Developing Perception Algorithms for Self-Driving Cars

ADAM GRZYWCZEWSKI | NVIDIA

Original content by Clement Farabet
ABOUT ME
Adam Grzywaczewski - adamg@nvidia.com

- Deep Learning Solution Architect @ NVIDIA - Supporting delivery of AI / Deep Learning solutions
- 10 years experience delivering Machine Learning of all scale (from embedded, mobile to Big Data)
- My past experience:
  - Capgemini: https://goo.gl/MzgGbq
  - Jaguar Land Rover Research: https://goo.gl/ar7LuU
NVIDIA AGX
EMBEDDED AI HPC

High-speed SerDes — 109 Gbps + 320 Gbps I/O
Up to 320 TOPS Tensor Ops
Up to 25 TFLOPS FP32
Up to 16 GIGA Rays
Starting from 15W
End to End System for AV

COLLECT DATA

- 1 PB+ per car per year

TRAIN MODELS

- 10+ DNNs
- 1M+ images per DNN

SIMULATE

- 10B+ miles to ensure safe driving

RESIMULATE

- Validate and verify for self-driving cars by 2020

MAPPING

- 5+ DNNs
- Create and update HD maps

Sources: NVIDIA, RAND Corporation
NVIDIA DRIVE: SOFTWARE-DEFINED CAR

Powerful and Efficient AI, CV, AR, HPC | Rich Software Development Platform
Functional Safety | Open Platform | 370+ partners developing on DRIVE

DRIVE IX

DRIVE AR

DRIVE AV

DRIVE OS

DRIVE AGX XAVIER
DRIVE AGX PEGASUS
NVIDIA DRIVE PLATFORM ADOPTION ACROSS TRANSPORTATION
NVIDIA JETSON
SOFTWARE-DEFINED AUTONOMOUS MACHINES
Powerful and efficient AI, CV, HPC | Rich Software Development Platform
Open Platform | 200K Developers
BILLIONS OF AUTONOMOUS MACHINES

Taking advantage of the progress in self driving cars
DEEP LEARNING
NEURAL NETWORKS ARE NOT NEW
And are surprisingly simple as an algorithm
NEURAL NETWORKS ARE NOT NEW
They just historically never worked well

Algorithm performance in small data regime

Andrew Ng, “Nuts and Bolts of Applying Deep Learning”, https://www.youtube.com/watch?v=F1ka6a13S9I
NEURAL NETWORKS ARE NOT NEW
They just historically never worked well

Algorithm performance in small data regime

Andrew Ng, “Nuts and Bolts of Applying Deep Learning”, https://www.youtube.com/watch?v=F1ka6a13S9I
NEURAL NETWORKS ARE NOT NEW
They just historically never worked well

Algorithm performance in small data regime

Andrew Ng, “Nuts and Bolts of Applying Deep Learning”, https://www.youtube.com/watch?v=F1ka6a13S9I
NEURAL NETWORKS ARE NOT NEW
Historically we never had large datasets or computers

The MNIST (1999) database contains 60,000 training images and 10,000 testing images.

Algorithm performance in small data regime

Andrew Ng, “Nuts and Bolts of Applying Deep Learning”, https://www.youtube.com/watch?v=F1ka6a13S9I
Historically we never had large datasets or compute.

- Transistors (thousands): 1.1X per year
- GPU-Computing perf: 1.5X per year
- Single-threaded perf: 1.5X per year

By 2025
CONTEXT
1.759 petaFLOPs in November 2009
CONTEXT

2 petaFLOPs - today
NEURAL NETWORKS ARE NOT NEW
But that changed and transformed the way we do machine learning

Andrew Ng, “Nuts and Bolts of Applying Deep Learning”, https://www.youtube.com/watch?v=F1ka6a13S9I
Neural networks are not new
Data and model size the key to accuracy
NEURAL NETWORK COMPLEXITY IS EXPLODING
To Tackle Increasingly Complex Challenges

- **2015 - Microsoft ResNet**
  Superhuman Image Recognition
  - 7 ExaFLOPS
  - 60 Million Parameters

- **2016 - Baidu Deep Speech 2**
  Superhuman Voice Recognition
  - 20 ExaFLOPS
  - 300 Million Parameters

