VISION SYSTEMS ARCHITECTURES AND SOLUTIONS

Dr TIAN Jing

tianjing@nus.edu.sg

© 2021 National University of Singapore. All Rights Reserved

Key considerations of vision systems (1/6)

- Analytics objective
- Acquisition
- Performance index
- Deployment
- User concerns
- 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

© 2021 National University of Singapore. All Rights Reserved

Key considerations of vision systems (2/6)

- Analytics objective
- Acquisition
- Performance index
- Deployment
- User concerns
- Business model

- Light-Sensitive Camera: Measures from the visible part of the electromagnetic spectrum, e.g., RGB camera.
- Multi/Hyper-Spectral Sensors: Measure from a broader part of the electromagnetic spectrum (than the light-sensitive cameras), e.g., infrared, thermal.
- Range Sensor: A device that measures the distance from the camera to a target, e.g., laser, radar, Lidar.
- Additional equipment needed (such as external lighting)

Lente	2.8mm	3.6mm	6mm	8mm	12mm	16mm
Angulo	90°	69°	50°	33°	22°	18°
Distancia	3~5 M	8 M	15 M	20~30 M	40~50 M	70~80 M

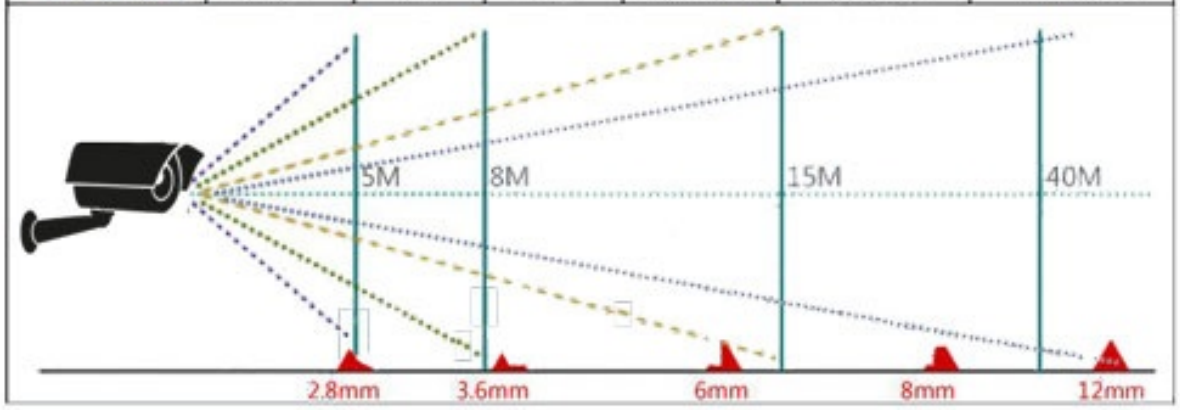


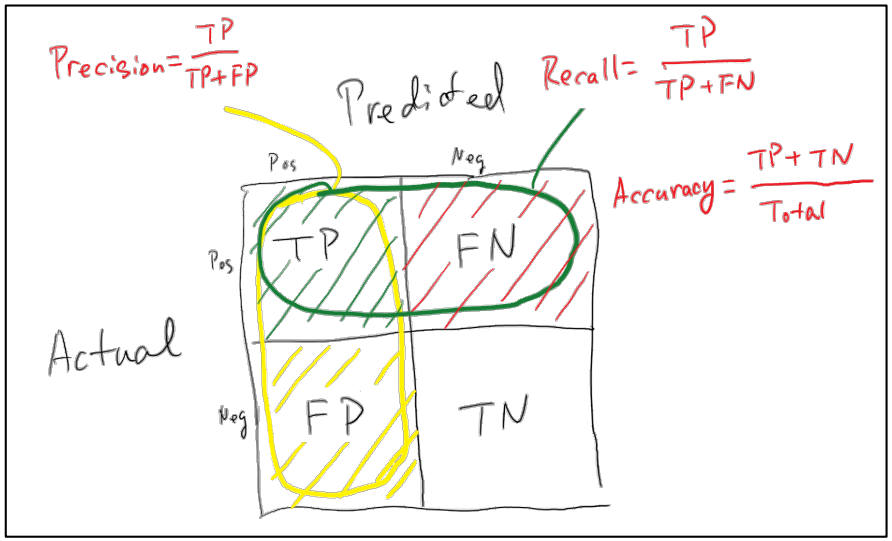
Photo: https://www.quora.com/What-is-the-average-angle-of-a-CCTV-camera

