The DEEP LEARNING Revolution

Gunter Roth
gunterr@nvidia.com
SA EMEA
October 2016
KEY DRIVERS FOR DEEP LEARNING

Big Data
- 350 millions images uploaded per day

Better Algorithms
- 2.5 Petabytes of customer data hourly

GPU Acceleration
- 300 hours of video uploaded every minute

“The Three Breakthroughs that have Finally Unleashed A.I. on the World”
The DL Software stack

NVIDIA DIGITS™
Interactively manage data and train deep learning models for image classification without the need to write code.

Deep Learning Frameworks
Design and train deep learning models using a high-level interface. Choose a deep learning framework that best suits your needs based on your choice of programming language, platform, and target application.

NVIDIA Deep Learning SDK
This SDK delivers high-performance multi-GPU acceleration and industry-vetted deep learning algorithms, and is designed for easy drop-in acceleration for deep learning frameworks.

Source: https://developer.nvidia.com/deep-learning-getting-started
EVERY INDUSTRY WANTS DEEP LEARNING

Cloud Service Provider
- Image/Video classification
- Speech recognition
- Natural language processing

Medicine
- Cancer cell detection
- Diabetic grading
- Drug discovery

Media & Entertainment
- Video captioning
- Content based search
- Real time translation

Security & Defense
- Face recognition
- Video surveillance
- Cyber security

Autonomous Machines
- Pedestrian detection
- Lane tracking
- Recognize traffic sign
CONVOLUTIONAL NEURAL NETWORKS

Long short-term memory (LSTM)

Hochreiter (1991) analysed vanishing gradient “LSTM falls out of this almost naturally”

Long time dependencies are preserved until input gate is closed (-) and forget gate is open (O)

Gates control importance of the corresponding activations

Fig from Vinyals et al, Google April 2015 NIC Generator

Fig from Graves, Schmidhuber et al, Supervised Sequence Labelling with RNNs
THE NEXT STEP – NATURAL LANGUAGE PROCESSING

Figure 4. Diagram of our multimodal Recurrent Neural Network generative model. The RNN takes a word, the context from previous images and sentence embeddings to generate the next word. The image-sentence score $S_{im}$ is calculated as the max over all words in the sentence. The score is then used to select the next word. The diagram shows how the model integrates visual and textual information to generate natural language descriptions.

Figure 3. Diagram for evaluating the image-sentence score $S_{im}$. Object regions are embedded with a CNN (left). Words (enriched by their context) are embedded in the same multimodal space with a BRNN (right). Pairwise similarities are computed with inner products (magnitudes shown in grayscale) and finally reduced to image-sentence score with Equation 8.
GOOGLE DEEPMIND ALPHAGOGO CHALLENGE
MICROSOFT: “SUPER DEEP NETWORKS”

Revolution of Depth

ImageNet Classification top-5 error (%)

Microsoft Deep ResNet

>6X MORE FLOPS

18 LAYERS
1.8 GF

152 LAYERS
11.3 GF
Deep Learning Hardware
END-TO-END PRODUCT FAMILY

TRAINING

FULLY INTEGRATED DL SUPERCOMPUTER
DGX-1

DESKTOP
Titan X

HYPERSCALE
Tesla M40
Tesla P40

INFERENCE

DATA CENTER
Tesla M4
Tesla P4

AUTOMOTIVE
Drive PX

EMBEDDED
Jetson TX1
NVIDIA DGX-1
WORLD’S FIRST DEEP LEARNING SUPERCOMPUTER

170 TFLOPS FP16
8x Tesla P100 16GB
NVLink Hybrid Cube Mesh
Accelerates Major AI Frameworks
Dual Xeon
7 TB SSD Deep Learning Cache
Dual 10GbE, Quad IB 100Gb
3RU - 3200W
Training organizations and individuals to solve challenging problems using Deep Learning

On-site workshops and online courses presented by certified experts

Covering complete workflows for proven application use cases
Image classification, object detection, natural language processing, recommendation systems, and more

www.nvidia.com/dli
DQN: deep Q-learning network

REINFORCEMENT LEARNING

Mastering Breakout
Pieter Abbeel
gym.openai.com
Questions?
Bon appetit!
gunterr@nvidia.com