VISION SYSTEMS ARCHITECTURES AND SOLUTIONS

Dr TIAN Jing

tianjing@nus.edu.sg

What are advantages of intelligent vision systems (security)

- Incident prevention can save significant costs related to asset damage, theft and business continuity disturbance. For instance, detecting a perimeter breach quickly and automatically can enable security personnel to respond to the threat and prevent it taking place.
- A reduction in the number of guards or personnel required. For example, automated remote video surveillance detecting perimeter intrusion as a service, which eliminates the need for on-site guards.
- Reduction in staff stress-related illnesses, particularly prevalent to those watching potentially disturbing video scenes.
- Video analytics can support other operational systems. For example, video analytics can be used to detect motion which can control lighting levels, eliminating the need to have lighting running at full brightness.

Reference: Video analytics adoption: Key considerations for the end-user, https://erncip-project.jrc.ec.europa.eu/sites/default/files/JRC102121_video_analytics_adoption-key_considerations_for_end_users.pdf

Key considerations of vision systems (1/7)

- Analytics
- Acquisition
- Development
- Performance index
- Deployment
- User interface
- Business model

	Low-level	Middle-level	High-level
Input	Image	Image	Image/attributes
Output	Image	Attributes	Understanding
Examples	• Noise removal • Image sharpening	• Object detection and segmentation	• Scene understanding • Event recognition

Q: What is the problem you are addressing? What is the function of the video analytic?

Image: https://www.researchgate.net/post/CCTV-is_there_a_way_or_a_new_method_of_improving_the_image_quality_for_facial_recognition

Key considerations of vision systems (2/7)

- Analytics
- Acquisition
- Development
- Performance index
- Deployment
- User interface
- Business model

Q: Are the image quality of your footage and camera position sufficient?

- Light-Sensitive Camera: measures from the visible part of the electromagnetic spectrum, typically red, green and blue dominance; the RGB camera is an example.
- Multi/Hyper-Spectral Sensors: measure from a broader part of the electromagnetic spectrum (than the light-sensitive cameras) with individual sensors tuned to specific bands.
- Range Sensor: is a device that measures the distance from the observer to a target. Methods include laser, radar, sonar, Lidar and ultrasonic range finding.
- Additional equipment needed (such as external lighting)

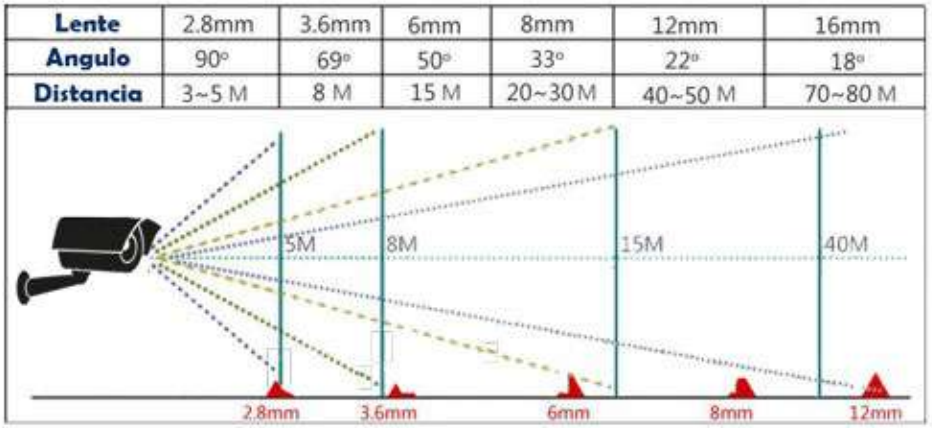


Photo: https://www.quora.com/What-is-the-average-angle-of-a-CCTV-camera

Key considerations of vision systems (3/7)

- Analytics
- Acquisition
- Development
- Performance index
- Deployment
- User interface
- Business model

Vision intelligence + machine learning system		
System	Data	Model
Deep learning framework	Evaluation script and metric	Model configuration & training strategy
Vision library	Data augmentation	Transfer / Ensemble

Development framework

- OpenCV (Python and C)
- Matlab
- Tensorflow (Python and Java)
- SimpleCV (Java)
- Amazon/Azure/Google/Intel Cloud vision intelligent services