© 2021 National University of Singapore. All Rights Reserved

Key considerations of vision systems (3/6)

- Analytics objective
- Acquisition
- Performance index
- Deployment
- User concerns
- 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/

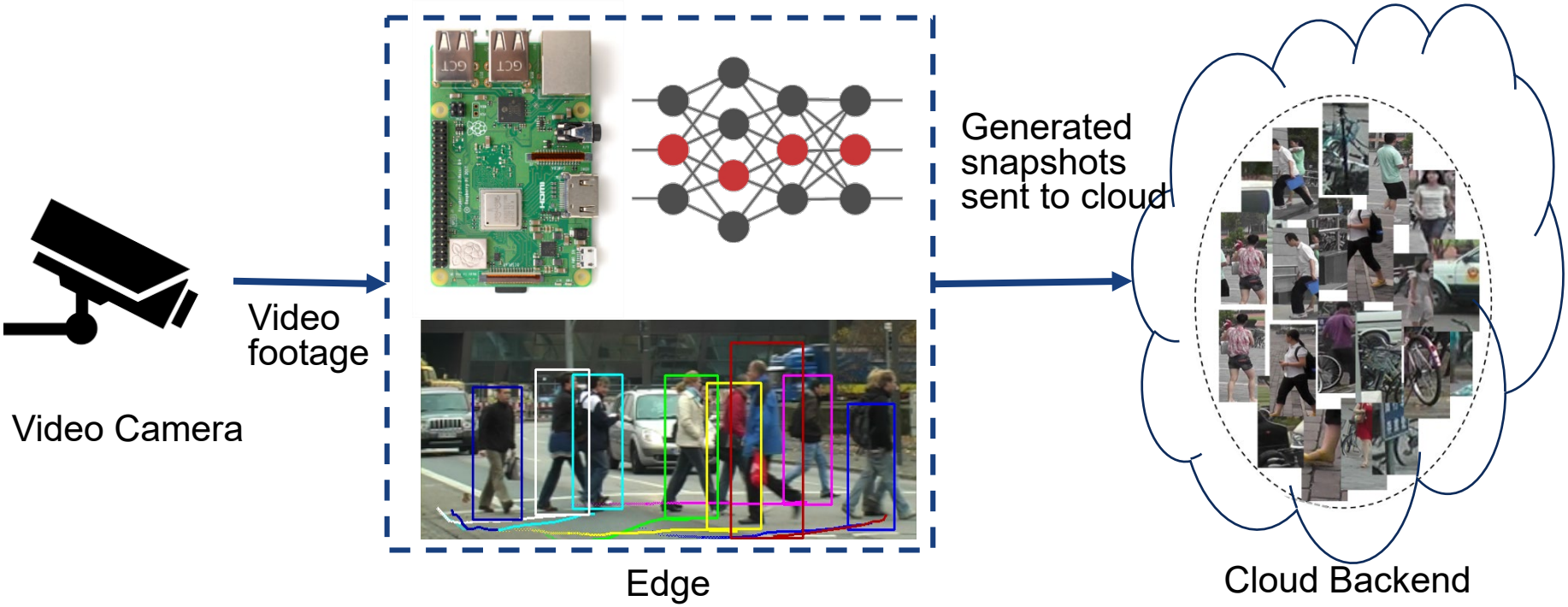
© 2021 National University of Singapore. All Rights Reserved

Key considerations of vision systems (4/6)

- Analytics objective
- Acquisition
- Performance index
- Deployment
- User concerns
- Business model

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

Example: 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.



© 2021 National University of Singapore. All Rights Reserved

Key considerations of vision systems (5/6)

- Analytics objective
- Acquisition
- Performance index
- Deployment
- User concerns
- Business model

- What are the legal restrictions which may exist in your country?
- Ethical considerations of data usage.

