Computer-on-Watch: Imagery Analysis

Greg Angelides

MIT ILP R&D Conference

14 November 2019

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Under Secretary of Defense for Research and Engineering.

Global AI Technology Race

"AI is probably the most important thing humanity has ever worked on. I think of it as something more profound than electricity or fire."

Google CEO Sundar Pichai, January 2018

"For years I've been telling people that the internet was the appetizer, and that AI is the main course. AI is the way of the future, something whose impact will be broader and deeper."

Baidu CEO Robin Li, November 2018

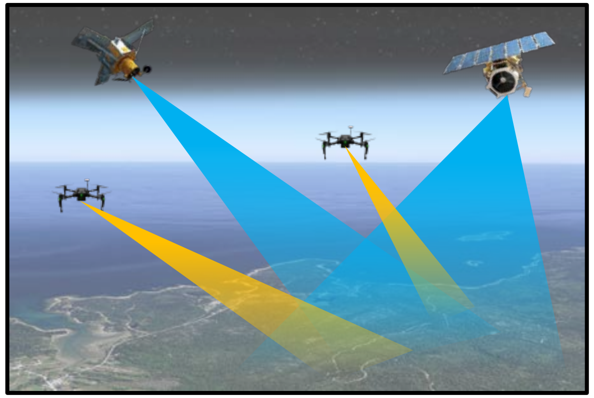
"Artificial intelligence is the future, not only for Russia, but for all humankind… Whoever becomes the leader in this sphere will become the ruler of the world."

Russian president Vladimir Putin, September 2017

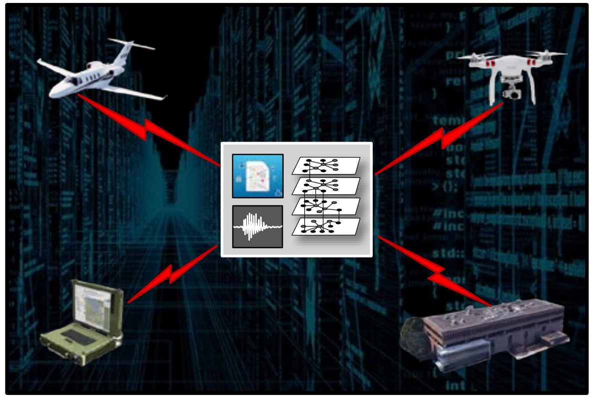
Significant global efforts to develop advanced AI systems and lead in this key field

Example Applications of AI

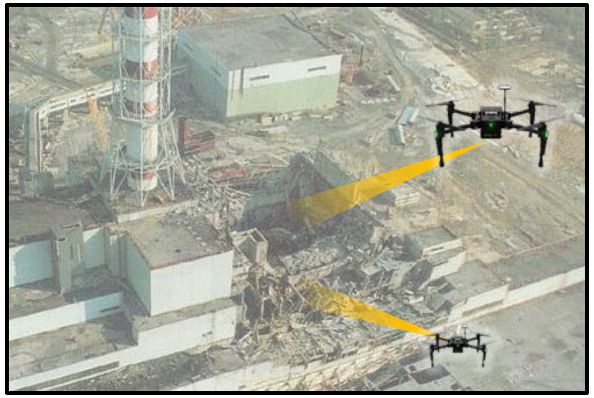
High-Throughput Data Collection
AI Supporting Data Analysis



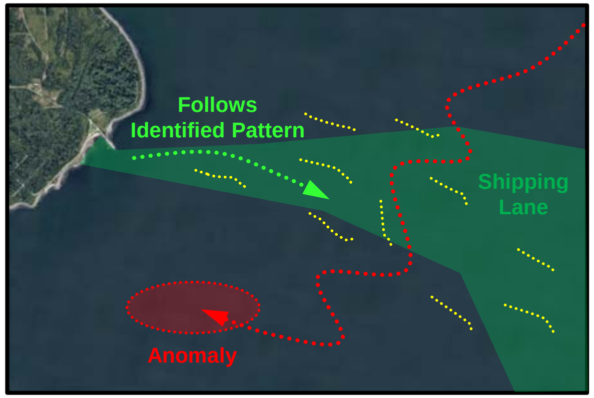
Communications
Organization and Retrieval



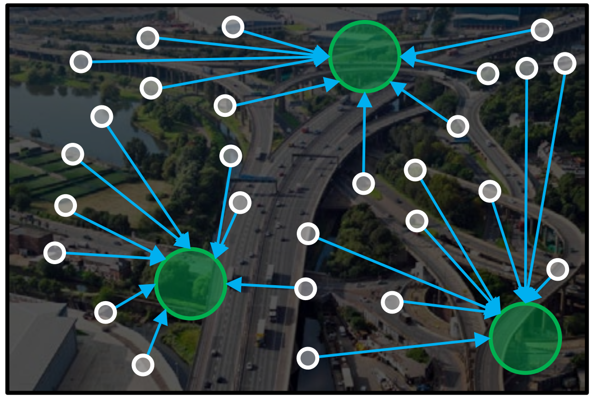
Hazardous Area Assessment
UAVs* Surveil Danger Zones



Pattern-of-Life
Anomaly Detection



Logistics
Resource Allocation

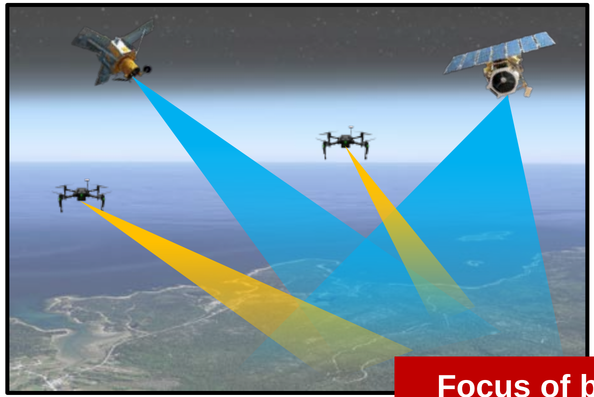


Autonomous Transport
UAV* Autonomous Resupply

* UAV – Unmanned Aerial Vehicle

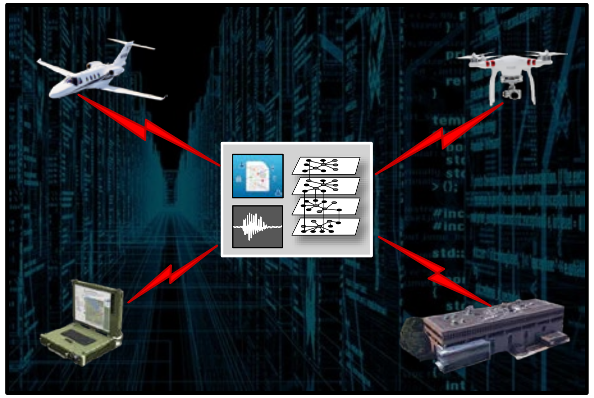
Example Applications of AI

High-Throughput Data Collection
AI Supporting Data Analysis

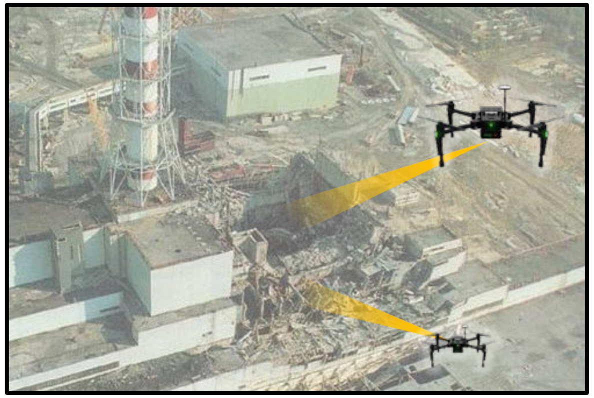


Focus of brief

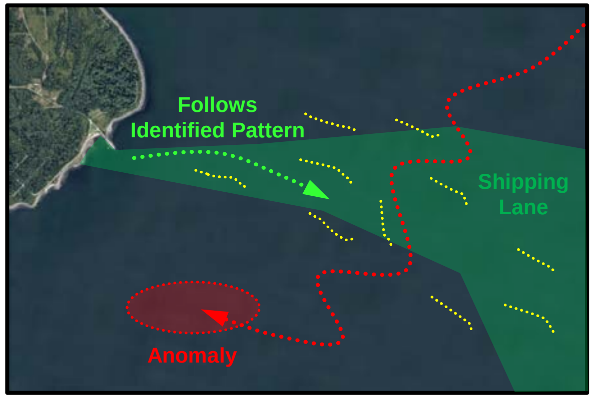
Communications
Organization and Retrieval