Machine learning framework

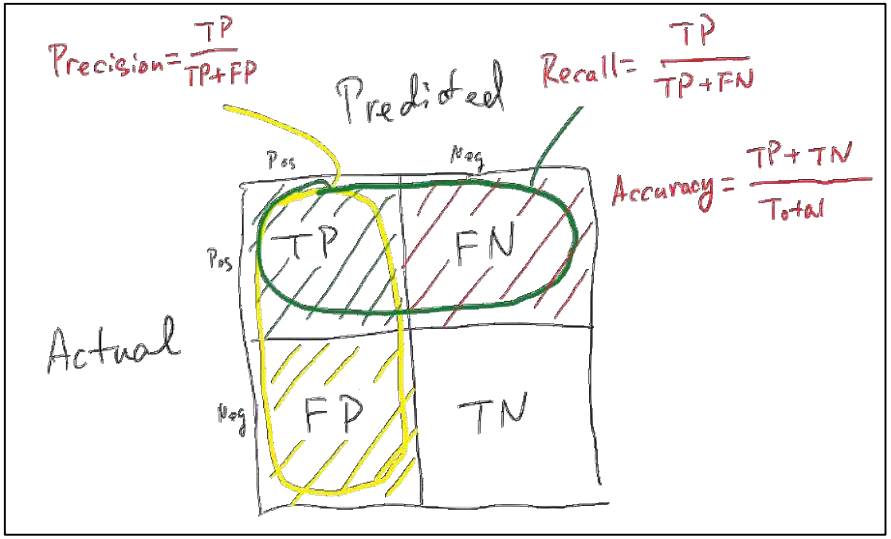
Caffe, Caffe2, Chainer, Cognitive Toolkit, MATLAB, mxnet, PaddlePaddle, PYTORCH, TensorFlow, ...

Reference:
- https://hub.packtpub.com/top-10-computer-vision-tools
- https://developer.nvidia.com/deep-learning

Key considerations of vision systems (4/7)

- Analytics
- Acquisition
- Development
- Performance index
- Deployment
- User interface
- Business model

- What detection rate and false alarm are acceptable?
- Test scripts, red teaming



True positives (TP)	The data that is correctly classified by a model as positive instance of the concept being modelled.
False positives (FP)	The data that is classified as positive instance by the model, but in fact are known not to be
True negatives (TN)	The data correctly classified by the model as not being instances of the concept
False negatives (FN)	The data that is classified as not being instances, but are in fact know to be

Reference: http://kflu.github.io/2016/08/26/2016-08-26-visualizing-precision-recall/

Key considerations of vision systems (5/7)

- Analytics
- Acquisition
- Development
- Performance index
- Deployment
- User interface
- Business model

Q: How do you want to deploy? edge vs cloud, real-time vs. post-event

API Category	API	Item	Price (USD / Item)
Facial Recognition	Detect API	Face Detection	0.0005
Facial Recognition	Compare API	Face Comparing	0.002
Facial Recognition	Search API	Face Searching	0.002
Facial Recognition	Face GetDetail API	Return Face Details	0.0001

SDK Product	Validity	Unit	Price (USD)
Face Landmark SDK	One year (365 days)	per platform	50K
Face Compare SDK	One year (365 days)	per platform	200K

Source: https://www.faceplusplus.com/pricing-details/

Key considerations of vision systems (6/7)

- Analytics
- Acquisition
- Development
- Performance index
- Deployment
- User interface
- Business model

- How much configuration is required?
- How do you want to be notified of alerts?
- What are the legal restrictions which may exist in your country?

An example in CCTV security system

Stakeholders	People (being watched)	System owner	System operator
Acquisition	• Consent • Signage • Anonymity	• No data missing • Video properties	• None
Storage	• Secure storage	• Secure storage • Deletion after retention period	• None
Transmission	• Confidentiality • Integrity • Authenticity	• Confidentiality • Integrity • Authenticity	• None
Monitoring	• Privacy safeguards • Authorized access • Public access to their data	• Continuous monitoring • Authorized access	• Data freshness • Timestamping • Easy to search

Reference: Q. Rajpoot, C. Jensen, "Security and Privacy in Video Surveillance: Requirements and Challenges," Int. Conf. on Information Security, 2014, https://hal.inria.fr/hal-01370363/document

Vision sensing and sense making (7/7)

- Analytics
- Acquisition
- Development
- Performance index
- Deployment
- User interface
- Business model

