

标识: SDK-Team

版本: 3.1.0

人脸抓拍识别服务器 API(V4)白皮书



目录

目录.....	I
1 公司简介.....	1
1.1 背景介绍.....	1
1.2 技术优势.....	1
2 产品介绍.....	2
2.1 产品概述.....	2
2.2 产品亮点.....	3
2.3 应用场景.....	3
2.4 软件环境.....	3
2.5 硬件标配.....	3
2.6 支持视频及图片格式.....	4
2.7 功能模块.....	4
2.7.1 人脸检测.....	5
2.7.2 人脸跟踪.....	5
2.7.3 属性分析.....	6
2.7.4 人脸识别.....	8
2.8 授权方式.....	8
3 性能指标.....	9
3.1 标准硬件配置.....	9
3.2 性能参数.....	9
4 集成方式.....	10

1 公司简介

1.1 背景介绍

Megvii 旷视是中国领军的人工智能产品公司。旷视以深度学习和物联传感技术为核心，立足于自有原创深度学习算法引擎 Brain++，深耕金融安全、城市安防、手机 AR、商业物联、工业机器人五大核心行业，致力于为企业级用户提供全球领先的人工智能产品和行业解决方案。发展至今，旷视已在北京、西雅图、南京设立独立研究院，并在十余个核心城市设立分部。在“赋能机器之眼，构建城市大脑”的愿景下，旷视人正在推动人工智能技术在中国及全球范围的产业落地，并通过打造 MegCity 城市大脑数据平台为构建智慧城市、平安城市基础设施而奋斗。

在中国科技部火炬中心“独角兽”榜单中，旷视排在人工智能类首位。2017 年 7 月，旷视曾作为唯一科技企业代表在政府半年经济会议中向李克强总理及国家各级领导做企业创新汇报。

1.2 技术优势

人才聚集，旷视科技首席顾问姚期智院士，图灵奖获得者、中科院院士。旷视科技首席科学家孙剑，前微软亚研院首席研究员。旷视美国研究院院长王珏，前 Adobe 首席科学家。

2017 年，在 COCO 及 Places 竞赛排名，获得了三项第一、一项第二的优异成绩，一举击败了来自微软、Facebook、Google、商汤科技等企业，成为第一个在 COCO 竞赛中获得冠军的中国企业。

2018 年 1 月 30 日，旷视包揽 2017 MIT Technology Review 榜单大满贯，旷视的“刷脸支付”技术入选“全球十大突破技术”；同年 6 月，旷视入选“全球 50 大最聪明公司”；旷视 CEO 印奇入选“35 岁以下科技创新 35 人”。

