Applications of Deep Learning to the GEOINT mission

June 2015

Overview

- Motivation
- Deep Learning Recap
- GEOINT applications:
  - Imagery exploitation
  - OSINT exploitation
  - Geospatial and activity based analytics
  - Sensor control
  - Deep Learning and the analyst
- Deep Learning deployment
- Questions/Discussion

Motivation

Rapid growth in remote sensing numbers and capability

facebook

350 Million Images Uploaded a Day

YouTube

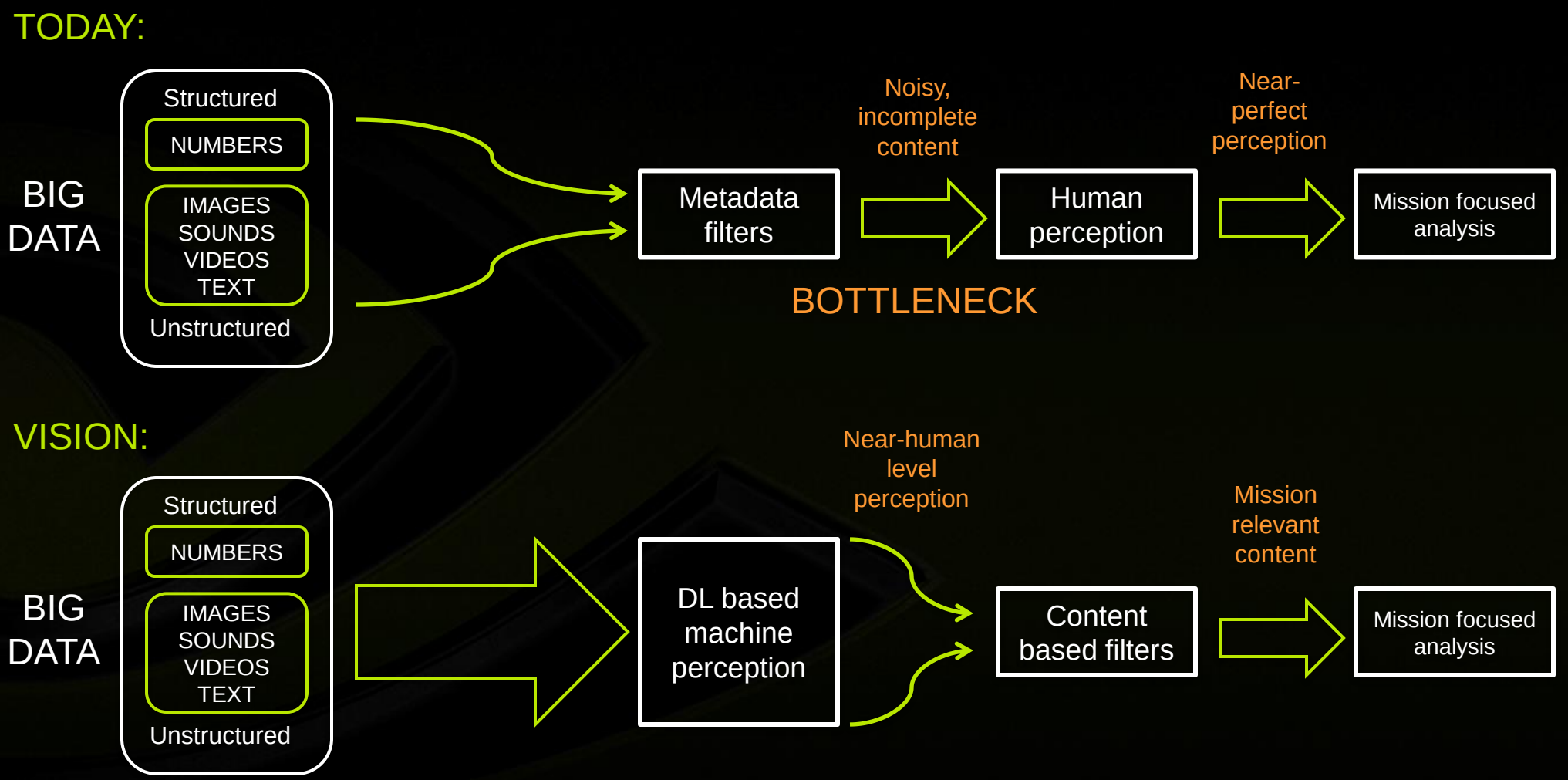
100 Hours Video Uploaded Every Minute

The GDELT Project

Tens of thousands of social and political events indexed daily

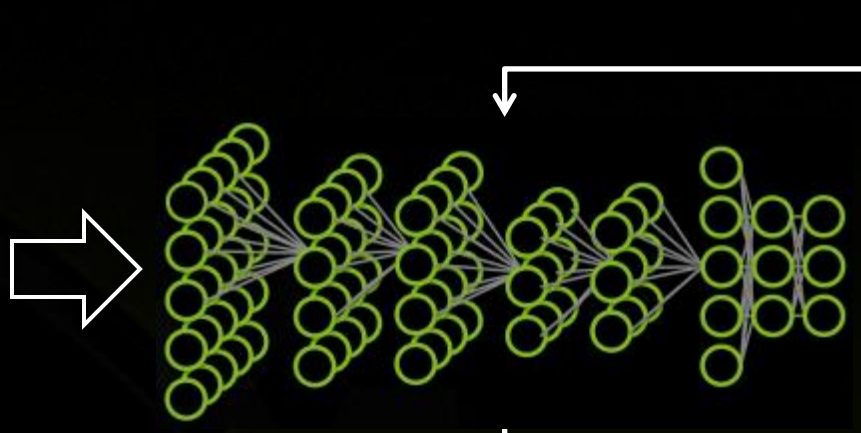
- There is not enough time or expertise to write algorithms for each individual information extraction task that needs to be performed
- Deep Learning provides general algorithms that identify mission-relevant content and patterns in raw data at machine speed

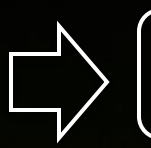
Motivation: GEOINT analysis workflow



Recap: What is Deep Learning?