Analysis and planning

Program management execution

Ongoing management

Consult

- Security and architecture assessments
- Operations consulting

Design

- Requirement development
- System architecture

Integrate

- Solution configuration
- Customization

Deploy

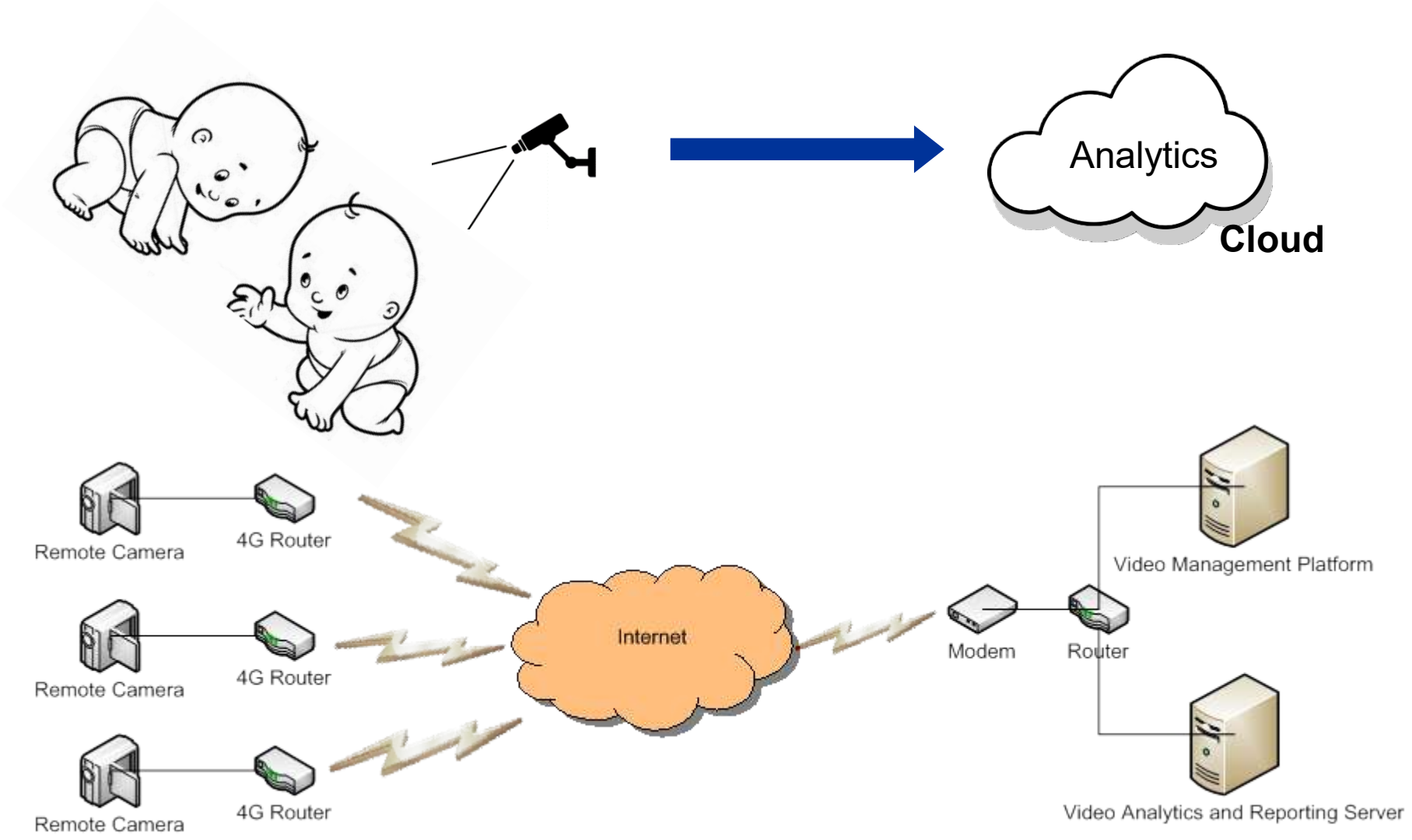
- Site surveys
- Civil works and installation
- Site acceptance testing

Maintain and operate

- Technical support
- Machine learning model update and tuning

Reference: Padma Kamath, Video Surveillance: The important basics, available at http://media.govtech.net/GOVTECH_WEBSITE/EVENTS/PRESENTATION_DOCS/2008/Los_Angeles_Tech_Forum/320VideoSurveillanceKAMATH.ppt

