SMU
SINGAPORE MANAGEMENT UNIVERSITY

School of Information Systems

The Cognitive Edge:
Promoting Energy-Efficient, Collaborative Sensing by IoT & Personal Devices

Archan Misra

SmartEdge 2019, March 14, 2019

My Research History

2010 - 2015 (2010 – 2015)

Mobile Sensing & Analytics
- Indoor Location
- Group Detection
- Queuing Detection

Key Research Thrusts
- Fusion of multi-modal sensing *(inertial)*
- Adaptive sampling & triggered sensing
- **Multiple live deployments** *(campus, malls, museums)* **+ licensing**

2014 - 2018 (2014 – 2018)

Wearable Sensing & Systems
- Eating *(Annapurna)*
- In-Store Shopping *(IRIS, I4S)*
- VR+ mobile *(Empath-D)*

Key Research Thrusts
- Optimize *(Energy, Accuracy, Latency)* tradeoffs
- Multi-modal sensor fusion *(inertial, image)*

2018-2022 (2022)

Wearable + IoT Systems
- Batteryless Wearables
- Wireless/RFID Sensing
- Fine-grained Gestural Tracking

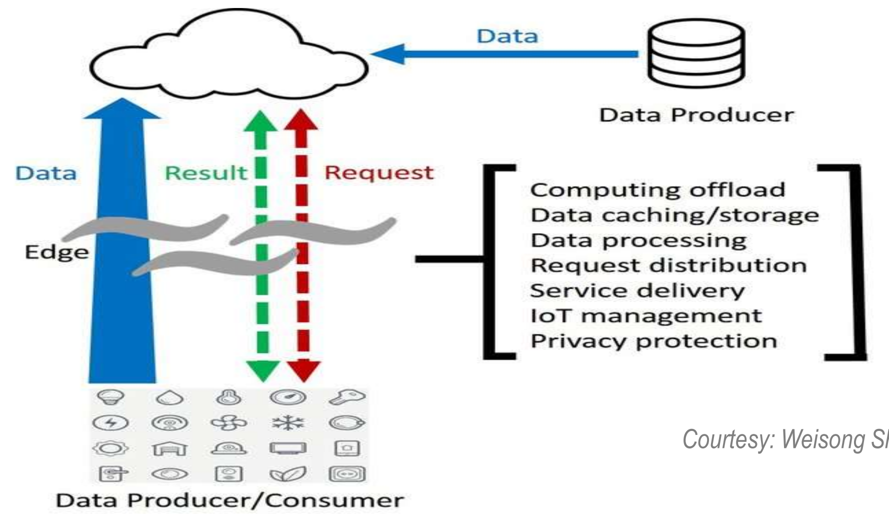
Key Research Thrusts
- Batteryless Sensing
- Wireless *(Commn, Power, Sensing)*
- Collaborative ML at edge *(voice, video, inertial, wireless)*

Edge Computing at Present

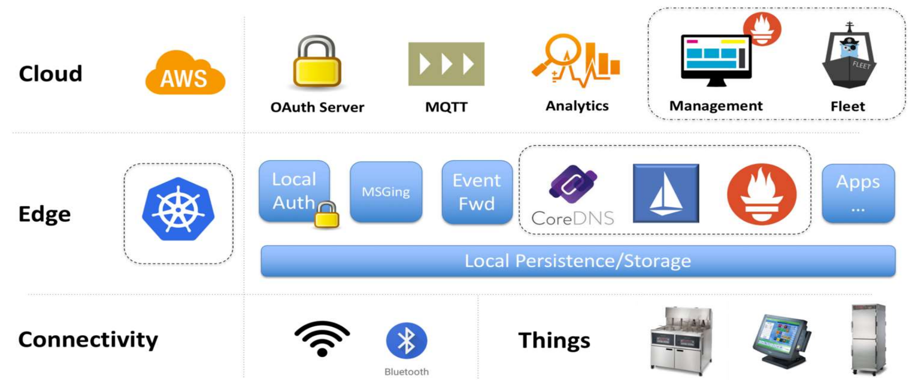
- Offload computation to a nearby, powerful-computational entity

- **Advantages**
  - Low-latency, real time ML pipelines
  - Data privacy
  - Energy-efficiency
  - Content caching

- **Isolated Interaction between individual device & "cloudlet"**



Courtesy: Weisong Shi



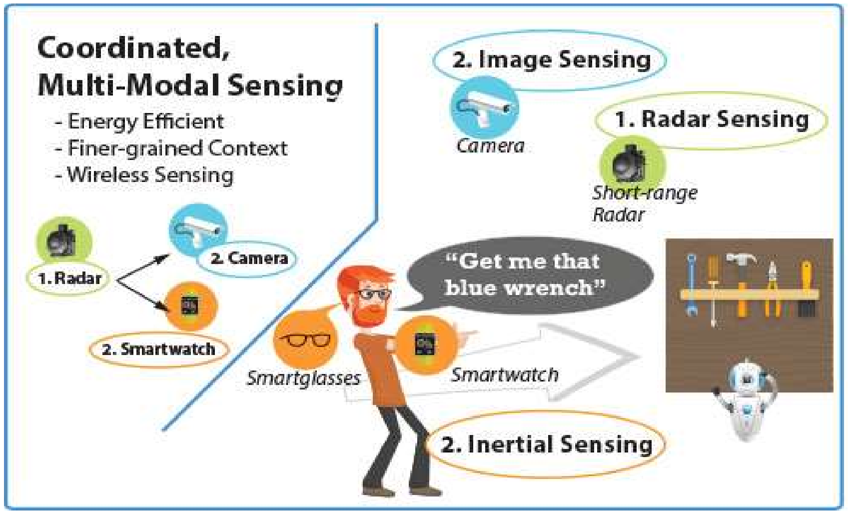
SMU Singapore Management University | School of Information Systems

This Talk: Edge for Coordination

Edge: A platform that allows real-time coordination between multiple IoT & wearable/mobile devices

DS: Distributed & Triggered Sensing

- Use Cheaper Sensors to Trigger More Expensive Sensors in tight coordination
- Optimize Sensing & Communication Overhead *(Stochastic Optimization)*

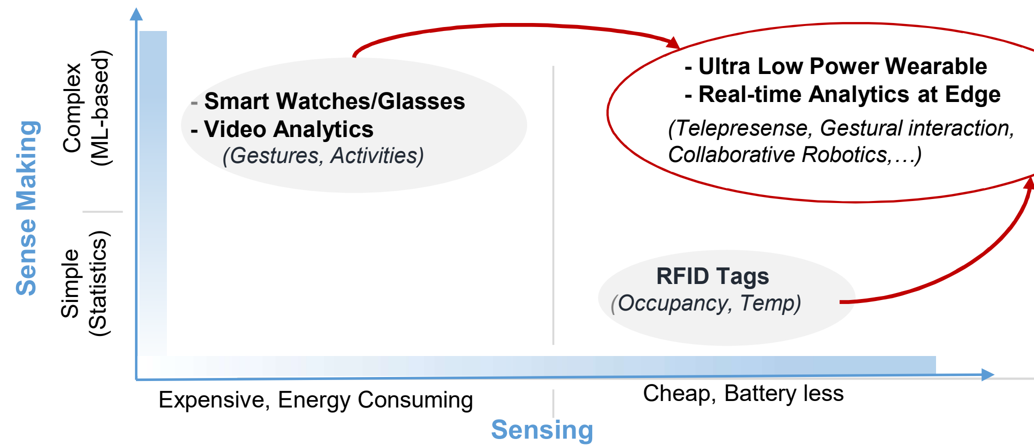


EA: Collaborative ML-based Edge Inferencing

- Distribute Inferencing Pipelines across multiple pervasive devices
- Optimize processing latency & Energy Consumption

Research Vision: Fine-grained, ultra-low power, sensing & sense-making

- Campuses, Malls, Homes, Factory, Warehouses, ….
- **New scenarios & uses!!** -- Multiple occupants, device-free interaction, super-cheap wearables….



Wearable Sensing

- Drive sensor power consumption down from current *(20-200mW)* to an average of **100-500µW** → usher in **batteryless wearables**
- Enable *fine-grained* gesture & limb tracking ***(2-5cm)***

Sense-making

- Migrate complex ML-based analytics from premises/ cloud servers to **edge devices**, to support *real time (**latency= 10-50 msecs**)*, *reduced power (**100mW**)* execution

Multi-disciplinary "Systems" approach
Embedded systems + Wireless + Signal Processing + Applied Machine Learning

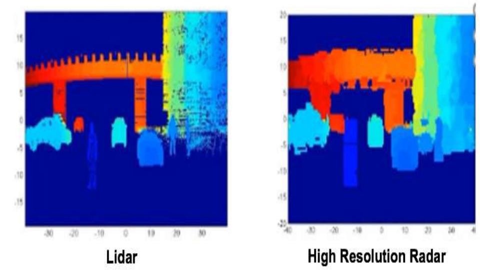
Need: Making Sensing "Cheap" & "Abundant"

2E

Two-order Reduction in Energy (2E)

Tens of thousands of sensors→ battery-powered operation (periodic replacement) is infeasible

- Campuses, factories, warehouses
- Cheap novel sensors becoming available *(e.g., short-range radar)*, but energy is a challenge

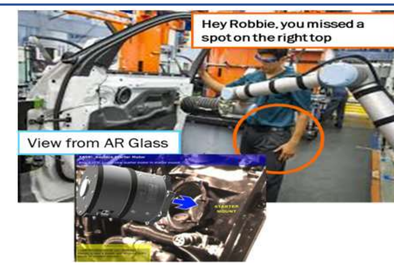


LL

Low-Latency Complex Sense-making (LL)