Train:

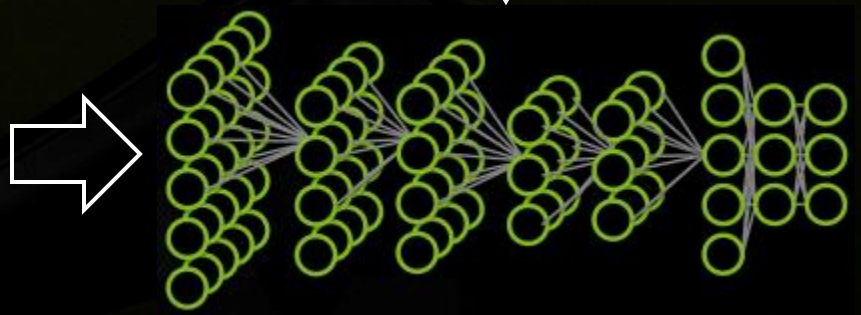


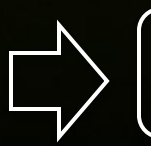


Errors

Dog
Cat
Raccoon

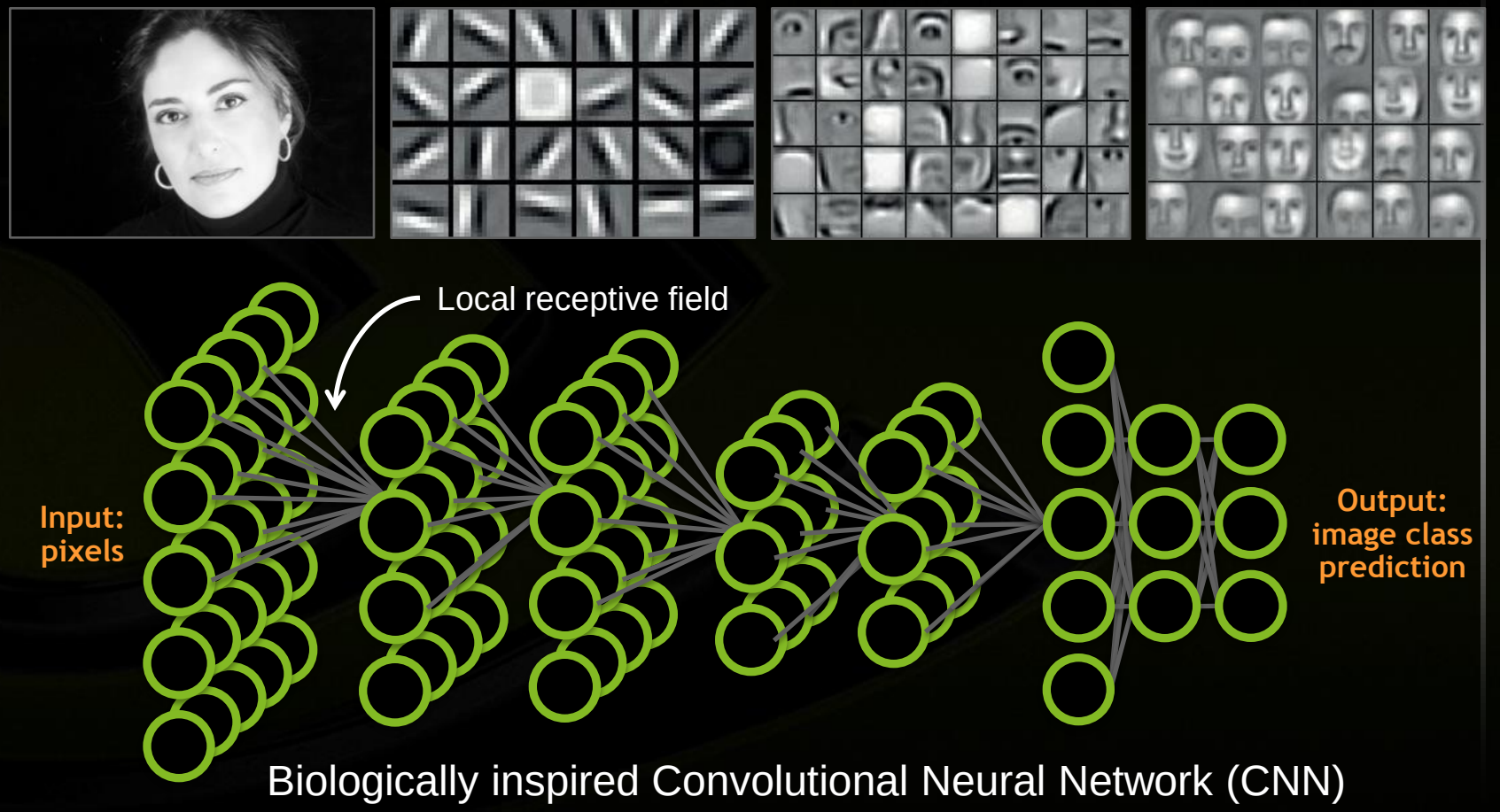
Deploy:





Dog

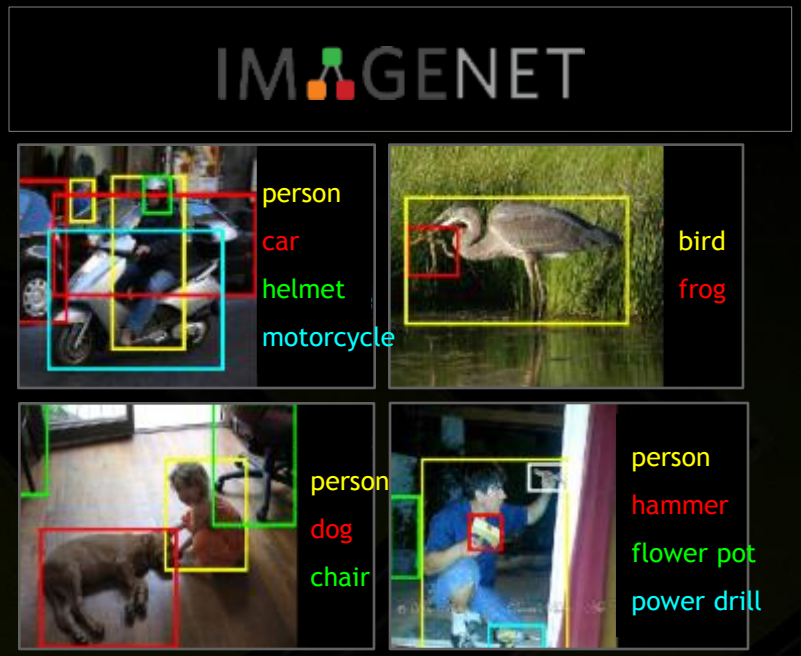
Deep Learning for Visual Perception



Application components:

- Task objective
  - e.g. Identify face
  - e.g. Classify age
- Training data
  - Typically 10K – 100M samples
- Network architecture
- Learning algorithm

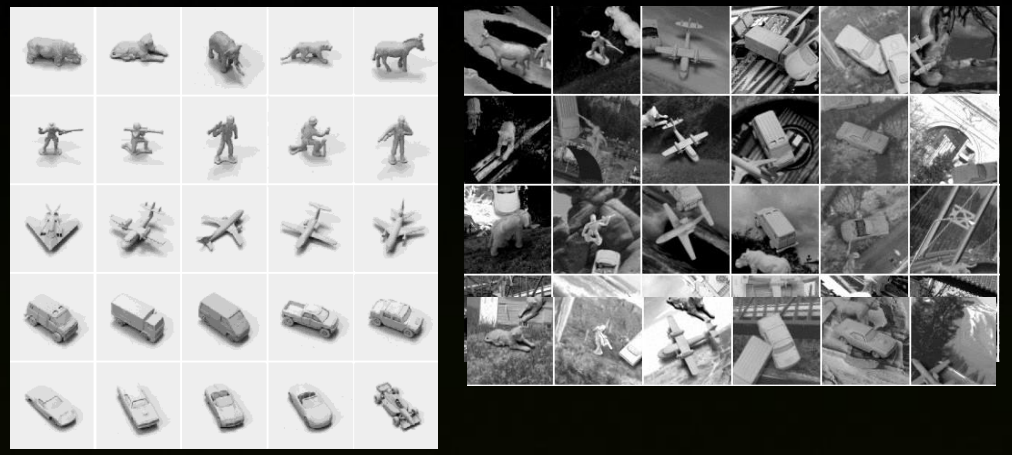
Visual Perception: DL State of the Art



1000 object classes
1.2 million training images [1]

Top-5 error (Google): 4.8%
Top-5 error (Human): 5.1%
Localization error (Oxford): 25.3%

NORB dataset (2004)

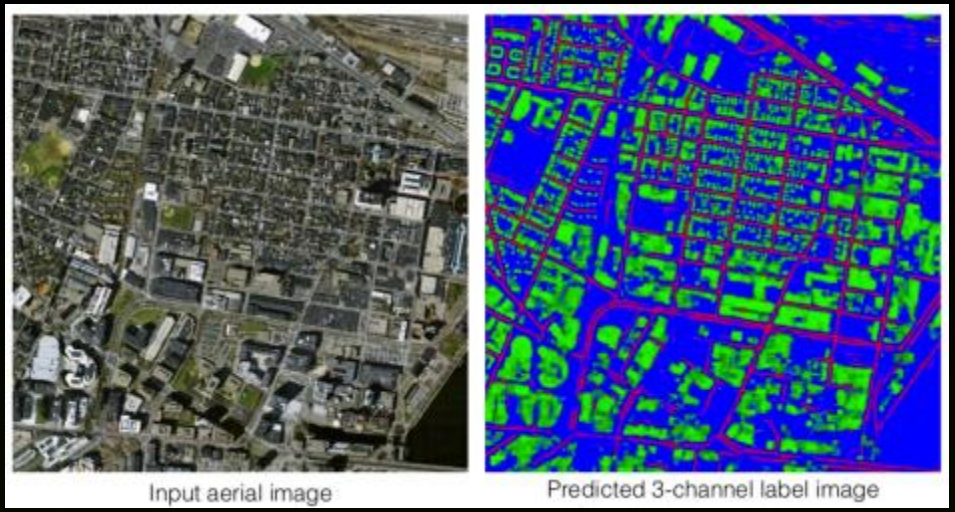


5 object classes
Multiple views and illuminations
291,600 training images
58,230 test images [2]

