

# COMPUTAÇÃO DE ALTO DESEMPENHO NA PONTA: CONSTRUINDO APLICAÇÕES DE IA COM A NVIDIA JETSON NANO

JOÃO MARCELO TEIXEIRA

JMXNT@CIN.UFPE.BR

ESSE/SBESC 2019



UNIVERSIDADE  
FEDERAL  
DE PERNAMBUCO



**DES**

DEPARTAMENTO DE ELETRÔNICA E SISTEMAS

# AGENDA

- Motivação
- A plataforma Jetson
- Jetpack SDK
- DeepStream SDK
- Isaac SDK
- Hello World AI

# MOTIVAÇÃO

Outubro de 2019

## NVIDIA with Microsoft Announces Technology Collaboration for Era of Intelligent Edge

By enabling closer integration between Microsoft Azure and NVIDIA's EGX platform, the companies are working together to advance edge-to-cloud AI computing capabilities, benefitting businesses worldwide.

# MOTIVAÇÃO



<https://www.aiconference.com.br/>

# MOTIVAÇÃO

(vídeo I Am AI)

<https://www.youtube.com/watch?v=GWL1HNHDSq4>

# MOTIVAÇÃO

## SEGUNDA-FEIRA, 16 DE SETEMBRO

- 10:05 - **Ray-tracing e NVIDIA RTX: futuro do rendering tempo real** - Esteban Clua - UFF
- 10:40 - **Classificação de risco na gravidez usando Deep Learning e NLP** - José Adenaldo - Hospital Albert Einstein
- 11:15 - **Seu comportamento é a sua identidade: como a IA está tornando nossa sociedade mais segura** - Rafael Libardi - No Leak
- 12:25 - **Inteligência Artificial na identificação de veículos para a engenharia de tráfego** - Raniere Machado - DBA Tecnologia
- 14:30 - **Tensor Pose: Tracking People in Real-Time** - Luiz Velho - IMPA
- 16:15 - **Aprimoramentos de dados de treinamento para a segmentação robusta de pólipos em imagens de colonoscopia** - Cesar Franco - Tecgraf

# MOTIVAÇÃO

TERÇA-FEIRA, 17 DE SETEMBRO

- 9:50 - **A IA que salva vidas: Como a computação cognitiva está revolucionando o cuidado ao paciente** - Cristian Rocha - Laura
- 11:35 - **Edge AI: cases e suas aplicações para segurança do trabalho** - Luiz Felipe Ribeiro - Displace
- 14:00 - **Sistemas de Inspeção de Linhas de Energia com I.A** - Diogo Caetano -Kasco P&D

# JETSON POWERS AUTONOMOUS MACHINES

## WAREHOUSE



## DELIVERY



## AGRICULTURE



## RETAIL



## INDUSTRIAL



# THE JETSON FAMILY

From AI at the Edge to Autonomous Machines



**JETSON NANO**

5–10W  
0.5 TFLOPS (FP16)  
45mm x 70mm  
\$129 / \$99 (Devkit)



**JETSON TX1 → JETSON TX2 4GB**

7–15W  
1–1.3 TFLOPS (FP16)  
50mm x 87mm  
\$299



**JETSON TX2 8GB | Industrial**

7–15W  
1.3 TFLOPS (FP16)  
50mm x 87mm  
\$399–\$749



**JETSON AGX XAVIER**

10–30W  
11 TFLOPS (FP16) | 32 TOPS (INT8)  
100mm x 87mm  
\$1099

AI at the Edge

Fully Autonomous Machines

Multiple Devices — Same Software

# A PLATAFORMA JETSON

- Jetson Xavier NX Module
  - Anunciada em Novembro de 2019
  - Disponível em Março de 2020



<b>AI Performance</b>	21 TOPS
<b>GPU</b>	384-core NVIDIA Volta™ GPU with 48 Tensor Cores
<b>CPU</b>	6-core NVIDIA Carmel ARM®v8.2 64-bit CPU 6MB L2 + 4MB L3
<b>Memory</b>	8 GB 128-bit LPDDR4x 51.2GB/s
<b>Power</b>	10W/15W

# JETSON NANO DEVKIT SPECS



## PROCESSOR

CPU	64-bit Quad-core ARM A57 @ 1.43GHz
GPU	128-core NVIDIA Maxwell @ 921MHz
Memory	4GB 64-bit LPDDR4 @ 1600MHz   25.6GB/s
Video Encoder	4Kp30   (4x) 1080p30   (2x) 1080p60
Video Decoder	4Kp60   (2x) 4Kp30   (8x) 1080p30   (4x) 1080p60