- **2017 - Google Neural Machine Translation**
  Near Human Language Translation
  - 100 ExaFLOPS
  - 8700 Million Parameters
  - 7 ExaFLOPS
  - 60 Million Parameters
  - 20 ExaFLOPS
  - 300 Million Parameters
  - 100 ExaFLOPS
  - 8700 Million Parameters
100 EXAFLOPS = 2 YEARS ON A DUAL CPU SERVER
NEURAL NETWORKS ARE NOT NEW

Exceeding human level performance

Algorithm performance in large data regime

Accuracy

Dataset Size

Small NN  ML1  ML2  ML3  Big NN  Bigger NN

Andrew Ng, “Nuts and Bolts of Applying Deep Learning”, https://www.youtube.com/watch?v=F1ka6a13S9I
EXPLODING DATASETS

Logarithmic relationship between the dataset size and accuracy

EXPLODING DATASETS

Logarithmic relationship between the dataset size and accuracy

- Translation
- Language Models
- Character Language Models
- Image Classification
- Attention Speech Models
EXPLODING DATASETS

Logarithmic relationship between the dataset size and accuracy

Making complex problems easy
Making unsolvable problems expensive
BUILDING AI FOR SDC IS HARD

Every neural net in our DRIVE Software stack needs to handle 1000s of conditions and geolocations.
WHAT TESTING SCALE ARE WE TALKING ABOUT?

We’re on our way to 100s PB of **real test data** = **millions of real miles** + 1,000s DRIVE Constellation nodes for **offline testing alone** & **billions of simulated miles**

- Target robustness per model (miles)
- Test dataset size required (miles)
- NVIDIA’s ongoing data collection (miles)

---

* DRIVE PEGASUS Nodes
<table>
<thead>
<tr>
<th>SDC SCALE TODAY AT NVIDIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-camera+Radar+Lidar RIG mounted on 30 cars</td>
</tr>
<tr>
<td>1PB collected/week</td>
</tr>
<tr>
<td>15PB active training+test dataset</td>
</tr>
</tbody>
</table>
TRAINING
“For any size of the data it’s a good idea to always make the data look small by using a huge model.”

Geoffrey Hinton
EXPLODING MODEL COMPLEXITY

Good news - model size scales sublinearly

Model Size Scales Sublinearly

EVIDENCE FROM IMAGE PROCESSING

Good news - model size scales sublinearly

IMPLICATIONS

Experimental Nature of Deep Learning - Unacceptable training time

- Experiment
- Idea
- Code
## IMPLICATIONS

### Automotive example

Majority of useful problems are too complex for a single GPU training

<table>
<thead>
<tr>
<th></th>
<th>VERY CONSERVATIVE</th>
<th>CONSERVATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fleet size (data capture per hour)</td>
<td>100 cars / 1TB/hour</td>
<td>125 cars / 1.5TB/hour</td>
</tr>
<tr>
<td>Duration of data collection</td>
<td>260 days * 8 hours</td>
<td>325 days * 10 hours</td>
</tr>
<tr>
<td>Data Compression factor</td>
<td>0.0005</td>
<td>0.0008</td>
</tr>
<tr>
<td>Total training set</td>
<td>104 TB</td>
<td>487.5 TB</td>
</tr>
<tr>
<td>InceptionV3 training time (with 1 Pascal GPU)</td>
<td>9.1 years</td>
<td>42.6 years</td>
</tr>
<tr>
<td>AlexNet training time (with 1 Pascal GPU)</td>
<td>2018 1.1 years</td>
<td>2018 5.4 years</td>
</tr>
</tbody>
</table>
CONCLUSIONS

What does your team do in the mean time

THE #1 PROGRAMMER EXCUSE
FOR LEGITIMATELY SLACKING OFF:
"MY CODE'S COMPILING."

HEY! GET BACK
TO WORK!

COMPILING!

OH. CARRY ON.
CONCLUSIONS

What does your team do in the meantime?

THE #1 PROGRAMMER EXCUSE FOR LEGITIMATELY SLACKING OFF:

“My DNN is training”
CONCLUSIONS

Need to scale the training process for a single job

<table>
<thead>
<tr>
<th></th>
<th>VERY CONSERVATIVE</th>
<th>CONSERVATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total training set</td>
<td>104 TB</td>
<td>487.5 TB</td>
</tr>
<tr>
<td>InceptionV3 (one DGX-1V)</td>
<td>166 days (5+ months)</td>
<td>778 days (2+ years)</td>
</tr>
<tr>
<td>AlexNet (one DGX-1V)</td>
<td>21 days (3 weeks)</td>
<td>98 days (3 months)</td>
</tr>
<tr>
<td>InceptionV3 (10 DGX-1V’s)</td>
<td>16 days (2+ weeks)</td>
<td>77 days (11 weeks)</td>
</tr>
<tr>
<td>AlexNet (10 DGX-1V’s)</td>
<td>2.1 days</td>
<td>9.8 days</td>
</tr>
</tbody>
</table>

