Recent Trends in Computer Vision and Deep Learning Systems

Yangqing Jia
Lead Researcher and Manager of AI Platform, Facebook
Computer Vision
So it begins.

AlexNet

11x11 conv, 96, /4, pool/2
5x5 conv, 256, pool/2
3x3 conv, 384
3x3 conv, 384
3x3 conv, 256, pool/2
fc, 4096
fc, 4096
fc, 1000
VGGNet

Punch it.
GoogLeNet
We must go deeper.
ResNet
And we took the word seriously
ResNet
And we took the word seriously
ResNeXT

We totally see it coming
Pushing the Performance

- ScSVM: 28.2
- AlexNet: 16.4
- VGGNet: 7.3
- GoogLeNet: 6.7
- ResNet: 3.57
- ResNeXT: 3.03
Why is it challenging?

Gradients, as one example

- exploding
- ideal
- vanishing

depth
Deep Learning Systems
"SAP"

- Scalability
Scalability

Run fast, run far

“How do I train on multiple GPUs and machines?"

- Probably the most question we got from Caffe users
Scalability
Run fast, run far

1.2 million =
(# of images in ImageNet1K)
(# of new images @FB every 5 mins in 2013)
(# of AI jobs per month @FB)
Scalability
Run fast, run far

L1 | L2 | L3 | L3b | L2b | L1b | U3 | U2 | U1
Scalability
Run fast, run far
Scalability
Run fast, run far
Scalability
Run fast, run far
The Return of MPI

"I’m your father", said Allreduce.

Allreduce
Tree based - $O(M\log N)$
Ring based - $O(M)$

etc.
And so we scale
"SAP"

Arithmetics
Quantized Computation

Forget about float, the world is bigger

- float: 8 bits, 23 bits
- fp16: 5 bits, 10 bits
- fixed16: 16 bits
- fixed8: 8 bits
Why do we care?

Battery life is life.

- float add: 0.9
- fp16 add: 0.4
- fixed16 add: 0.05
- fixed8 add: 0.03
- float mul: 4.0
- fp16 mul: 1.0
- fixed8 mul: 0.2
How does it perform?

Why does it matter for cars?

250 watts
10 $\rightarrow$ 20 TFlops

10 watts
0.7 $\rightarrow$ 1.5 TFlops
"SAP"

- Portability
Portable System
One software to rule them all, and...

AI Math and Algorithms

Deployment Platforms
Portable System
Cloud, Mobile, IoT, Cars, Drones, Coffee makers

Model

auto predictor = caffe2::Predictor(model_file)

public class Predictor implements Caffe2ModelInterface;
The Land of Deep Learning System
Not as complex as a car, but still.
Thank you!

Recent Trends in Computer Vision and Deep Learning Systems
Yangqing Jia