Require sense-making at the edge → Deep ML algorithms are too heavyweight

- **Telepresence, augmented reality, collaborative robotics**
- *Hardware co-processors still consume too much power (700mW+)*

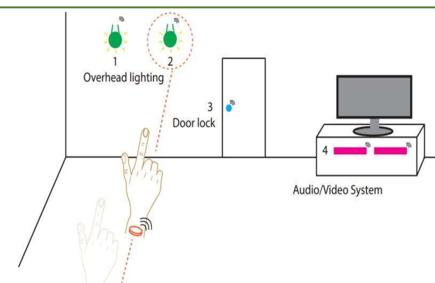


HR

High-resolution user & gesture tracking (HR)

Track multiple humans & gestures at O(cm) level precision

- Touchless **device-free interaction** with smart objects (e.g., Google's Project Soli) requires O(10cm) level tracking
- *Pure-wearable based approaches: expensive & cumbersome*



DS: Distributed & Triggered Sensing

- **RF+ Wearable Monitoring in Smart Spaces**
  - Batteryless Wearables
  - RF+ Wearables for Gym Activity Monitoring

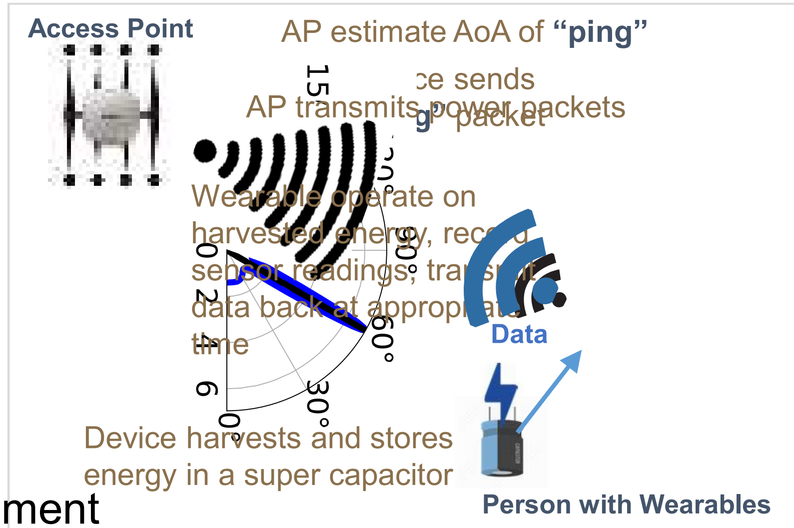
- **Human-Robot Naturalized Interaction**

DS1. Battery Free Wearable/IoT Sensors

Percom 2019

Vision

- Utilize battery-free sensors on wearables & IoT devices to provide fine-grained tracking
- **Key breakthrough**: Charge devices wirelessly via WiFi "power packet" transmissions

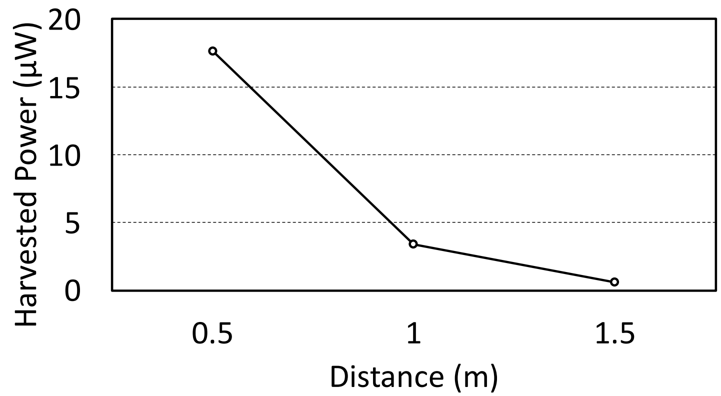


Applications

- Activity Tracking of Workers & Moving Equipment
- Product Monitoring in Warehouses
- Elderly Monitoring in smart homes

Challenges

- Low energy density using omnidirectional WiFi antenna *(< 1μW at 1.5m)*
- WiFi AP coordination to charge multiple devices



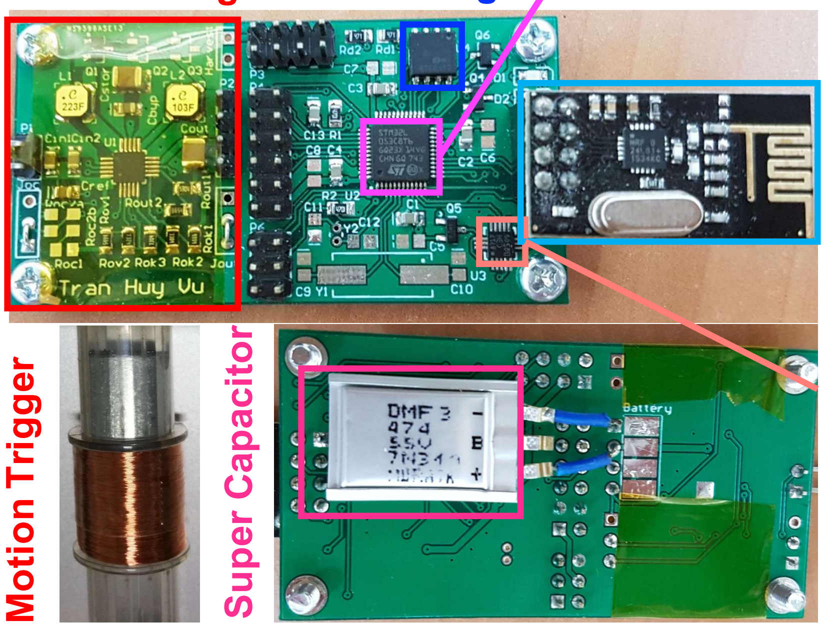
The Wearable + AP System

Power Management Storage Micro-controller

RF Comm.

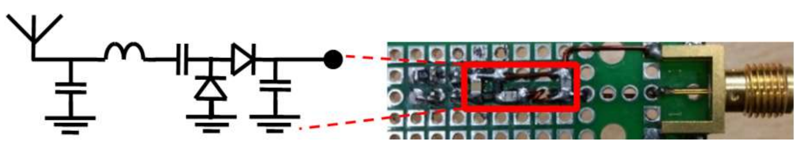
Accelerometer

Motion Trigger Super Capacitor

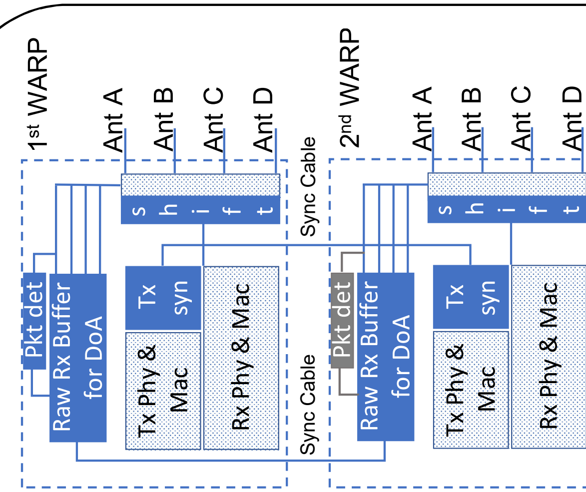


The Wearable

The Harvester:
- Matching Circuit
- Rectifier

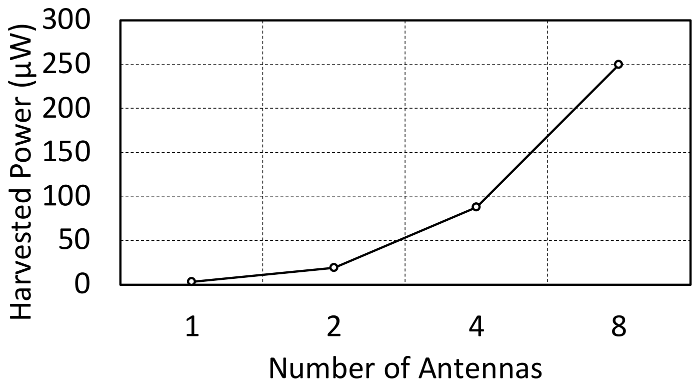


The Beamforming AP



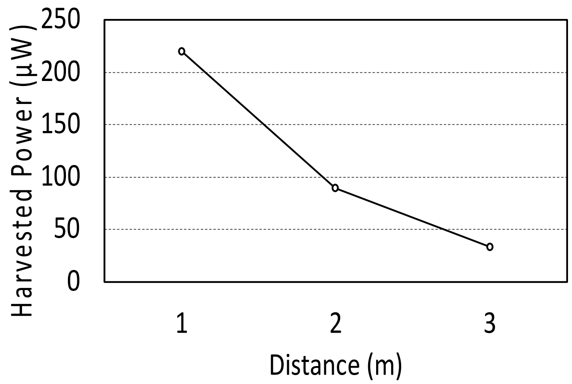
- Ping detection (nRF24L01+)
- Rx Buffer for AoA
- Tx Phase Sync for Beamforming

The WiWear Wearable: Results



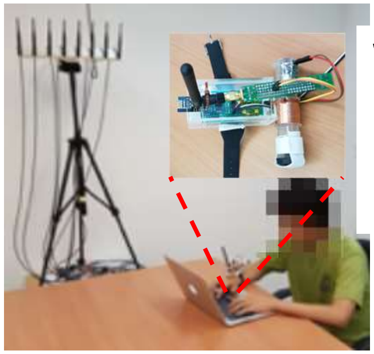
Higher efficiency:

- 1 antenna x 20dBm: ~3μW
- 8 antenna x 20dBm: > 200μW



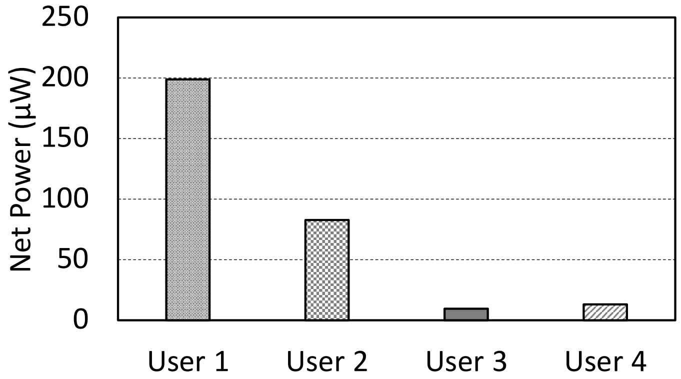
Long distance:

- 3 meters: 33μW
- Previous study (30dBm): 100μW at 2 feet, 10μW at 10 feet



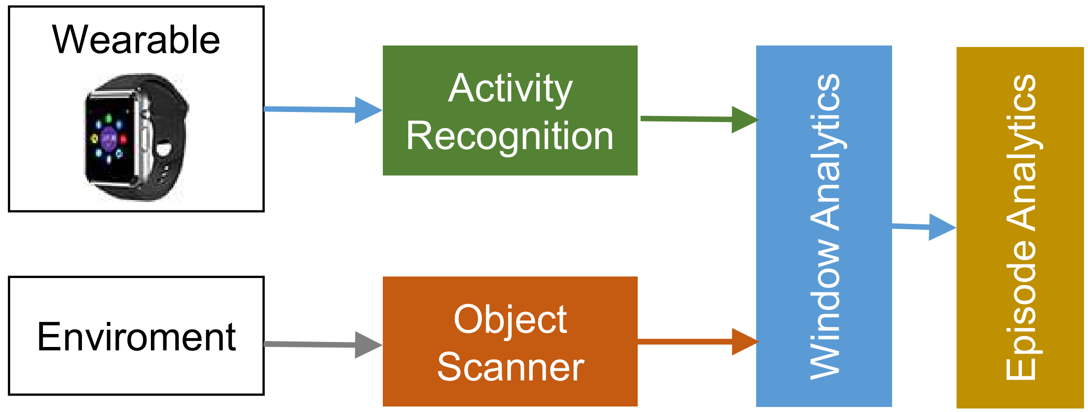
Work with users:

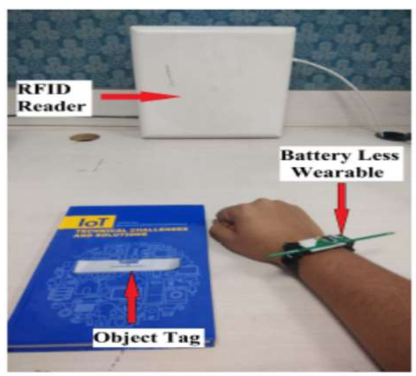
- Room: 4.5m x 3.5m
- Table: 1.1m x 1.9m
- Each user at a corner



Batteryless RFID: Table-Of-Interest

- User wears batteryless RFID tags *(with accelerometer)*
- RFID scanner *(edge device)* operates in two modes
  - **Regular**: omnidirectional, low-power
  - **Activity**: directional, high-power
- RF perturbations of object-mounted tags provides trigger *(Regular→Activity)*





Accurate Batteryless Activity Tracking

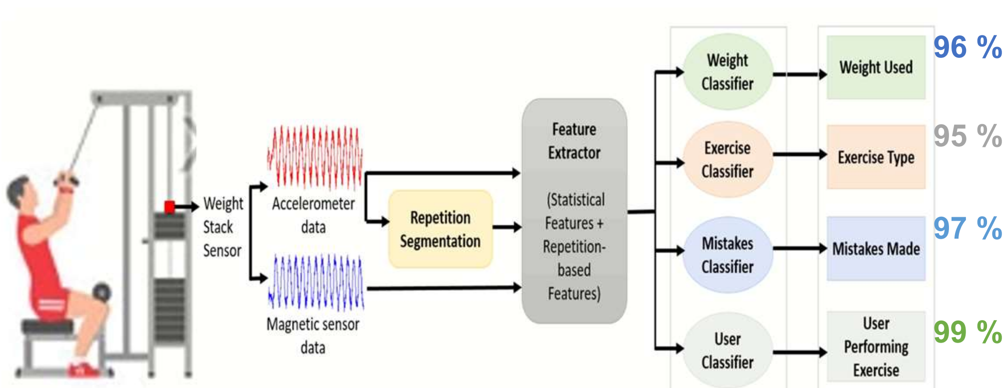
	Eating	Reading	Typing
Eating	531	13	7
Reading	6	543	19
Typing	2	19	514

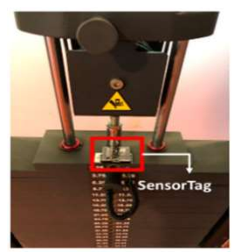
W8-Scope: IoT-based Exercise Monitoring

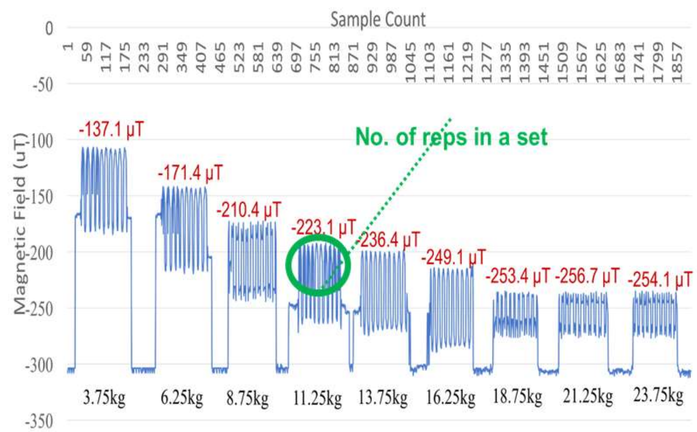
Goals:
- Quantified insights on weight stack-based exercises→ provide personalized digital coaching

Techniques:
- Simple weight stack sensor *(accelerometer+ magnetometer)* to track & understand exercises

Results:
- Longitudinal Data Collection at 2 gyms→95+% accuracy & adaptation to medium-term evolutionary behavior







Magnetic Sensor on Wt. Stack→ {Weight, Type, User}

(Radar+ Wearable) Sensing of Exercise Activities

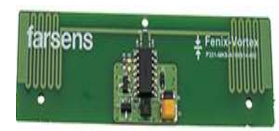
Goals:

- Fine-grained, accurate monitoring of exercise gestures of **multiple users**
- Help track exercise type, intensity (rep/set count), exercise correctness

Techniques:

- Sensor fusion: Multiple **short-range radars** plus batteryless tags
- Dynamic gestures and body movements recognition

Short-range radar

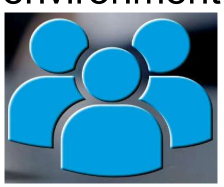


Batteryless RFID tags with accelerometer

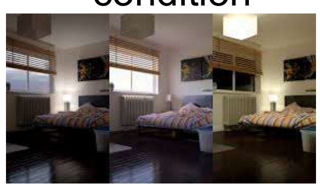
Advantages:

Non-invasive

Multi-user environment



Any lighting condition



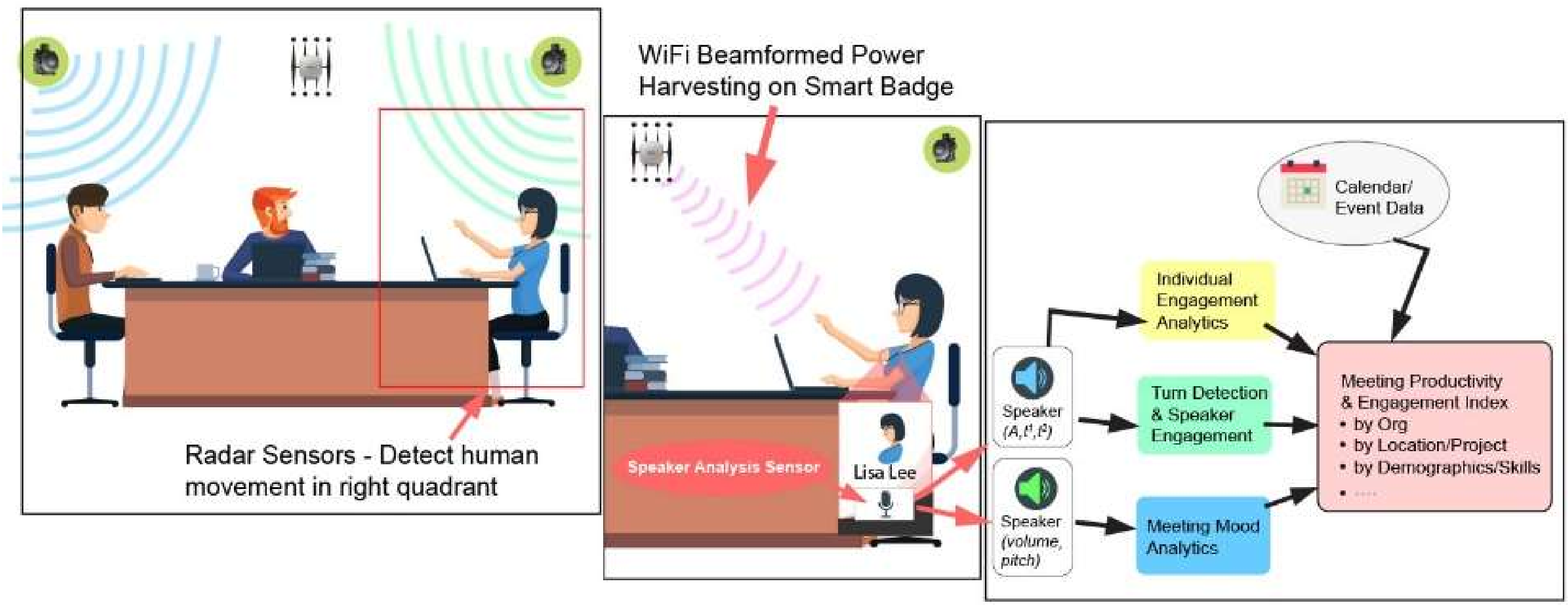
Challenges:

- Multiple reflecting objects
- Proper placement and no. of sensors
- Low sensor sampling rate

Next: Batteryless “WorkStyle” Analytics

Wearable+ Infrastructure for Office Activity Sensing:

- Batteryless mic+ accel on employee badge
- Short-range Radar+ WiFi AP
- Applications: employee productivity | wellness and work interaction analytics



Takeaways

- **New opportunities**: Edge-Coordinated Activation of wearable devices.

DS2: Human-Robot Interaction

Vision:

- Robots providing more **natural assistance** to workers
- Based on deeper understanding of worker context, using inputs from wearable & IoT devices

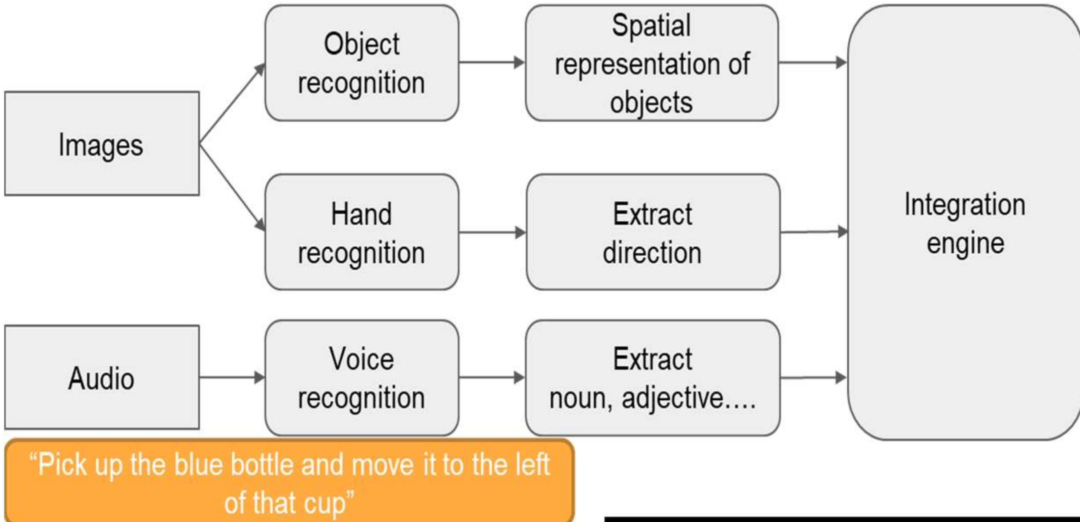
Key Novelties

- Augment vision-based analysis with other pervasive sensors *(sound, inertial sensors, etc.)*

New Opportunities

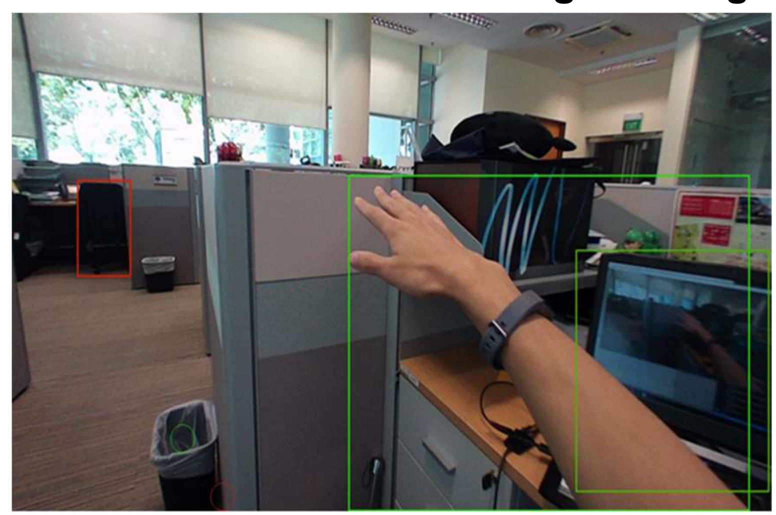
Augmental robotics: co-working by robot & human

AR+ Robot: Multi-modal Interaction

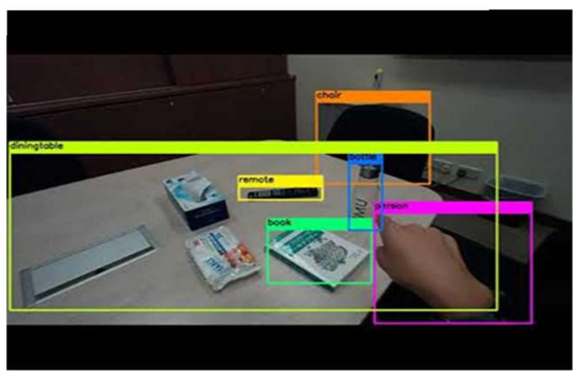


Meta-2: 2560*1440

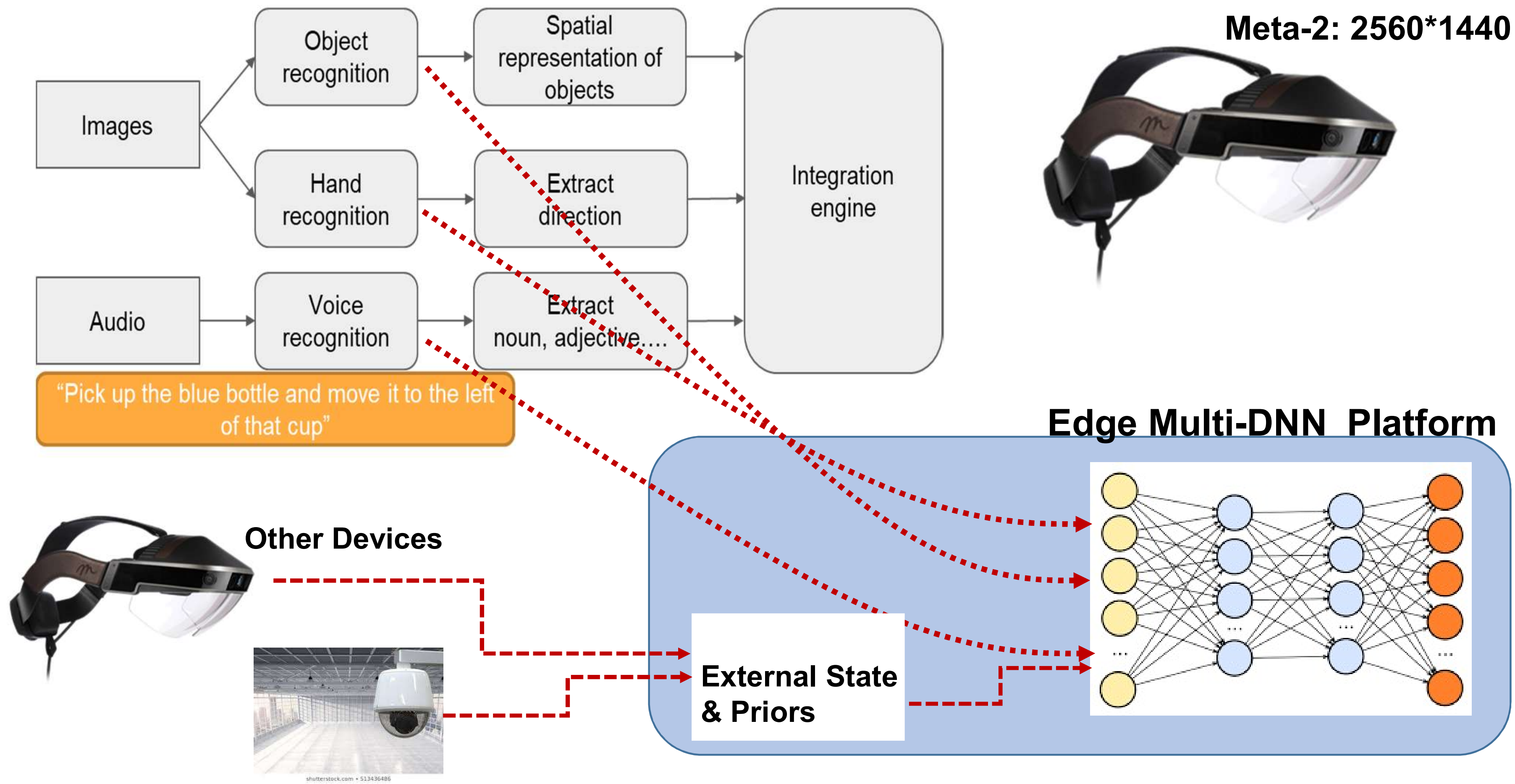
Gesture & Pointing Tracking



Vision-based Object Recognition



Edge-Supported Multi-Modal Interaction (Collaborative)



Takeaways

- **New opportunities**: Edge-Coordinated Activation of wearable devices.

