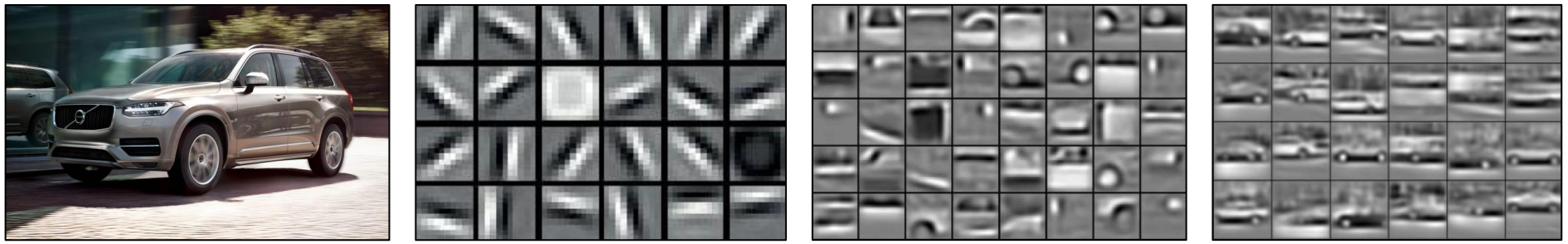
GPU TECHNOLOGY CONFERENCE

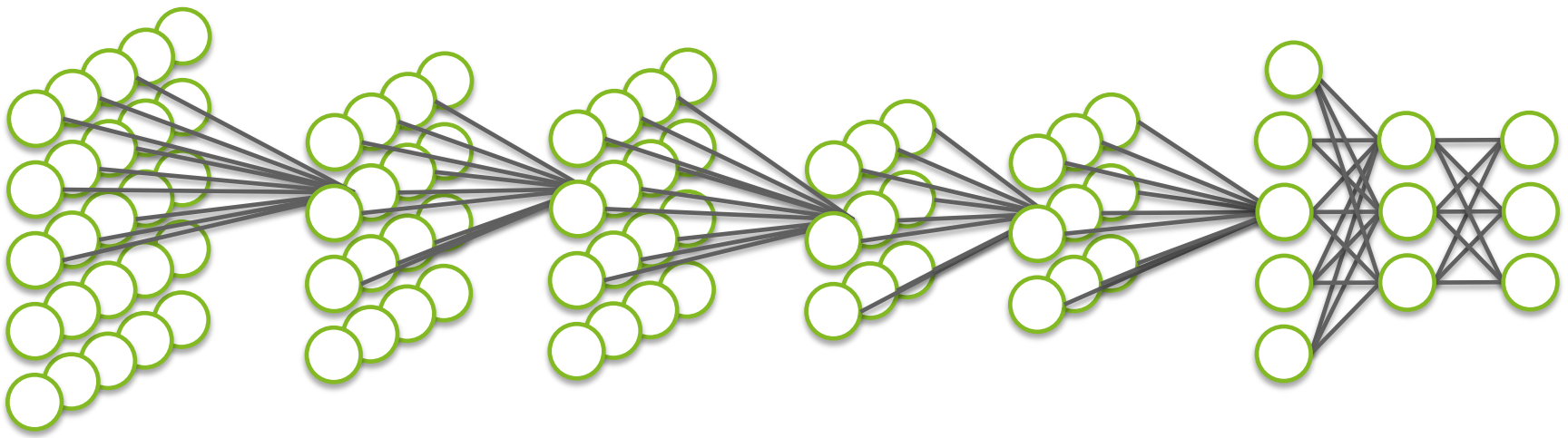
DEEP LEARNING WITH NVIDIA GPUS

JONATHAN COHEN, DIRECTOR OF ENGINEERING
DEEP LEARNING SOFTWARE, NVIDIA

What is Deep Learning?

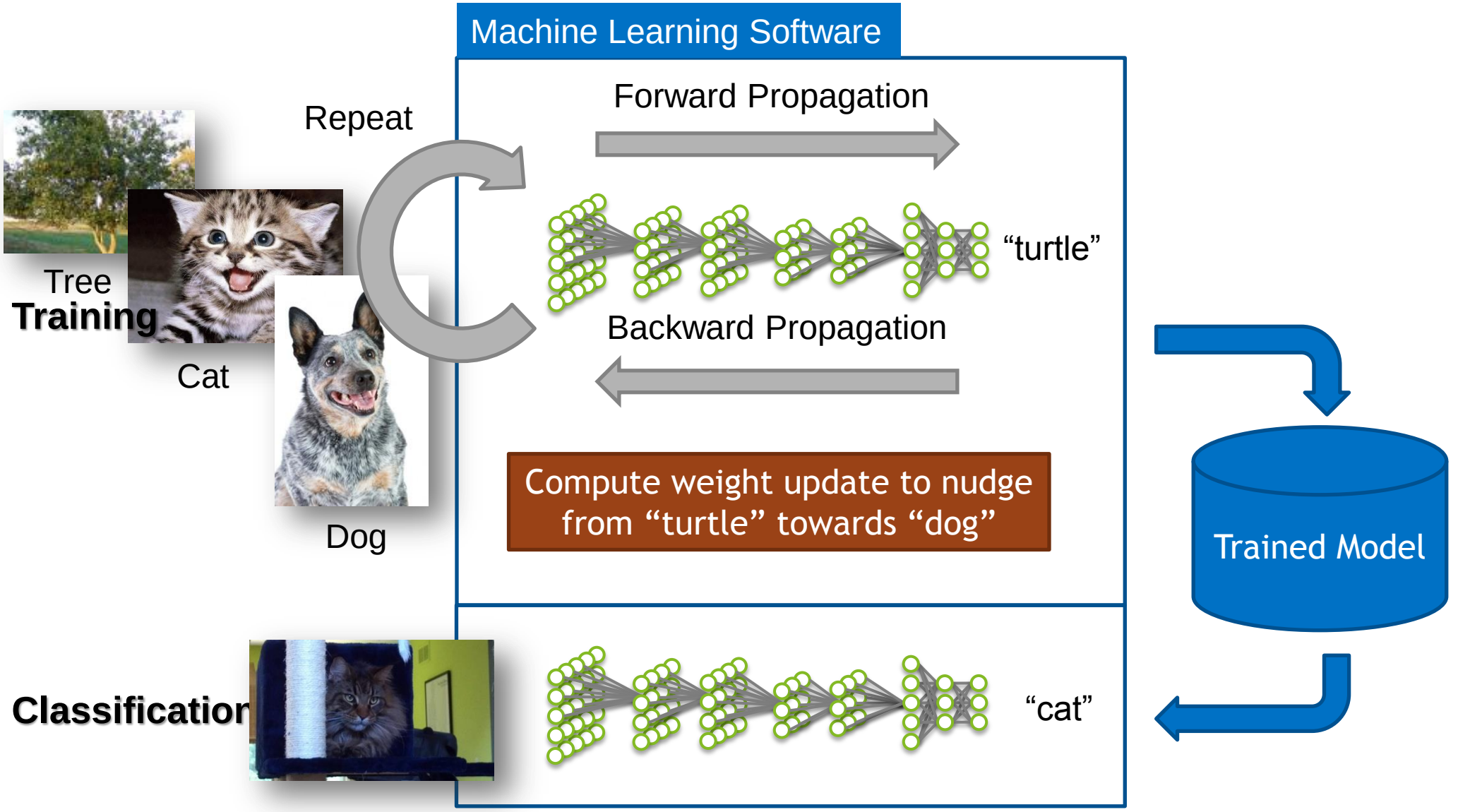


Image



“Volvo XC90”