## INTERFACES

USB	(4x) USB 3.0 A (Host)   USB 2.0 Micro B (Device)
Camera	MIPI CSI-2 x2 (15-position Flex Connector)
Display	HDMI   DisplayPort
Networking	Gigabit Ethernet (RJ45, PoE)
Wireless	M.2 Key-E with PCIe x1
Storage	MicroSD card (16GB UHS-1 recommended minimum)
40-Pin Header	UART   SPI   I2C   I2S   Audio Clock   GPIOs
Power	5V DC (µUSB, Barrel Jack, PoE) - 5W   10W
Size	80x100mm

Distributors Include:



# JETPACK SDK

- JetPack
  - Imagens de Sistemas Operacionais
  - Bibliotecas e APIs
  - Ferramentas de desenvolvimento
  - Exemplos
  - Documentação

# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

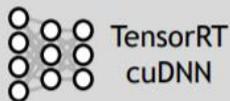
Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks  
OpenCV

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

Nsight Developer Tools

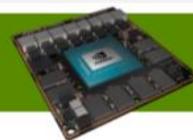
CUDA-X • Linux for Tegra • ROS



Jetson Nano



Jetson TX1/TX2



Jetson AGX Xavier

# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks  
OpenCV

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

Nsight Developer Tools

CUDA-X • Linux for Tegra • ROS



Jetson Nano



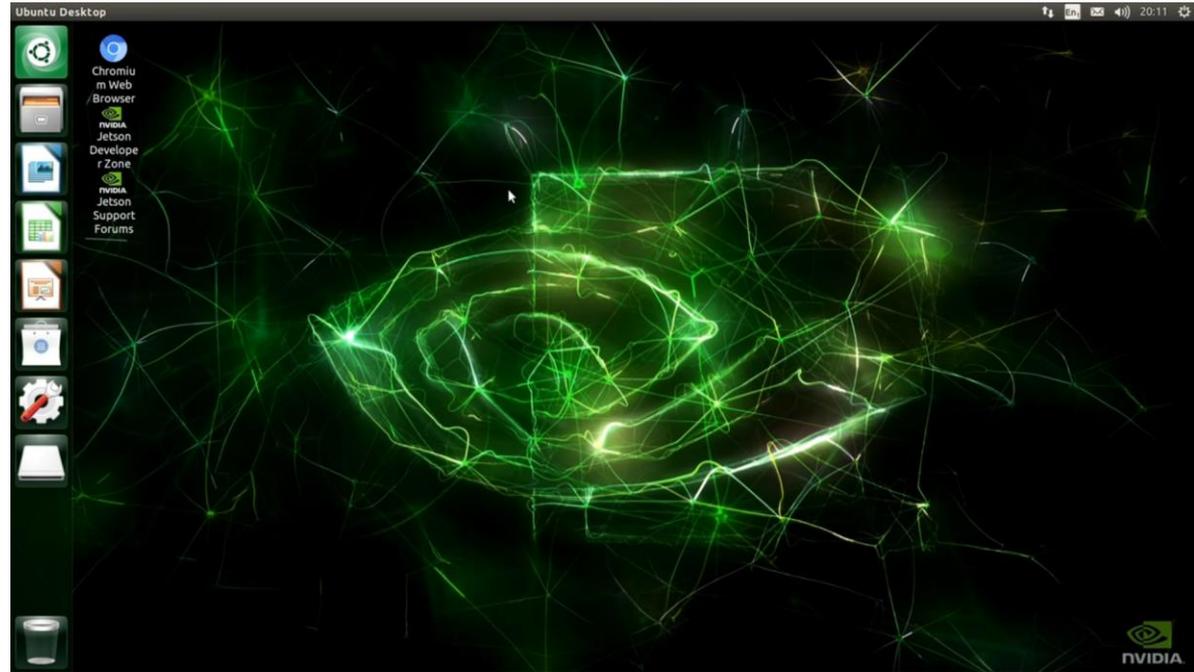
Jetson TX1/TX2



Jetson AGX Xavier

# JETPACK SDK :: SISTEMA OPERACIONAL

- Bootloader
- Kernel do Linux
- Firmwares
- Drivers da NVIDIA
- Sistema de arquivos de exemplo



# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

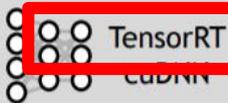
Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT



VisionWorks  
OpenCV



cuBLAS  
cuFFT



Vulkan  
OpenGL



libargus  
Video API



Drivers  
Ecosystem

Deep Learning

Computer Vision

Accel. Computing

Graphics

Multimedia

Sensors

CUDA-X • Linux for Tegra • ROS

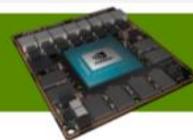
Nsight Developer Tools



Jetson Nano



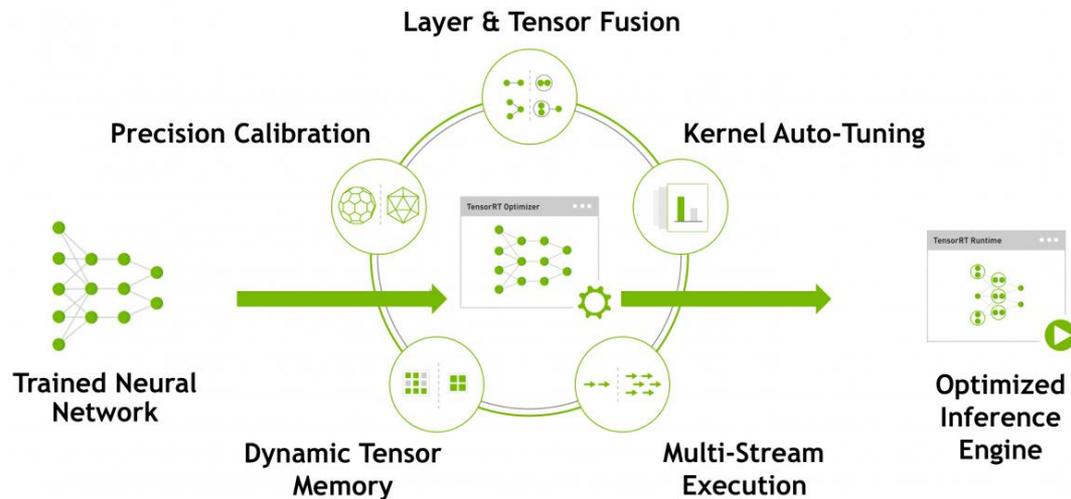
Jetson TX1/TX2



Jetson AGX Xavier

# JETPACK SDK :: TENSORRT

- Runtime de inferência de aprendizagem profunda que suporta redes neurais para classificação, segmentação e detecção de objetos
- Acelera o processo de inferência (~40x mais rápido que CPU)
- Reduz os requisitos de memória para redes neurais convolucionais e deconvolucionais



# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

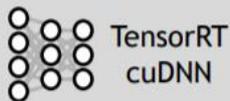
Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks  
OpenCV

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

Nsight Developer Tools

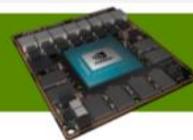
CUDA-K • Linux for Tegra • ROS



Jetson Nano



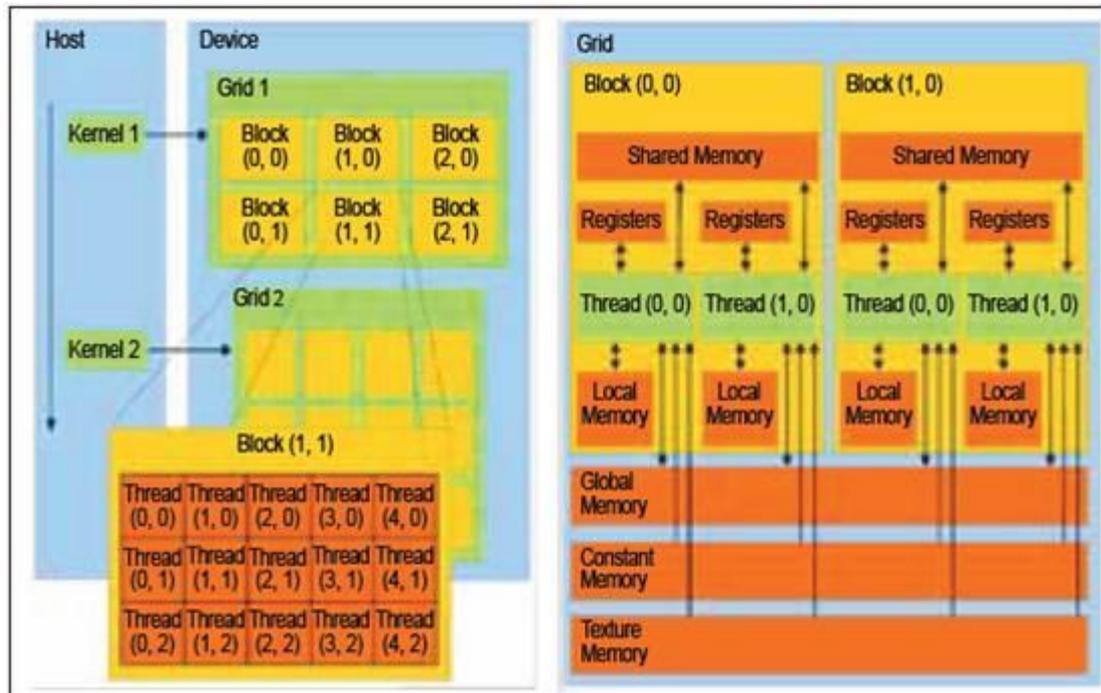
Jetson TX1/TX2



Jetson AGX Xavier

# JETPACK SDK :: CUDA

- Inclui:
  - Compilador (C/C++) para GPUs da NVIDIA
  - Bibliotecas matemáticas
  - Ferramentas para depuração (debugging) e otimização de desempenho



# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

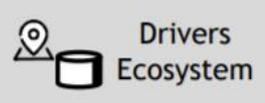
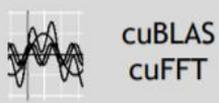
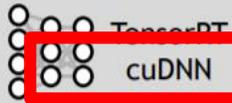
Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



Deep Learning

Computer Vision

Accel. Computing

Graphics

Multimedia

Sensors

CUDA-X • Linux for Tegra • ROS

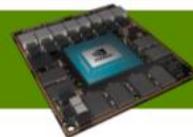
Nsight Developer Tools



Jetson Nano



Jetson TX1/TX2



Jetson AGX Xavier

# JETPACK SDK :: CUDNN

- Biblioteca de redes neurais profundas em CUDA
- Suporta várias primitivas de deep learning:
  - Convoluções
  - Funções de ativação
  - Transformações de tensores
- Utilizada por vários frameworks de deep learning:
  - Caffe, Caffe2, Chainer, Keras, MATLAB, MxNet, TensorFlow e PyTorch

# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks  
OpenCV

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

Nsight Developer Tools

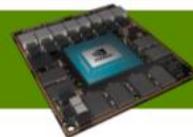
CUDA-X • Linux for Tegra • ROS



Jetson Nano



Jetson TX1/TX2



Jetson AGX Xavier

# JETPACK SDK :: MULTIMEDIA API

- APIs de baixo nível para desenvolvimento de aplicações
- Camera application API:
  - Controle de parâmetros de câmera a cada frame
  - Suporte a múltiplas câmeras
  - Manipulação de dados RAW (sem compactação)
- Sensor driver API:
  - Video encoding/decoding
  - Conversão de formatos
  - Scaling
  - Controle de bit rate, qualidade, etc.

# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

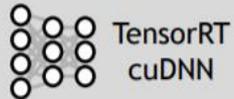
Gesture Recognition

Path Planning

Autonomous Navigation

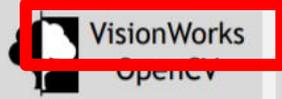
Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

Nsight Developer Tools

CUDA-X • Linux for Tegra • ROS



Jetson Nano



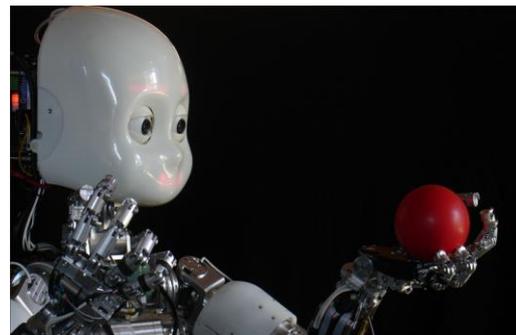
Jetson TX1/TX2



Jetson AGX Xavier

# JETPACK SDK :: VISIONWORKS

- Implementa e estende o padrão Khronos OpenVX
- Usado em soluções para:
  - Robótica e drones
  - Carros autônômicos
  - Análise inteligente de vídeos
  - Realidade Aumentada



## VisionWorks includes the following primitives:

### IMAGE ARITHMETIC

- Absolute Difference
- Accumulate Image
- Accumulate Squared
- Accumulate Weighted
- Add / Subtract / Multiply +
- Channel Combine
- Channel Extract
- Color Convert +
- CopyImage
- Convert Depth
- Magnitude
- MultiplyByScalar
- Not / Or / And / Xor
- Phase
- Table Lookup
- Threshold

### FLOW & DEPTH

- Median Flow
- Optical Flow (LK) +
- Semi-Global Matching
- Stereo Block Matching
- IME Create Motion Field
- IME Refine Motion Field
- IME Partition Motion Field

### GEOMETRIC TRANSFORMS

- Affine Warp +
- Warp Perspective +
- Flip Image
- Remap
- Scale Image +

### FILTERS

- BoxFilter
- Convolution
- Dilation Filter
- Erosion Filter
- Gaussian Filter
- Gaussian Pyramid
- Laplacian3x3
- Median Filter
- Scharr3x3
- Sobel 3x3