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

© 2021 National University of Singapore. All Rights Reserved

Key considerations of vision systems (6/6)

- Analytics objective
- Acquisition
- Performance index
- Deployment
- User concerns
- 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

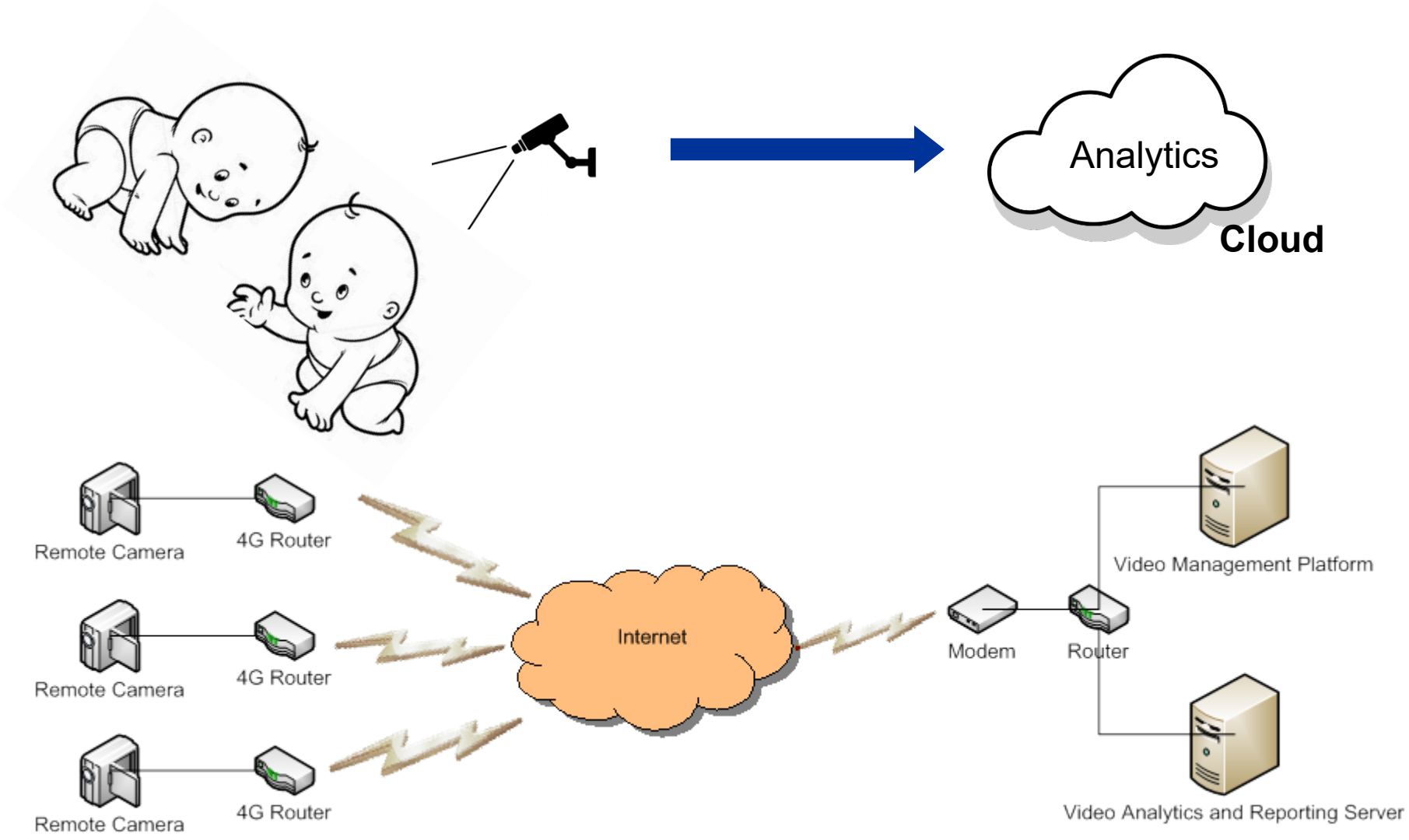