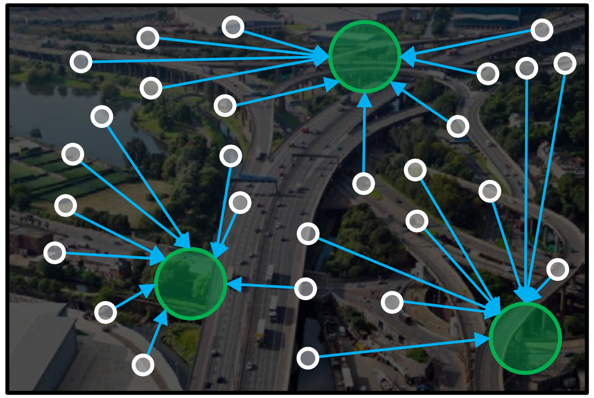
Hazardous Area Assessment
UAVs* Surveil Danger Zones



Pattern-of-Life
Anomaly Detection



Logistics
Resource Allocation



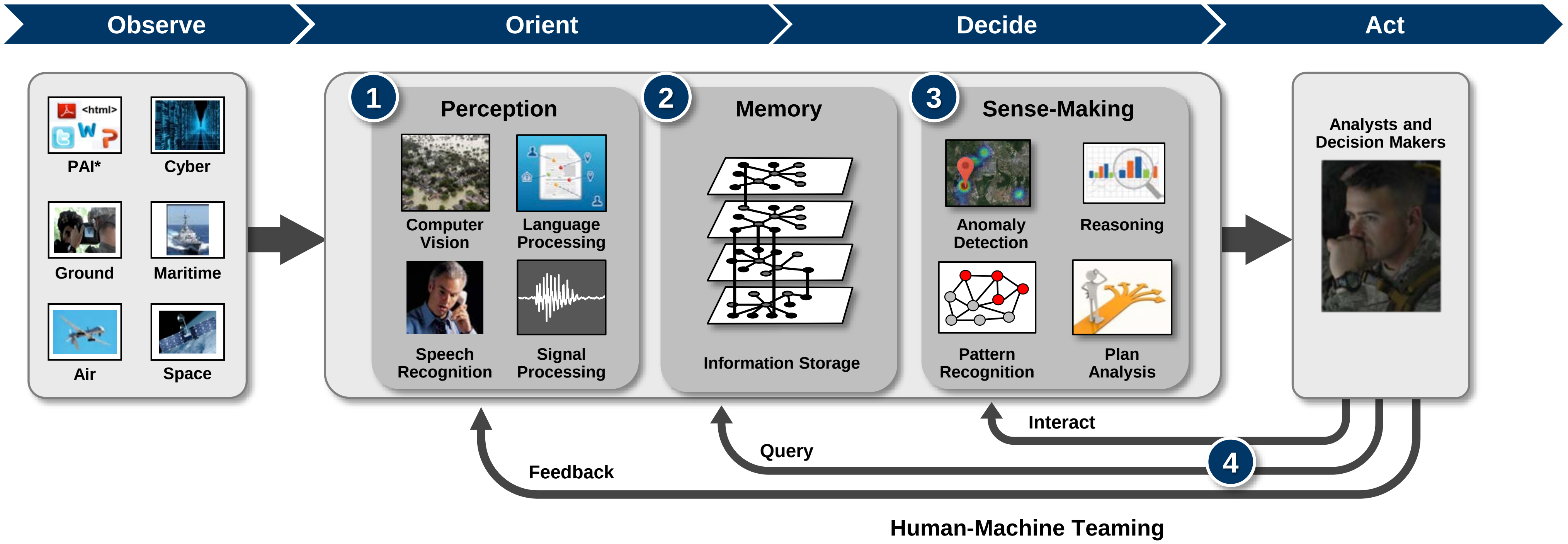
Autonomous Transport
UAV* Autonomous Resupply

* UAV – Unmanned Aerial Vehicle

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Cognitive Computer on Watch Framework

A cognitive assistant to support 24/7 data analysis



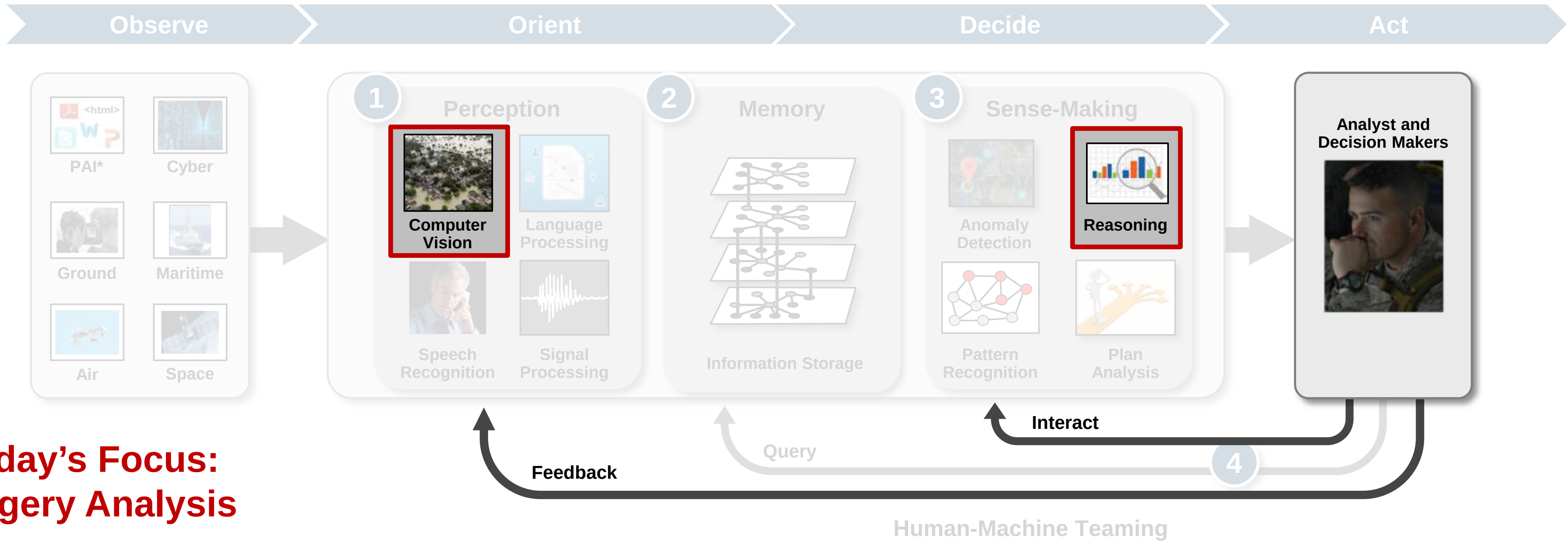
Key Technical Components:

1. *Perception: process sensor data to turn unstructured content into structured information*
2. *Memory: information is indexed and stored to facilitate search and retrieval of historical data*
3. *Sense-Making: advanced analytics predict, reason, and support decision making*
4. *Human-Machine Teaming: effective teaming requires efficient interaction and a natural language interface*

* PAI – publicly available information

Cognitive Computer on Watch Framework

A cognitive assistant to support 24/7 data analysis



Today's Focus: Imagery Analysis

Key Technical Components:

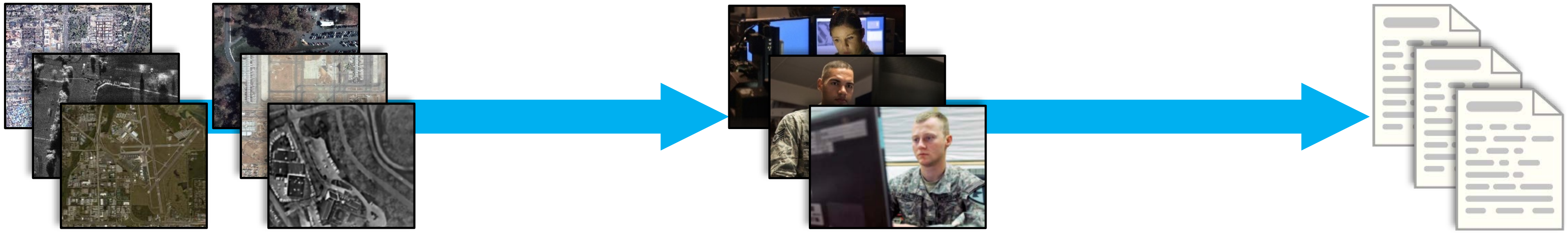
1. *Perception: process sensor data to turn unstructured content into structured information*
2. *Memory: information is indexed and stored to facilitate search and retrieval of historical data*
3. *Sense-Making: advanced analytics predict, reason, and support decision making*
4. *Human-Machine Teaming: effective teaming requires efficient interaction and a natural language interface*

Outline

- Introduction
- Accelerating Imagery Analysis
- Interactive and Interpretable AI
- Summary

Current Image Processing Pipeline

Data | **Imagery Analysis** | **Data Products**

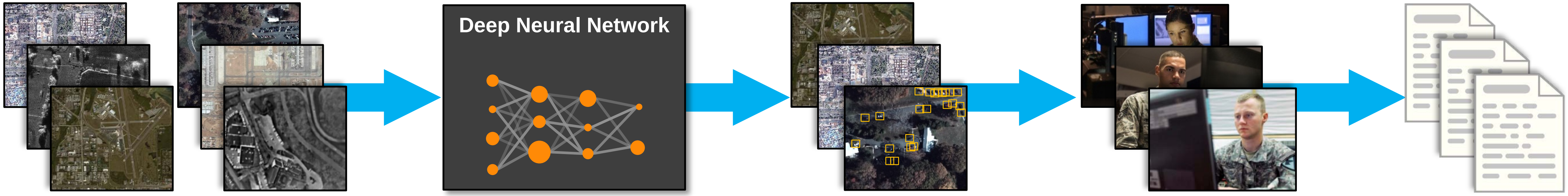


- Large volume with varying resolution, modality, etc.
- Skilled analysts required
- Distilled information

- **Challenging for limited numbers of skilled analysts to assess large volume of collected imagery**

Image Processing Pipeline Leveraging AI

Data | **AI Algorithms** | **Prioritized and Annotated Data** | **Imagery Analysis** | **Data Products**



Data
- Large volume with varying resolution, modality, etc.

AI Algorithms
- Robust automated systems

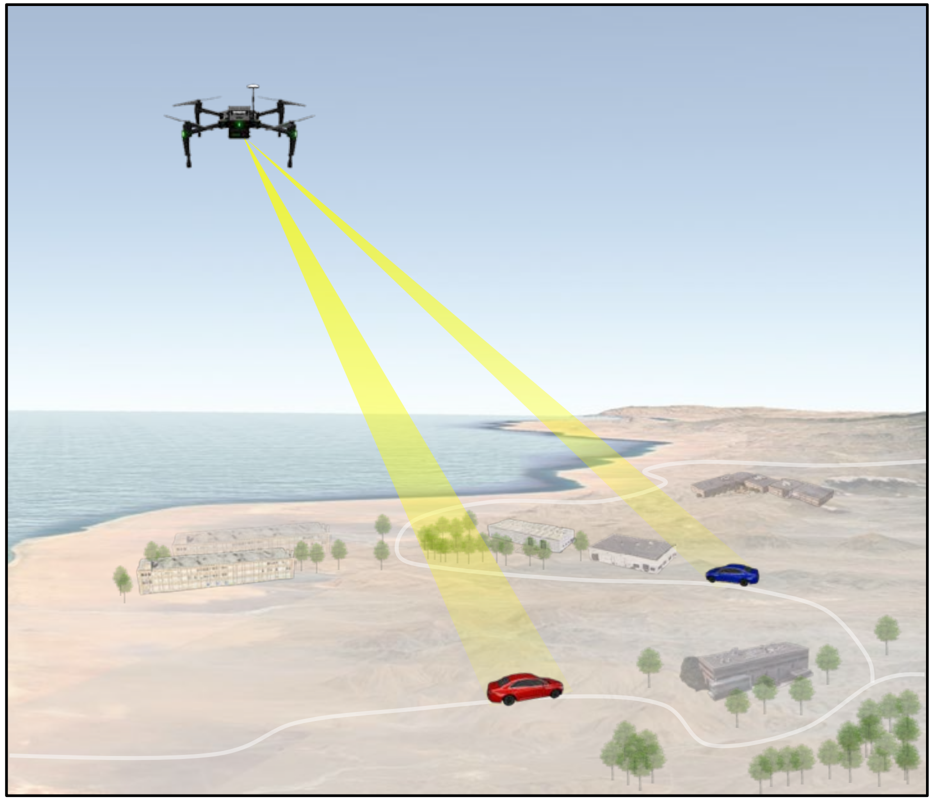
Imagery Analysis
- Analysts perform high-level reasoning

Data Products
- Distilled information

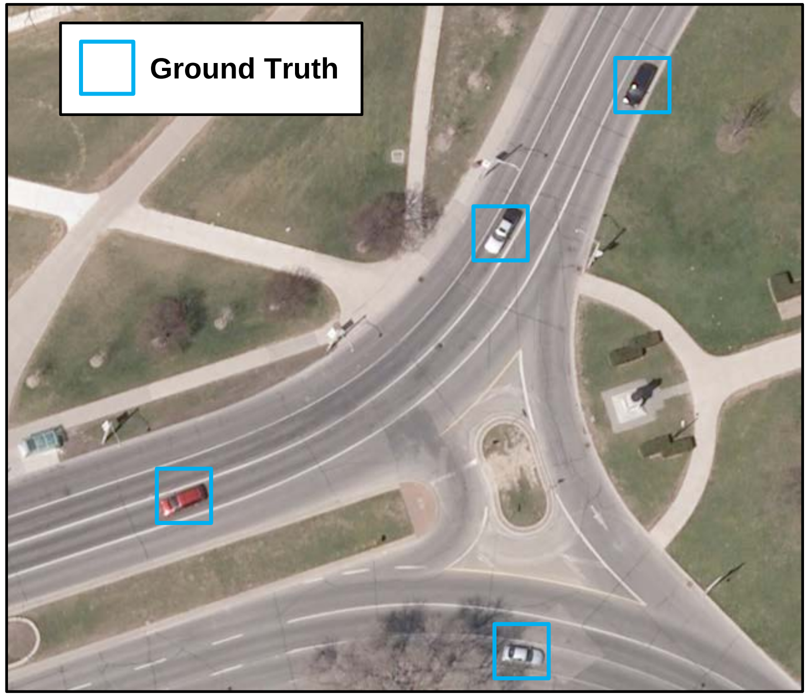
- Challenging for limited numbers of skilled analysts to assess large volume of collected imagery
- Goal: Leverage advances in computer vision to extract relevant information from traditional sensors

Example Application

Vehicle Classification and Detection



Training Data: Cars Overhead with Context (COWC) Research Dataset*



- Overhead imagery with 15 cm resolution
- Data collected from six locations, across four countries
- 5,284 100x100m images
- 32,716 annotated cars, 58,247 annotated negative examples

* Mundhenk, T. Nathan, et al. "A Large Contextual Dataset for Classification, Detection and Counting of Cars with Deep Learning." European Conference on Computer Vision. Springer International Publishing, 2016.