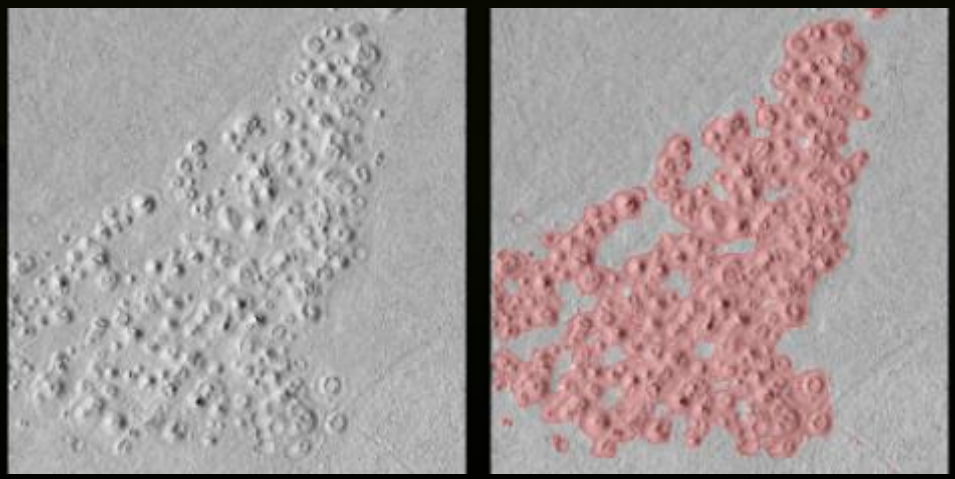
<6% classification error on test set with cluttered backgrounds (NYU)

Remote Sensing Imagery Exploitation

- Object detection and classification
- Scene segmentation
- Land usage classification
- Geologic feature classification
- Change detection
- Crop yield prediction
- Surface water estimation
- Population density estimation
- Super-resolution
- Photogrammetry

