Advanced Technology Labs – Brasil

Challenges in real-time event detection in video

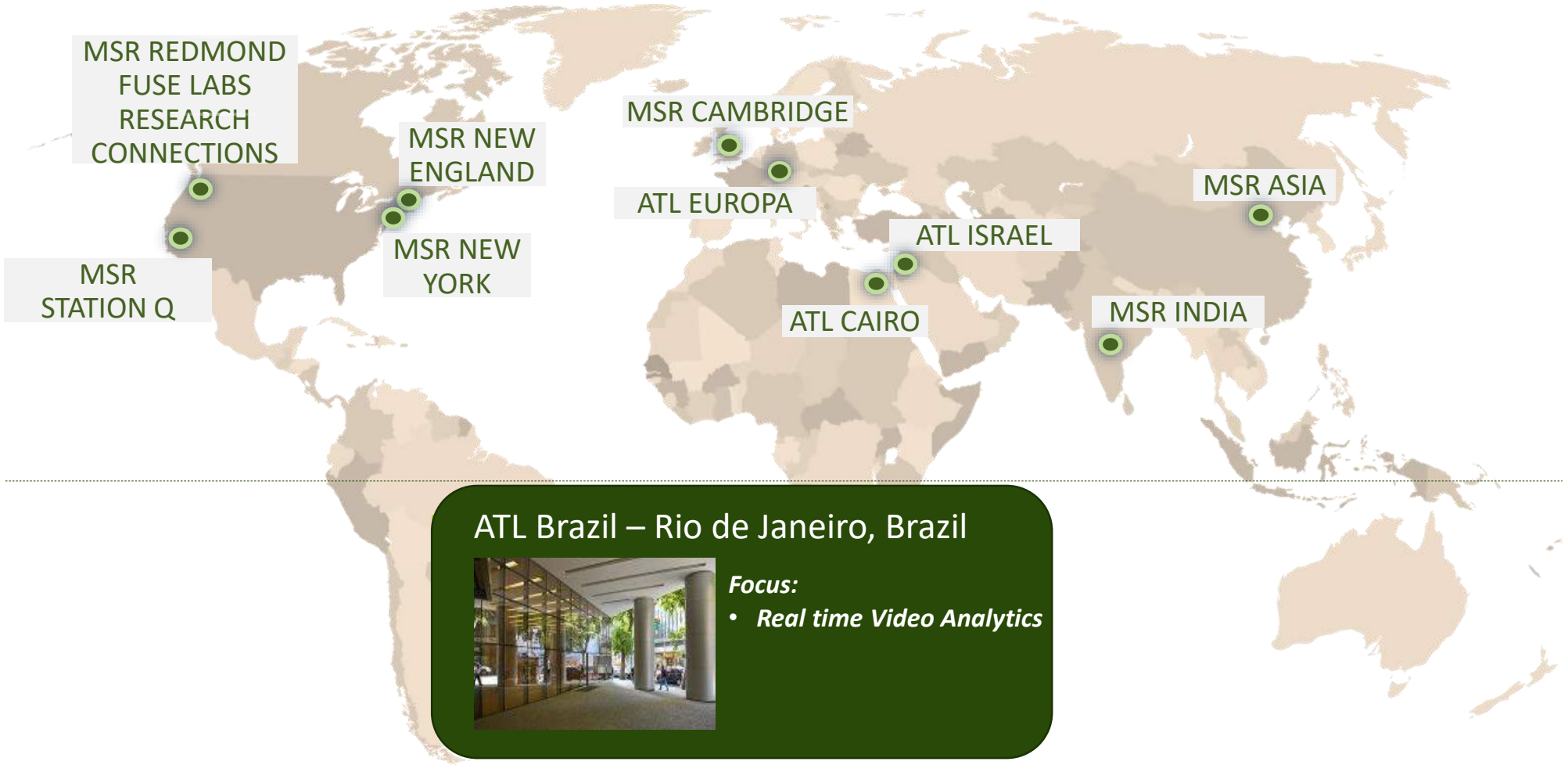
Leonardo Nunes
lnunes@microsoft.com

Microsoft

Objective

- Motivation on real-time event detection in video
- Challenges when analyzing real-world data
- Some lessons learned
- Future research directions