Classification Model

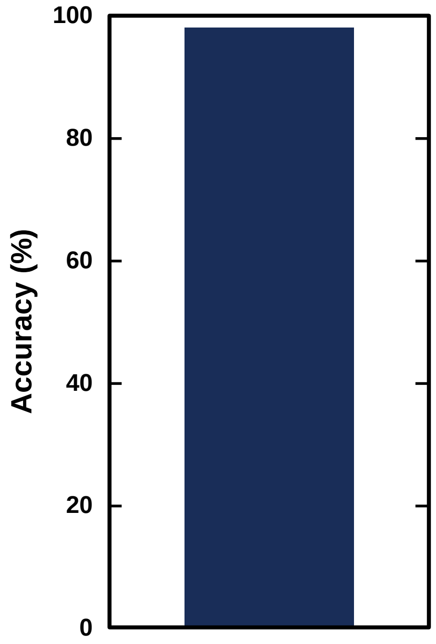
- **Goal: Differentiate cars from car-like objects (boats, trailers, etc.)**

Example Manually Annotated Training Images

Training Data: Manually labeled 90,963 car and background examples

– **Estimated labeling time: >250 hours***

Classification Model Performance



- **Very high classification accuracy achievable with sufficient training data**

*Based on a rate of 10 seconds per sample

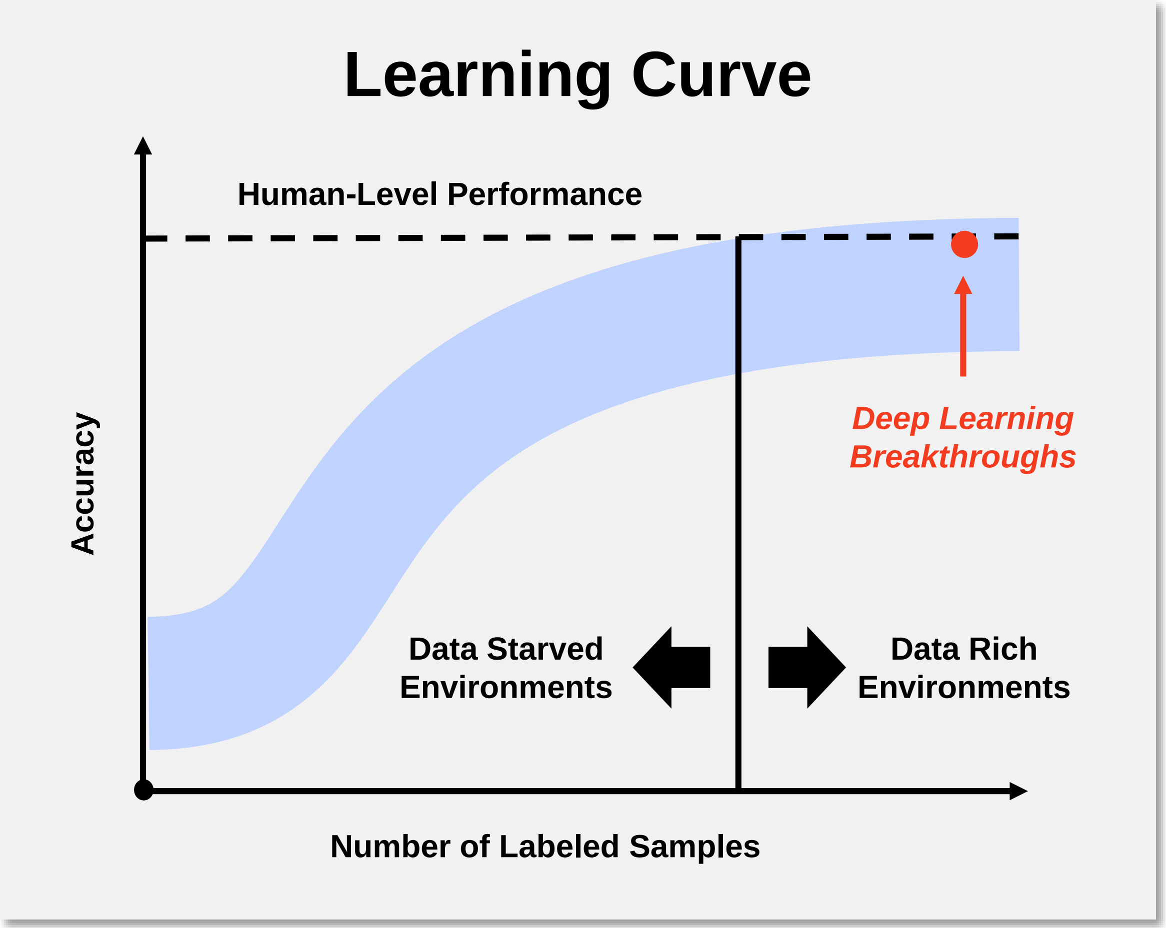
AI Data Challenges

Data Rich Environments

- Labels are free or crowd sourced
- Data is easy to collect

Data Starved Environments

- Data may be challenging to label for machine learning applications
- Data are difficult to collect because content of interest is rare