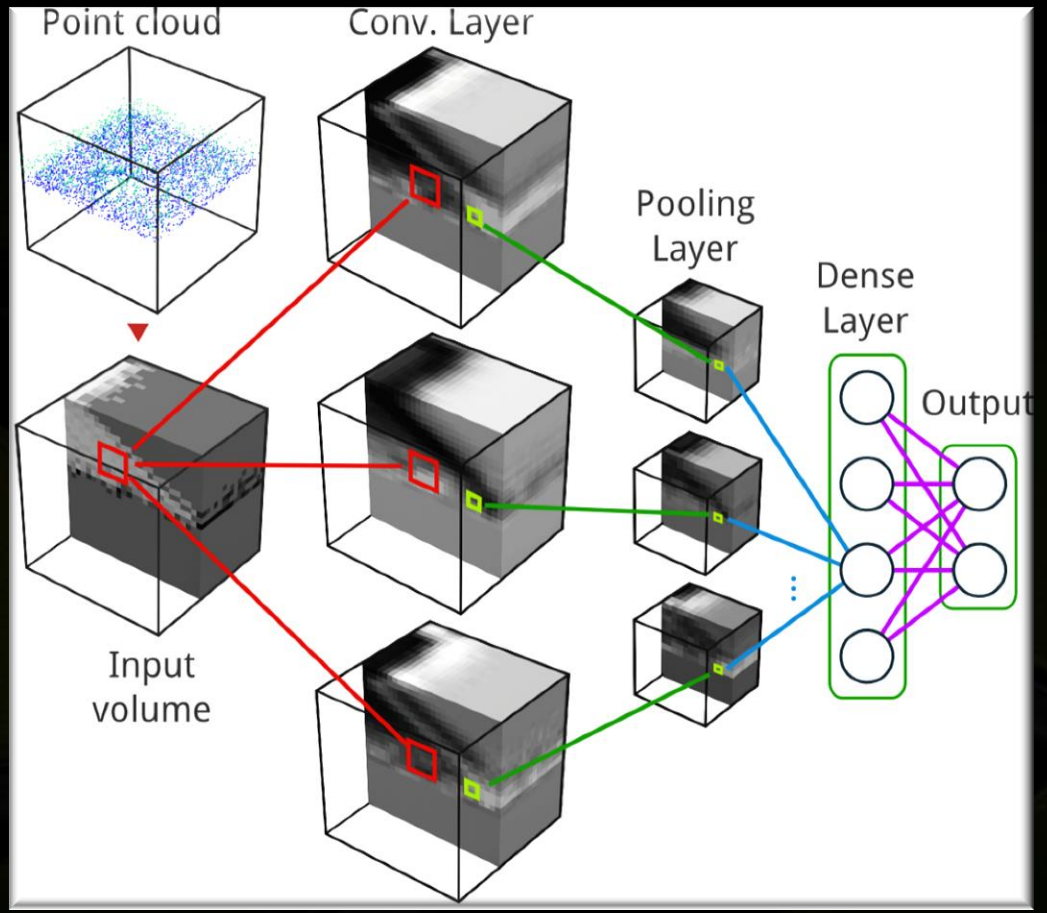


[3] Keio University, Japan – SPIE EI 2015



[4] University of Arizona

Advanced Imaging Modalities

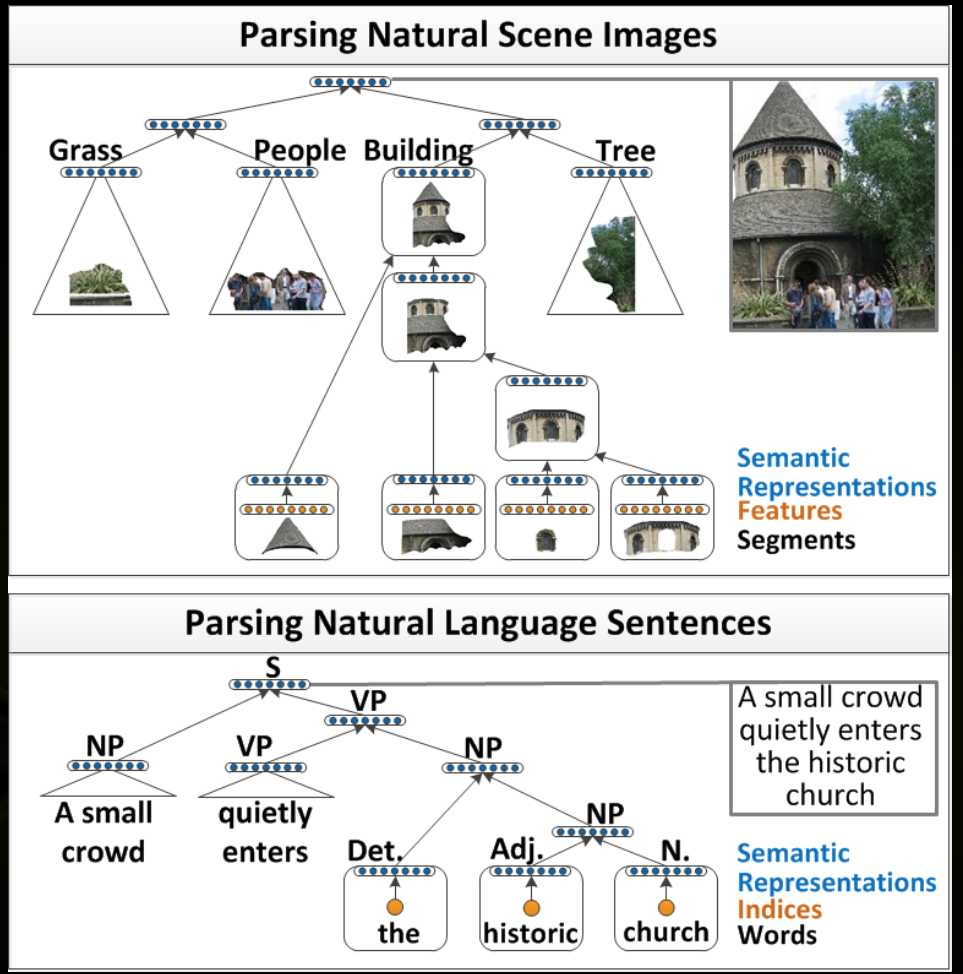


- CNN architecture supports:
  - MSI/HSI data cubes
  - SAR imagery
  - Volumetric data, e.g. LIDAR

[5] D. Maturana and S. Scherer. *3D Convolutional Neural Networks for Landing Zone Detection from LiDAR*. In ICRA. 2015

Open-source Imagery Exploitation

- Object detection
- Scene labeling
- Face recognition
- Image geo-location estimation
- Text extraction from images
- Geographic property estimation
- Image de-noising



[6] Stanford University, NLP group

Video and Audio Understanding

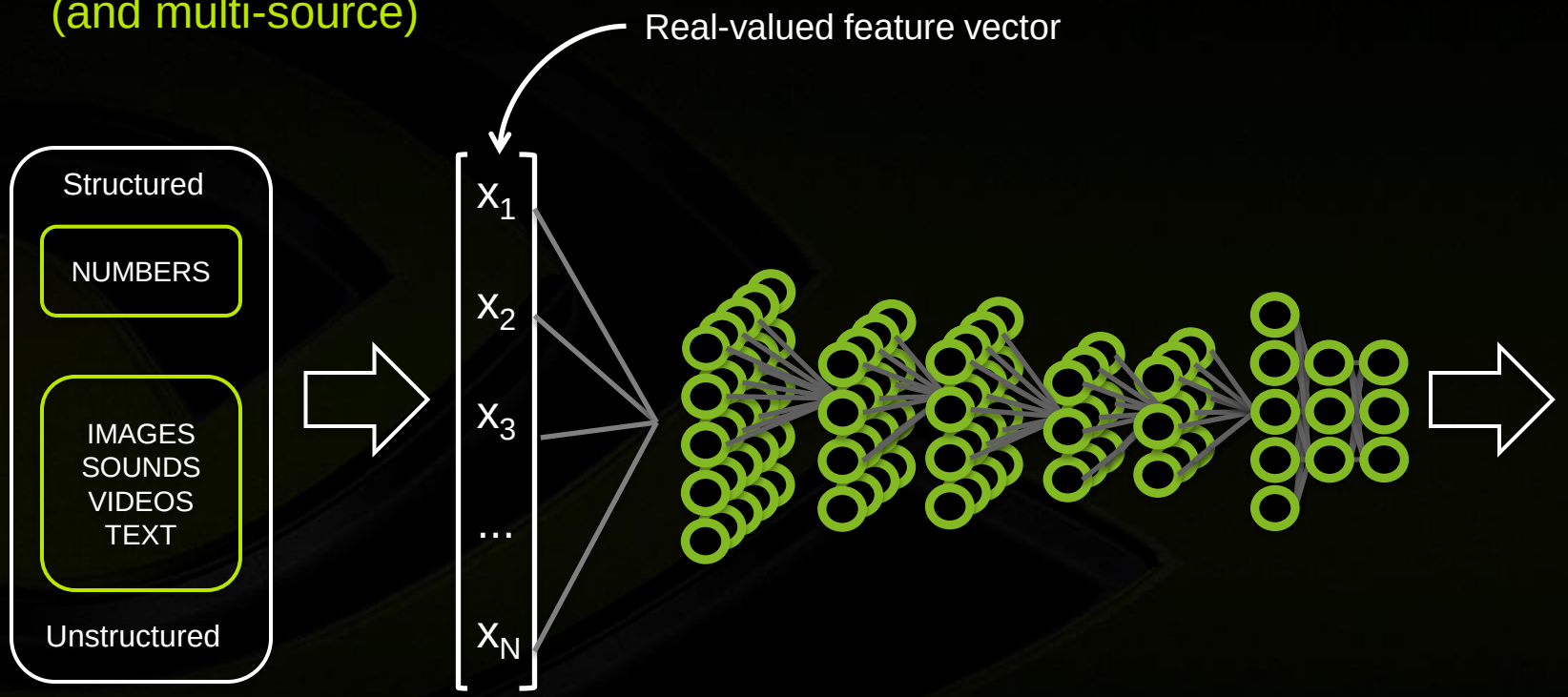
NVIDIA

[7] Google/Stanford

- **Video:**
  - Object tracking
  - Scene tagging
  - Human activity recognition
  - Occlusion prediction
  - 3D layout prediction
  - Augmented reality
- **Audio:**
  - Speech recognition
  - Speaker identification
  - Real-time translation

