




2016 杭州·云栖大会
THE COMPUTING CONFERENCE

云栖社区
yq.aliyun.com

DeepStream: GPU加速海量视频数据智能处理

赖俊杰, 资深技术经理
NVIDIA

2016
The Computing Conference

主办单位:  杭州

 Alibaba Group
阿里巴巴集团

战略合作伙伴: 

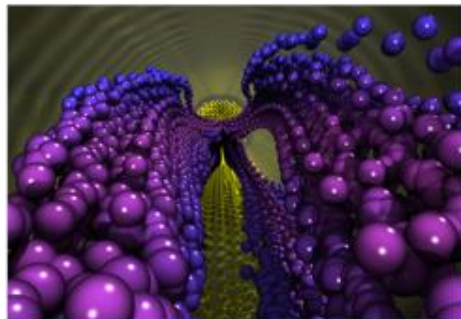


扫码观看大会视频



TESLA ACCELERATES DISCOVERY AND INSIGHT

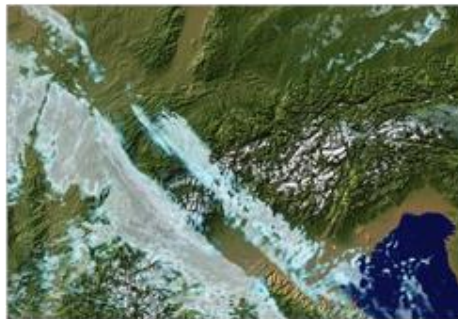
SIMULATION



DATA SCIENCE



VISUALIZATION

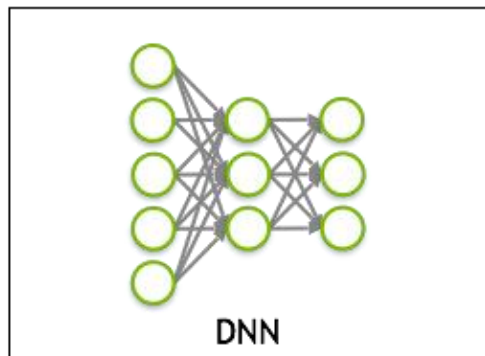