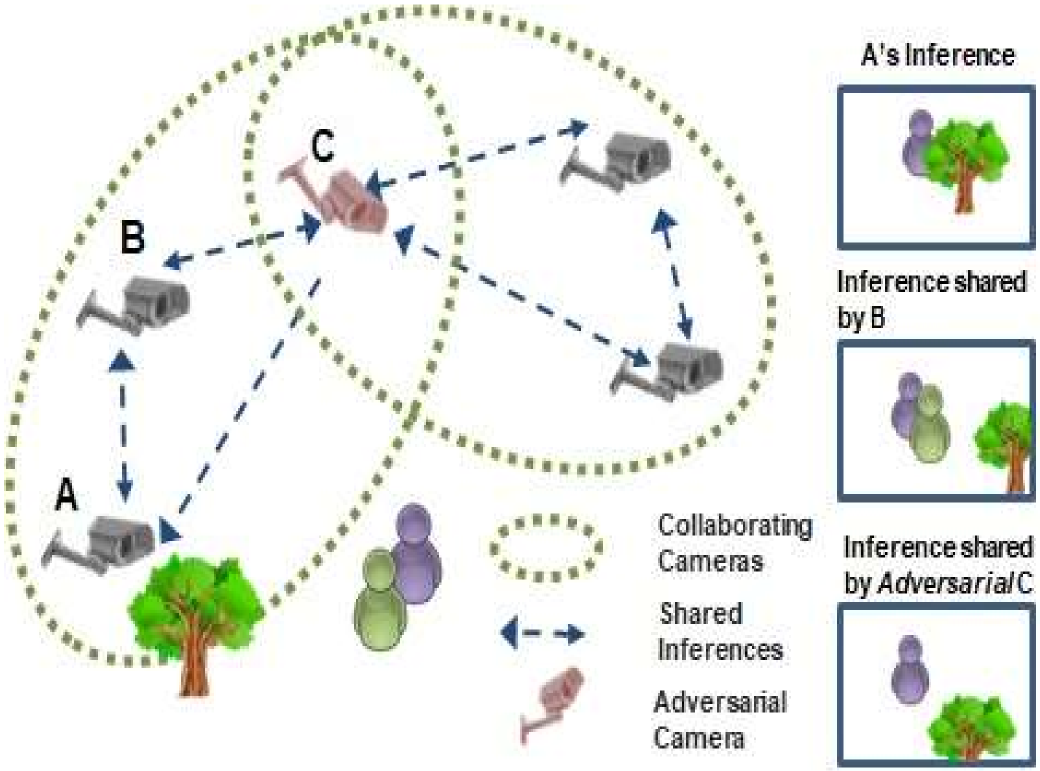
- Edge devices augmenting the sensing capabilities of wearable devices

EA: Edge-Augmented Analytics

- **Collaborative ML-based Inferencing among IoT Devices**
  - People Counting & Tracking over Multiple Cameras

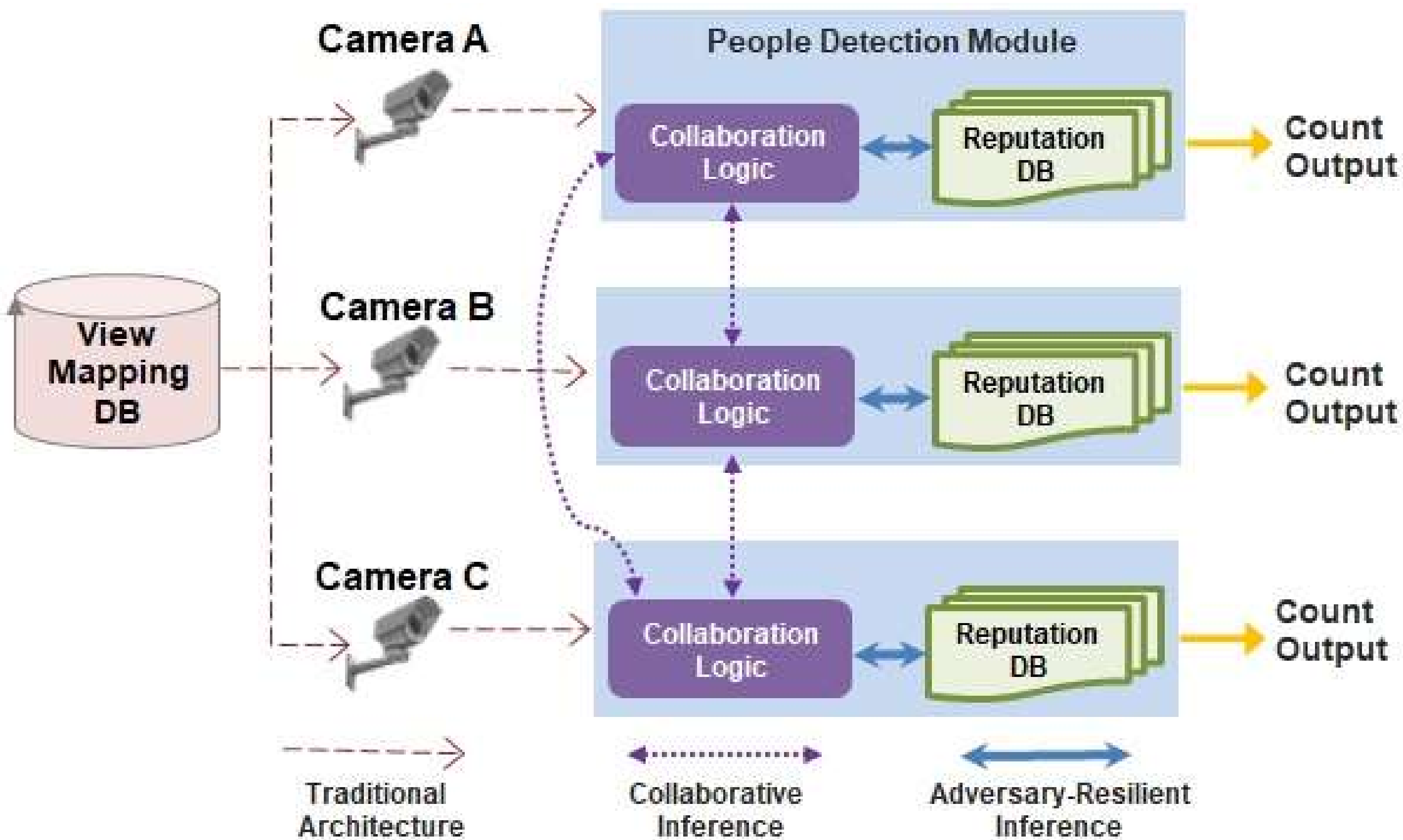
- **Cognitive Edge: Matchmaking IoT Devices**

EA1: Collaborative IoT & The Edge



- **Collaborative Sensing**
  - Adjust Inferencing Pipeline
- **Dependable Systems**
  - Resilience to Adversarial Attacks

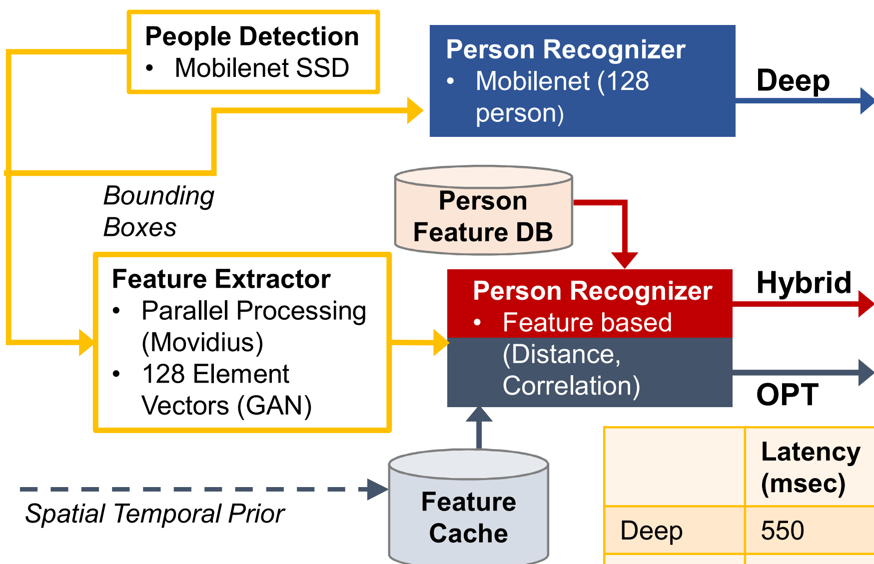
Collaborative Inferencing (Impl.)