Deep Learning is adaptable

Varied data types
(and multi-source)



Varied tasks

Classification

Regression

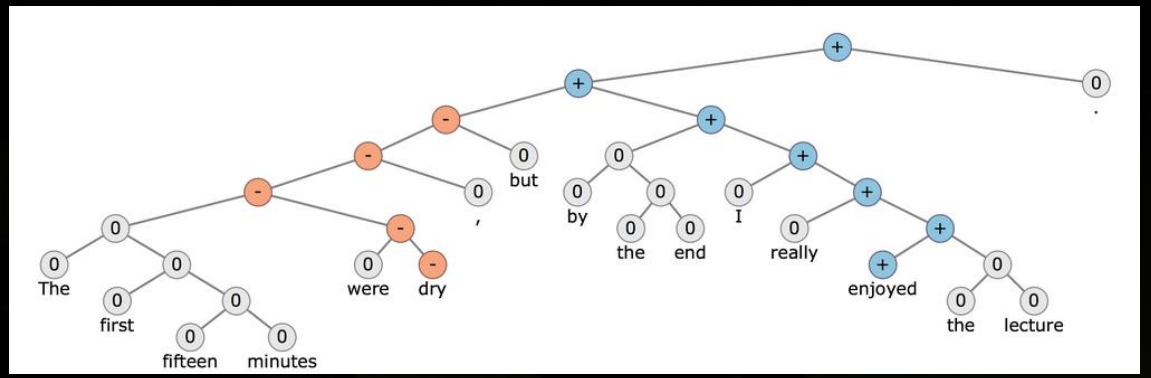
Unsupervised learning
- Clustering
- Topic extraction
- Anomaly detection