TESLA ACCELERATED COMPUTING



扫码观看大会视频

THE BIG BANG IN MACHINE LEARNING

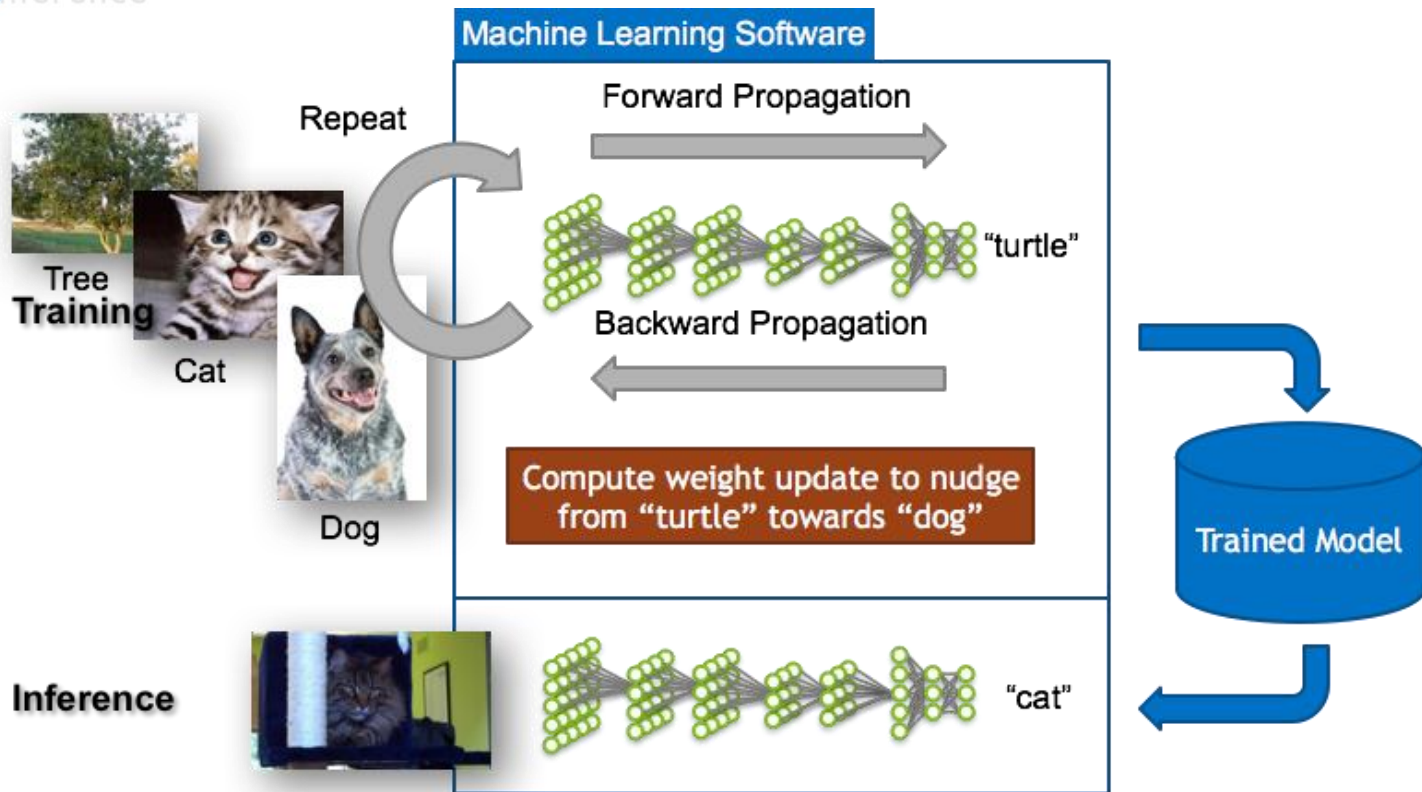


“Google’s AI engine also reflects how the world of computer hardware is changing. (It) depends on machines equipped with GPUs... And it depends on these chips more than the larger tech universe realizes.”

WIRED



扫码观看大会视频



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 ”

