APPLICATIONS OF DEEP LEARNING TO GEOINT

Jon Barker, Solutions Architect

August 2015
Overview

- Motivation
- Introduction to Deep Learning
- GEOINT applications
- Deep Learning deployment
- Questions
Motivation

Rapid growth in remote sensing numbers and capability

- 350 Million Images Uploaded a Day
- 100 Hours Video Uploaded Every Minute
- 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: Multi-INT analysis workflow**

**TODAY:**

BIG DATA
- NUMBERS
- IMAGES
- SOUNDS
- VIDEOS
- TEXT

**VISION:**

BIG DATA
- NUMBERS
- IMAGES
- SOUNDS
- VIDEOS
- TEXT

**BOTTLENECK**

Metadata filters → Human perception → Mission focused analysis

- Noisy content
- Near-perfect perception

**VISION:**

Automated machine perception → Semantic content based filters → Mission focused analysis

- Near-human level perception
- Mission relevant content
Deep Learning has become the most popular approach to developing Artificial Intelligence (AI) - machines that perceive and understand the world.

The focus is currently on specific perceptual tasks, and there are many successes.

Today, some of the world’s largest internet companies, as well as the foremost research institutions, are using GPUs for deep learning in research and production.
Practical Deep Learning Examples

- Image Classification, Object Detection, Localization, Action Recognition, Scene Understanding
- Pedestrian Detection, Traffic Sign Recognition
- Speech Recognition, Speech Translation, Natural Language Processing
- Breast Cancer Cell Mitosis Detection, Volumetric Brain Image Segmentation
Traditional Machine Perception – hand crafted features

- **Raw data**
  - [Image: Stapler]
  - [Image: Sound wave]
  - [Image: Document]

- **Feature extraction**
  - [Image: 3D scan]
  - [Image: Heat map]
  - [Image: Syntax tree]

- **(Linear) Classifier**
  - e.g. SVM
  - e.g. HMM
  - e.g. LSA

- **Result**
  - Speaker ID, speech transcription, …
  - Topic classification, machine translation, sentiment analysis…
Deep Neural Network (DNN)

- Modern reincarnation of Artificial Neural Networks
- A very large collection of simple, trainable mathematical units
- Collectively they can learn very complex functions mapping raw data to decisions
- Loosely inspired by biological brains

Raw data ➔ “dog” ➔ Output decision
Deep Learning approach

Train:

Dog

Cat

Honey badger

Feature extraction

(Linear) Classifier

Dog
Cat
Raccoon

Valid

Invalid
Deep Learning approach

Train:
- Dog
- Cat
- Honey badger

Errors:
- Dog
- Cat
- Raccoon

NVIDIA
Deep Learning approach

Train:
- Dog
- Cat
- Honey badger

Deploy:
- Dog
- Cat
- Raccoon

Errors:
- Dog
- Cat
- Raccoon

[Diagram showing the process of training and deploying a deep learning model with examples of correct and incorrect classifications.]
Deep Learning for Visual Perception

Input: pixels

Output: image class prediction

Biologically inspired Convolutional Neural Network (CNN)

Application components:

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

Local receptive field
Visual Perception: DL State of the Art

**IMAGENET**

- 1000 object classes
- 1.2 million training images [1]

**NORB dataset (2004)**

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

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

<6% classification error on test set with cluttered backgrounds (NYU)
Deep Learning Dominates at Visual Perception

1000 object classes
1.2 million training images [1]

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


[4] University of Arizona
Deep Learning supports the analyst
Advanced Imaging Modalities

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

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
Deep Learning Dominates at Visual Perception
Deep Learning supports the analyst

We can leverage the visual knowledge encoded in a Deep Neural Network to enable unlabelled data exploration.
Deep Learning generalizes across problems

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
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
Geospatial activity data

- Deep Neural Networks (DNNs) naturally ingest structured data
- Modern networks can learn complex predictive patterns including temporal sequences

Real-time destination prediction for taxis using DNN
Montreal Institute for Learning Algorithms (MILA), 2015
Sensor/Platform Control

Reinforcement learning:

\[ \Delta(\text{predicted future reward, actual reward}) \]

Applications:
- Sensor tasking
- Autonomous vehicle navigation

Why is Deep learning hot *now*?

Three Driving Factors...

**Big Data Availability**
- Facebook: 350 millions images uploaded per day
- Walmart: 2.5 Petabytes of customer data hourly
- YouTube: 100 hours of video uploaded every minute

**New DL Techniques**

**GPU acceleration**
Why are GPUs good for deep learning?

<table>
<thead>
<tr>
<th>Neural Networks</th>
<th>GPUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherently Parallel</td>
<td>✓</td>
</tr>
<tr>
<td>Matrix Operations</td>
<td>✓</td>
</tr>
<tr>
<td>FLOPS</td>
<td>✓</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>✓</td>
</tr>
</tbody>
</table>

**GPUs deliver --**
- same or **better** prediction accuracy
- faster results
- smaller footprint
- lower power
- lower cost
Deep learning with COTS HPC systems
A. Coates, B. Huval, T. Wang, D. Wu, A. Ng, B. Catanzaro
ICML 2013

“Now You Can Build Google’s $1M Artificial Brain on the Cheap”

GOOGLE DATACENTER
1,000 CPU Servers
2,000 CPUs • 16,000 cores
600 kWatts
$5,000,000

STANFORD AI LAB
3 GPU-Accelerated Servers
12 GPUs • 18,432 cores
4 kWatts
$33,000

GPUs make deep learning accessible
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
- Integrates into analytical workflows
  - Semantic content based filtering and search
  - Drives data exploration and visualization
  - Models improve based on analyst feedback

Scales across problems
- Models improve with more, varied data
- Models from one dataset can be leveraged in new problems
- Compact models can be easily shared and deployed
GPU accelerated Deep Learning is:

- **Revolutionizing** machine perception accuracy
- **Adaptable** to many varied GEOINT workflows and deployments scenarios
- **Scalable** – thrives on complex raw data
- **Available** to apply in production and R&D today
Popular DL frameworks:
- Caffe (UC Berkeley)
- Theano (U Montreal)
- Torch
- DIGITS

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
- [9] Kaggle Taxi Trajectory Prediction Competition
- [10] Kaggle San Francisco Crime Classification Competition