Advanced Technology Lab, Brazil



ATL Brazil – Rio de Janeiro, Brazil

Focus:

- ***Real time Video Analytics***

Motivation

There are over 245 million surveillance cameras around the world[1] (and growing), a market of nearly US$38 Billion/year

Real-time visual monitoring is not practical, so cameras are almost always only used for after-the-fact analysis

Recent advances in image understanding are finally making truly smart automated VA possible and economically viable

Video analytics market will *grow from US$1.5B to >US$4B/year in 5 years (>20% CAGR)* [2]

Source: [1] IHS Technology; [2] MarketsandMarkets

Event Detection

- Event/Action recognition:
  - Triggered when something of interest happens in a video
  - Caused by either a human, vehicle, or an object
- Major applications:
  - Surveillance
  - Mobility
  - Retail

Constraints

Real-time

Cloud-based

COTS sensors

Microsoft Cognitive Services

Cognitive Services API Collection

Vision

Computer Vision | Emotion | Face | Video

Speech

Computer Recognition | Speaker Recognition Speech | Translator

Language

Bing Spell Check | Language Understanding Linguistic Analysis | Text Analytics | Web Language Model

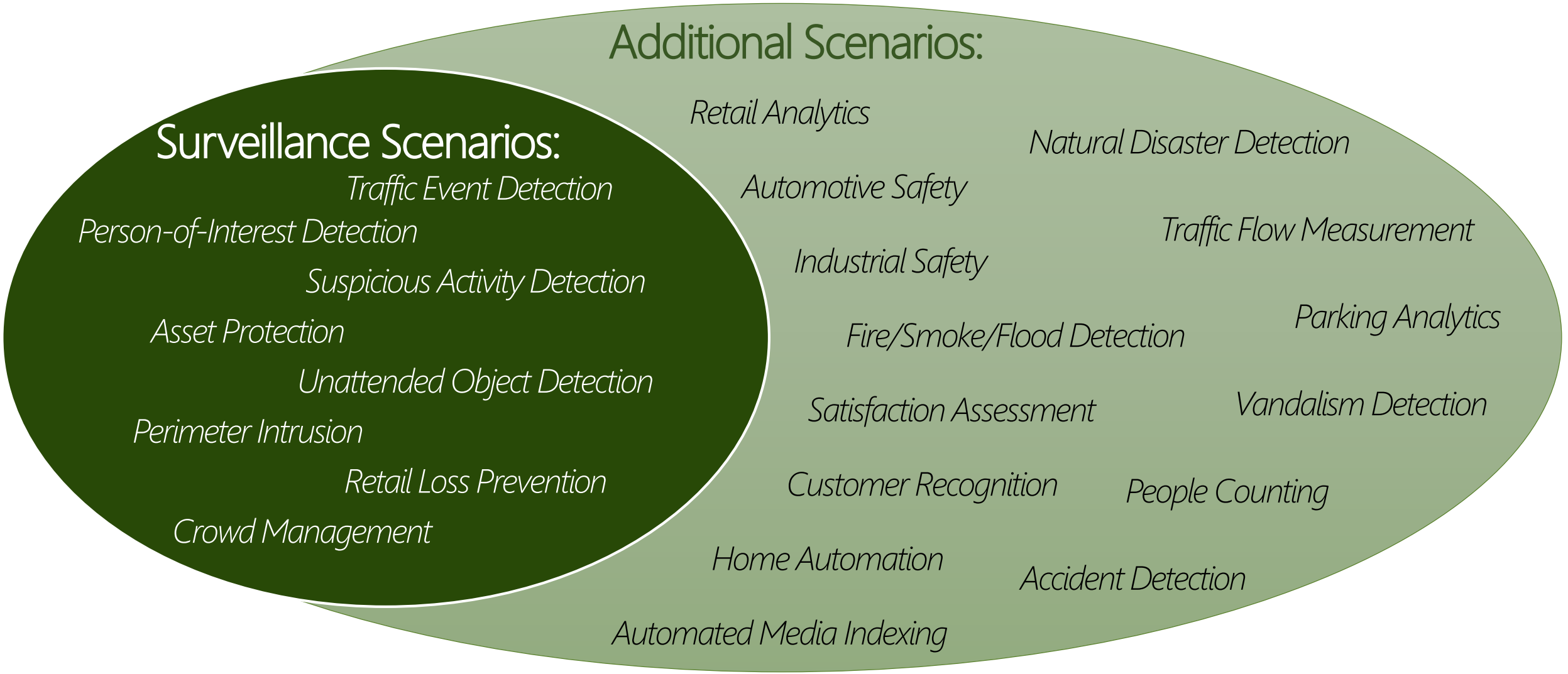
Knowledge

Academic Knowledge | Entity Linking Knowledge Exploration | Recommendations

Search

Bing Autosuggest | Bing Image Search | Bing News Search Bing Video Search | Bing Web Search