WIRED

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

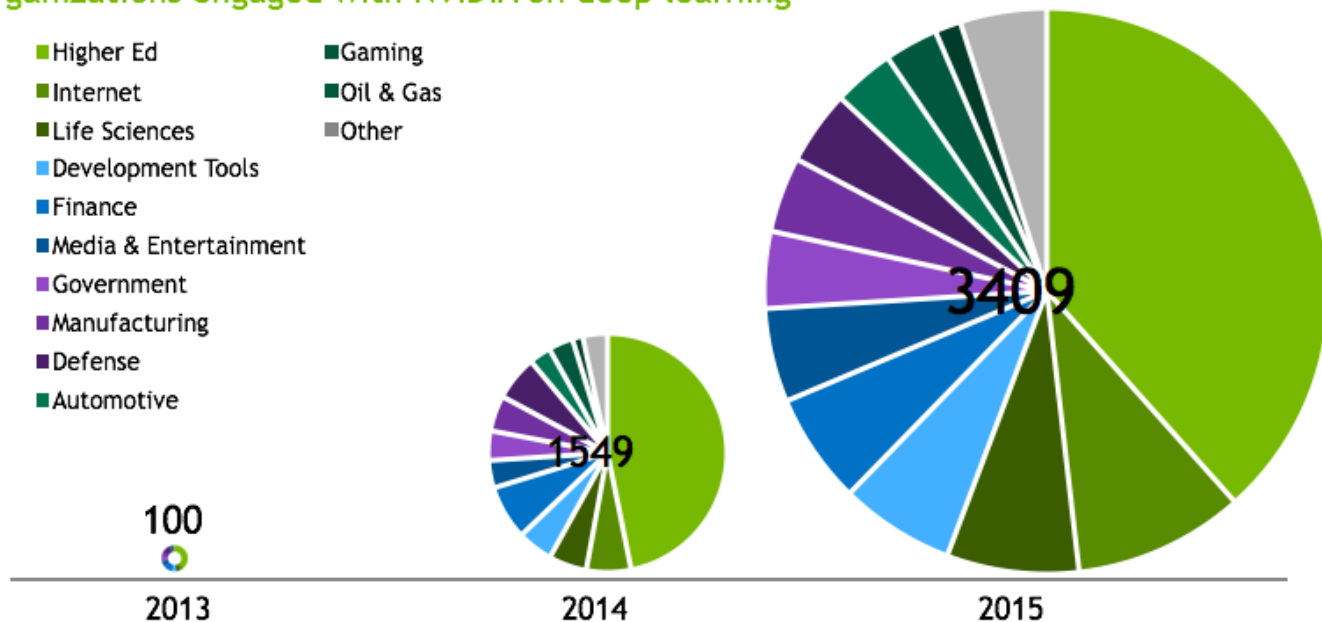
4 kWatts
\$33,000



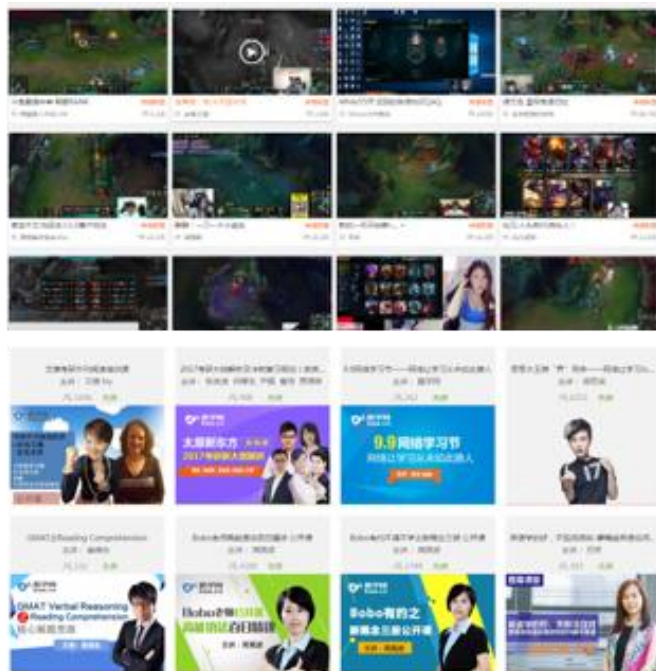
扫码观看大会视频

EVERY INDUSTRY WANTS INTELLIGENCE

Organizations engaged with NVIDIA on deep learning



扫码观看大会视频



NVIDIA DeepStream SDK

Delivering Video Analytics at Scale

DeepStream SDK simplifies deployment of high performance video analytics applications powered by deep learning

Delivers responsive and interactive AI-powered video services such as real-time video categorization

Simple high-level C++ API for GPU-accelerated transcoding, image resizing and scaling and deep learning inference powered by TensorRT™