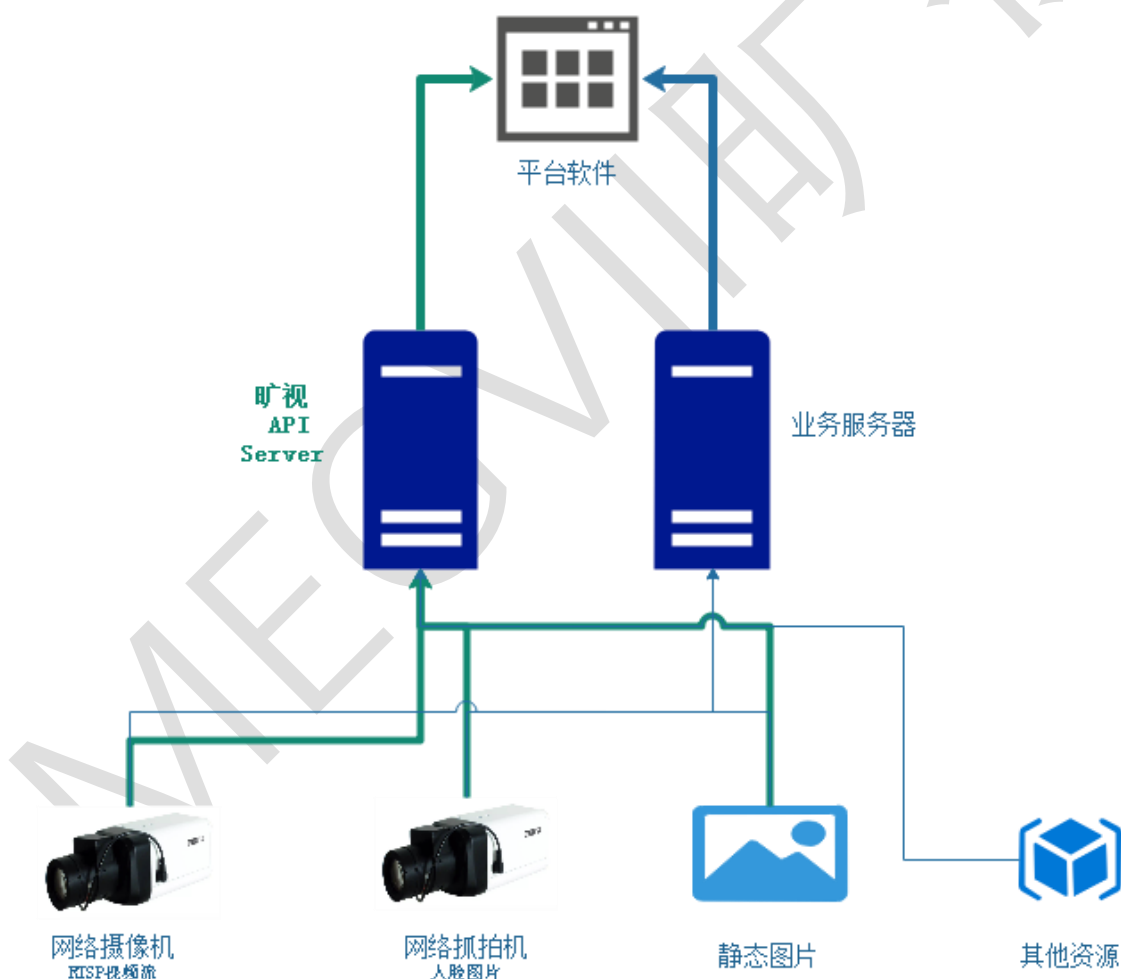


2 产品介绍

2.1 产品概述

旷视科技人脸抓拍识别服务器 SDK 产品，以旷视科技完全自主知识产权的深度学习算法为基础，供高识别率高效率的人脸实时比对、人脸特征结构化等功能，将“三站一场”等重点防控区域的前端设备采集到的实时数据进行解析、抽取特征、比对识别、告警；通过以图搜图的方式，实现在人像库中的检索、身份确认。

被集成使用时，网络拓扑及数据流方向请见下图：



2.2 产品亮点

快速，依托旷视业界领先的分析算法与技术实力，特殊优化的抓取人脸算法，极速处理并返回结果，满足各行业的业务需求。

准确，旷视的人脸检测技术能够很好地适应各类常见影响，如脸部遮挡、光照情况不理想、头部朝向不正等。

领先，基于旷视强大的深度学习算法，人脸检测与识别算法在多项世界权威比赛中夺冠。

2.3 应用场景

典型应用场景包括火车站、汽车站、地铁站、飞机场等人流量大的区域内，并且具有天然通道型的场景，如人员闸机、安检通道、进出门口、ATM 防护仓、过街天桥、小区出入口等。

2.4 软件环境

SDK	BootOS
人脸抓拍识别服务器 API	旷视基于 ubuntu16.04 的专属操作系统

2.5 硬件标配

人脸识别服务器型号	详细参数
KSS-CS3010-H1	Intel XEON E5-2630 V4*2
	16G DDR4*8
	4T 7.2k 3.5" SATA 6G(希捷)*4(做 RAID5) +
	Intel S4500 480 GB SSD*1(直连主板)
	板载 1000M 网卡*2
	冗余双电 800W
	GPU: RTX 2070(colorful turbo)*2
	LR382A/8 口/SAS 12Gb/半高/PCIe 3.0 x8/1GB 缓存/