Applications



Example Applications

People

- Person identification
- Attribute description
- Tracking

Vehicles

- Vehicle identification
- Model and maker ID
- Speed and trajectory estimation

Objects

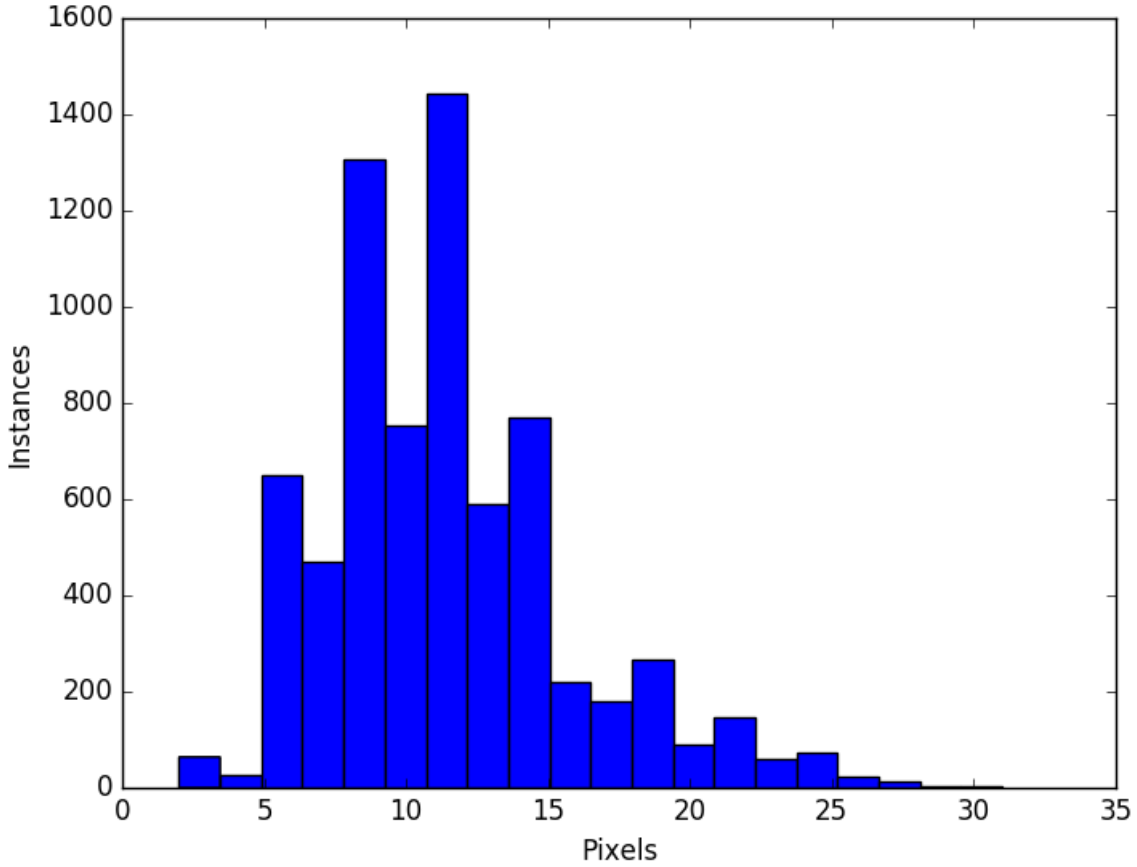
- Classification

Challenges

Video Quality

- Low-quality sensors
- Severe compression artifacts
- Video information is not reliable!
  - Repeated frames
  - Upsampled images

Small Targets



- Usually small targets
- Example:
  - Motorcycles
  - Single camera