Video Decoding & GIE Inference

Overview of the Sample

The classification result
of this frame

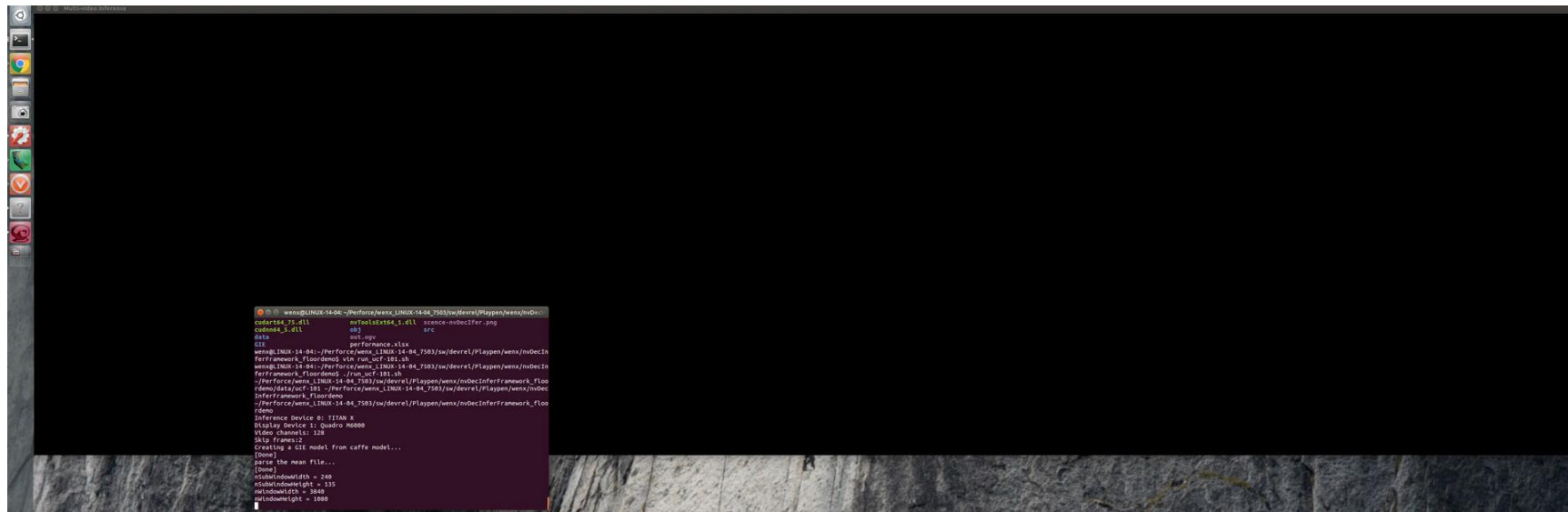
OpenGL frame
rendering



Log message

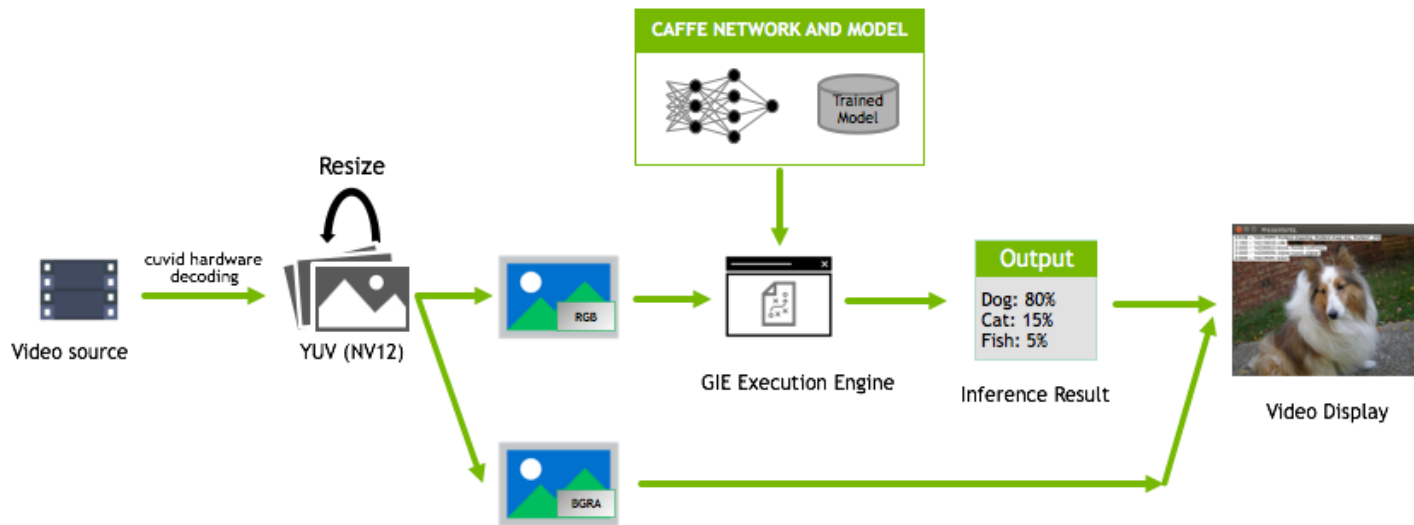


扫码观看大会视频



Video Decoding & GIE Inference

Workflow of the Sample



2016 The
Computing
Conference
THANKS