Image source: “Unsupervised Learning of Hierarchical Representations with Convolutional Deep Belief Networks” ICML 2009 & Comm. ACM 2011. Honglak Lee, Roger Grosse, Rajesh Ranganath, and Andrew Ng.


Machine Learning Software
Repeat
Forward Propagation
"turtle"
Backward Propagation
Compute weight update to nudge from "turtle" towards "dog"
Tree
Training
Cat
Dog
Trained Model
Classification
"cat"

Why is Deep Learning Hot Now?

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 ML Techniques

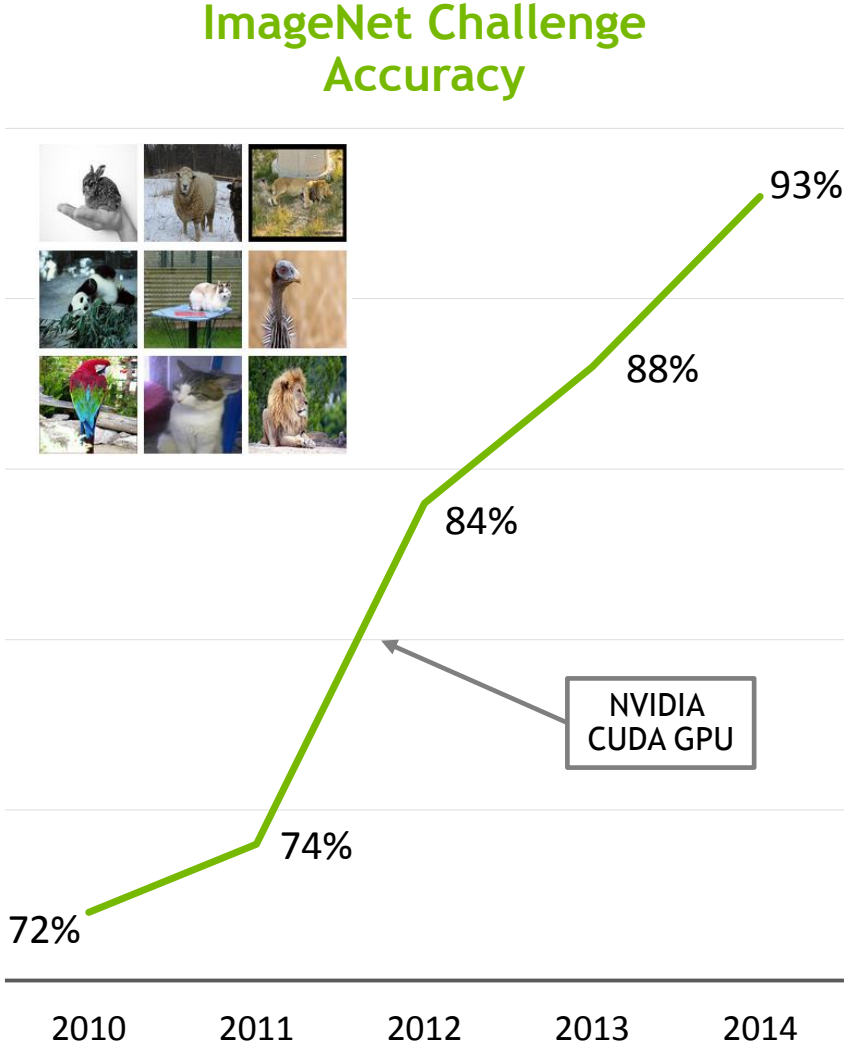
GPU Acceleration

GPUs and Deep Learning

	NEURAL NETWORKS	GPUS
Inherently Parallel	✓	✓
Matrix Operations	✓	✓
FLOPS	✓	✓
Bandwidth	✓	✓

GPUs deliver --

- *same or better prediction accuracy*
- *faster results*
- *smaller footprint*
- *lower power*

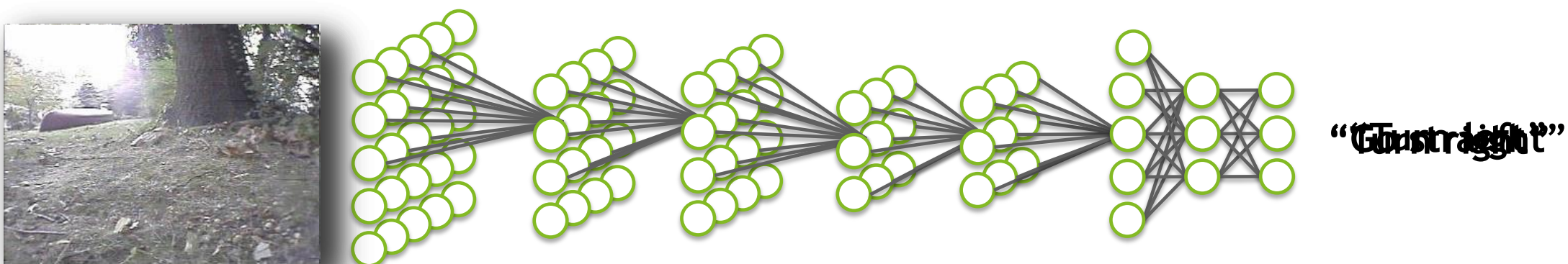


DAVE

DARPA Autonomous Vehicle (2004)

Deep Learning approach to robot navigation

Deep Neural Network “watches” human drivers , learns how to react



The Theory of DAVE

Learn: Visual Input => Action

DAVE in Action

DRIVE PX

An advanced computing platform based on NVIDIA Tegra processors for autonomous driving cars