- Off the shelf detectors might not work

Enviromental Conditions

Events of Interest are Rare

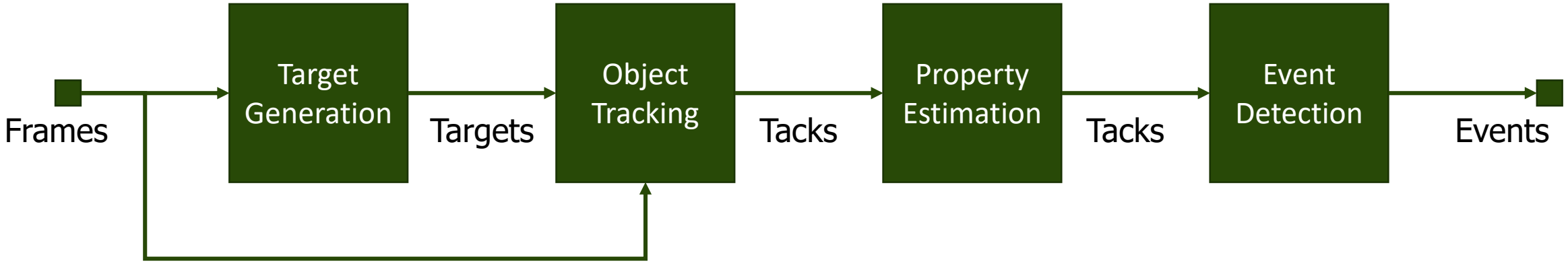
- Events of interest are rare
  - E.g.: crimes, accidents, person identification
- No training example available
- False alert rate vs Probability of detection
- Example (mobility):
  - Jaywalking: 3 events/hour
  - Motorcycle stopped: 0.2 events/hour

Infrastructure Issues

- Variable bit-rate
  - As low as 200 kbps
- Network issues
- Camera movement
- Variable CPU time
- … all leading to lost frames