### FEATURES

- Canny Edge Detector
- FAST Corners +
- FAST Track +
- Harris Corners +
- Harris Track
- Hough Circles
- Hough Lines

### ANALYSIS

- Histogram
- Histogram Equalization
- Integral Image
- Mean Std Deviation
- Min Max Locations

# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

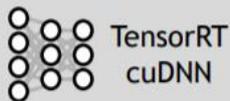
Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks  
OpenCV

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

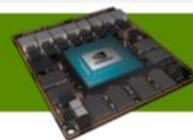
CUDA-X • Linux for Tegra • ROS



Jetson Nano



Jetson TX1/TX2

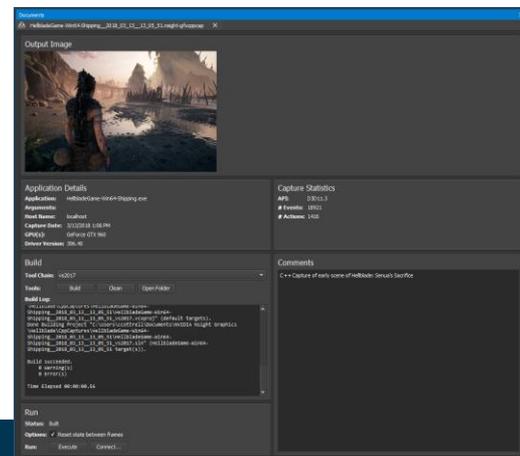


Jetson AGX Xavier

Nsight Developer Tools

# JETPACK SDK :: FERRAMENTAS DE DESENVOLVIMENTO

- CUDA Toolkit
  - Nsight Eclipse Edition
  - Ferramentas de Debugging e Profiling (Nsight Compute, além de ferramentas para cross-compiling)
- NVIDIA Nsight Systems é uma ferramenta de profiling de baixo overhead que permite otimizar desempenho de software rodando na GPU
- NVIDIA Nsight Graphics é uma aplicação standalone que permite realizar debugging e profiling de aplicações gráficas



# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

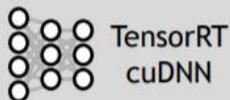
Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks  
OpenCV

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

Nsight Developer Tools

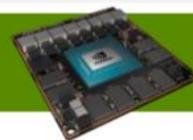
CUDA-X • Linux for Tegra • ROS



Jetson Nano



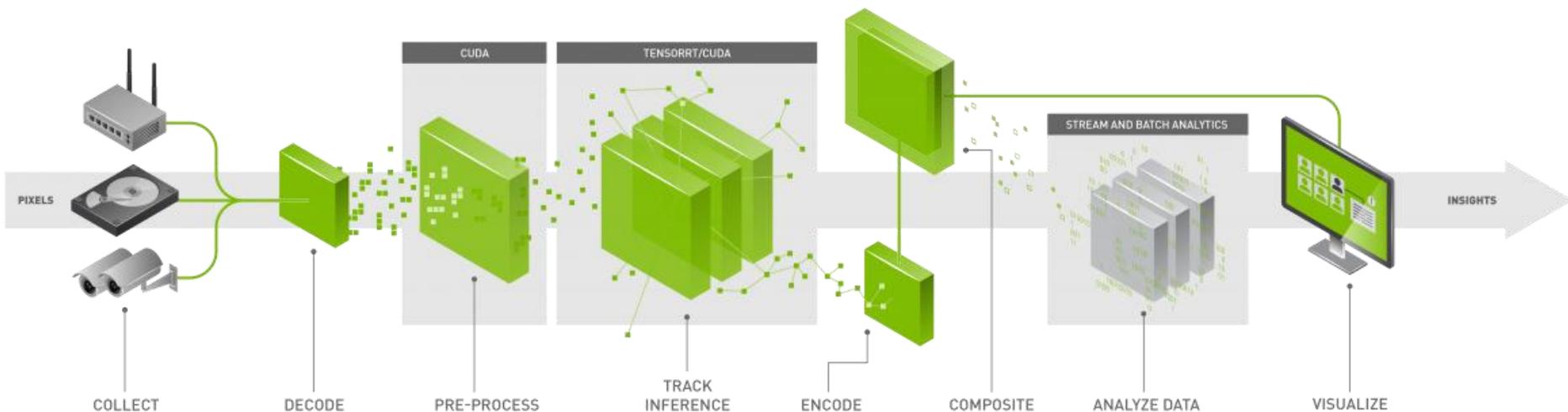
Jetson TX1/TX2



Jetson AGX Xavier

# DEEPCONSTREAM SDK

Permite analisar dados de câmeras, sensores e gateways IoT em tempo real



# DEEPCONVOLUTIONAL SDK

- <https://www.youtube.com/watch?v=Y43W04sMK7I&t=12s>

# DEEPSTREAM SDK

<b>NVIDIA Products</b>	<b>H.264</b>	<b>H.265</b>
Jetson Nano	8	8
Jetson TX1	8	8
Jetson TX2	14	14
Jetson AGX Xavier	32	49
T4	35	68