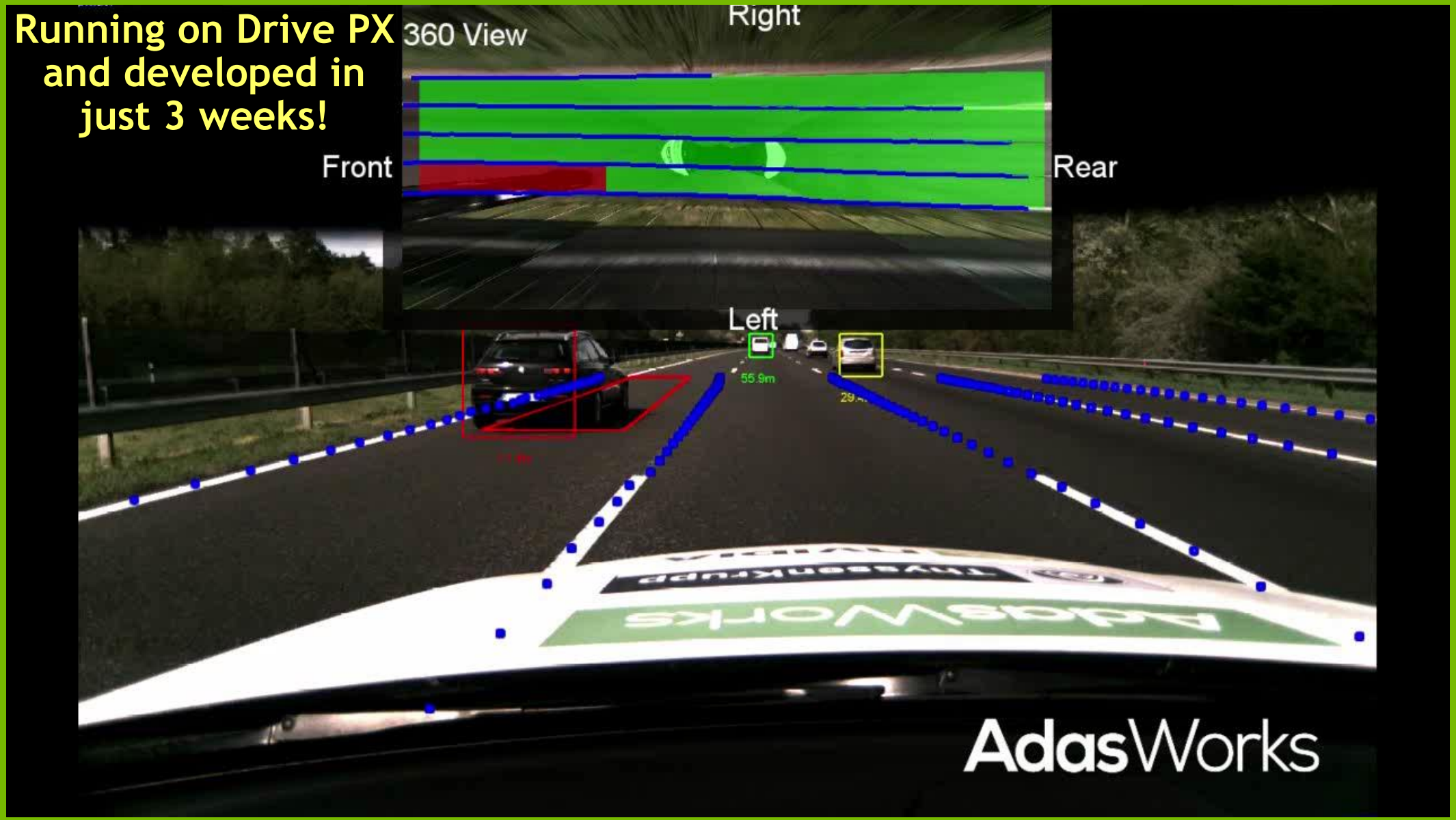
FEATURES

The ability to capture and process multiple HD camera and sensor inputs

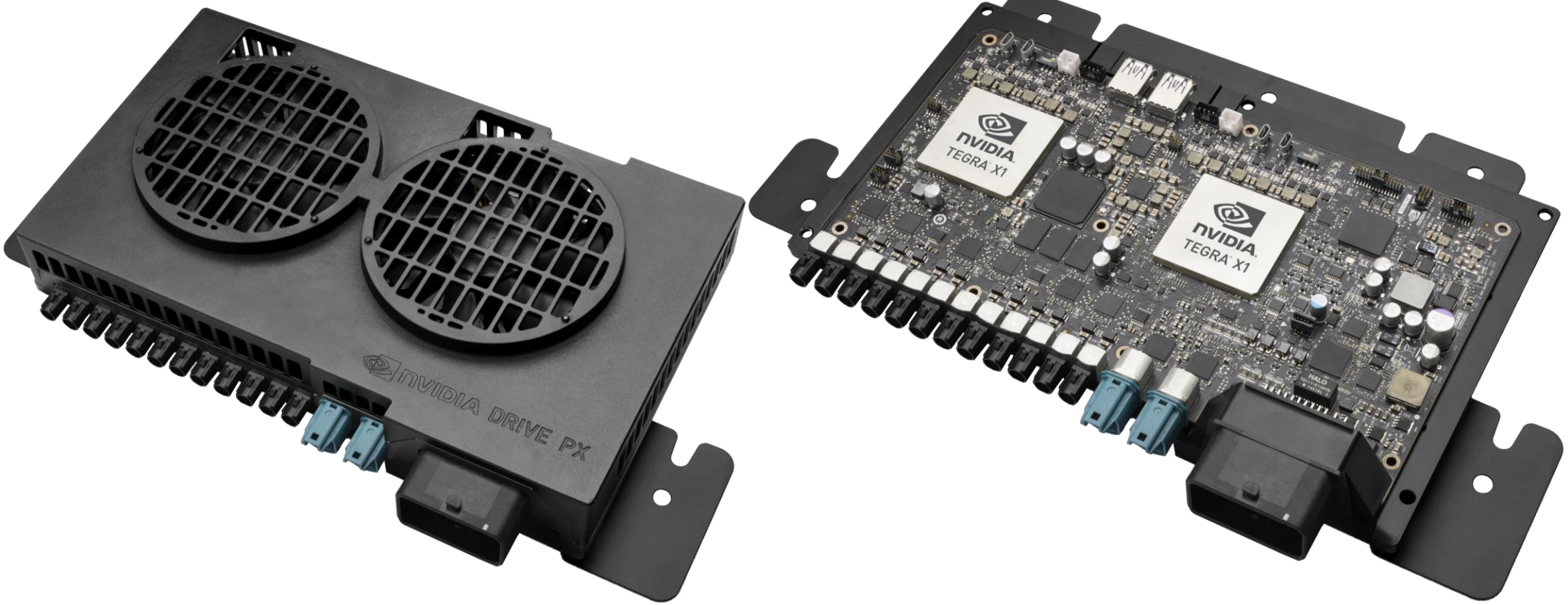
A rich middleware for computer graphics, computer vision and deep learning

A powerful and easy to develop platform for algorithm research and rapid prototyping

Running on Drive PX and developed in just 3 weeks!