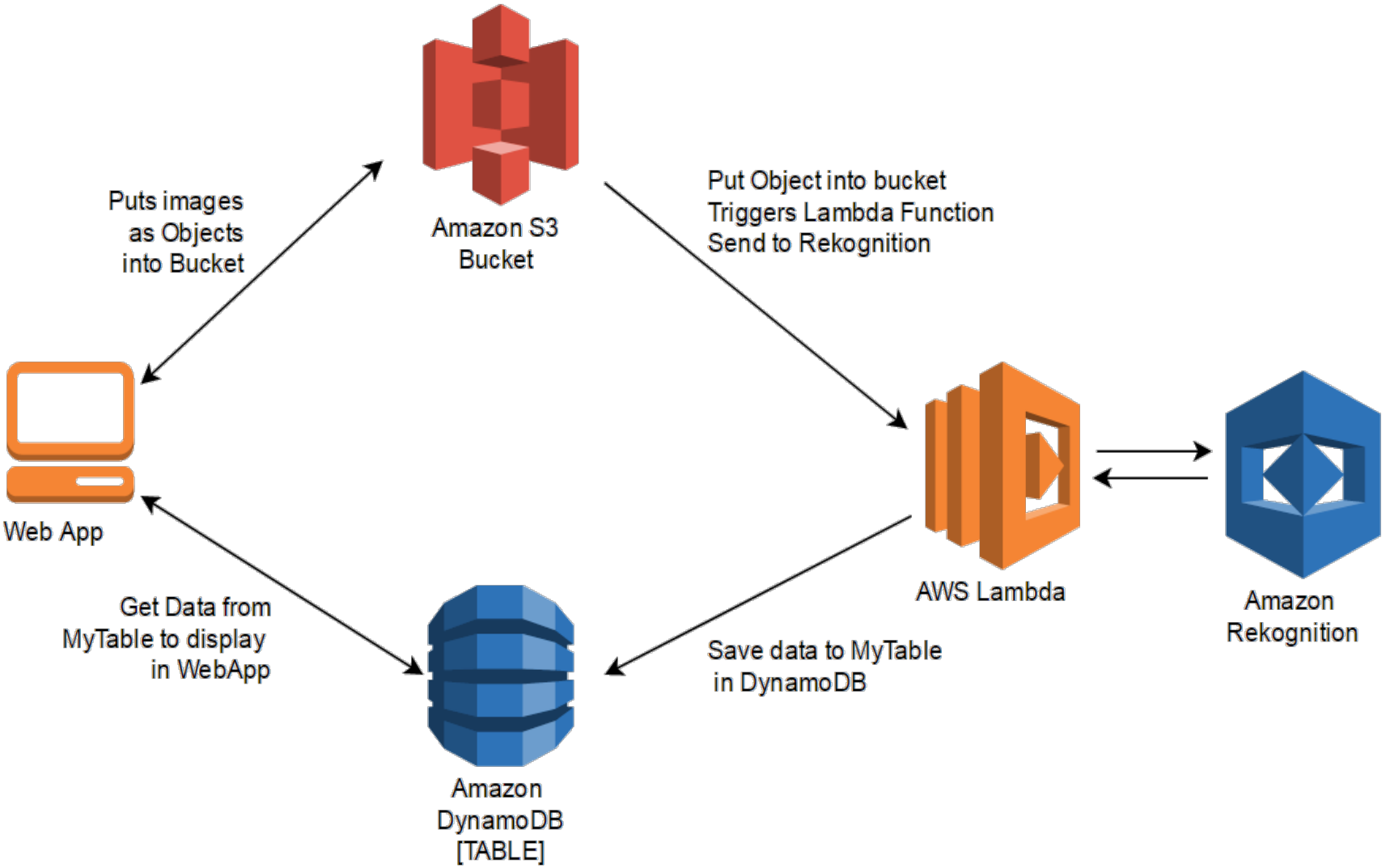
Vision intelligence: Cloud



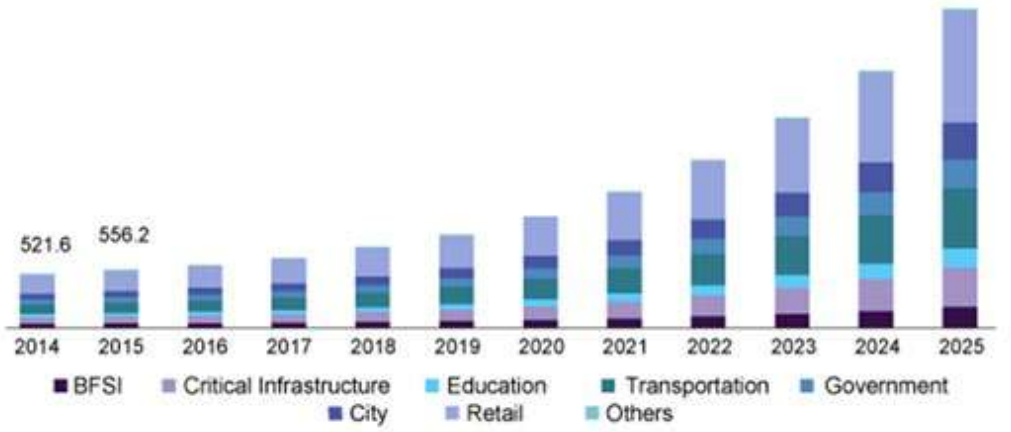
Vision intelligence: Cloud

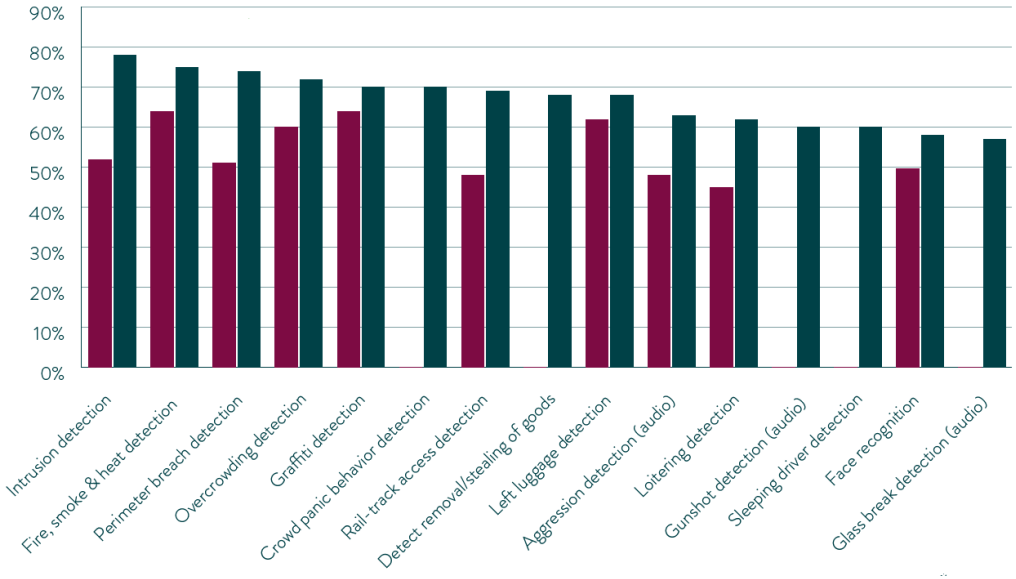
- Agility and Affordability: No capital investment of large-size infrastructures for analytical workloads.
- Data Analytics Platforms in Clouds: Leveraging cloud-enabled and ready platforms are fast and easy.
- Databases and Data Warehouses in Clouds: All kinds of database management systems and data warehouses in cloud speed up the process of data analytics.
- Enterprise-class Applications in Clouds: All kinds of customer-facing applications are cloud-enabled and deployed in highly optimized and organized cloud environments.
- Sensor/Device-to-Cloud Integration are available to transmit ground-level data to cloud storages and processing.
- Latency and Response time is often a critical part, especially when you deal with emergency procedure.
- Bandwidth Cost and Capacity: If you want to use a set of smart devices requiring each one to communicate certain bytes of data then you can quickly reach huge bandwidth requirements reaching Mbit/s or even Gbit/s at a gateway level.
- Security and Privacy: Transmitting device data over any open and public network is risky.
- Offline usages versus only-online usages: Pure cloud services do not allow offline usages. It is a major shortcoming since smart cities and industry 4.0 applications require a dual offline/online paradigm.