支持 RAID 0,1,5,6,10,50,60,JBOD

备注：性能指标是基于上述硬件环境测试得出。目前支持的显卡型号为：

NVIDIA P4

NVIDIA P40

NVIDIA T4

NVIDIA TITAN RTX

Geforce RTX 2080 Ti

Geforce RTX 2080

Geforce RTX 2070

NVIDIA TITAN Xp

NVIDIA TITAN X

GeForce GTX 1080 Ti

GeForce GTX 1080

GeForce GTX 1070

2.6 支持视频及图片格式

对于视频方面，支持对标准 H.264 视频流进行解码及后续分析，使用方可以使用 VLC、FFmpeg 来做视频解码验证。

对于图片方面，支持 JPEP/JPG、PNG、TIFF、BMP 图片格式。

2.7 功能模块

人脸抓拍识别服务器的功能模块如下图所示：

功能模块	输入	输出	其他说明
人脸检测	视频帧	人脸位置数据（若干人脸框等）	检测出其中所有人脸，并定位出人脸的位置
视频跟踪	连续视频帧	人脸位置数据（若干人脸框等） track ID	检测出每一帧图像中的所有人脸，并在后续视频帧中跟踪住人脸（保持 id 不变），且定位出每一个人脸的位置

属性分析	视频帧/静态图片	角度 模糊度 年龄、性别等	人脸角度, 分为 yaw/pitch/roll 三个方向, 角度越大, 越不合适做人脸识别。 人脸模糊度越大, 越不合适做人脸识别。
人脸识别	人脸图片	比对到的前 N 个人脸信息 比对分值	人脸识别结果的比对分值建议, 需根据实际场景具体分析。
图片入库	人脸静态图片	图片在底库中的唯一 ID	将图片、获取到的图片特征值, 全部存储, 为人脸比对做准备
底库管理	用户的 API 操作	底库的增删改查	底库: 即需要进行 1:N 比对的 N 张人脸信息