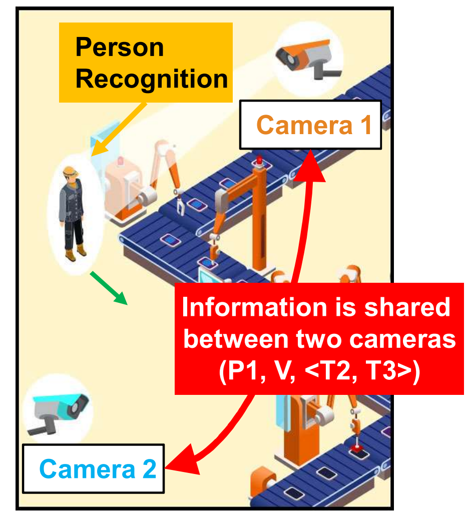
Cooperative ML@Edge: Utilize Priors

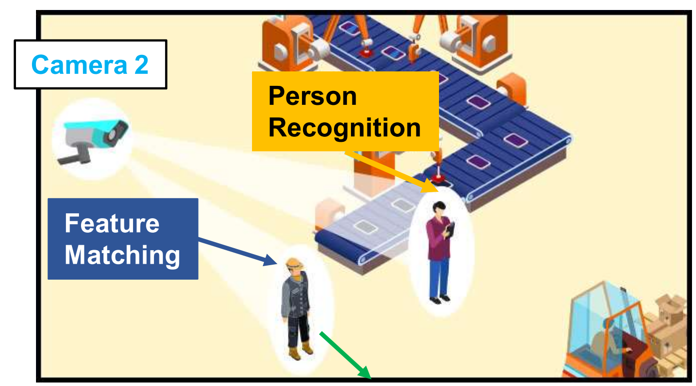
Spatiotemporal Priors

- Share Object Trajectory & Features with peer cameras *(sensors)*
- Optimize ML execution on frames with multiple objects *(different priors)*



	Latency (msec)	Accuracy
Deep	550	92%
Hybrid	170	54%
OPT	25	81%



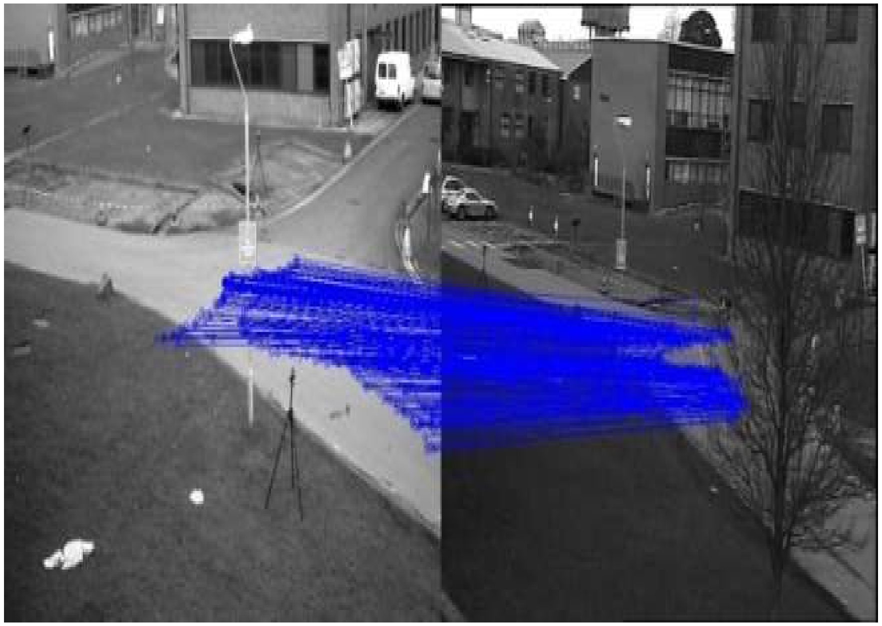


Ongoing Results

Initial analyses on the Performance Evaluation of Tracking and Surveillance 2009 benchmark dataset

Procedure

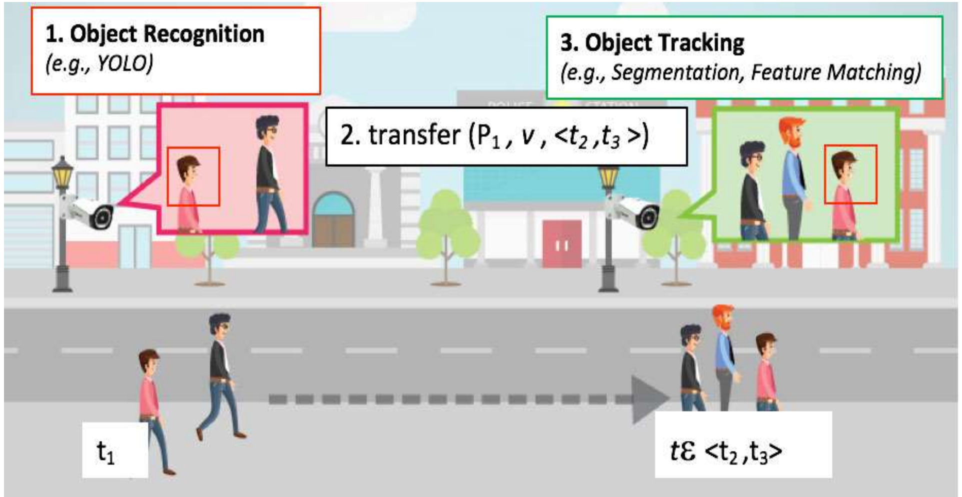
- Spatial relationship between two overlapping views "learned" from shared objects
- Homography matrix for mapping coordinates to a common plane
- Combine detections from multiple views using Non-Maximum Suppression (consider as same bounding box if overlap ratio > 0.3)



SMU Singapore Management University | School of Information Systems

Resilient, Collaborative Edge

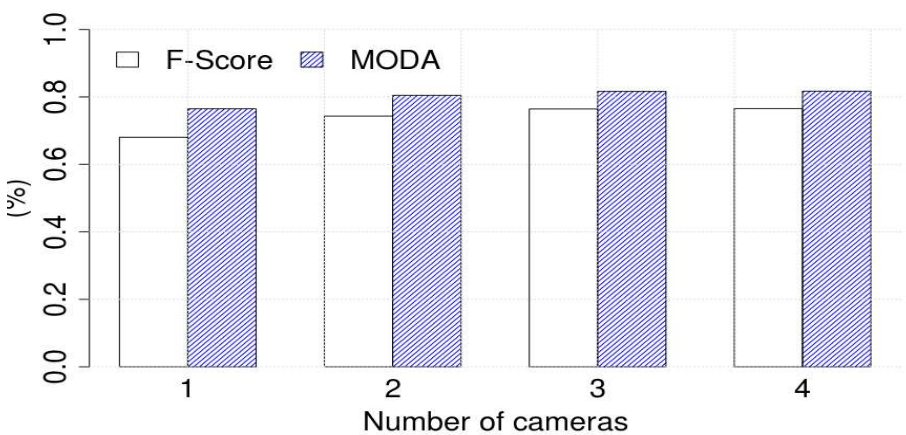
- Cascade tamper-proof "state synopsis" across IoBT devices
  - E.g., each camera passes on a time-stamped, signed set of priors
- Bayesian distribution estimation of each device's reliability/ reputation
  - Prob (device j was incorrect | observations at device i)
  - Time & Space-dependent conditional probabilities
- Adapt the pipeline optimization logic
  - Incorporate device's reputation in optimization framework

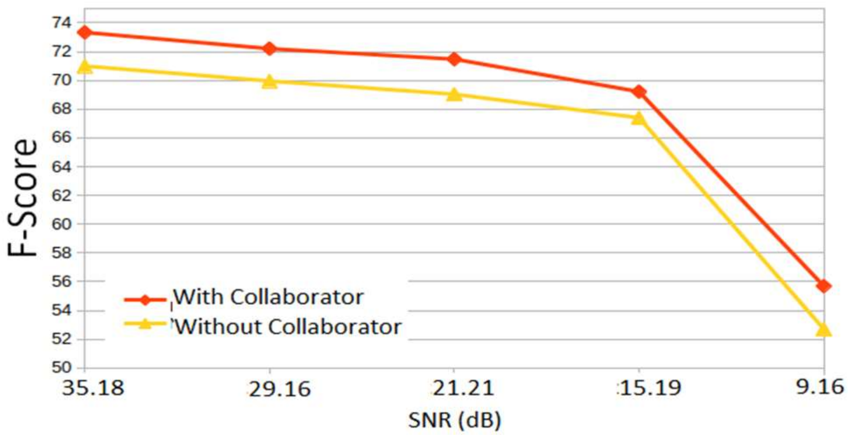


Compromised/Malicious Camera 1 can lead to cascaded failures

Early Results

PETS Dataset
(8 cameras)