Drive PX Development Platform

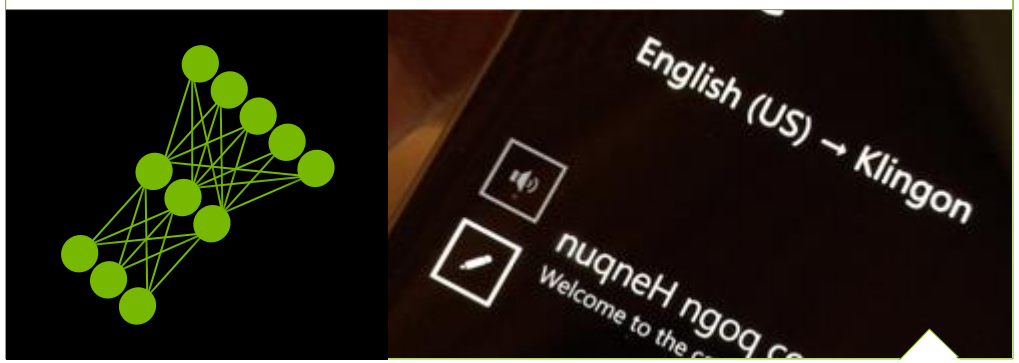


Todo: better version of this slide

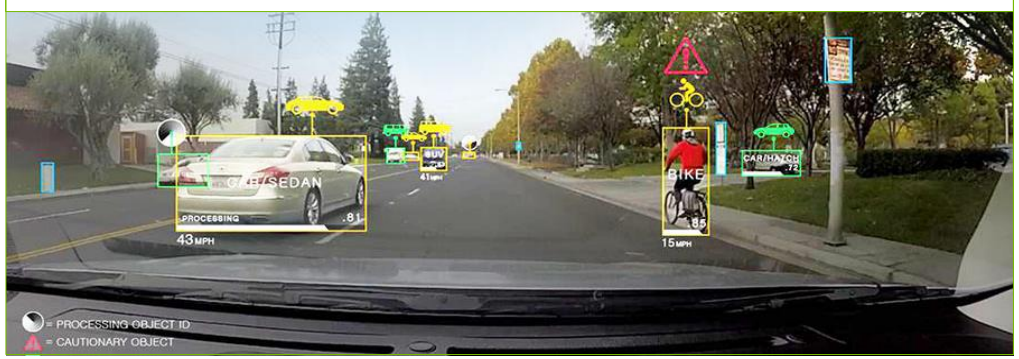
Practical Examples of Deep Learning

Image Classification, Object Detection, Localization, Action Recognition

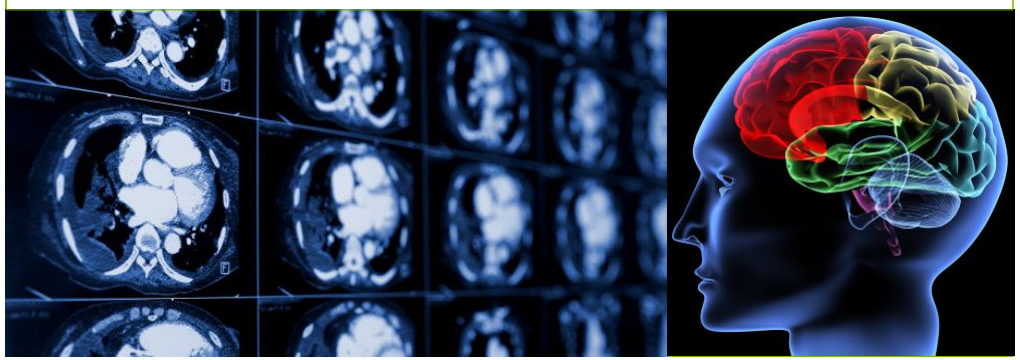
Speech Recognition, Speech Translation, Natural Language Processing



