This material is based upon work supported by the Assistant Secretary of Defense for Research and Engineering under Air Force Contract No. FA8721-05-C-0002 and/or 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 Assistant Secretary of Defense for Research and Engineering.


© 2018 Massachusetts Institute of Technology.

Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (Feb 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work.
Examples of Artificial Intelligence Applications

2014

- Intelligent assistant capable of voice interaction
- Speech recognition is performed with deep neural networks trained on large data

2016

- Defeated top ranked Go players
- AlphaGo’s supervised learning drew on 160,000 games containing 29.4 million positions. It then played itself millions of times to get better and better

2017

- Testing autonomous cars without a driver
- Scene understanding is powered by deep neural networks learning on 2.5 million real-world miles and 1 billion virtual miles in 2016

Waymo
What Makes AlphaGo Go?

Access to Data
- AlphaGo’s supervised learning drew on 160,000 games (played by 6–9 dan players) containing 29.4 million positions
- It then played itself millions of times to get better and better

Computing Power
- Distributed version of AlphaGo used 40 search threads running on 1202 CPUs and 176 GPUs
  - Google Tensor Processing Unit (TPU) used when playing Lee Sedol

Algorithm Advances
- Two deep neural networks
  - Value: 13 layers, Policy: 15 layers
- Monte-Carlo tree search provided the means to heuristically prune the huge move space

Availability of data and advances in computing hardware and algorithms have led to machines approaching or exceeding human performance in some domains
Applying AI to National Security

Commercial Space is Data Rich
- Data is easy to collect
- Labels are free or crowd source
- Rich datasets like ImageNet, COCO, and others.
Applying AI to National Security

DoD Problem Space is Data-Starved

- Data has not been labeled
- Data is difficult to collect because content of interest is rare or adversary makes it hard
Data Starved AI Challenges

Not Enough Labeled Data

Not Enough Data

National Security Interest is often in the tail of distribution

Objects / Events of Interest

DIUx Challenge Dataset
Xviewdataset.org

Not Enough Labeled Data

Number of Examples

DIUx Challenge Dataset
Xviewdataset.org

Number of Examples

National Security Interest is often in the tail of distribution

Objects / Events of Interest
Applying AI to National Security

Data Rich
- Data is easy to collect
- Labels are free or crowd source

Data-Starved
- Insufficient labeled data
- Data is difficult to collect
- Content of interest is rare

Example Research Thrusts
1. Develop Gold-standard datasets
2. Efficient data labeling at scale
3. Develop algorithms that require less training data
4. Pursue Cognitive Science research to inform machine learning
5. Hybrid learning that merges deep learning with model-based learning

More sophisticated algorithms are needed in a data-starved environment

*Vehicle detection in low-res FMV; an example of AI applied to data-rich military domain
** Identification of camouflaged military targets: an example of a low-resourced and adversary-countered AI task
Data-Starved AI Session Talks

Computer Vision
- Object Detection
  - Subset Prioritized by Uncertainty
  - Active Learning Cycle
  - AI for Imagery Analysis in Low Resource Domains

Cyber Warrior
- CHARIOT
  - Detecting Online Cyber Discussions
  - TF-IDF Features
  - Logic Regression Classifier
  - Miss Probability (%)
  - False Alarm Probability (%)
  - AI to Aid Rapid Response to Cyber Attacks

Inferencing
- Probabilistic Computing for Data-Starved AI
  - Start location, Observed locations, Inferred goal locations
  - Gold standard inference algorithm
  - Target inference algorithm

Data-Starved AI - 9
SM 03/05/18
Data-Starved AI Session Posters

Computer Vision in Low Resource Environments

Mr. David Mascharka, MIT Lincoln Laboratory

Teaming with the AI Cyber Warrior

Dr. William Streilein, MIT Lincoln Laboratory

Interpretable Machine Learning

Dr. Jonathan Su, MIT Lincoln Laboratory

Threat Network Detection: Countering Weaponization of Social Media

Dr. Olga Simek, MIT Lincoln Laboratory
Keynote: Prof. Antonio Torralba

Research Interests

• Building systems that can perceive the world like humans do. A system able to perceive the world through multiple senses might be able to learn without requiring massive curated datasets.

MIT-IBM Watson Lab

• The Lab is focused on advancing four research pillars: AI Algorithms, the Physics of AI, the Application of AI to industries, and Advancing shared prosperity through AI
Recent advances in hardware, algorithms, and the availability of large training data have led to machines approaching or exceeding human performance in some domains.

Challenge in applying AI for National Security: How do we gain understanding of the world to enable time-critical decisions in an environment that is adversarial and data starved.

Advances in data-starved AI are needed to meet national needs

- MIT Lincoln Laboratory is actively working in this area
- Looking forward to collaborating with you to improved the state of the art