2.7.1 人脸检测

给定一张图像, 检测出其中所有人脸, 并定位出人脸的所有关键点, 用于后续属性判断和识别。

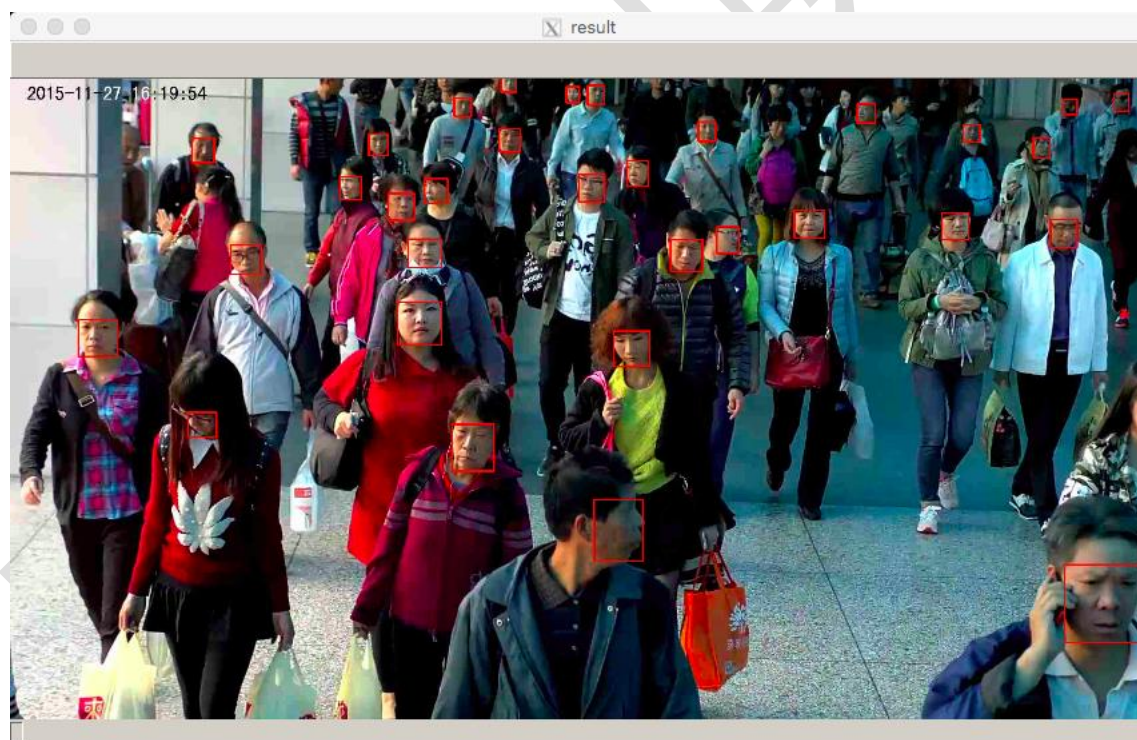


图 2-1 动态人脸检测

2.7.2 人脸跟踪

给定一段视频, 检测出每一帧图像中的所有人脸, 并在后续视频帧中跟踪人脸 (保持 id 不变), 且定位出每一个人脸的关键点。跟踪相比于检测优势在于:

- (1)减少每帧检测的带宽
- (2)保证一个人轨迹的连续性，以抽取更高质量的人脸图像用于识别，提升识别性能，减小识别带宽。



图 2-2 动态人脸跟踪

2.7.3 属性分析

表 2-3 属性功能

属性名	备注	属性名	备注
age	年龄	gender	性别
mouthStatus	嘴部遮挡	eyeStatus	眼部遮挡
noseOcclusionStatus	鼻子遮挡	headwearStatus	头部装饰
beardStatus	胡子状态	blurriness	人脸模糊度
Pose	人脸角度		

2.7.3.1 POSE

表 2-4 pose 三维介绍

参数名	类型	参数说明
-----	----	------

roll	float	以鼻子为中心的旋转角度，取值[0-180], 0 为正脸
pitch	float	上下仰角度值，取值[0-180], 0 为正脸
yaw	float	左右偏转角度值，取值[0-180], 0 为正脸

人脸正对摄像机的角度大小，分为 3 个轴向，如角度过大会影响到后期识别效果，所以会对抓拍的人脸有一定角度的限制，过滤掉过大角度人脸图片。

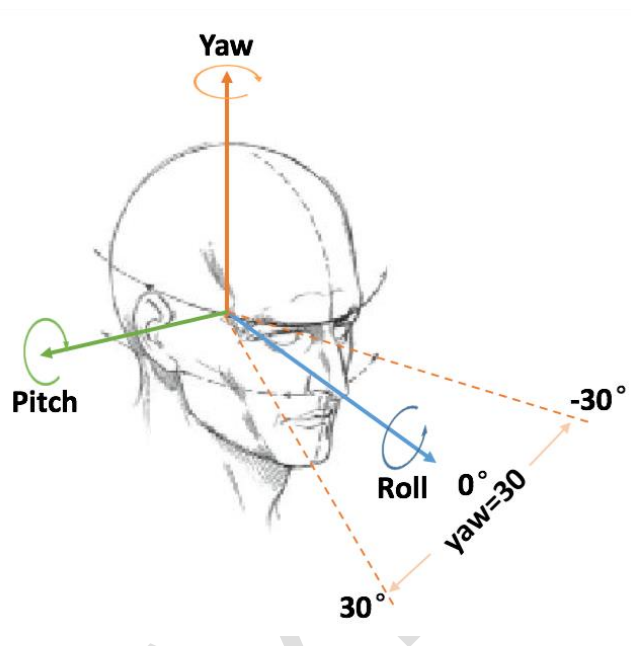


图 2-3 人脸角度与变量关系

2.7.3.2 BLURNESS

对图像质量的判断方法，通过检测图像的模糊程度判断待检测（或者识别）图片的质量。定义范围 0-1，其中越靠近 0 图像越清晰，越靠近 1 图像越模糊。如下图所示，左边图像靠近 0，右边图像靠近 1。