Example: AWS Rekognition



Vision meets edge computing



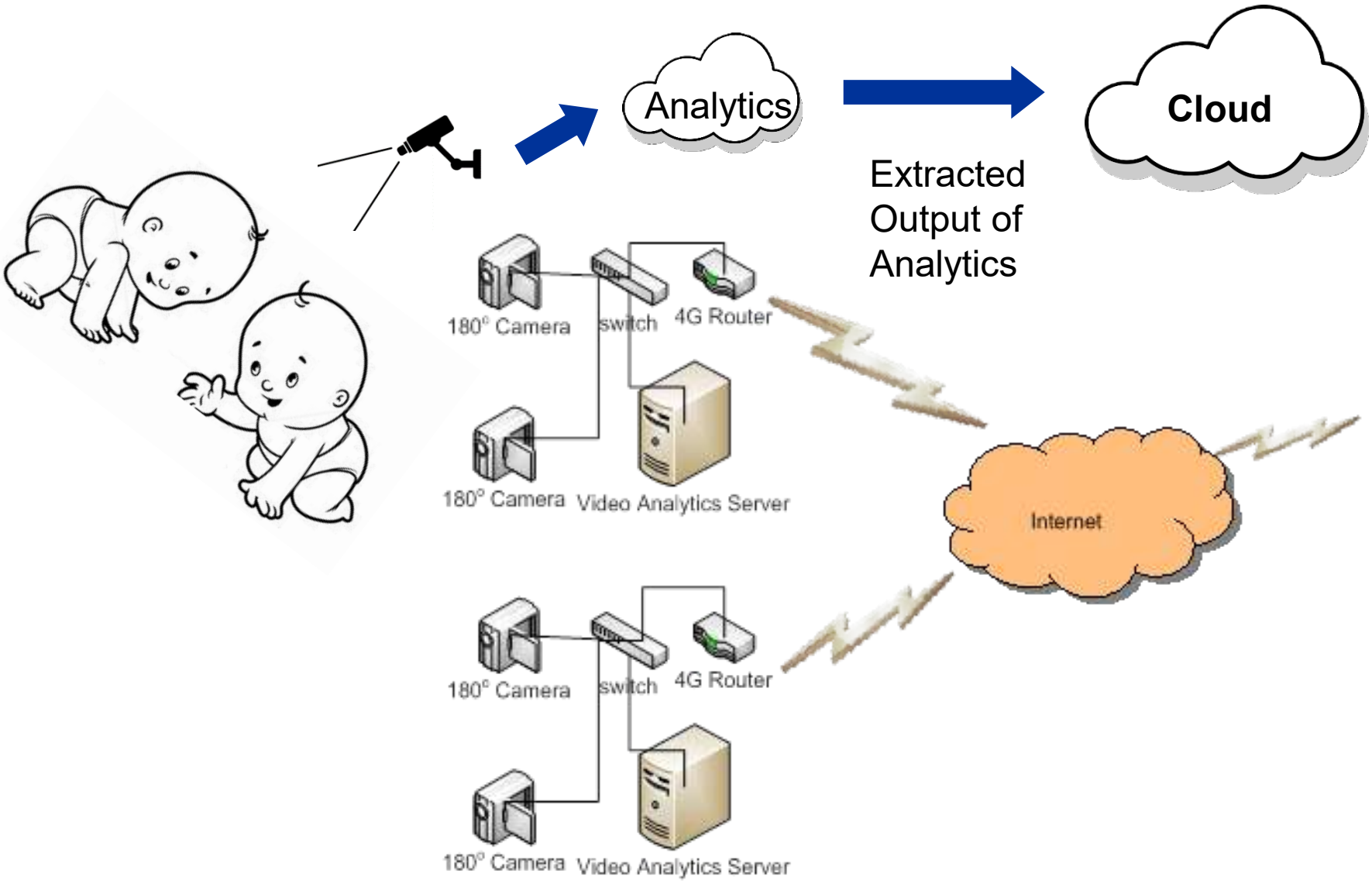


Reference

- The overview of North America real time video surveillance market revenue by application, from year 2014 to year 2025 (USD Million), available at Video analytics market trend, https://www.grandviewresearch.com/industry-analysis/video-analytics-market.
- The overview of real time video analytics applications of future interests, *International trends in video surveillance*, https://www.uitp.org/sites/default/files/cck-focus-papers-files/1809-Statistics%20Brief%20-%20Videosurveillance-Final.pdf.

AI at the Edge: The Little Ripper Lifesaver UAV, https://www.youtube.com/watch?v=QJOMfDyhUyo

Vision intelligence: Edge



Vision intelligence: Edge

- Volume and Velocity: processing and storing such huge amounts of data which is gathered in real-time.
- Security: devices can be located in sensitive environments, control vital systems or send private data.
- Bandwidth: if devices constantly send the sensor and video data, it will hog the internet and cost a fortune.
- Real-time Data Capture, Storage, Processing, Analytics, Knowledge Discovery, Decision-making and Actuation.
- Less Latency and Faster Response.

Deep learning library for edge computing

Library	CNTK [56]	Chainer [57]	TensorFlow [58]	DL4J [59]	TensorFlow Lite [60]	MXNet [61]	(Py)Torch [62]	CoreML [63]	SNPE [53]	NCNN [64]	MNN [65]	Paddle-Mobile [66]	MACE [67]	FANN [68]
Owner	Microsoft	Preferred Networks	Google	Skymind	Google	Apache Incubator	Facebook	Apple	Qualcomm	Tencent	Alibaba	Baidu	XiaoMi	ETH Zrich
Edge Support	×	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Android	×	×	×	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	×
iOS	×	×	×	×	×	✓	✓	✓	×	✓	✓	✓	✓	×
Arm	×	×	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓
FPGA	×	×	×	×	×	×	✓	×	×	×	×	✓	×	×
DSP	×	×	×	×	×	×	×	×	✓	×	×	×	×	×
GPU	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×
Mobile GPU	×	×	×	×	✓	×	×	✓	✓	✓	✓	✓	✓	×
Training Support	✓	✓	✓	✓	×	✓	✓	×	×	×	×	×	×	✓

Reference:
Convergence of Edge Computing and Deep Learning: A Comprehensive Survey, https://arxiv.org/pdf/1907.08349.pdf

AI edge computing hardware and systems

	Owner	Production	Feature
Integrated Commodities	Microsoft	Data Box Edge [29]	Competitive in data preprocessing and data transmission
	Intel	Movidius Neural Compute Stick [30]	Prototype on any platform with plug-and-play simplicity
	NVIDIA	Jetson [31]	Easy-to-use platforms that runs in as little as 5 Watts
	Huawei	Atlas Series [32]	An all-scenario AI infrastructure solution that bridges "device, edge, and cloud"
AI Hardware for Edge Computing	Qualcomm	Snapdragon 8 Series [33]	Powerful adaptability to major DL frameworks
	HiSilicon	Kirin 600/900 Series [34]	Independent NPU for DL computation
	HiSilicon	Ascend Series [35]	Full coverage from the ultimate low energy consumption scenario to high computing power scenario
	MediaTek	Helio P60 [36]	Simultaneous use of GPU and NPU to accelerate neural network computing
	NVIDIA	Turing GPUs [37]	Powerful capabilities and compatibility but with high energy consumption
	Google	TPU [38]	Stable in terms of performance and power consumption
	Intel	Xeon D-2100 [39]	Optimized for power- and space-constrained cloud-edge solutions
	Samsung	Exynos 9820 [40]	Mobile NPU for accelerating AI tasks
Edge Computing Frameworks	Huawei	KubeEdge [41]	Native support for edge-cloud collaboration
	Baidu	OpenEdge [42]	Computing framework shielding and application production simplification
	Microsoft	Azure IoT Edge [43]	Remotely edge management with zero-touch device provisioning
	Linux Foundation	EdgeX [44]	IoT edge across the industrial and enterprise use cases
	Linux Foundation	Akraino Edge Stack [45]	Integrated distributed cloud edge platform
	NVIDIA	NVIDIA EGX [46]	Real-time perception, understanding, and processing at the edge
	Amazon	AWS IoT Greengrass [47]	Tolerance to edge devices even with intermittent connectivity
	Google	Google Cloud IoT [48]	Compatible with Google AI products, such as TensorFlow Lite and Edge TPU