Some Lessons Learned

Pipeline System



Reliable Detectors

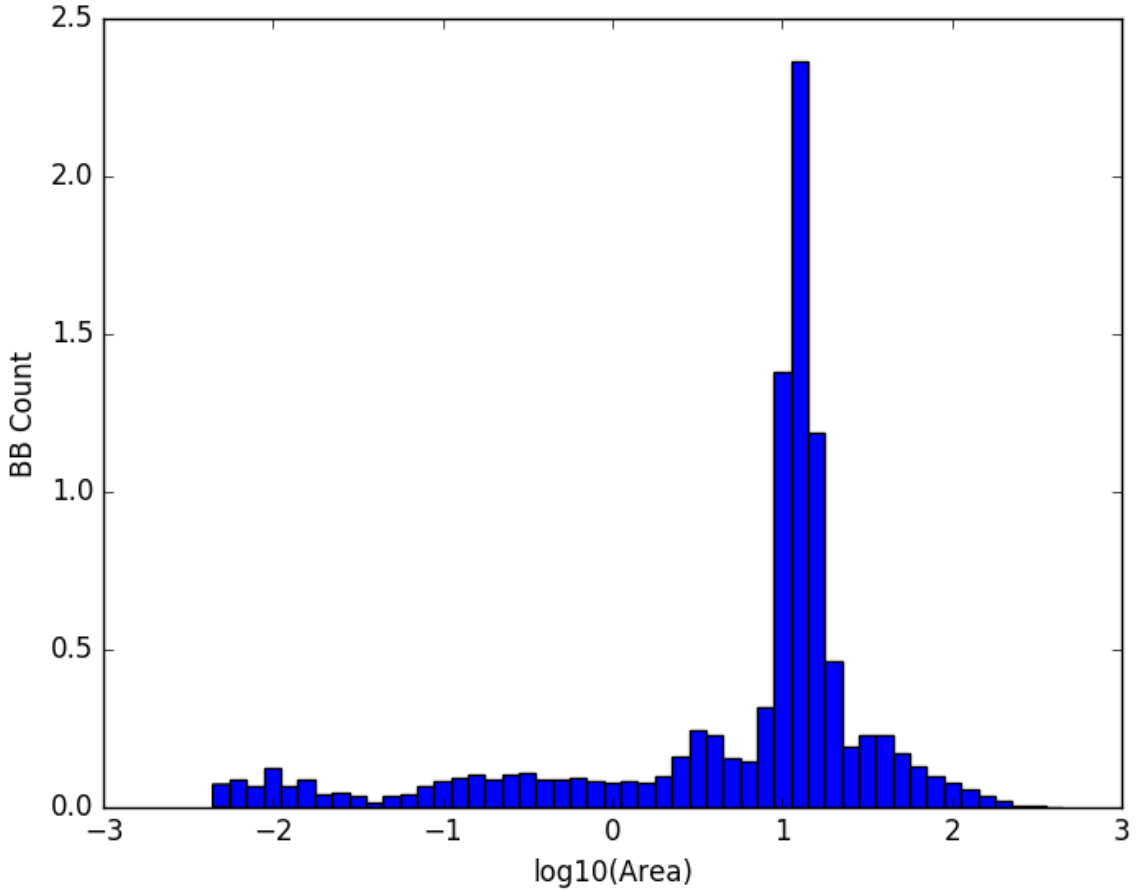
- Low-quality / small objects
  - Robust background subtraction algorithm
- Better quality / larger objects
  - Deep object recognition algorithms
- Most expensive component!
  - Not executed every frame: new targets are infrequent
  - Only when and where there is motion

Tracking by Detection

- Object detector executed:
  - Old tracks are updated with new targets
  - Similarity: distance and appearance
- Object detector not executed:
  - Tracks updated using Kernelized Correlation Filter
- Net effect: not every pixel is analyzed every frame

Ancillary information helps

- Perspective calibration
- Unsupervised learning of target sizes



Need to monitor network/hardware conditions

- Scene conditions detection:
  - Anomalous luminosity
  - Number of discarded frames
  - Camera movement
- Need paths to halt and re-initialize different components

Future

Microsoft

Video vs Image Algorithms

- When to use image vs video detectors?
- Deep object recognizers for video
- Improvements for small objects
- End to end learning

Unsupervised adaptation

- When to run target detectors?
- How to adapt detectors to camera?
- How to adapt tracker parameters?
- How to integrate deep architectures into the pipeline?

Summary

- Lots of challenges
- Specific tools, models for different tasks
- Lack of data, heuristics are still needed
- Unsupervised adaptation for different conditions

Thank you!

Leonardo Nunes

lnunes@microsoft.com