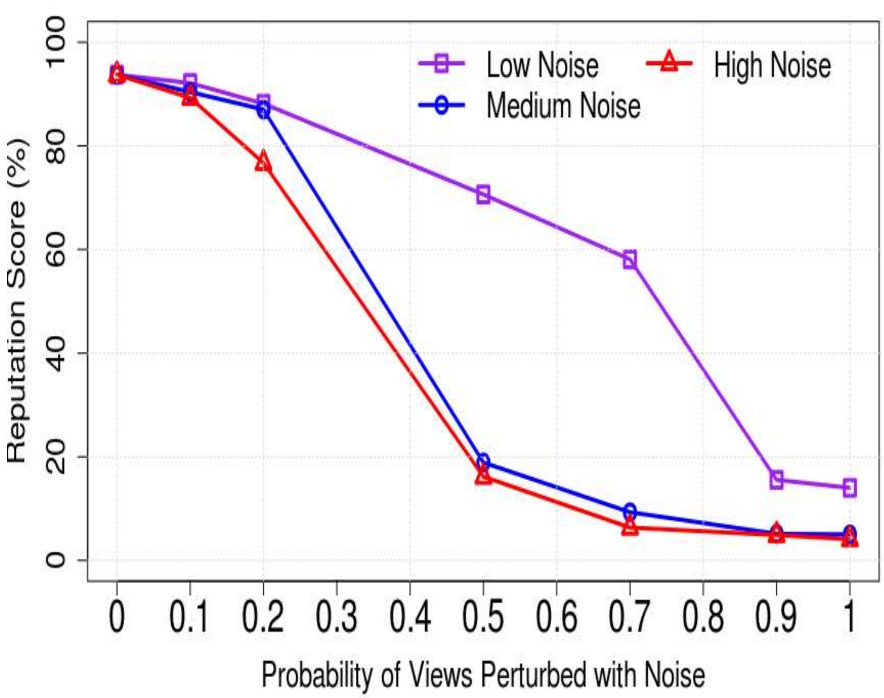


Not Resilient to Errors/Noise

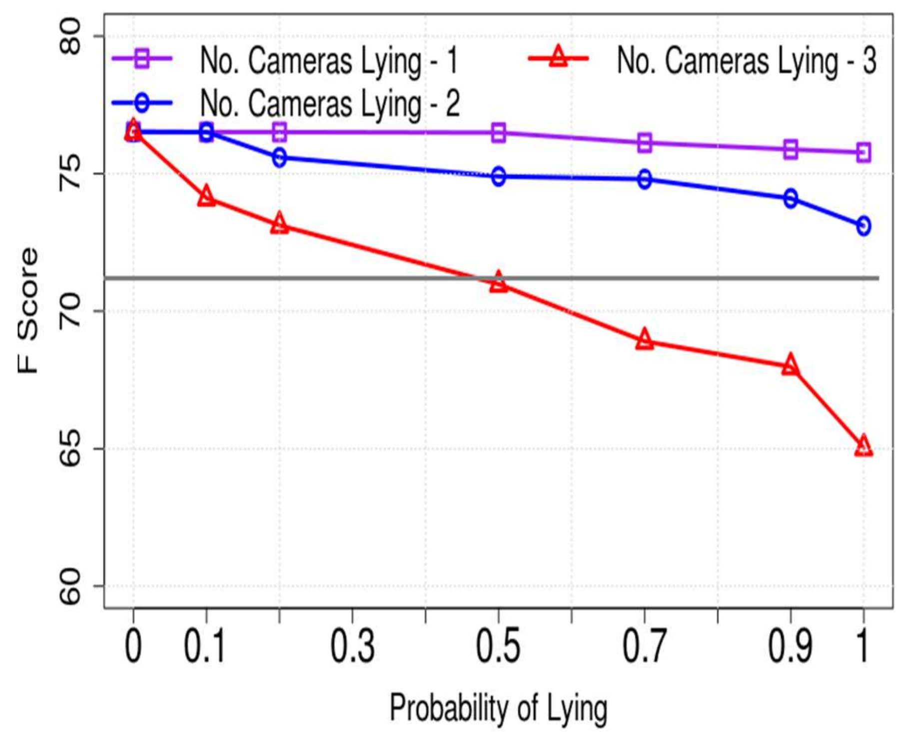
SMU Singapore Management University | School of Information Systems

Early Results (Robustness)

Rapid Drop in Reputation Score with Increasing Malicious Behavior (p)



Robust to Multiple Cameras Lying



Cooperative ML@Edge: Real-time ML State

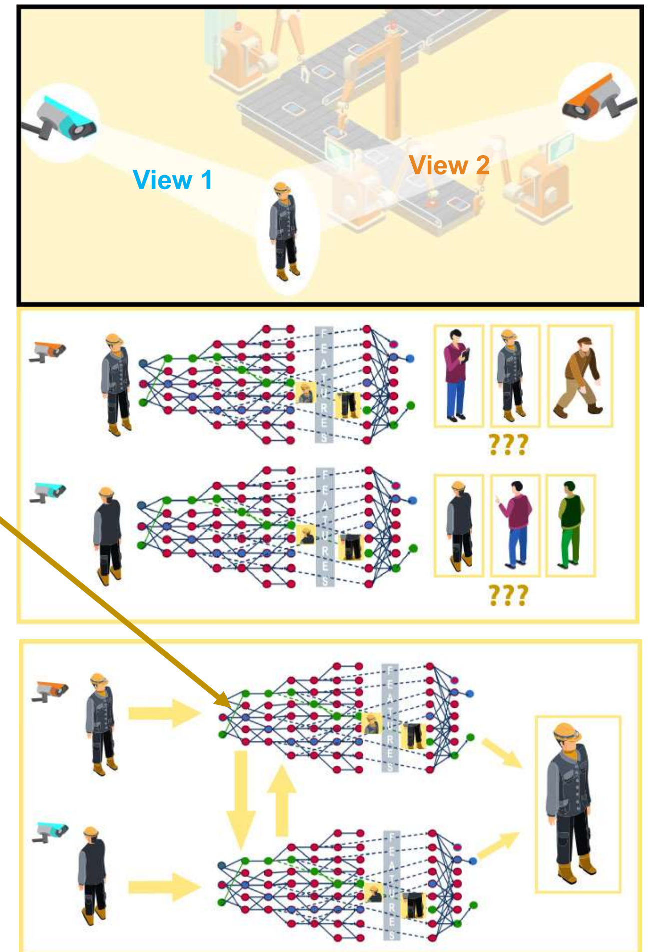
Collaborative ML Execution

- Take advantage of multiple "correlated" observers (***sensor multiplicity***)
- Adapt the ML pipelines "on the fly" to take advantage of correlated intermediate states
  - Exchange intermediate ML state
- **Communication ↓   Computation ↑**

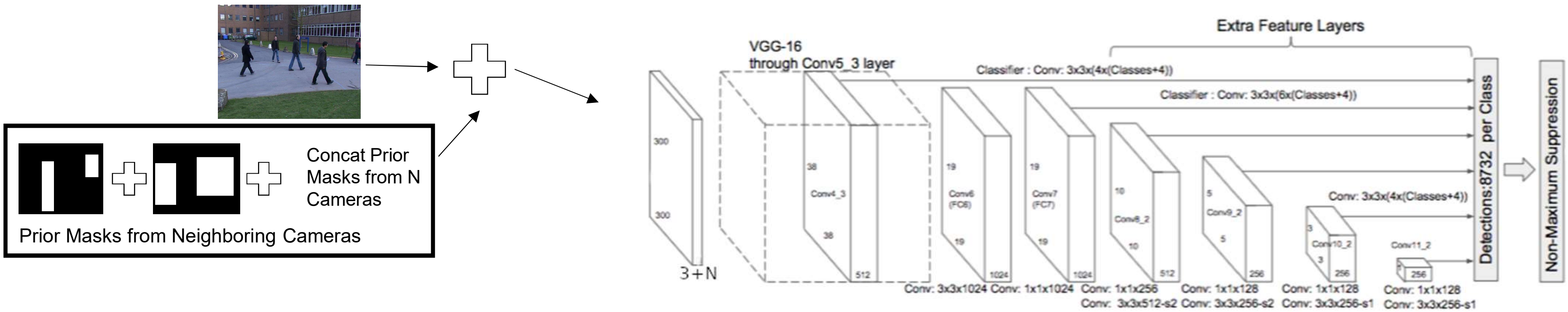
Statistics-driven Dynamic Optimization

- Exchange *(distribution, FFT)* of activation values → fuse ML+ signal proc.
- Change structure of downstream proc.
  - Quantization, Pruning, Factorization
- **Real-time exchange and short-circuiting of processing pipeline**
  - Shave off O(100msec) latency in computation→ exploit power of distributed IoBT devices

Processing time is shorter and more accurate if information is shared



Cooperative ML: Modified SSD



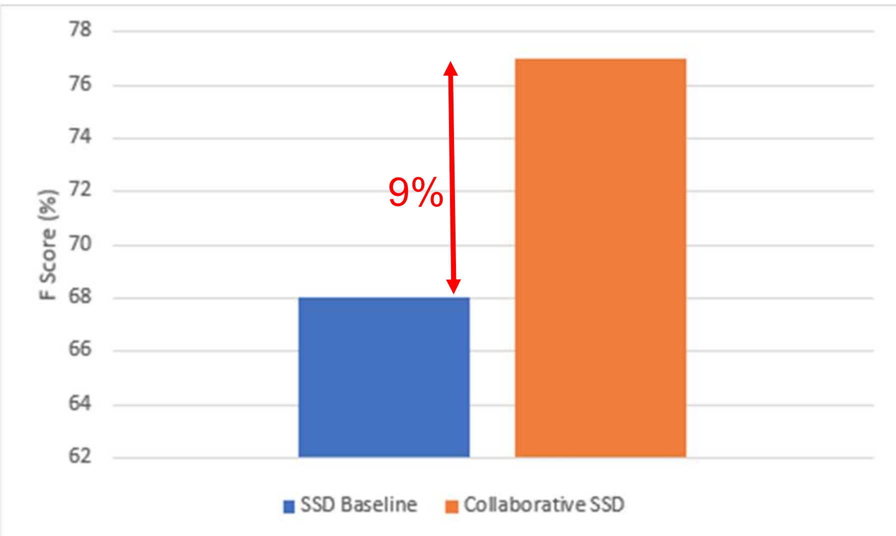
Calibration Stage

- Compute homography transformation matrices among cameras→ spatial mapping among the views

Collaborative Inference Stage

- Receive bounding box detections and associated confidence scores from each neighboring cameras
- Use respective homography transformation matrix to calculate approximate regions of interests of people and derive probability distribution mask (PDM)
- Concatenate prior masks from all neighboring cameras and feed to Collaborative SSD DNN along with the video frame from the current camera

Modified SSD: Results



Evaluated on PETS Multi-Camera Benchmark Dataset

High accuracy improvement with minimal latency and power overhead

	Inference Time	Power / 12.5 FPS per Movidius stick
SSD Baseline	80ms	1.1W
Collaborative SSD	85ms	1.1W