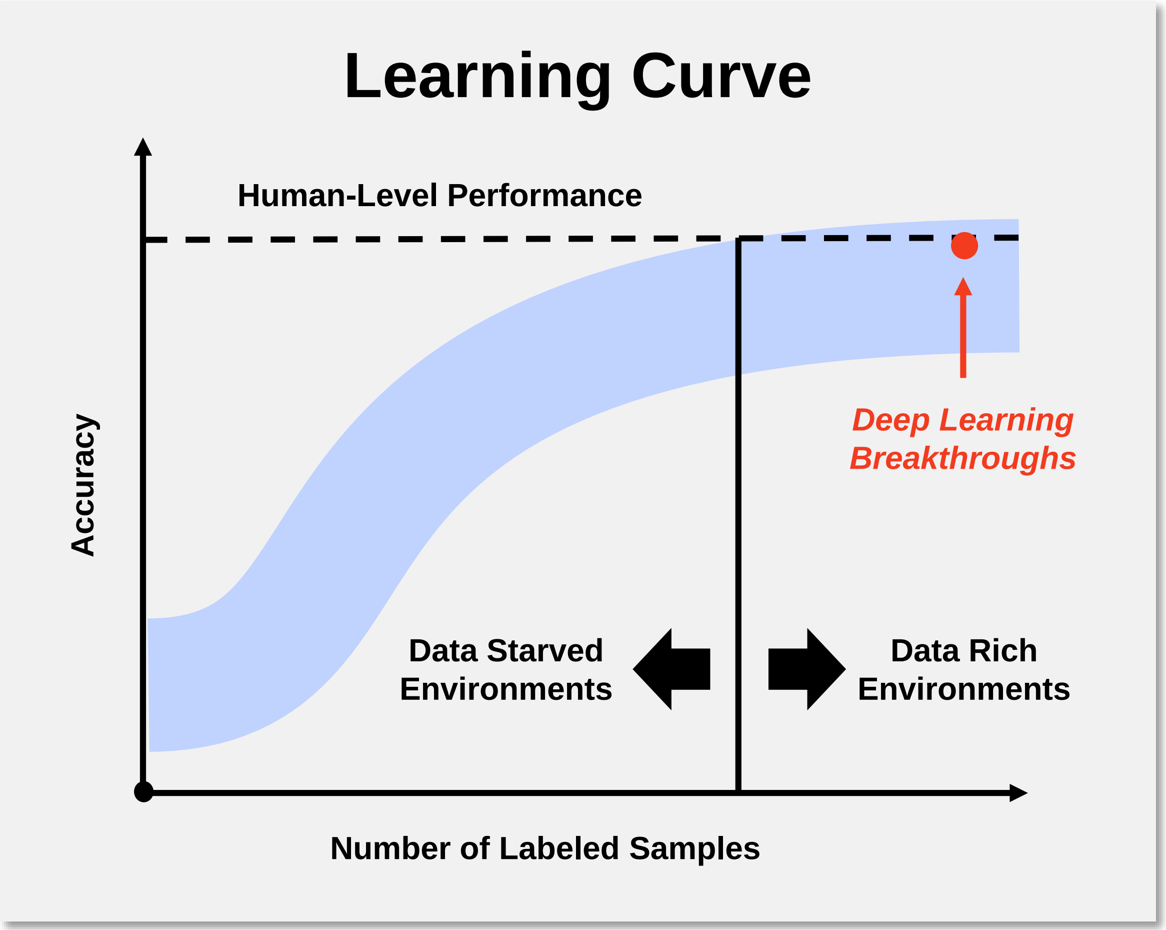
AI Data Challenges

Data Rich Environments

- Labels are free or crowd sourced
- Data is easy to collect

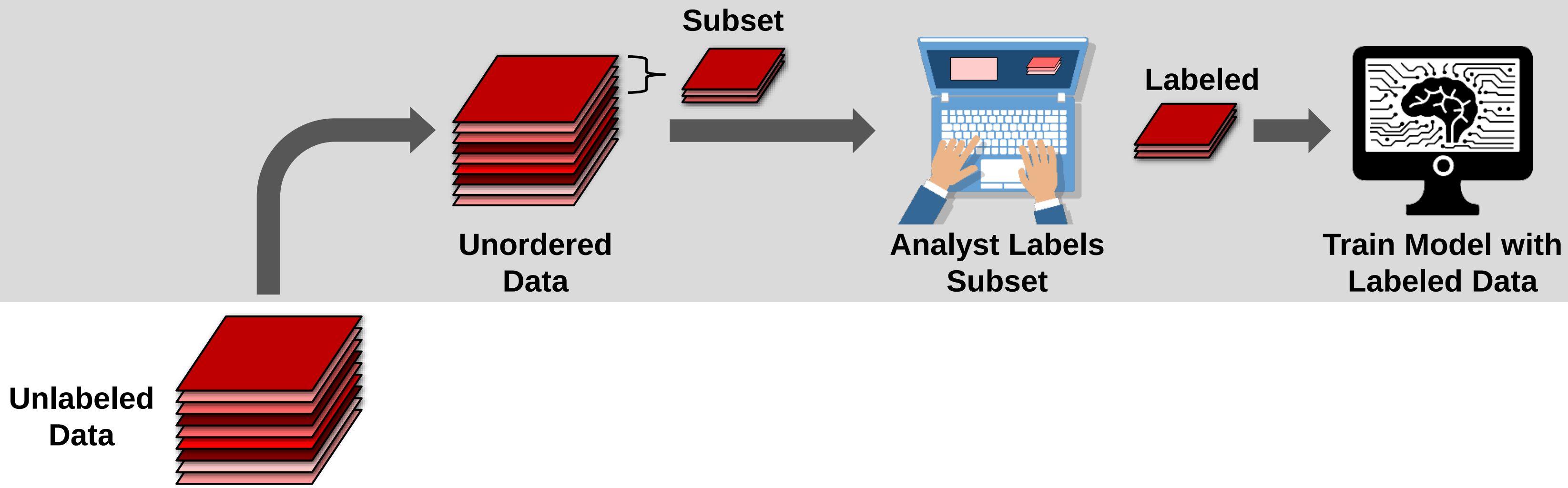
Data Starved Environments

- Data may be challenging to label for machine learning applications
- Data are difficult to collect because content of interest is rare

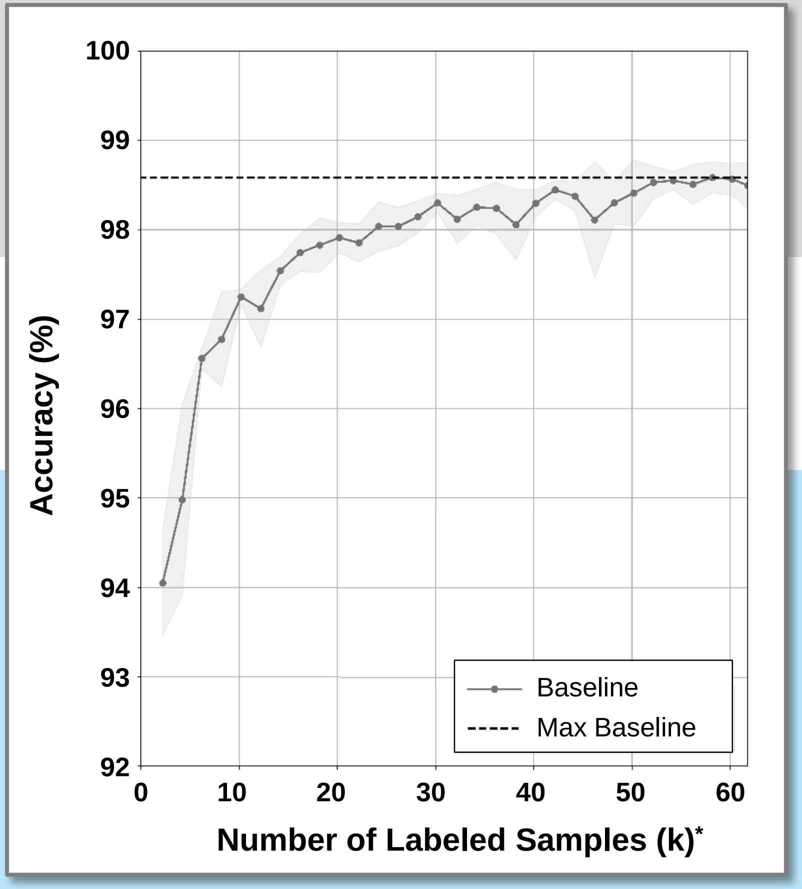


Impact of Data Quantity on System Performance

Baseline Labeling Approach



Learning Curve



*Labeled data augmented 5x for training

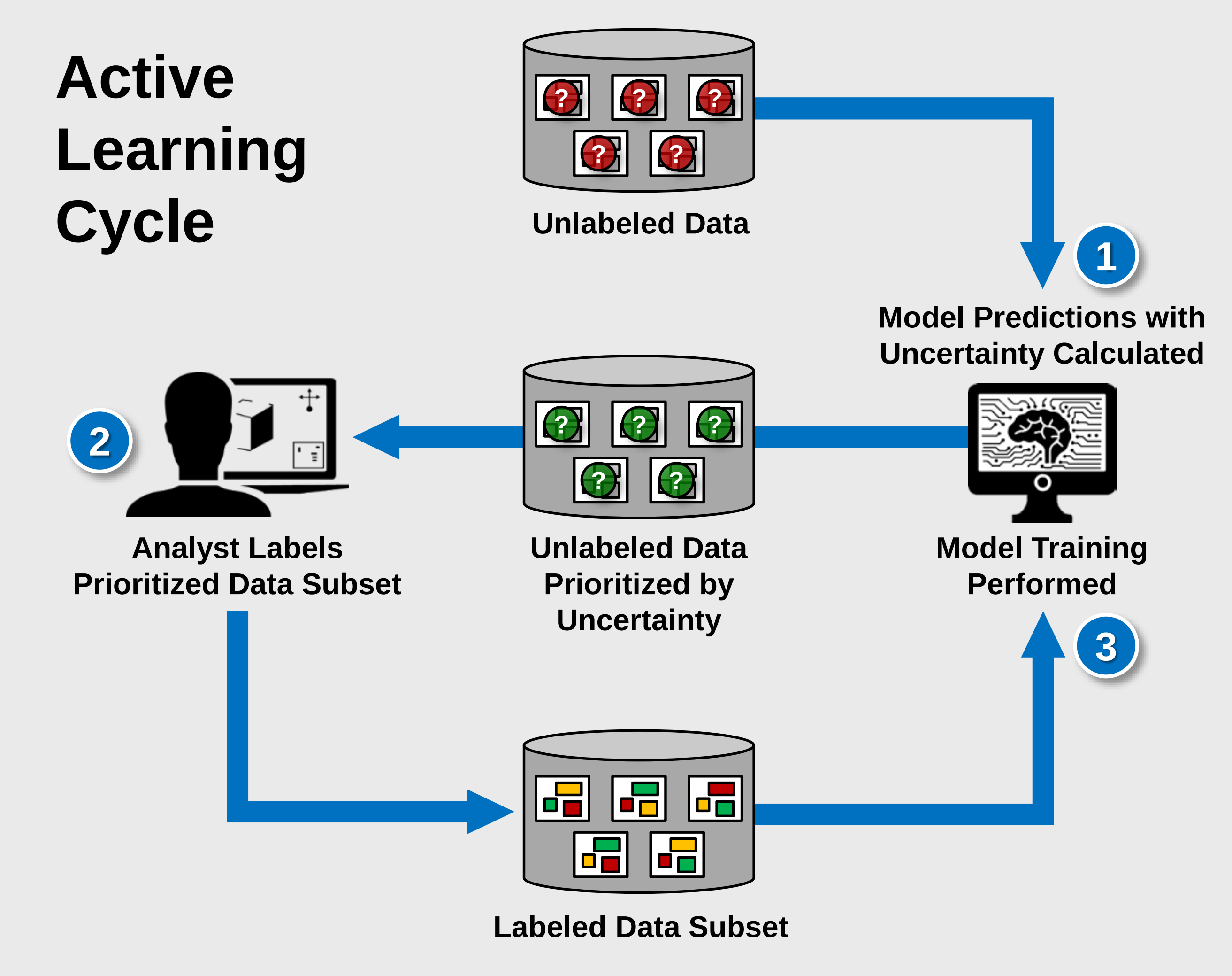
Efficient Approach to Labeling Data: Active Learning

Overview:

- In an active learning framework, training samples are prioritized for labeling according to model uncertainty
- This prioritizes samples that will have the largest impact on model performance

Technical Challenge:

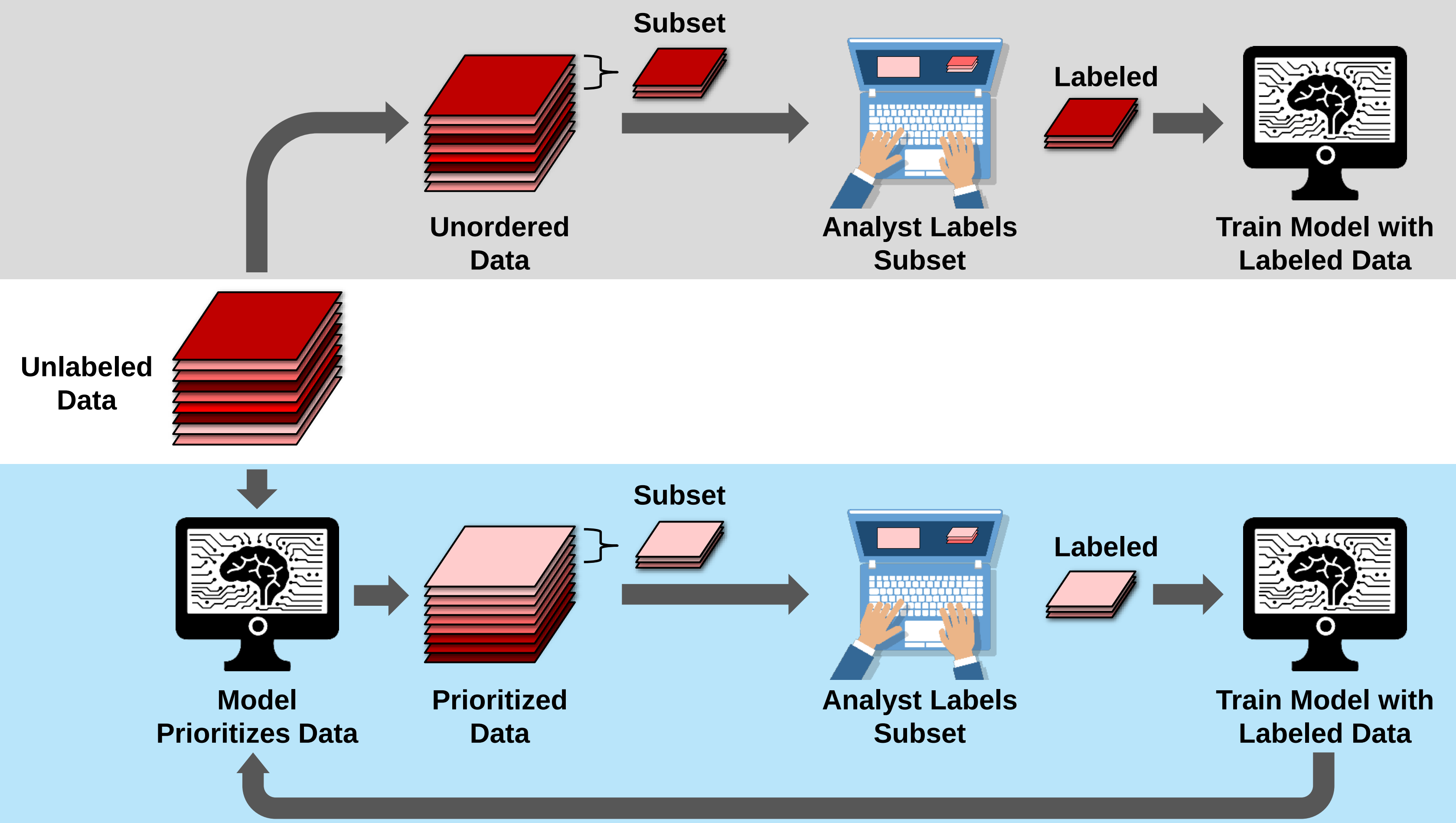
- *Limitation:* Accurate estimation of model uncertainty is challenging
- *Solution Employed:* Uncertainty estimated by sampling multiple model architectures*



* Gal, Yarin, and Zoubin Ghahramani. "Bayesian convolutional neural networks with Bernoulli approximate variational inference." arXiv preprint arXiv:1506.02158 (2015).

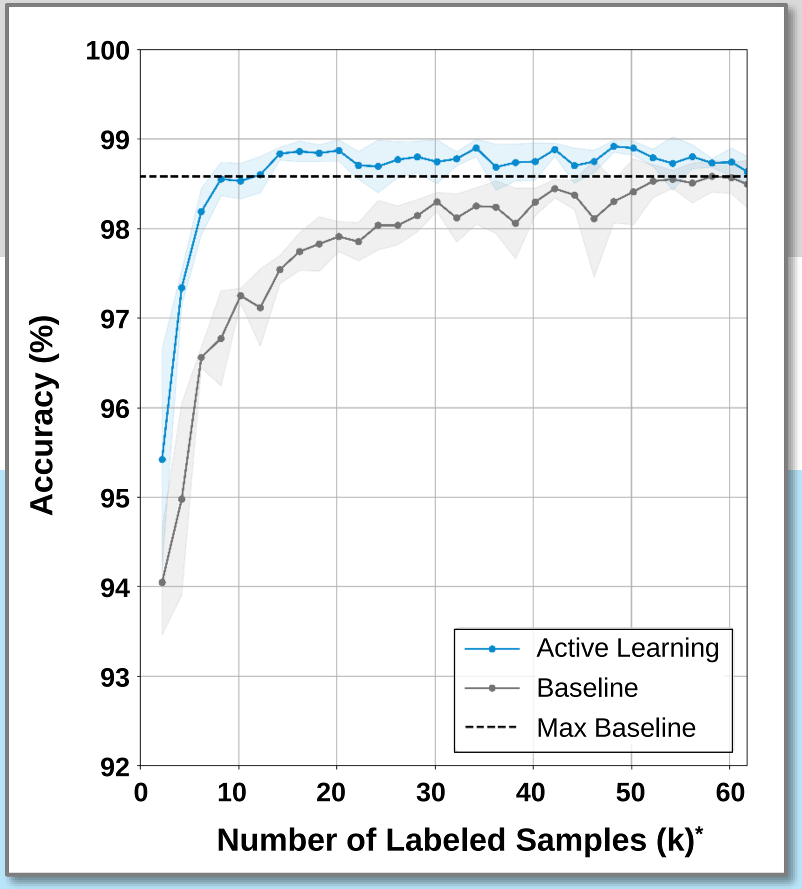
Impact of Data Quantity on System Performance

Baseline Labeling Approach



Efficient Labeling with Active Learning

Learning Curve



*Labeled data augmented 5x for training

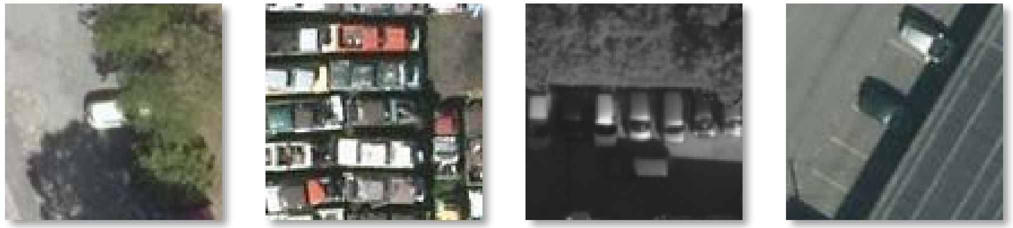
Impact of Active Learning on Data Requirements

Highlights

- **Efficient labeling reveals that only ~1/5 of the data need to be labeled to achieve maximum baseline performance**