Reference: Convergence of Edge Computing and Deep Learning: A Comprehensive Survey, https://arxiv.org/pdf/1907.08349.pdf

Case study: Face-based access control

Face landmark detection with cheaper HOG
Facial recognition via comparison of 2x 128D mapped face and euc dist (~ <0.4)
Frame optimization

External, low-res webcam with manual focus capability

1. **Card detection**
2. **Info capture** with barcode/QR scanning

Central app server with exposed API

Basic face spoof detection

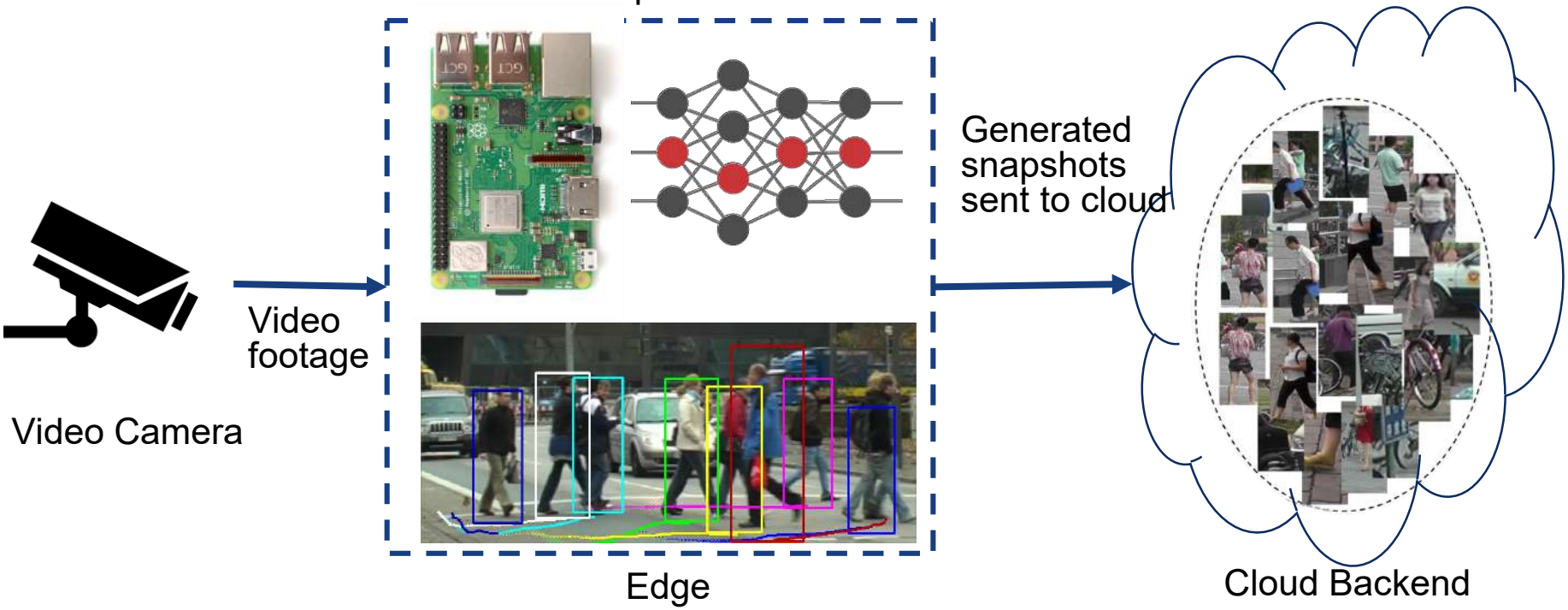
Case study: Person re-identification

Person re-identification is the task of recognizing an individual who has previously been observed over a camera network, e.g., *on-line tracking* of individuals over different cameras, and *off-line retrieval* of the video sequences containing an individual of interest, whose image is given as a query.