Sequence prediction

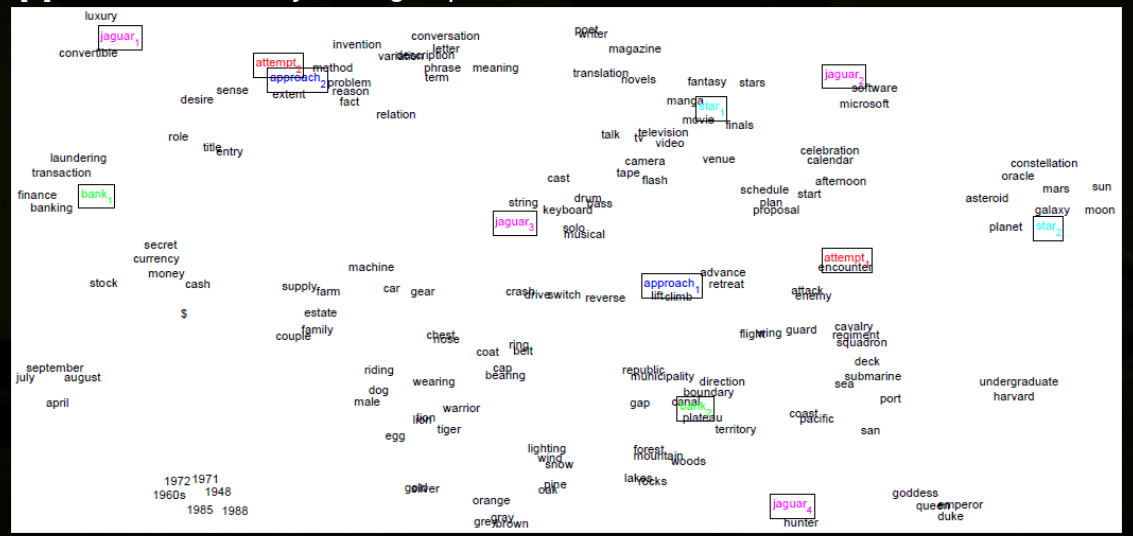
Control policy learning

Constants: Big (high dimensional) Data + a complex function to learn

Language Understanding



[8] Stanford University, NLP group

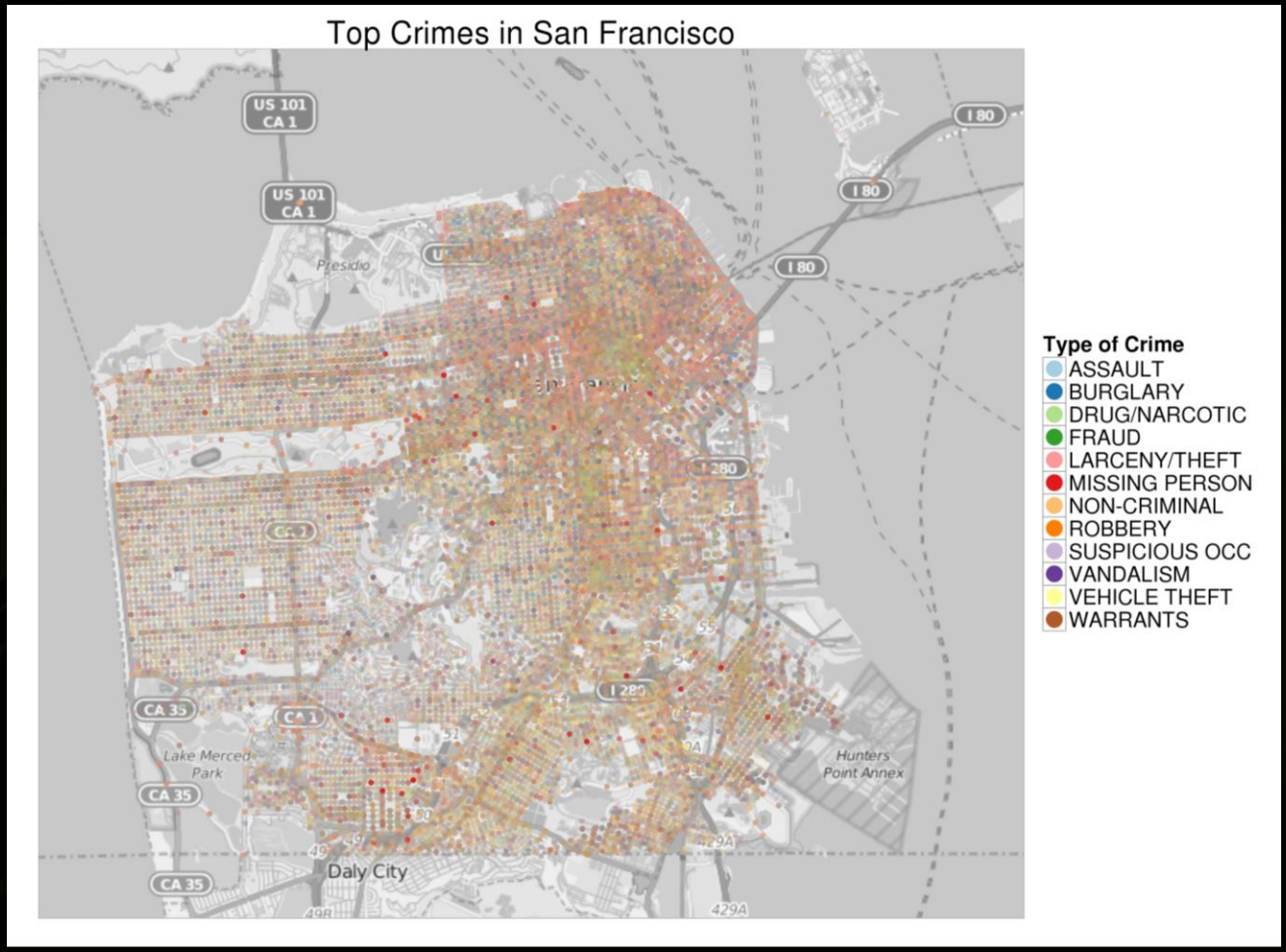


[9] Richard Socher (Metamind/Stanford NLP group)

- Topic extraction
- Named entity recognition
- Question Answering
- Author identification
- Language tagging and translation
- Image captioning
- Sentiment analysis

Geospatial Analytics

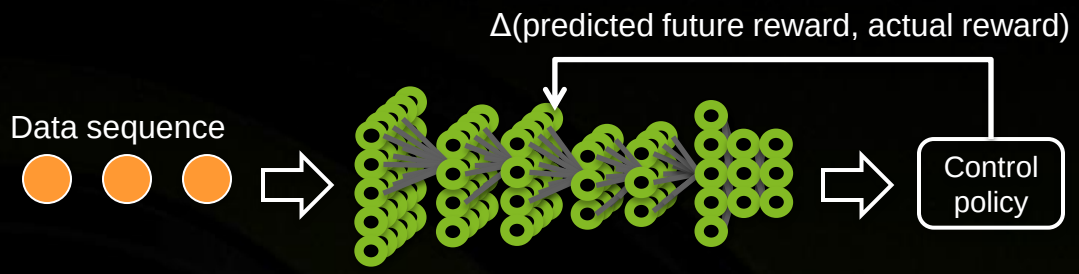
- 12 years of San Francisco crime reports
- Given date, time and location DL model predicts crime:
  - Top-5 error: 59%
- ~4 hours work (including training) using open source tools



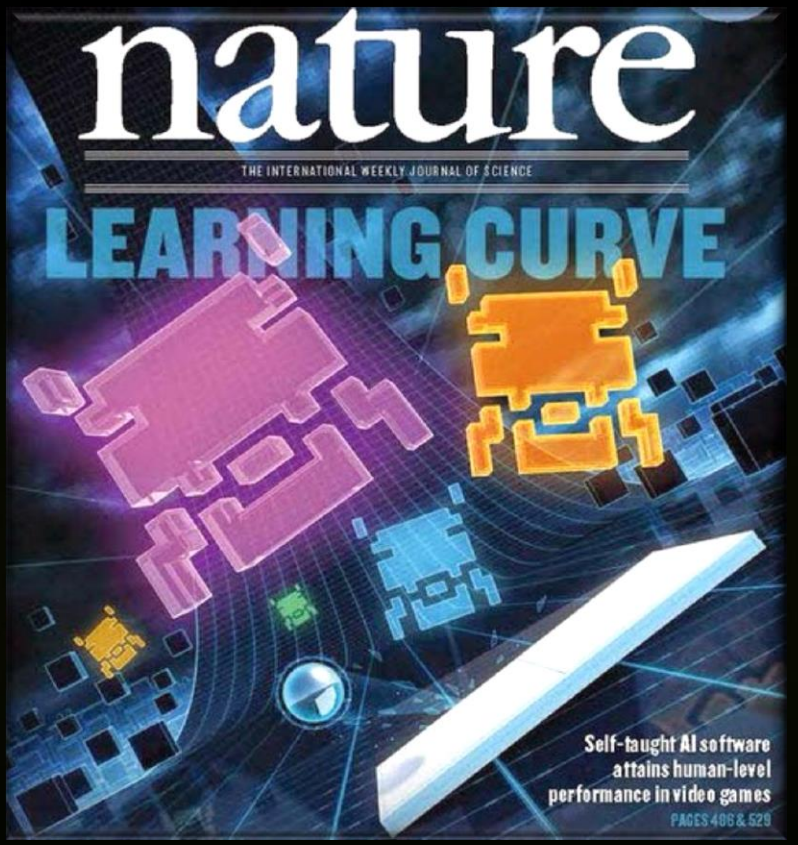
[10] Kaggle San Francisco Crime Classification Competition