Maintain and operate

- 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

© 2021 National University of Singapore. All Rights Reserved

Vision intelligence: Cloud



© 2021 National University of Singapore. All Rights Reserved

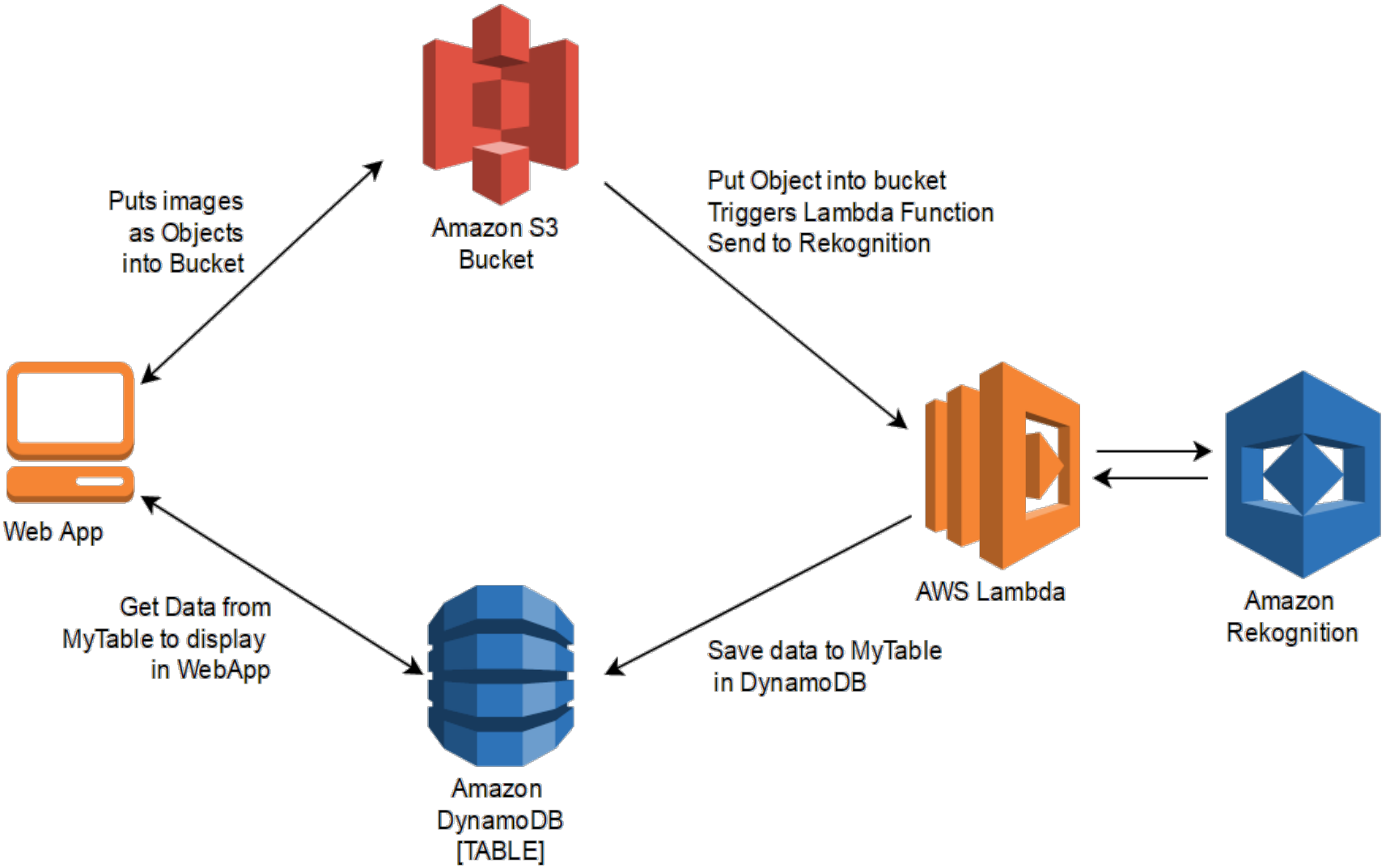
Vision intelligence: Cloud

- 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.