Pedestrian Detection, Lane Detection, Traffic Sign Recognition



Breast Cancer Cell Mitosis Detection, Volumetric Brain Image Segmentation



The Deep Learning Community: Detecting Diabetic Retinopathy

kaggle

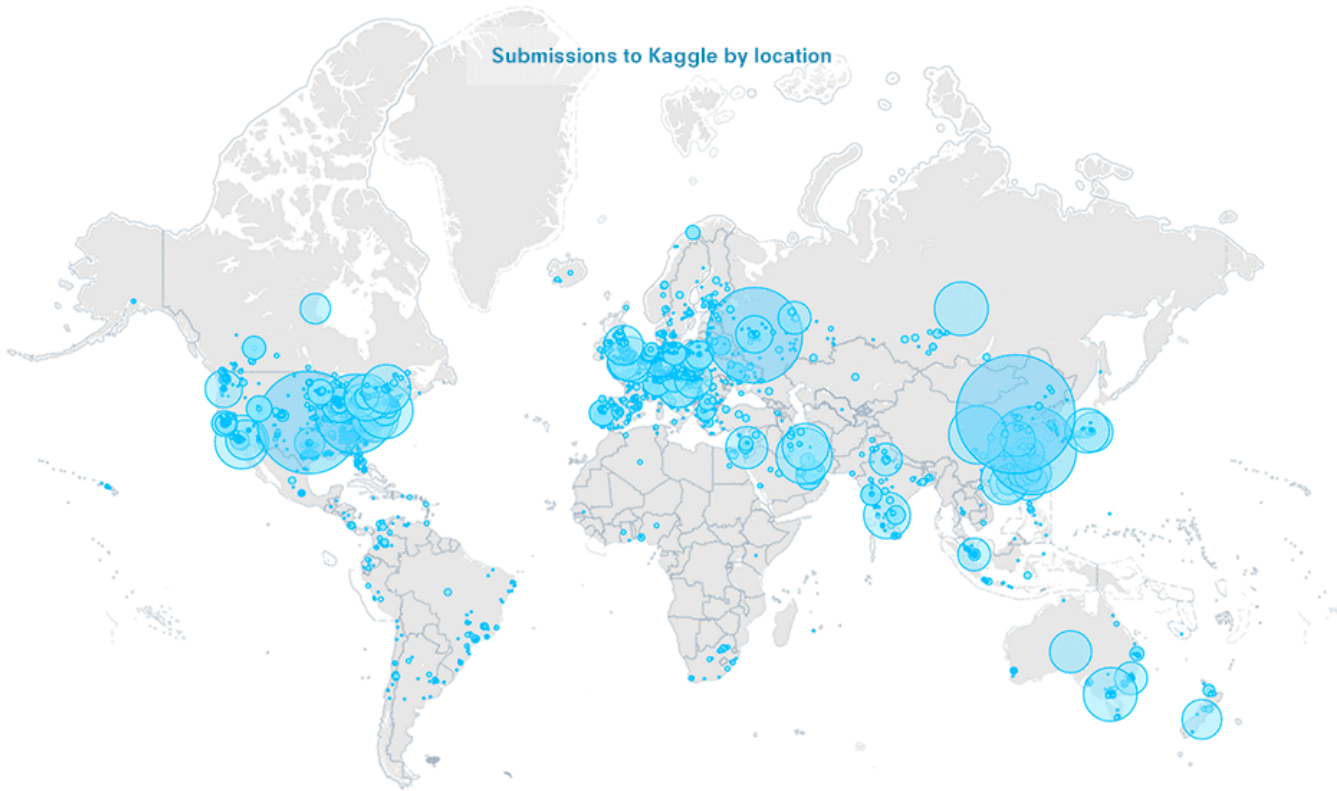
Data Scientist Community

Founded in 2010

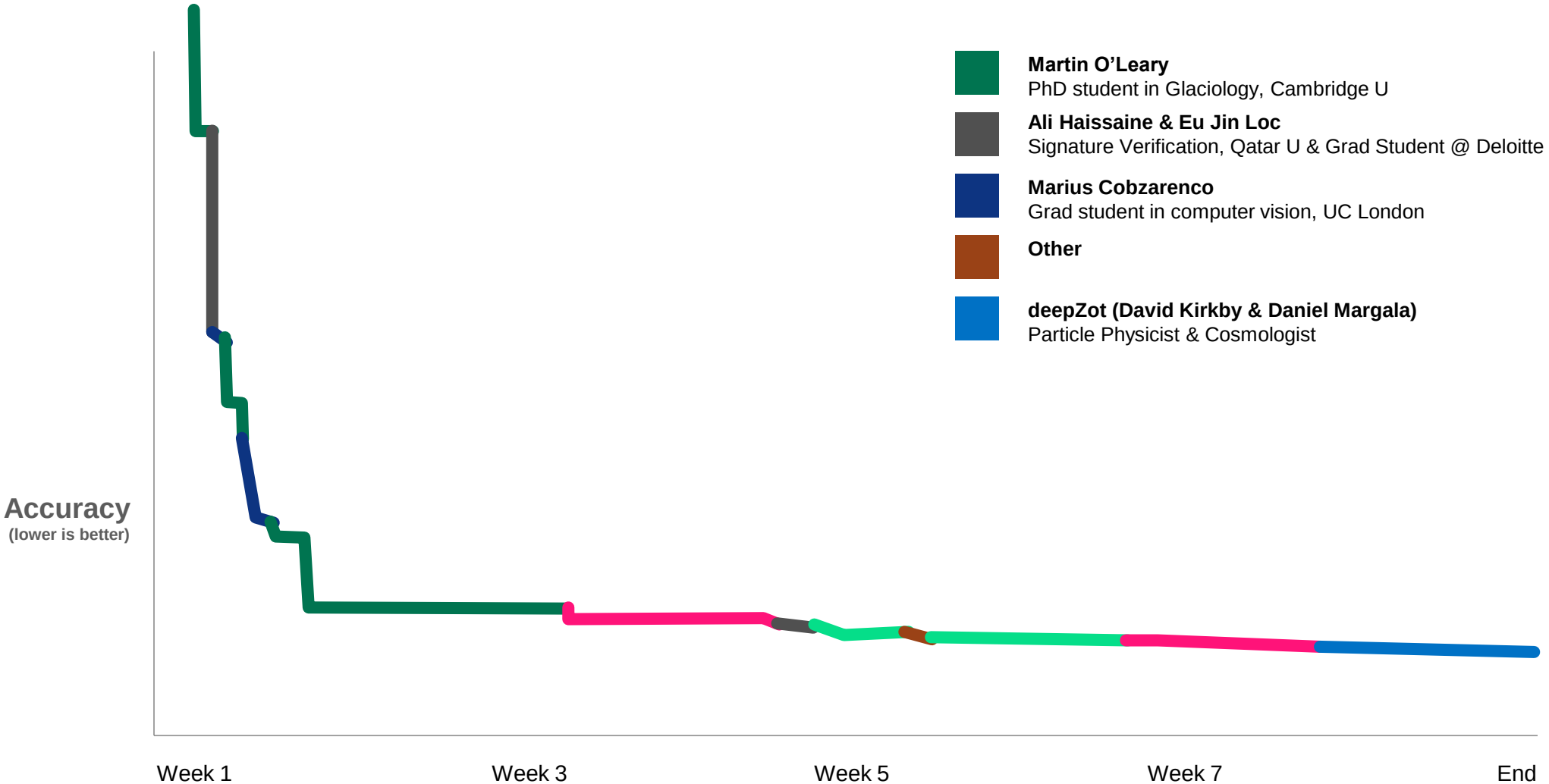
Sponsor contests to spur collaborative problem solving