*[Data measured using deepstream-app from DeepStream SDK 4.0.1](#)*

# JETSON SOFTWARE

DeepStream SDK

Isaac Robotics Engine

Modules

Depth Estimation

Object Detection

Pose Estimation

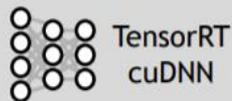
Gesture Recognition

Path Planning

Autonomous Navigation

Ecosystem Modules

JetPack SDK



TensorRT  
cuDNN

Deep Learning



VisionWorks  
OpenCV

Computer Vision



cuBLAS  
cuFFT

Accel. Computing



Vulkan  
OpenGL

Graphics



libargus  
Video API

Multimedia



Drivers  
Ecosystem

Sensors

Nsight Developer Tools

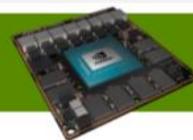
CUDA-X • Linux for Tegra • ROS



Jetson Nano



Jetson TX1/TX2



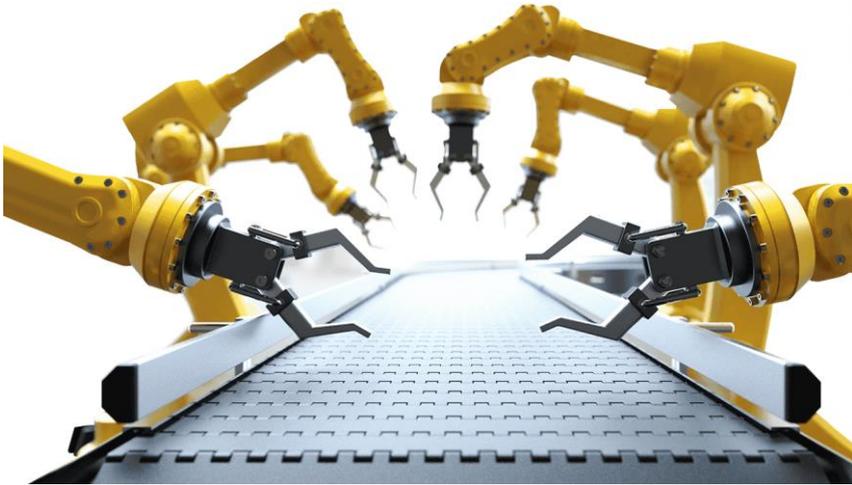
Jetson AGX Xavier

# ISAAC SDK

- Que tipo de problema a robótica ajuda a resolver?

# ISAAC SDK

- **Indústria; Agricultura; Saúde; Inspeção; Serviços**



# ISAAC SDK

- Indústria; Agricultura; **Saúde; Inspeção;** Serviços



# ISAAC SDK

- Indústria; Agricultura; Saúde; Inspeção; **Serviços**



# ISAAC SDK

- Tarefas Inteligentes

<http://www.youtube.com/watch?v=jqpWRmgM1oc>

# ISAAC SDK

- Você sabe **como** desenvolver uma aplicação para um robô?
  - Para um robô específico, pelo menos?
- E como desenvolver uma aplicação para resolver uma tarefa **mais complexa**, como por exemplo:
  - Detectar um objeto em uma imagem e realizar uma ação baseada nisso
  - Detectar movimento em tempo real e evitar colisão ou dar avisos
  - Navegar em um ambiente desconhecido

# ISAAC SDK

- Que problemas há no vídeo a seguir?

# ISAAC SDK

- Que problemas há no vídeo a seguir?