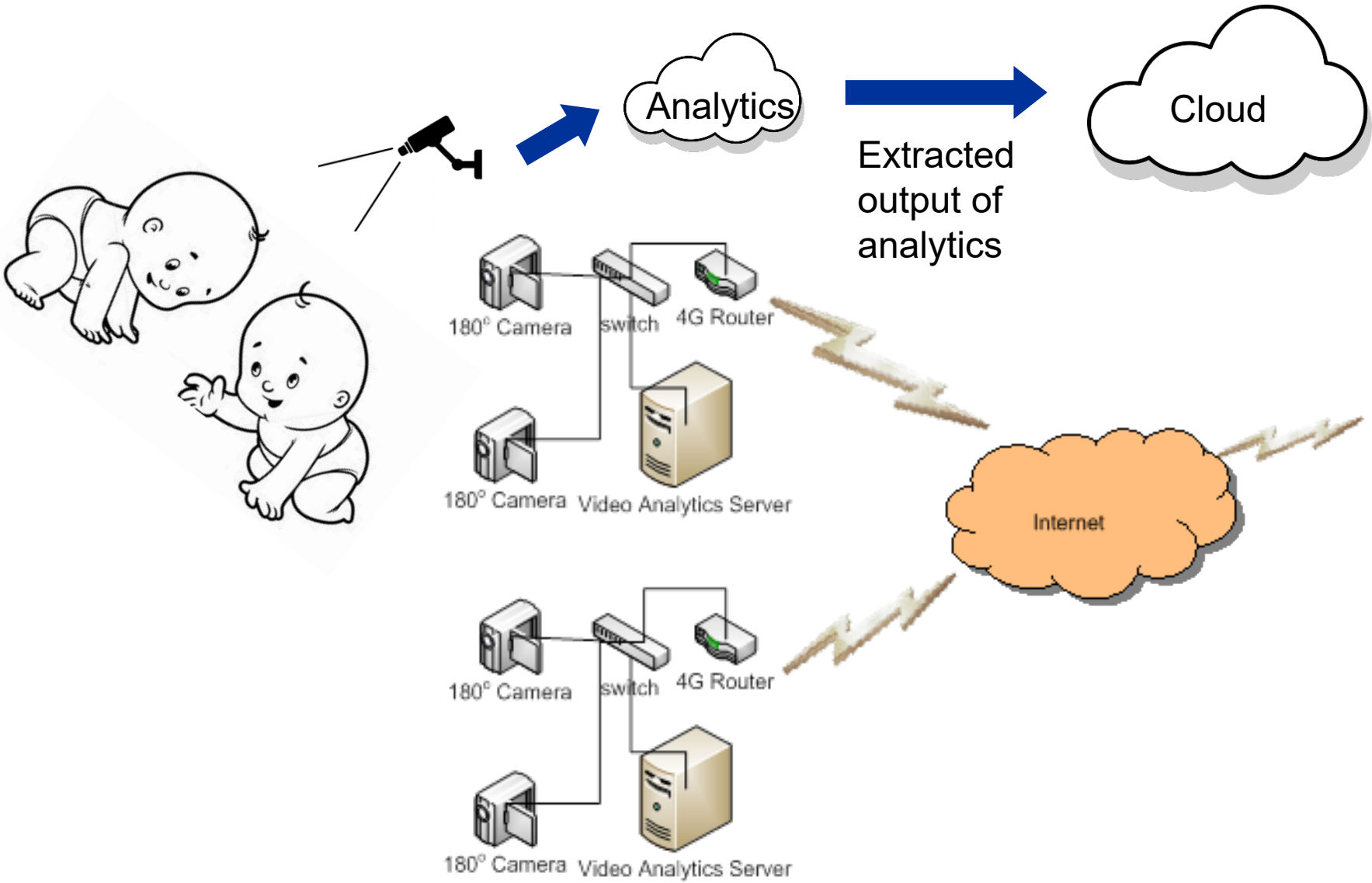
© 2021 National University of Singapore. All Rights Reserved

Example: AWS Rekognition



© 2021 National University of Singapore. All Rights Reserved

Vision intelligence: Edge



© 2021 National University of Singapore. All Rights Reserved

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.

