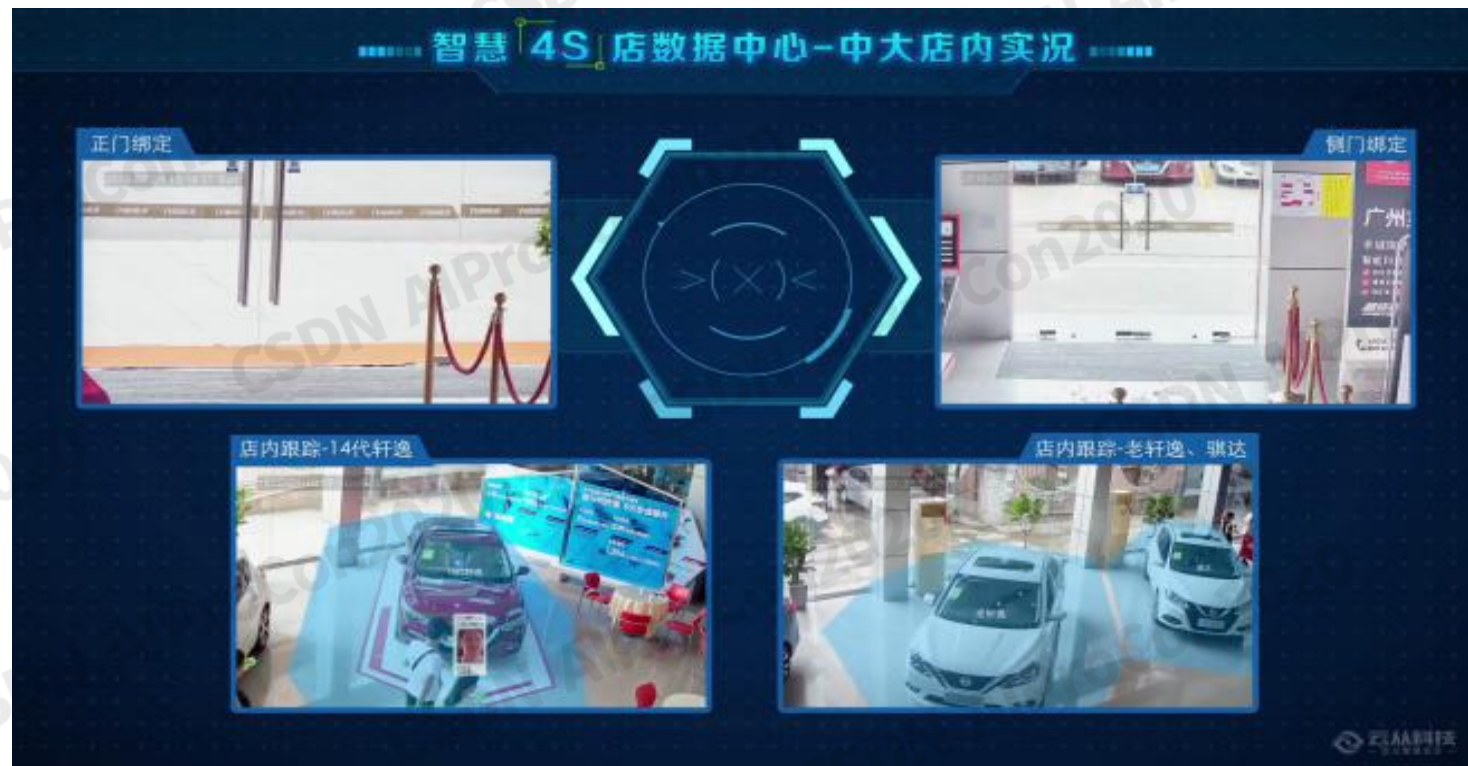
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