Inference Latency and Power usage for SSD Models in Intel Movidius

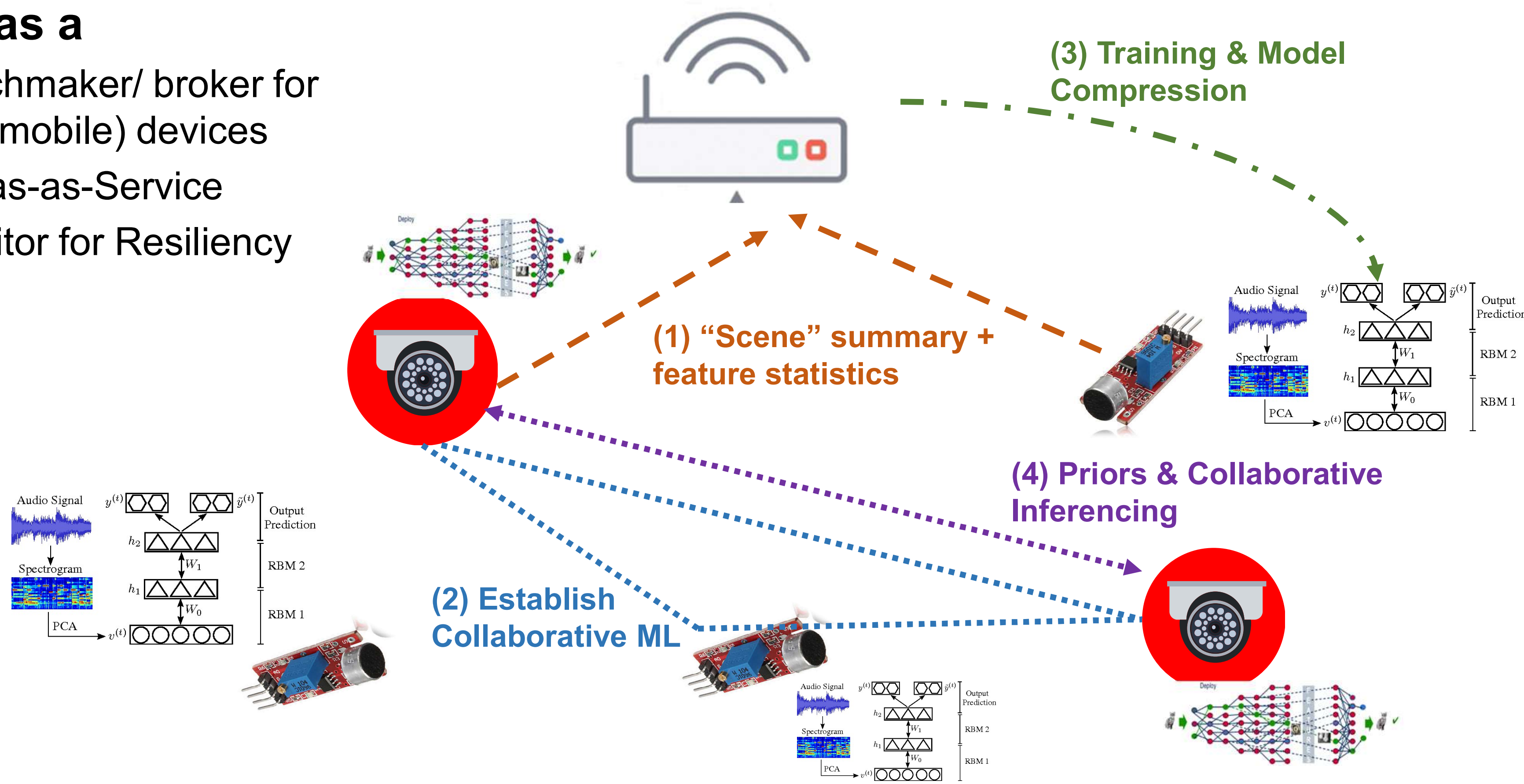
Takeaways

- **New opportunities**: Edge-Coordinated Activation of wearable devices.
- Edge devices augmenting the sensing capabilities of wearable devices
- ML Coordination between a set of distributed edge (IoT) & wearable devices

EA2: Vision: Cognitive Edge for IoT

Edge as a

- Matchmaker/ broker for IoT (mobile) devices
- ML-as-as-Service
- Monitor for Resiliency



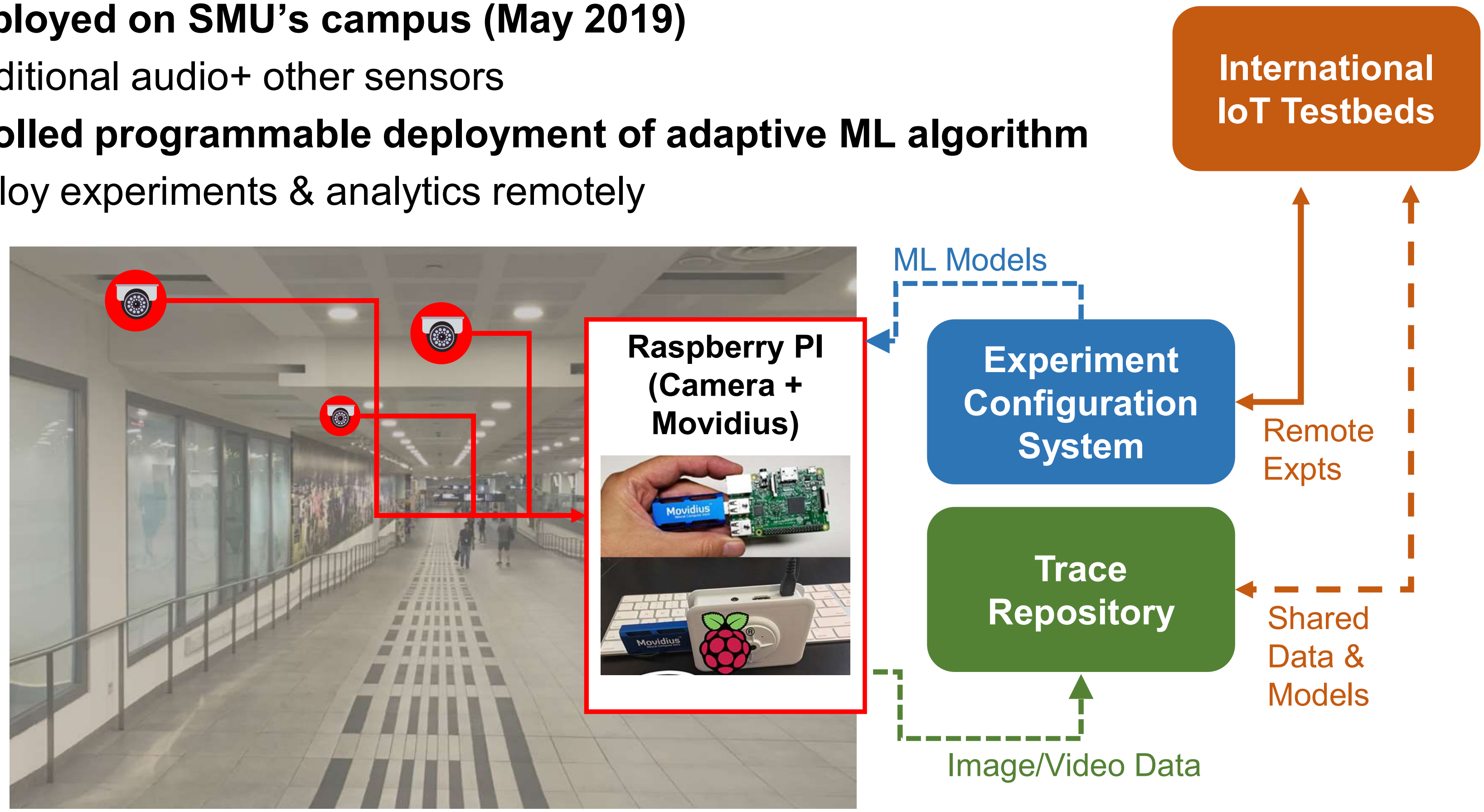
Challenges for The Cognitive Edge

- **Find Useful Spatiotemporal Correlations among Devices**
  - Minimizing Communication Overhead
  - Handling Disparate Sensing Modalities
  - Handle Redundancy in Dense IoT Deployments
- **Enable trusted interactions among Devices**
  - Find Correlations from non-sensitive Metadata/Features
  - Identify and isolate malicious/non-conformant devices
- **Handle Dynamic Workloads**
  - Mobile devices that temporarily reside in specific areas
  - Changes in spatiotemporal human/event patterns

SMU School of Information Systems

SMU "Cognitive IoT" Testbed

- **30 Cameras deployed on SMU's campus (May 2019)**
  - Extend to additional audio+ other sensors
- **Centrally-controlled programmable deployment of adaptive ML algorithm**
  - Ability to deploy experiments & analytics remotely



Takeaways

- **New opportunities**: Edge-Coordinated Activation of wearable devices.
- Edge devices augmenting the sensing capabilities of wearable devices
- ML Coordination between a set of distributed edge (IoT) & wearable devices
- Edge as a Dynamic Matchmaker between Dumb IoT Devices

Conclusion

- Need for greater interaction between wearable devices & edge computing/network entities
  - Latency is key → ensure that the 'edge' is both nearby and has high compute capability
  - Many VR/AR applications will need such 'edge augmentation' for both computation and sensing

- Need for inferencing orchestration among edge devices
  - Significant opportunities for scaling up ML-based applications
  - **Need for standardized models for distributing computational state**

(E) archanm@smu.edu.sg
(U) https://sites.google.com/view/archan-misra