<http://www.youtube.com/watch?v=iaF43Ze1oel>

# ISAAC SDK

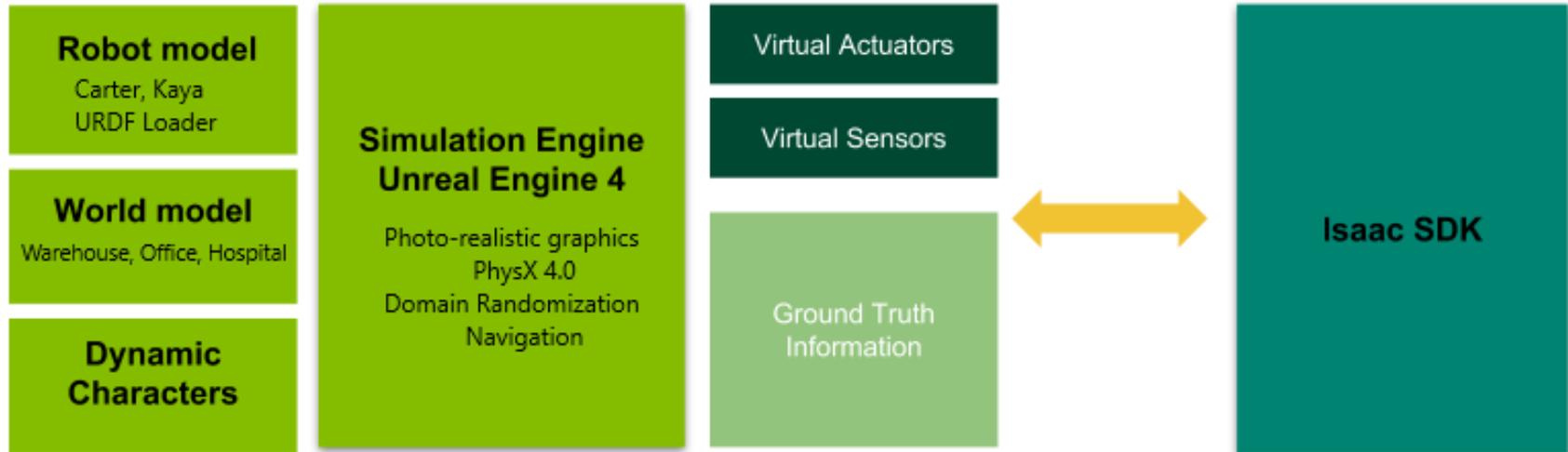
- O que é **importante** em uma ferramenta de desenvolvimento de aplicações robóticas?
- Quais são as **vantagens/desvantagens** de usar uma ferramenta virtual para desenvolver aplicações robóticas?

# ISAAC SDK

Uma ferramenta de desenvolvimento de aplicações robóticas deveria ser/ter:

1. Modular
2. Modificável
3. Sem “reality gap” (alta coerência da simulação física com a realidade)
4. Montagem do ambiente simplificada
5. Curva de aprendizado muito pequena
6. Baixo custo
  - a. Computacional
  - b. Financeiro

# ISAAC SDK



# ISAAC SDK

[https://www.youtube.com/watch?v=oa\\_wkSmWUw](https://www.youtube.com/watch?v=oa_wkSmWUw)

# ISAAC SDK :: JETBOT

<http://www.youtube.com/watch?v=byGZt5ZYup0>

# ISAAC SDK :: KAYA

<http://www.youtube.com/watch?v=5c9hJHmMKAs>

# ISAAC SDK :: CARTER

<http://www.youtube.com/watch?v=SnyFcQ6LR3s>

HELLO AI WORLD



# HELLO AI WORLD

## Getting Started with Deep Learning

### 1. Download and Build the GitHub Repo

```
git clone http://github.com/dusty-nv/jetson-inference
```

### 2. Classifying Images from Command Line

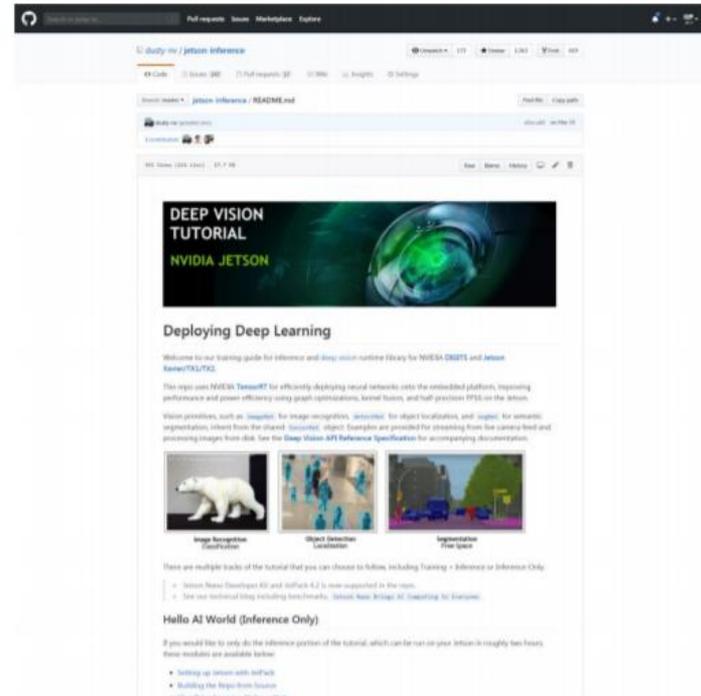
### 3. Coding Your Own Recognition Program