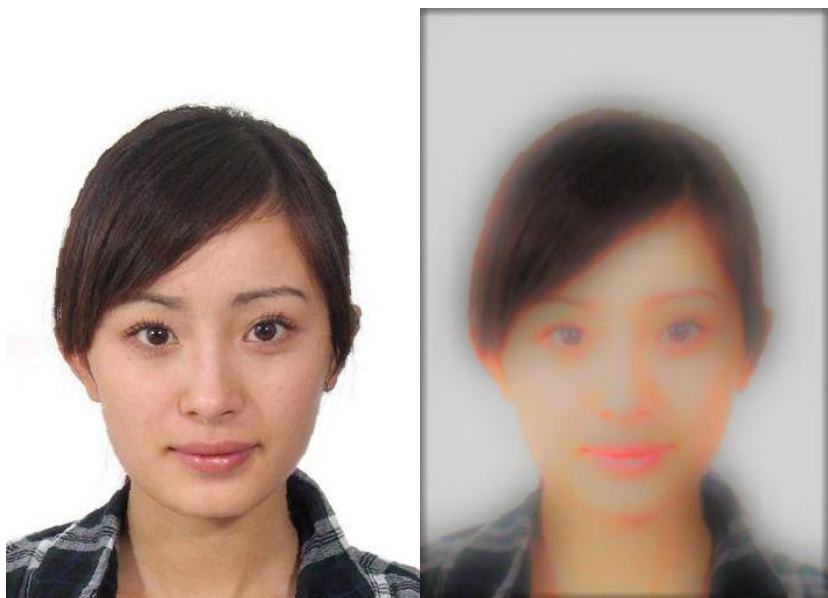


图 2-4 模糊度对比图

2.7.4 人脸识别

快速检测两张静态照片的人脸特征，判定是否为同一个人。



图 2-5 静态人脸识别

2.8 授权方式

通过硬狗加密狗来完成授权，加密狗中会对于接入的视频流路数，图片的并发数量等

内容来做权限控制。

3 性能指标

3.1 标准硬件配置

见本文档 2.5 《硬件标配》内容。

3.2 性能参数

功能	性能指标	备注
底库管理	单机：1000 库	库 = 相机 + 底库
入库	入库成功率：99.5%	
	入库速度单机双卡：240	基于 640*480 分辨率的图片测试结果
	入库最小人脸：40*40	像素
	入库支持图片分辨率：60*60 像素~4K 分辨率	
	入库延迟： 低 QPS (≤ 10 QPS) 情况下 < 500ms 高 QPS (~400 QPS) 情况下 < 10s	
实时视频流	支持分辨率：720P~4K (resize 至 1080P 处理)	
	支持编码格式：H.264, H.265 兼容 SVAC (路数/性能有损失)	
	支持属性：10 种	人脸角度、人脸模糊度、年龄、性别、眼部遮挡、嘴部遮挡、鼻子遮挡、头部装饰、胡子状态等

	单卡 16 路	单机最多可扩展至 8 卡，掉点≤5%
	集群 2000 路	
	布控底库容量大小：100 万	
图片流（抓拍机） 接入	支持属性：10 种	人脸角度、人脸模糊度、年龄、性别、眼部遮挡、嘴部遮挡、鼻子遮挡、头部装饰、胡子状态
	单卡指标：150QPS	分辨率不超过 640*480，每张图片只分析一个人脸
	单机最大支持：7 卡，1000QPS	剩下 1 卡做检索
	布控底库容量大小：100w	
图片 1:1 比对	单机双卡指标：120 QPS	基于 640*480 分辨率的图片测试结果
图片分析	单机双卡指标：240 QPS	基于 640*480 分辨率的图片测试结果
离线视频分析	支持的最大时长：≤ 2h	
	支持的文件大小：4GB	
布控库 1:N 检索	检索速度：95%<500ms	1 个库，单机 100 万，Top 100，10 并发，20QPS
	检索速度：95%<1s	1 个库，单机 100 万，Top 100，200 并发，100QPS
静态库 1:N 检索	检索速度：95%<1s	1 个库，单机 7000 万，Top 100，10 并发，10QPS
	检索速度：95%<3s	1 个库，单机 7000 万，Top 100，100 并发，30QPS
库库碰撞	单机双卡，100 万*100 万，耗时 5 分钟	

4 集成方式

本版本人脸抓拍识别服务器 API 在功能方面提供 Web Restful API 接口的方式，方便使用方采用任何一种语言进行二次开发集成。如下图所示：