© 2021 National University of Singapore. All Rights Reserved

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

© 2021 National University of Singapore. All Rights Reserved

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

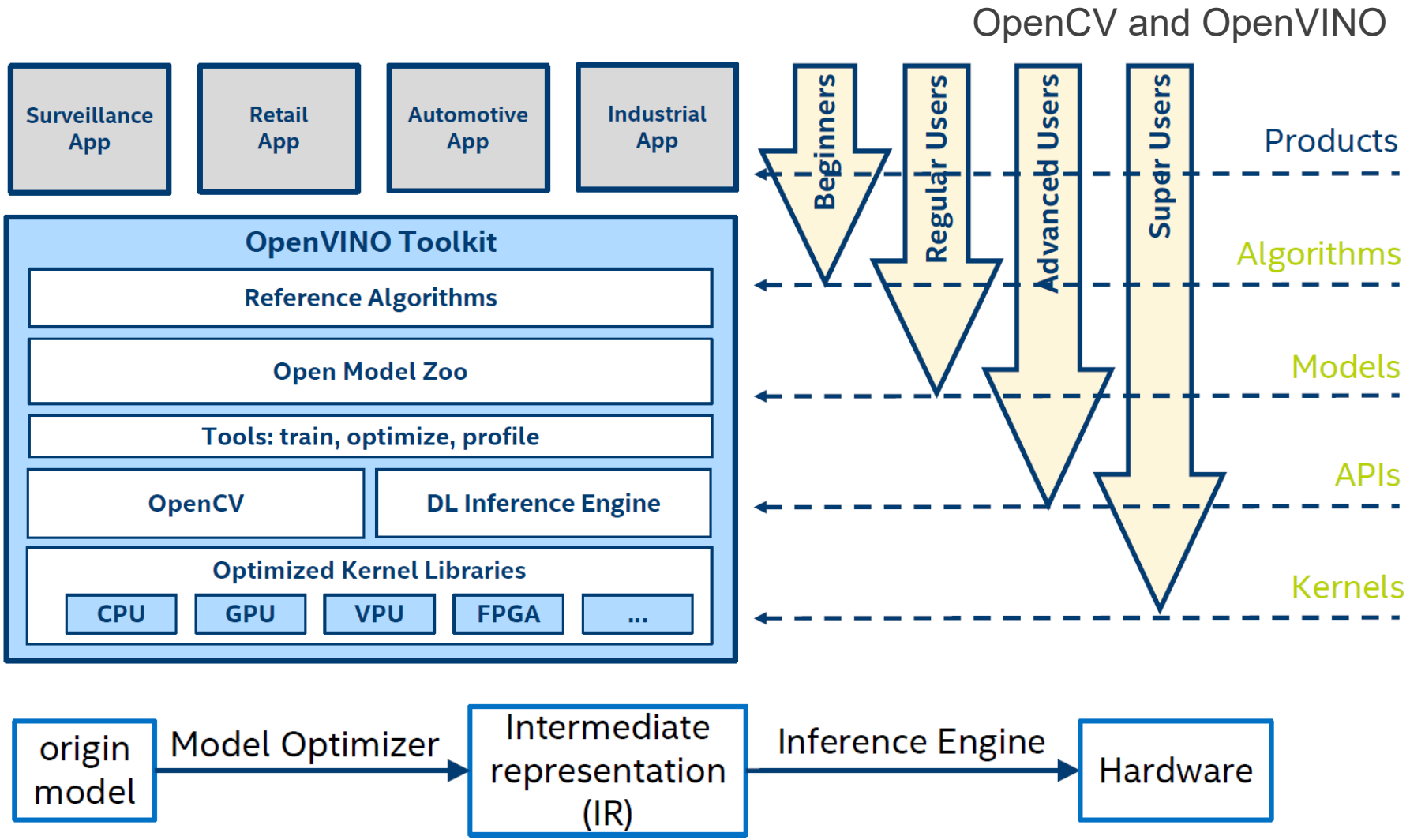
Reference: Convergence of Edge Computing and Deep Learning: A Comprehensive Survey, https://arxiv.org/pdf/1907.08349.pdf

© 2021 National University of Singapore. All Rights Reserved

Case study: Face-based access control

© 2021 National University of Singapore. All Rights Reserved

Example: OpenVINO (Intel)



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

© 2021 National University of Singapore. All Rights Reserved

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

© 2021 National University of Singapore. All Rights Reserved

Thank You.

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

© 2021 National University of Singapore. All Rights Reserved