Sensor/Platform Control

Reinforcement learning:



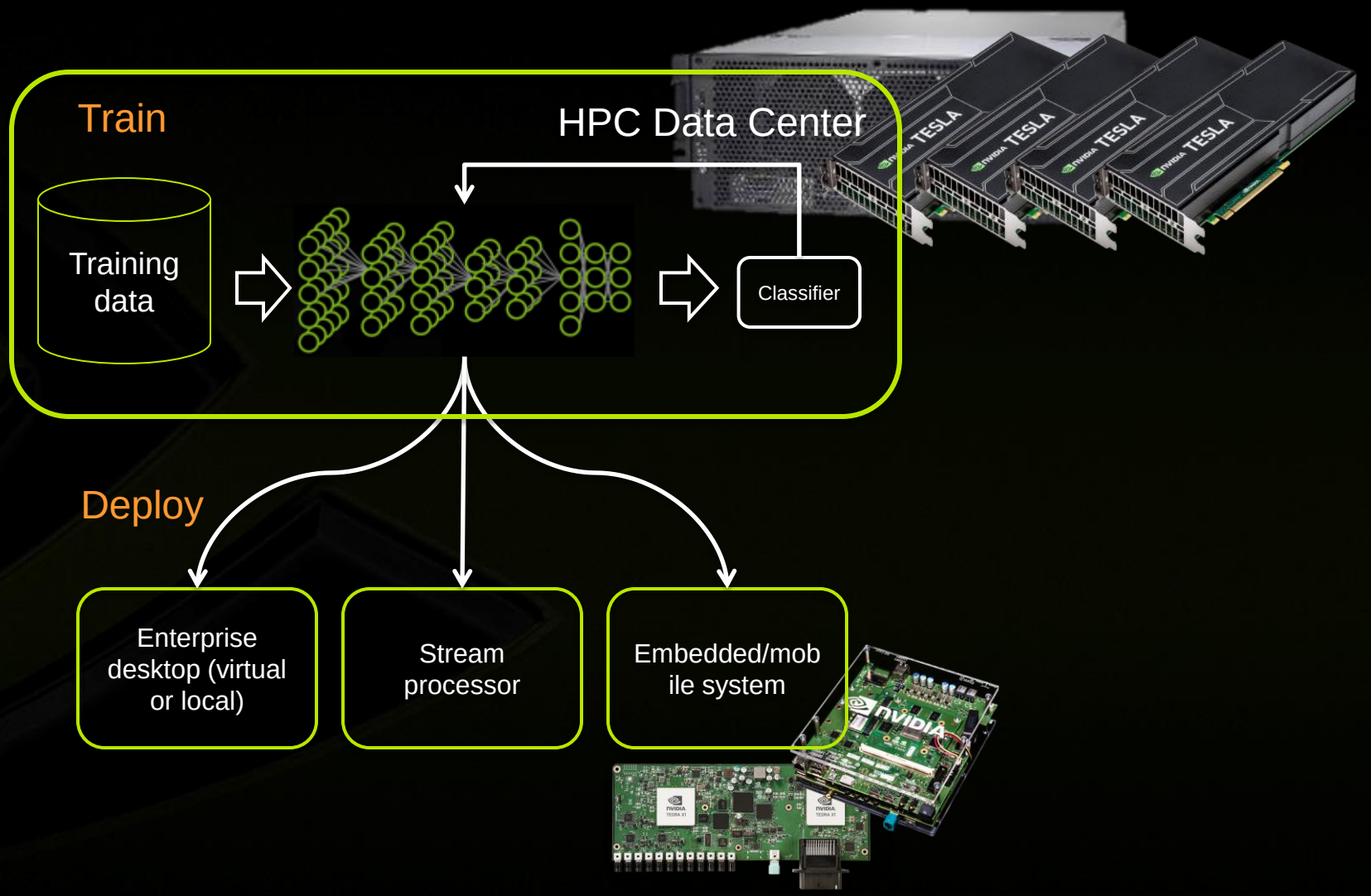
- **Applications:**
  - **Sensor tasking**
  - **Autonomous vehicle navigation**



[11] Google DeepMind in Nature

Deep Learning deployment options

Long training (hours to days), batch updates, leverage GPU acceleration



~100ms response for new data sample, model interactivity

Deep Learning is a GEOINT force multiplier

- **Managing Big Data**
  - Real-time near-human level perception at web-scale
- **Data exploration and discovery**
  - Semantic and similarity based search
  - Dimensionality reduction
  - Transfer learning
- **Model sharing**
  - Compact model representations
  - Models can be fine-tuned based on multiple analysts feedback

Summary

- **Deep Learning is:**
  - **Adaptable** to many varied GEOINT workflows and deployments scenarios
  - **Available** to apply in production and R&D today
  - **Approachable** using open-source tools and libraries

Resources

- Popular DL frameworks:
  - Caffe (UC Berkeley)
  - Theano (U Montreal)
  - Torch
- Examples from talk:
  - [1] Imagenet Large Scale Visual Recognition Challenge
  - [2] NORB dataset
  - [3] Keio University, Japan - Aerial image segmentation
  - [4] University of Arizona - Geographic feature detection
  - [5] D. Maturana and S. Scherer. *3D Convolutional Neural Networks for Landing Zone Detection from LiDAR*. In ICRA. 2015
  - [6], [8] Stanford NLP group Deep Learning research
  - [7] Google/Stanford Large Scale Video Classification with CNNs
  - [9] Richard Socher's word embedding research
  - [10] Kaggle San Francisco Crime Classification Competition
  - [11] Google DeepMind Nature article


Thank you
NVIDIA