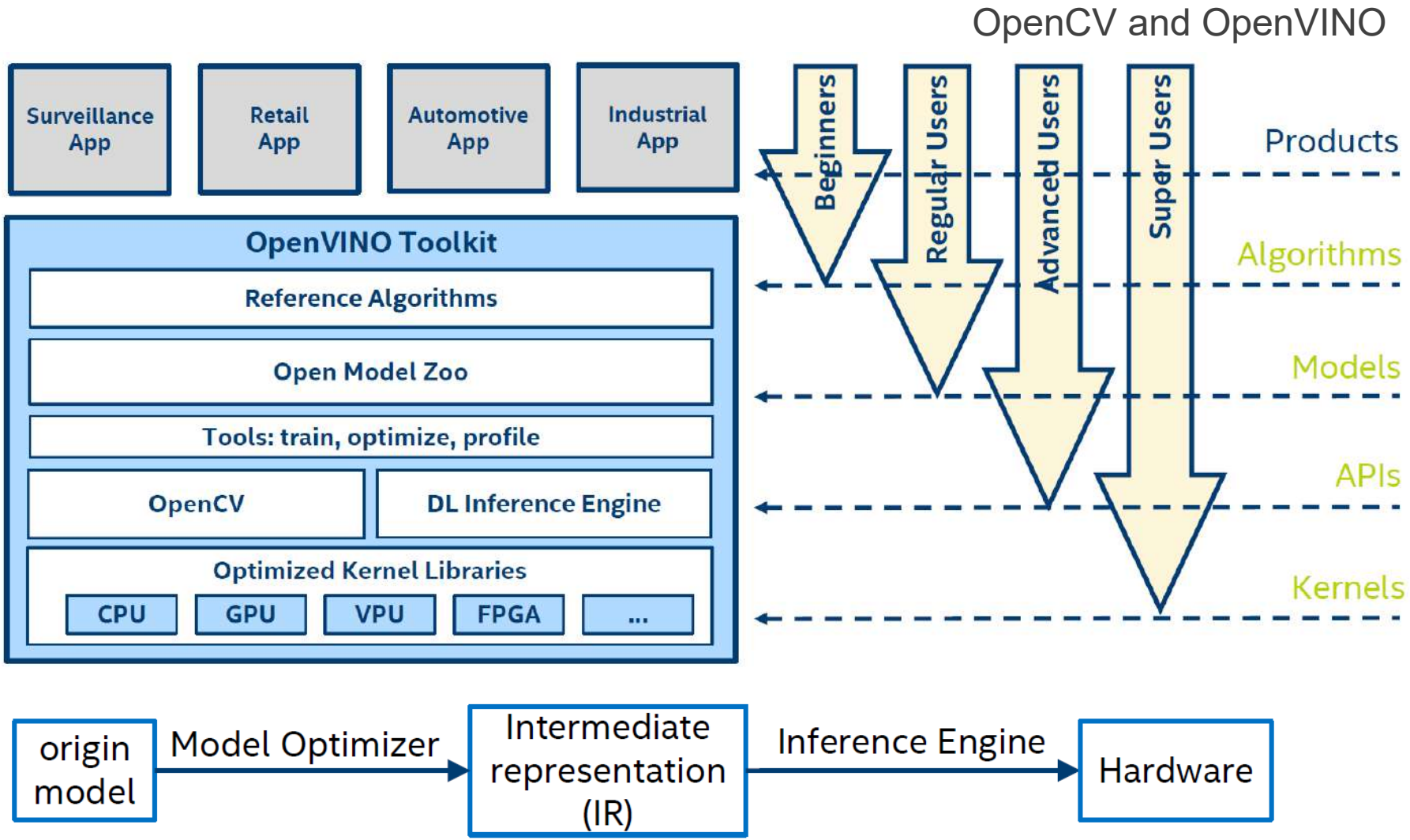
Step1: Video camera records surveillance footage

Step 2: Edge performs person detecting and tracking using compressed model and generates image snapshots of detected people with time stamp and camera location

Step3: Backend processes re-identification query based on gallery of image snapshots.



Example: OpenVINO (Intel)



Reference: OpenCV tutorial in CVPR 2019, https://opencv.org/cvpr-2019-tutorial.html/

Example: Atlas 200DK (Huawei)

Atlas 200 DK AI Developer Kit

Powered by the Huawei Ascend 310 processor, the Atlas 200 DK AI Developer Kit helps AI application developers quickly get familiar with the development environment. The device provides external ports for developers to quickly and easily access and use the powerful processing capability of the Ascend 310 processor. Equipped with the Atlas 200 AI accelerator module as its core component, the Atlas 200 DK AI Developer Kit provides the main service ports of the Ascend 310 processor on the expansion baseboard through a high-speed connector. Thanks to the full-stack, all-scenario capability of the Ascend processor, programs can be developed on the Atlas 200 DK AI Developer Kit in a one-off manner and deployed in device-edge-cloud scenarios with zero code modification. It can be used in various fields such as safe cities, drones, robots, video servers, and smart gates.

Reference: https://e.huawei.com/sg/products/cloud-computing-dc/atlas/atlas-200

Thank You.

Dr TIAN Jing
Email: tianjing@nus.edu.sg

Some slides are also used in my half-day module “Intelligence for IoT" in course "NICF- Architecting IOT Solutions (SF)", offered by ISS, NUS (https://www.iss.nus.edu.sg/executive-education/course/detail/nicf--architecting-iot-solutions-(sf)/software-systems).