### 4. Realtime Recognition from Live Camera

### 5. Detecting Objects in Images from Disk

### 6. Object Detection from Live Camera

[github.com/dusty-nv/jetson-inference](http://github.com/dusty-nv/jetson-inference)



# HELLO AI WORLD

## Getting Started with Deep Learning

1. Download and Build the GitHub Repo
2. Classifying Images from Command Line  
`./imagenet-console bear_0.jpg output_0.jpg`
3. Coding Your Own Recognition Program
4. Realtime Recognition from Live Camera
5. Detecting Objects in Images from Disk
6. Object Detection from Live Camera

[github.com/dusty-nv/jetson-inference](https://github.com/dusty-nv/jetson-inference)



# HELLO AI WORLD

## Getting Started with Deep Learning

1. Download and Build the GitHub Repo
2. Classifying Images from Command Line
3. Coding Your Own Recognition Program  
./my-recognition test-image.jpg
4. Realtime Recognition from Live Camera
5. Detecting Objects in Images from Disk
6. Object Detection from Live Camera

[github.com/dusty-nv/jetson-inference](https://github.com/dusty-nv/jetson-inference)

```
#include <jetson-inference/imageNet.h>
#include <jetson-utils/loadImage.h>

int main( int argc, char** argv )
{
    // load the image recognition network with TensorRT
    imageNet* net = imageNet::Create(imageNet::GOOGLENET);

    // this variable will store the confidence of the classification (between 0 and 1)
    float confidence = 0.0;

    // classify the image with TensorRT on the GPU (hence we use the CUDA pointer)
    // this will return the index of the object class that the image was recognized as
    const int classIndex = net->Classify(imgCUDA, imgWidth, imgHeight, &confidence);

    // make sure a valid classification result was returned
    if( classIndex >= 0 )
    {
        // retrieve the name/description of the object class index
        const char* classDescription = net->GetClassDesc(classIndex);

        // print out the classification results
        printf("image is recognized as '%s' (class #i) with %f%% confidence\n",
            classDescription, classIndex, confidence * 100.0f);
    }

    // free the network's resources before shutting down
    delete net;
    return 0;
}
```

# HELLO AI WORLD

## Getting Started with Deep Learning

1. Download and Build the GitHub Repo
2. Classifying Images from Command Line
3. Coding Your Own Recognition Program
4. Realtime Recognition from Live Camera
  - `./imagenet-camera googlenet`
5. Detecting Objects in Images from Disk
6. Object Detection from Live Camera

[github.com/dusty-nv/jetson-inference](https://github.com/dusty-nv/jetson-inference)



# HELLO AI WORLD

## Getting Started with Deep Learning

1. Download and Build the GitHub Repo
2. Classifying Images from Command Line
3. Coding Your Own Recognition Program
4. Realtime Recognition from Live Camera
5. Detecting Objects in Images from Disk
  - `./detectnet-console dogs.jpg output.jpg coco-dog`
  - `./detectnet-console peds.jpg output.jpg multiped`
6. Object Detection from Live Camera

[github.com/dusty-nv/jetson-inference](https://github.com/dusty-nv/jetson-inference)



# HELLO AI WORLD

## Getting Started with Deep Learning

1. Download and Build the GitHub Repo
2. Classifying Images from Command Line
3. Coding Your Own Recognition Program
4. Realtime Recognition from Live Camera
5. Detecting Objects in Images from Disk
6. Object Detection from Live Camera

```
./detectnet-camera <model-name>
```

[github.com/dusty-nv/jetson-inference](https://github.com/dusty-nv/jetson-inference)



### Object Detection Models

facenet	(faces)	multiped	(humans)
coco-dog	(dogs)	coco-bottle	(bottles)
coco-chair	(chairs)	coco-airplane	(airplanes)

## REFERÊNCIAS

- <https://ssl.lvl3.on24.com/event/19/77/14/4/rt/1/documents/resourceList1556647149485/jetsonnanowebinar1556647147958.pdf>
- Youtube: NVIDIA Jetson Partner Stories

# COMPUTAÇÃO DE ALTO DESEMPENHO NA PONTA: CONSTRUINDO APLICAÇÕES DE IA COM A NVIDIA JETSON NANO

JOÃO MARCELO TEIXEIRA

JMXNT@CIN.UFPE.BR

ESSE/SBESC 2019



UNIVERSIDADE  
FEDERAL  
DE PERNAMBUCO



**DES**

DEPARTAMENTO DE ELETRÔNICA E SISTEMAS