354K data scientists in Kaggle community

92K machine learning models submitted to Kaggle competitions each month



Strength of Community-based Data Science – “Mapping Dark Matter” results



Diabetic Retinopathy

Affects 347 million people worldwide

Leading cause of blindness among working age population in developed world

Changes to blood vessels in the retina lead to aneurisms and fluid leaks

If no treated early, can cause blindness

Requires regular screenings

Fundus photography with interpretation by trained physician

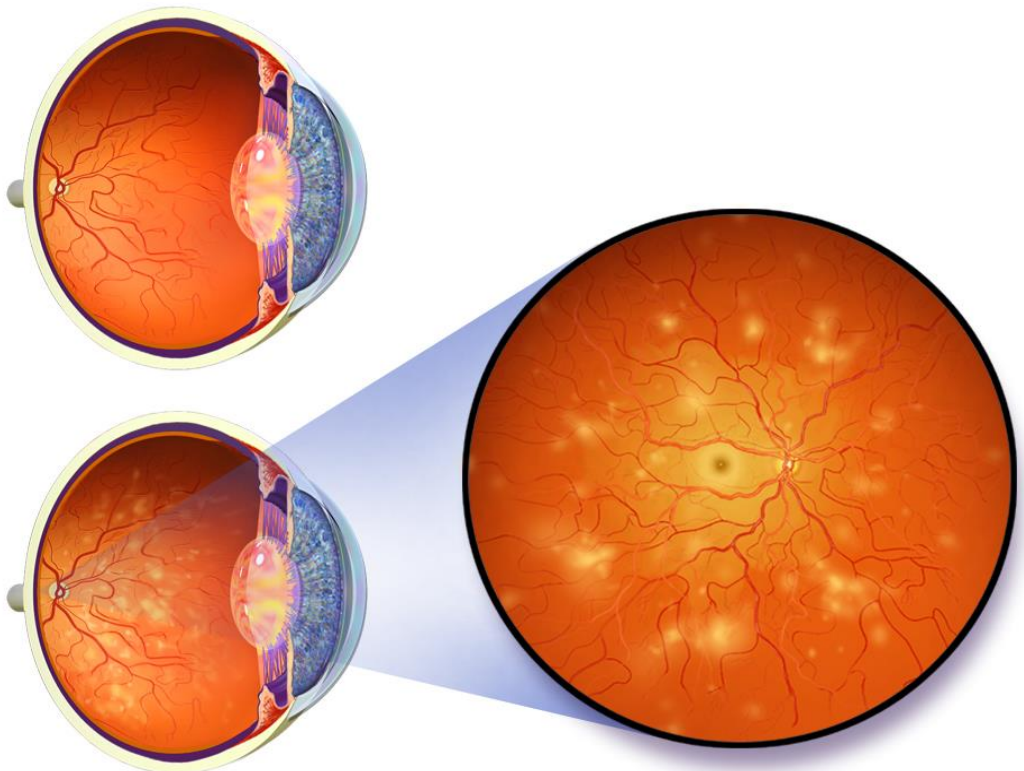


Image credit: Blausen.com staff. "Blausen gallery 2014". *Wikiversity Journal of Medicine*. DOI:10.15347/wjm/2014.010. ISSN 20018762.

Kaggle Diabetic Retinopathy Contest

$100,000 award sponsored by the California Healthcare Foundation

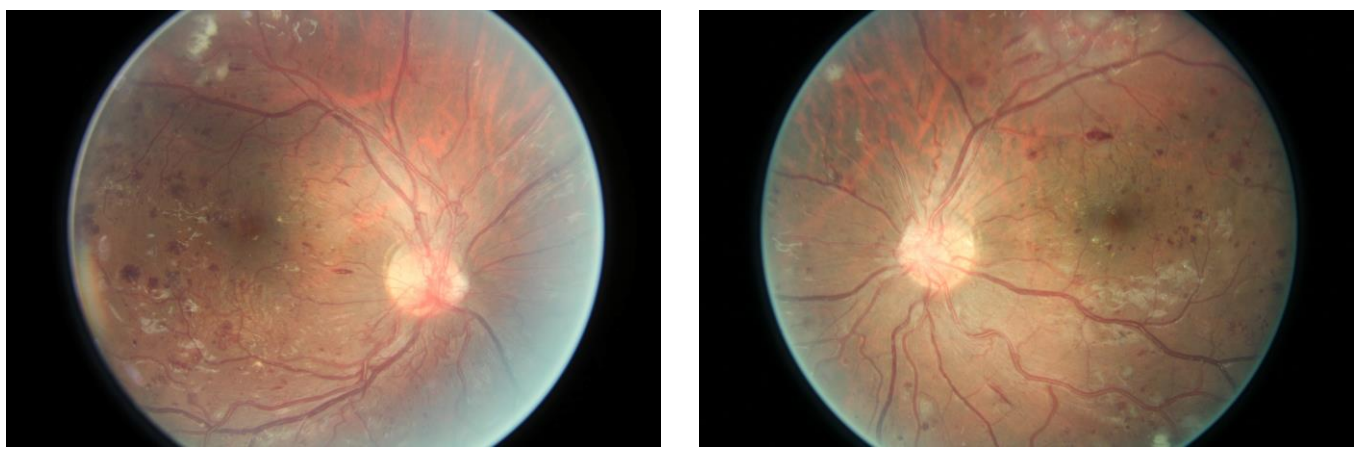
Contestants provided 17,000 left/right images with score: 0 (healthy) to 4 (diseased)

Typical clinician scores 0.83 (1.0 = perfect agreement with another clinician)

661 teams entered

Winning score 0.84958

4 teams above 0.83



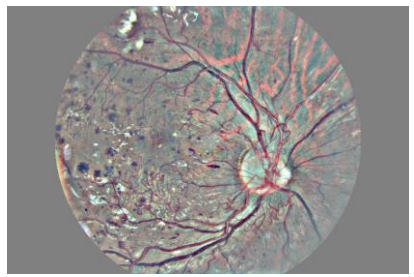
Benjamin Graham – Finished #1!

Assistant Professor in Stats and Complexity, University of Warwick

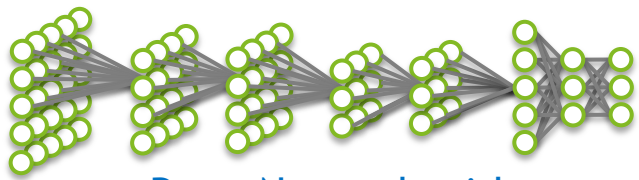
SparseConvNet
(written in CUDA)

NVIDIA CUDA

NVIDIA GPU



SparseConvNet



Deep Network with
fractional pooling

Random Forest
(scikits-learn)