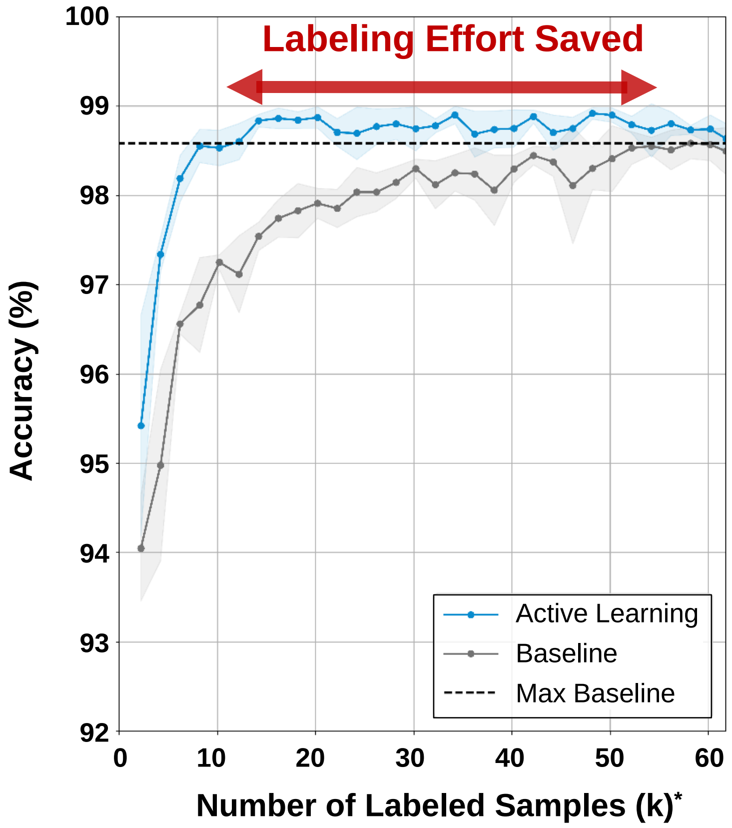
Samples Prioritized with Active Learning

Challenging Positives



Challenging Negatives

Learning Curve



*Labeled data augmented 5x for training

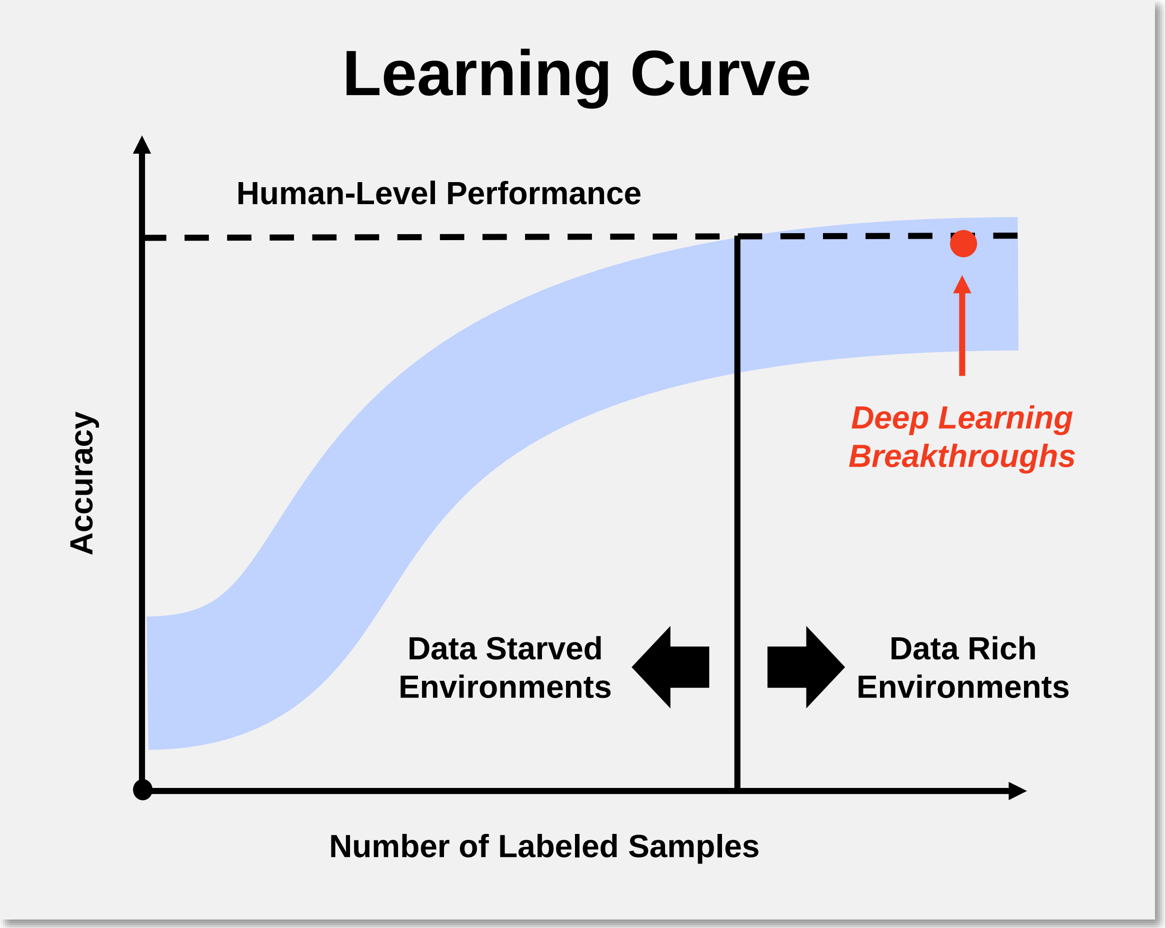
AI Data Challenges

Data Rich Environments

- Labels are free or crowd sourced
- Data is easy to collect

Data Starved Environments

- Data may be challenging to label for machine learning applications
- Data are difficult to collect because content of interest is rare



Data Simulation

- Limited real data available for training AI systems on many targets of interest
- Simulation can be used to generate training data on these targets

Example Simulation Environments Developed

Toronto, Ontario

- 0.21 km^2

Ft. Devens (Ayer, MA)

- 8.3 km^2

Joint Base Cape Cod (JBCC)

- 340 km^2

- Unity game engine simulations incorporate real-world information
  - Height maps
  - Tree locations
  - Road locations
  - Building locations
- Implemented semi-automated procedural pipeline to build real-world environments

Simulation Data Products

- Training data: EO imagery, LWIR imagery*
- Object level truth data: category, positions, detection boxes
- Pixel level truth data: category (segmentation masks), range, incident angle to ground

EO Image

Ground Truth

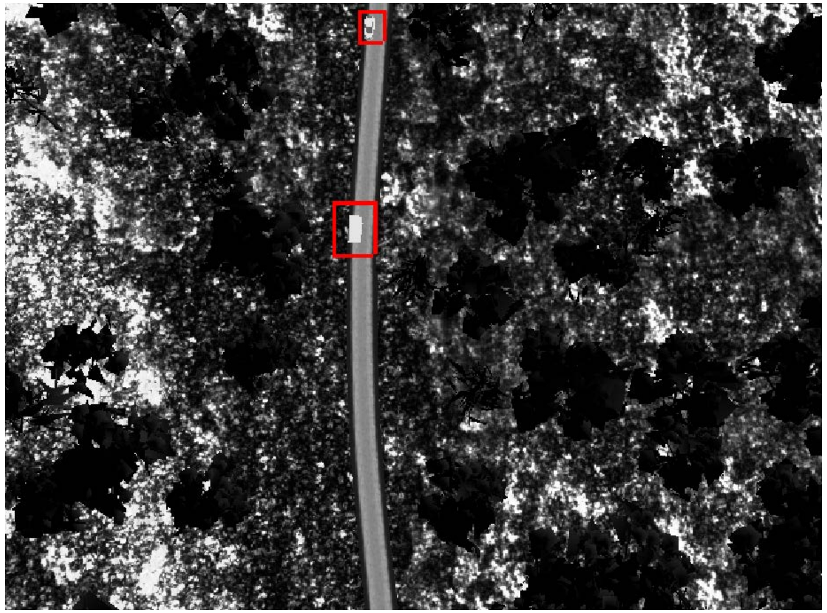
- Simulation enables generation of massive amounts of accurately labeled and diverse training data
- Utilizing simulated data with the training of real-world models is an open area of research

* Hsu et al. Empirical LWIR Scene Simulation Based on E/O Satellite and Airborne LWIR Imagery

Target Detection in Simulation

Example EO Vehicle Detections

Example IR Vehicle Detections



- **Vehicle detection model trained on large amounts of simulated overhead imagery**
- **Very high performance achieved on simulation environments (~90% precision and recall)**
  - **Simulation currently leveraged for initial algorithm development and hardware-in-the-loop testing**
  - **Approaches for transferring simulation trained models to real-world applications currently being investigated**