应用举例-4S店



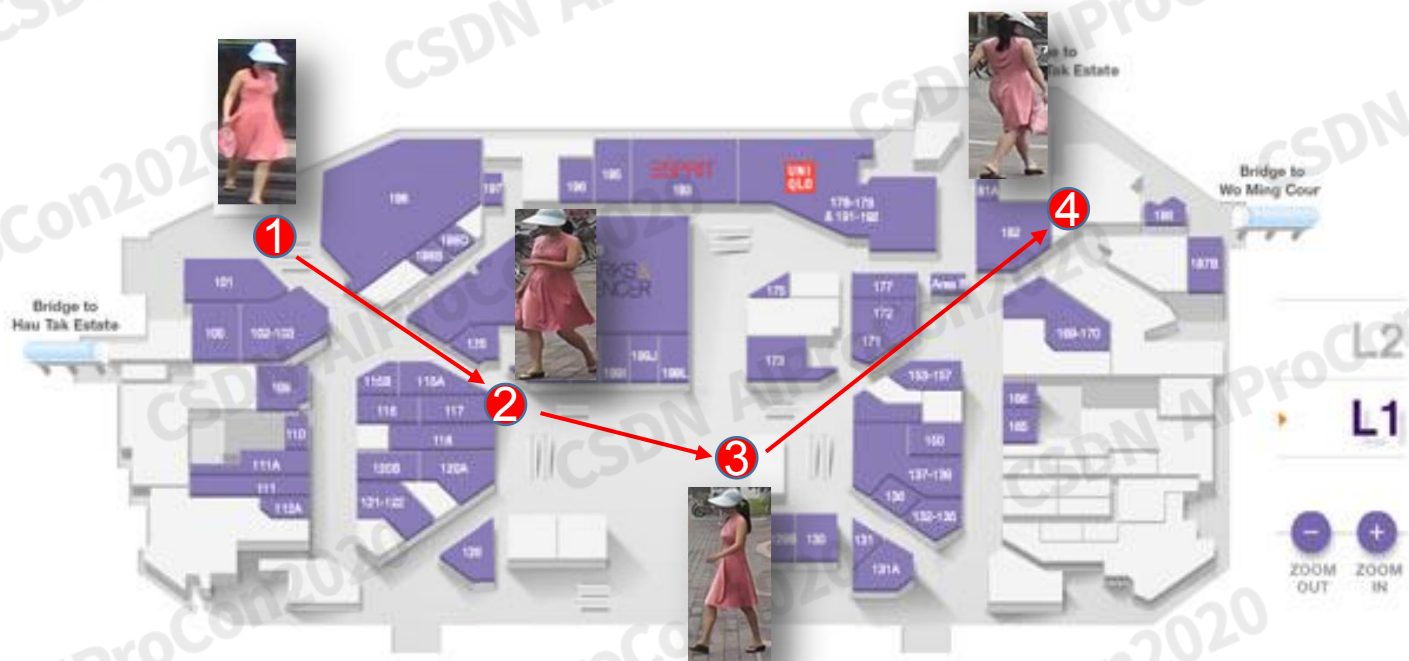
百万
人学AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

应用举例-商场



识别行进与停留轨迹，智能理解用户需求（公共场合）



7月3日-4日

CSDN

云从人脸和人体分析技术

人脸识别

人脸三维重建

动作识别

活体检测

人脸属性解析

行人属性解析

表情识别

人体ReID

客流统计

模型训练平台

模型发布引擎

底层平台：CWOS

百万
AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

云从模型训练平台

研究院训练云平台

Dashboard / 任务 / 新建任务

- 任务
- 文件
- 虚拟机
- 关于我
- 集群状态
- 私有池
- 用户管理

* 任务名称: resnet100f

* 任务描述: resnet100f

* 任务框架: caffe-mn

* CPU: 300

* GPU: 12x8

* 内存: 1200

* solver_file: /workspace/SE.LResNet100F.IR/config/solver.prototxt

* log_file: /workspace/SE.LResNet100F.IR/logs/caffe-mn.log

* 数据卷 + 添加数据卷

* 卷名	jianyi	* 挂载点	/workspace	* 创建者	jianyi
* 卷名	facelist	* 挂载点	/facelist	* 创建者	jianyi
* 卷名	idcard380w	* 挂载点	/ssd/idcard380w	* 创建者	jianyi
* 卷名	idcard170w	* 挂载点	/ssd/idcard170w	* 创建者	jianyi
* 卷名	foreigner	* 挂载点	/ssd/foreigner	* 创建者	jianyi
* 卷名	ms1m	* 挂载点	/ssd/ms1m	* 创建者	jianyi
* 卷名	face72w	* 挂载点	/ssd/face72w	* 创建者	zhouyafel



7月3日-4日

CSDN

云从模型部署平台 (IBIS)

集成生物识别平台 (IBIS)



百万AI

2020
ProCon
万人开发者大会

7月3日-4日

CSDN

THANKS

云从科技集团

广州市·南沙区·香江国际科创中心

重庆市·渝北区·数字经济产业园

上海市·浦东新区·张江人工智能岛

北京市·东城区·环球贸易中心

苏州市·工业园区·人工智能产业园

成都市·天府新区·天府菁蓉中心

南京市·雨花台区·宁双路软件谷

杭州市·萧山区·诺德财富中心

400-151-5992

商务合作: business@cloudwalk.cn

媒体合作: Media@cloudwalk.cn

渠道合作: business_partner@cloudwalk.cn

百万
人学AI

 **2020**
ProCon
万人开发者大会

7月3日-4日



CSDN
扫码关注