Antreas Antoniou – Finished in top 3rd

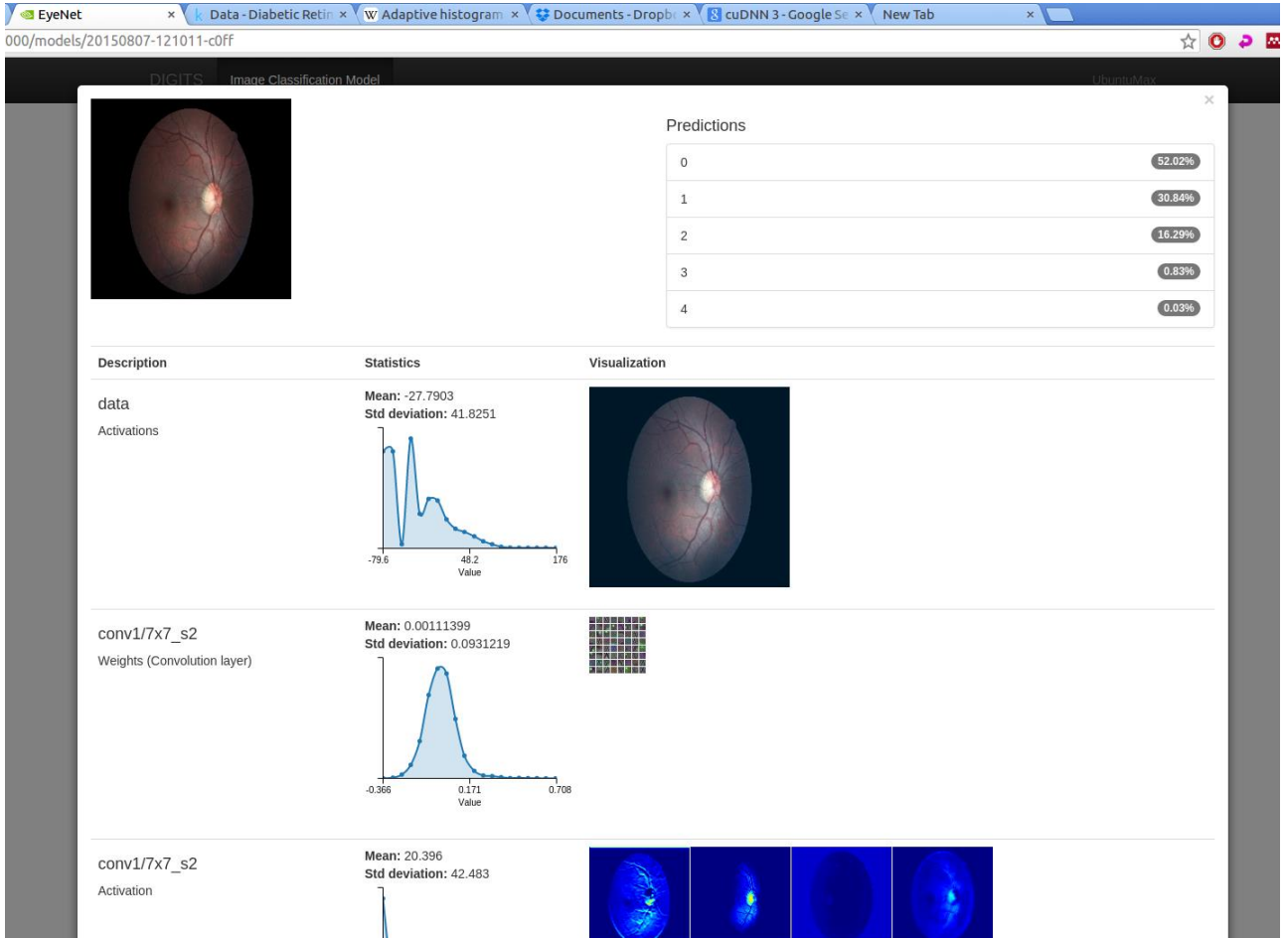
Master's Data Science student, University of Lancaster

NVIDIA DIGITS

NVIDIA cuDNN

NVIDIA CUDA

NVIDIA GPU



Deep Learning Platform

DEVELOPMENT

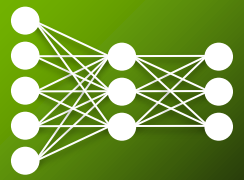
DEPLOYMENT

Applications

DIGITS Tools

Deep Learning Frameworks

Software



cuDNN

Systems

DIGITS DevBox

Hardware

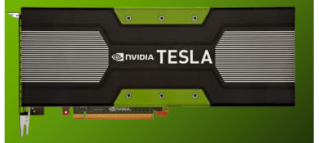
Titan X

Software

System Management

Systems

Hardware



Tesla

NVIDIA DEEP LEARNING PLATFORM

NVIDIA cuDNN

GPU Acceleration for Deep Learning Frameworks

High performance neural network training

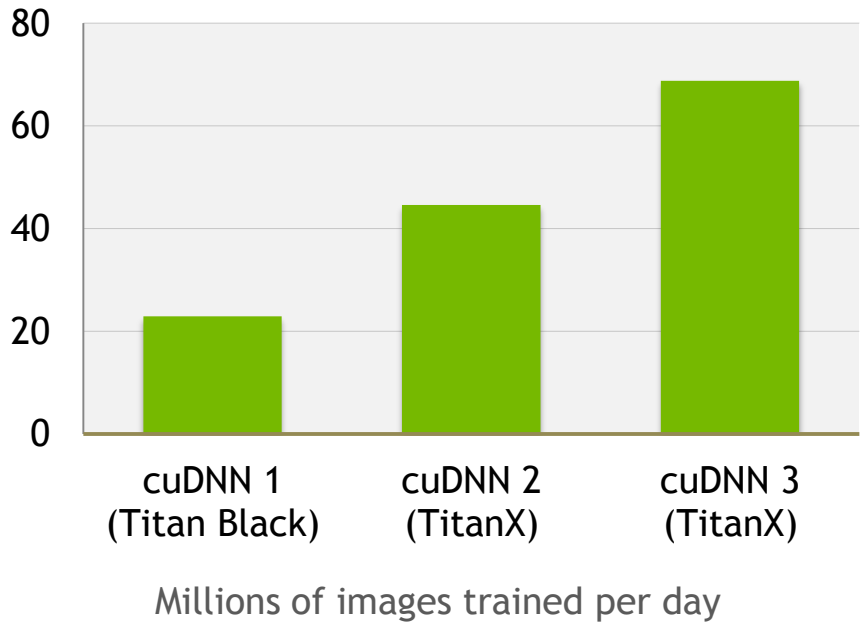
GPU acceleration for Caffe, Theano, Torch and other deep learning frameworks

Support for widely-used layer types, including pooling, ReLU, sigmoid, softmax and TANH

Performance optimized for the latest NVIDIA GPU architectures

Linux, Windows, OSX and Linux for Tegra (ARM)