Outline

- Introduction
- Accelerating Imagery Analysis
- Interactive and Interpretable AI
- Summary

Example Human-Machine Teaming: Visual Question Answering Problem

Today

Ask questions by querying a database

```
SELECT COUNT(*)
FROM cars, buildings
WHERE ST_Contains(ST_GeomFromGeoHash('9qqj7nmxncg'),
                  ST_GeomFromGeoHash(building.geo))
AND ST_Contains(ST_GeomFromGeoHash('9qqj7nmxncg'),
                ST_GeomFromGeoHash(car.geo)))
AND ST_Azimuth(ST_GeomFromGeoHash(building.geo),
               ST_GeomFromGeoHash(car.geo)) < 3*pi()/2
AND ST_Azimuth(ST_GeomFromGeoHash(building.geo),
               ST_GeomFromGeoHash(car.geo)) > pi()/2
```

Desired Future

Ask questions through natural language

1. ***How many cars are south of the large building?***
2. ***Are there any airplanes in the region?***
3. ***How many helicopters are pointing east?***

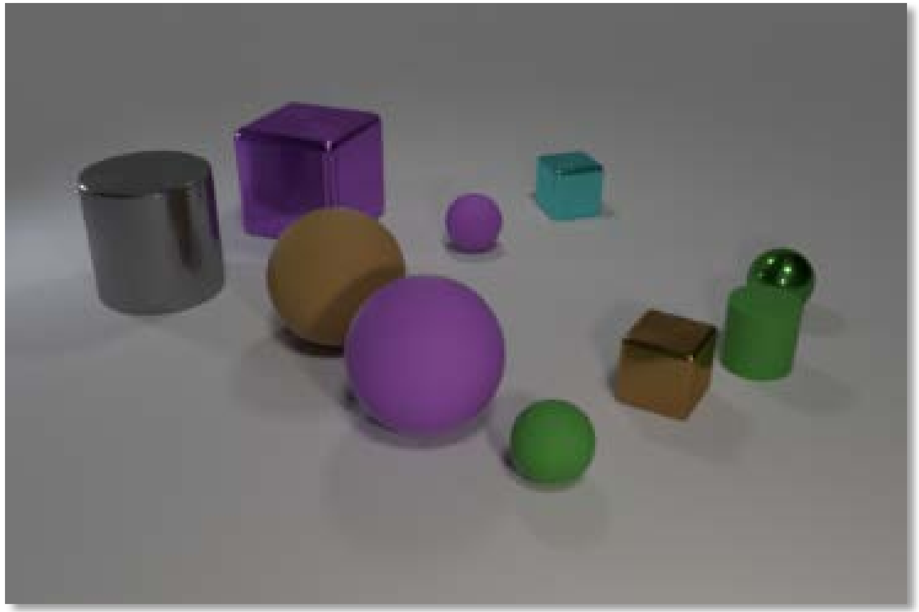
LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Visual Question Answering with Machine Learning

- **CLEVR Visual Reasoning Research Dataset**

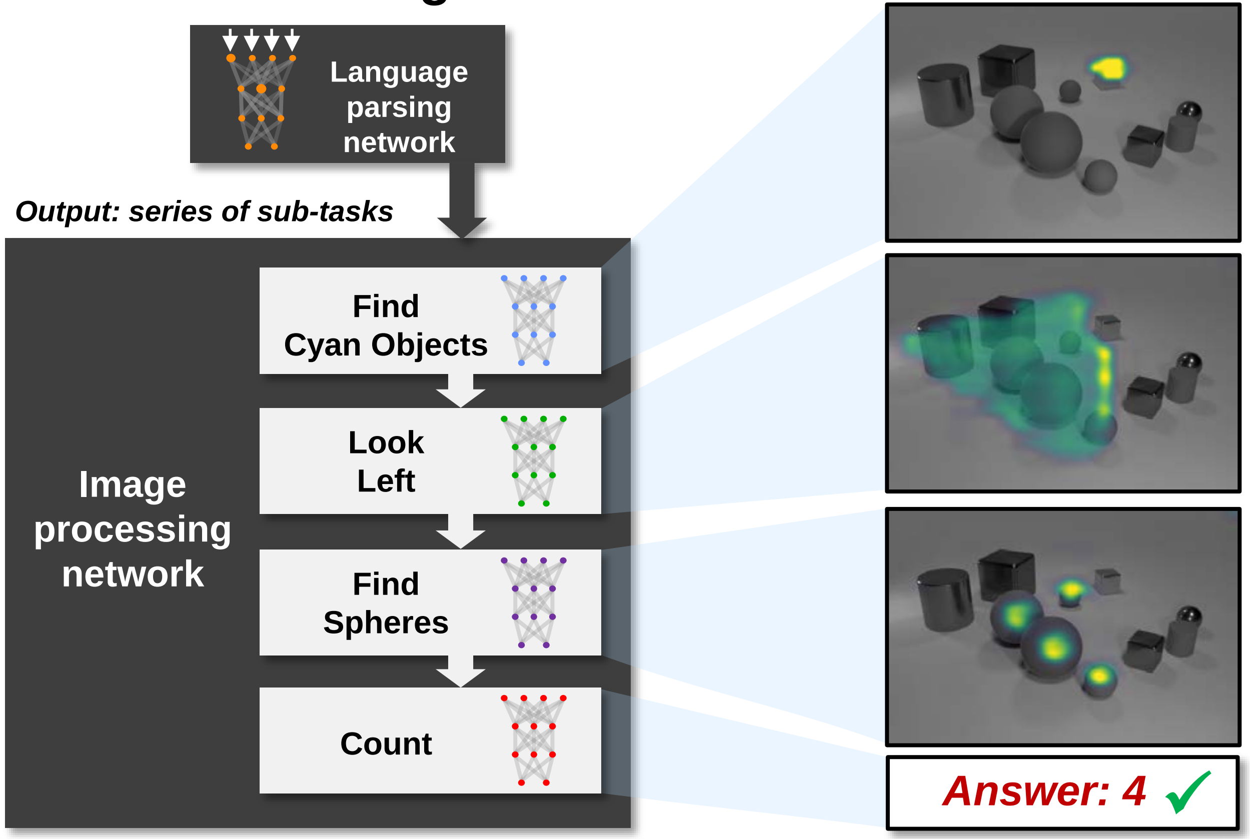
Question:

How many spheres are left of the cyan object?



Transparency by Design Visual Reasoning Network

Outputs



- **Transparency by Design[1] (TbD) networks are performant (99.1% accuracy) and produce interpretable outputs**

1. David Mascharka, Philip Tran, Ryan Soklaski, and Arjun Majumdar. "Transparency by Design: Closing the Gap Between Performance and Interpretability in Visual Reasoning." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), 2018.

Transparency by Design Approach to Visual Question Answering

Question: How many cars are south of the large building?

Parse question and identify sub-tasks:

1. Find large building
2. Look south of the large building
3. Find cars in region south of the large building
4. Count the cars

Answer: 23

- **Transparency by Design networks can be combined with advanced analytics to provide rich access to imagery data**
- **Interpretable decisions allow for review and understanding of algorithm results**

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Summary

- Modern machine learning techniques provide a means of extracting information from high-throughput sensors on relevant time scales
  - Automated exploitation for rapid analysis
- Methods for developing AI systems in environments with limited training data is an active area of research
  - Active learning approaches can make more efficient use of data, reducing labeling requirements
  - Simulation can provide large volume of supplemental data when real data is limited
    - Research ongoing into best way to leverage simulated environments for real-world use
- Continued development of interactive and interpretable machine learning systems key for providing advanced decision support tools
  - Transparency by Design networks can be combined with advanced analytics to provide rich access to imagery data

Acknowledgements

- Bob Bond
- Curt Davis
- Constantine Frost
- Vijay Gadepally
- Dan Griffith
- Rick Heinrichs
- Bernadette Johnson
- Samantha Jones
- Alicia Kendall
- Ben Landon
- Arjun Majumdar
- David Mascharka
- Paul Metzger
- Sanjeev Mohindra
- Paul Monticciolo
- Tommy O'Connell
- Bob Shin
- Ben Smith
- Michael Snyder
- Ryan Soklaski
- Kim Tanguay
- Philip Tran
- Marc Viera
- Joseph Zipkin