Training From Months or Years To Weeks or Days
BALANCED HARDWARE

DGX-1 as a reference point for solution design
DGX POD
Reference architectures

- Hiding the complexity of hardware
- Hiding complexity of devops
- Hiding the complexity of software toolkit management
- Partner reference architectures with every major storage provider
ITERATION TIME

Short iteration time is fundamental for success

ResNet 50 Training Time in minutes

- Microsoft (Dec 2015): 1740 minutes
- Preferred Networks (Feb 17): 264 minutes
- Facebook (Jun 17): 60 minutes
- IBM (Aug 17): 48 minutes
- SURFsara (Sep 17): 40 minutes
- UCB (Oct 17): 24 minutes
- Preferred Networks (Nov 17): 15 minutes
- Sony (Nov 18): 3.7 minutes
VALIDATION
SAFE AV REQUIRES A COMPREHENSIVE VALIDATION APPROACH

Large Scale | Millions of Miles
Diverse Vehicle and World Conditions
Data Driven | Scenario based
Repeatable and Reproducible
End-to-End System Level Test
THE AV VALIDATION GAP

COMPONENT LEVEL SIL
Low Fidelity | Scalable

ON ROAD TESTING
High Fidelity | Doesn’t Scale

No Coverage for Extreme & Dangerous Scenarios
NVIDIA DRIVE VALIDATION METHODOLOGY

Three Pronged Approach

FIDELITY

MILES

On Road Testing
Large Scale System Level HIL
Component Level SIL
HARDWARE IN THE LOOP SIMULATION

Bit Accurate & Timing Accurate

PERCEPTION
Camera | Radar | Lidar | IMU

CONTROL
Steering | Throttle | Brake
DRIVE CONSTELLATION ARCHITECTURE

Environment Model
Traffic Model
Vehicle Model
Sensor Model
Scenario Model

DRIVE Sim API

DRIVE Sim

DRIVE Constellation OS

DRIVE Constellation Simulator

Perception
Mapping
Planning

DRIVE AV

DRIVE CORE | DRIVE NETWORKS

DRIVE OS

DRIVE IX

DRIVE Constellation Vehicle
ENVIRONMENT MODEL
Highly Detailed Environments
ENVIRONMENT MODEL
Change Time of Day
ENVIRONMENT MODEL

Add Traffic Scenarios
ENVIRONMENT MODEL
Change Weather
DRIVE CONSTELLATION FOR 3RD PARTY AV VALIDATION

Open | Accessible | Available at Scale

Open Platform | Wide Ecosystem Support
Cloud Based Solution | Scalable
Accessible to OEMs and Researchers
Demonstrate Best Practice for AV Validation
# The Drive Sim Ecosystem

<table>
<thead>
<tr>
<th>Mapping</th>
<th>Vehicle Dynamics</th>
<th>Environment Model</th>
<th>System Integrators</th>
</tr>
</thead>
<tbody>
<tr>
<td>3Dmapping</td>
<td>carsIM</td>
<td>51 VR</td>
<td>AVL</td>
</tr>
<tr>
<td>OpenDRIVE</td>
<td></td>
<td>parallelDomain</td>
<td>dSPACE</td>
</tr>
<tr>
<td>TOMTOM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZENRIN</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensor Model</th>
<th>Scenario Model</th>
<th>Traffic Model</th>
<th>3rd Party Certification</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSYS</td>
<td>AAi</td>
<td>cognata</td>
<td>CANTARC</td>
</tr>
<tr>
<td>ECHODYNE</td>
<td></td>
<td>cognata</td>
<td>UHSLDI</td>
</tr>
<tr>
<td>QVI Innoviz</td>
<td></td>
<td>PTV VISUM</td>
<td>TUV</td>
</tr>
<tr>
<td>Sony</td>
<td></td>
<td>SUMO</td>
<td></td>
</tr>
<tr>
<td>SIEMENS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velodyne</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ifpro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vires</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
BEST PRACTICES FOR AV VALIDATION
Planning the Path to Safety

Partnerships with leading safety organizations
Public | Private | Worldwide
Creating best practices and standards
THANK YOU

twitter.com/nvidiaAI