Performance continues to improve

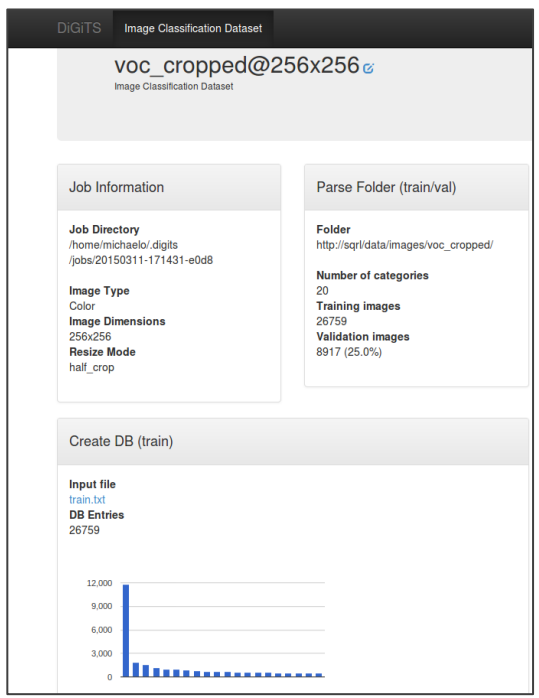


http://developer.nvidia.com/cuDNN

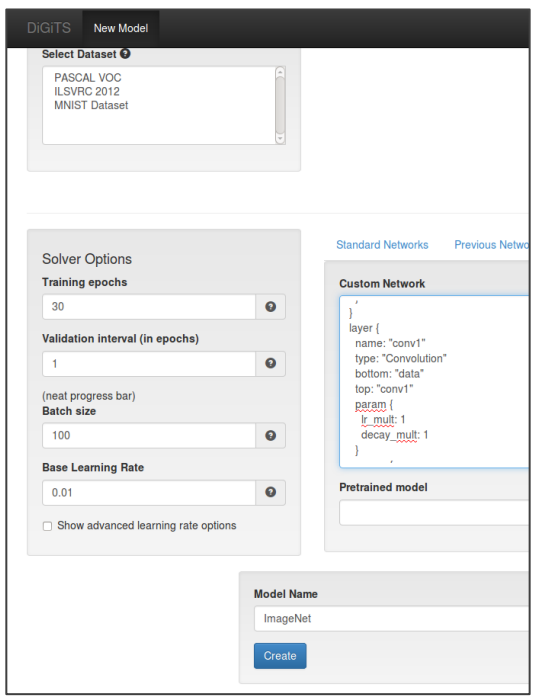
NVIDIA DIGITS

Interactive Deep Learning GPU Training System

Process Data

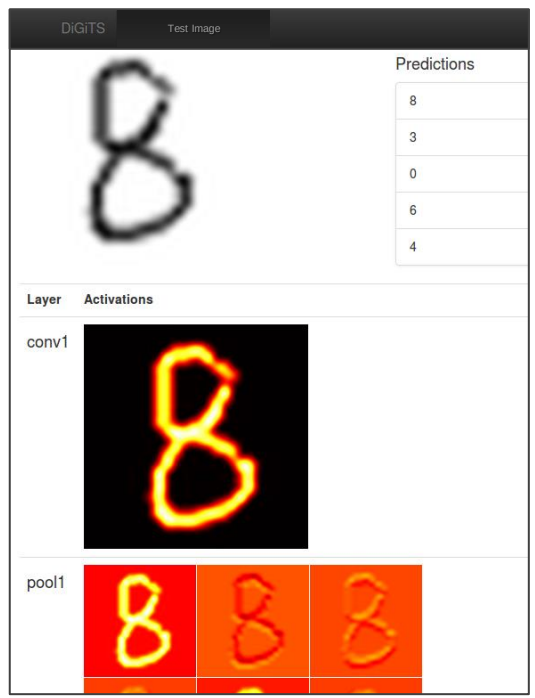


Configure DNN



Monitor Progress

Visualize Layers



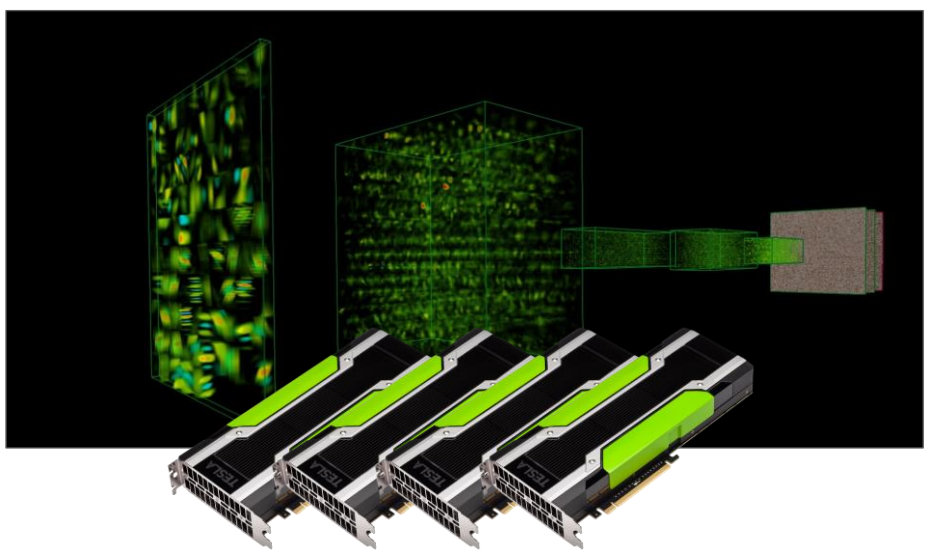
http://developer.nvidia.com/digits

Automatic Multi-GPU Training

DIGITS 2 interactive deep learning training system

Automatic multi-GPU scaling

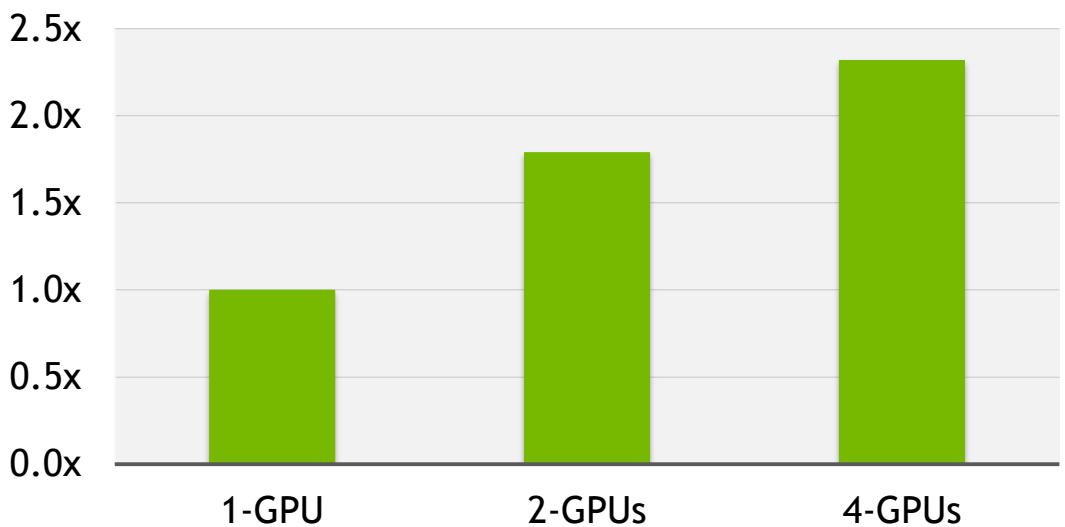
Up to 4 GPUs



developer.nvidia.com/digits

DIGITS 2 trains models up to 2x faster

with Multi-GPU Scaling



DIGITS 2 performance vs. previous version on an NVIDIA DIGITS DevBox system

Learn More: Introduction to Deep Learning

Free 10-week Online Course

Get Started with Deep Learning

- DIGITS, Caffe, Theano, Torch

5 units – available worldwide

- Live classes (recordings available)
- Hands-on labs (no GPU required)
- Office hours with NVIDIA experts

Wednesdays, 9-10am Pacific Time

developer.nvidia.com/deep-learning-courses

GPU TECHNOLOGY CONFERENCE

THANK YOU

JOIN THE CONVERSATION

#GTC15