인공 지능의 역사

조 영 혁

노다시스템

인공지능

앨런 튜링

-미국의 신경외과의 워렌 맥컬록(Warren Mc Cullonch)과 논리학자 월터피츠(Walter Pitts)가 전기 스위치처럼 온, 오프 하는 기초기능의 인공신경을 그물망 형태로 연결하면 사람의 뇌에서 동작하는 아주 간단한 기능을 흉내 낼 수 있다는 것을 이론적으로 증명
- 헵은 생물학적 신경망 내에서 반복적인 시그널이 발생할 때 신경세포들은 그 시그널을 기억하는 일종의 학습효과가 있음을 증명
-코넬대 심리학자 프랭크 로센블래트의 연구에서 퍼셉트론(Perceptron: 뇌 신경을 모사한 **인공 신경 뉴런**) 탄생

인공지능

2 세대 - 퍼지이론

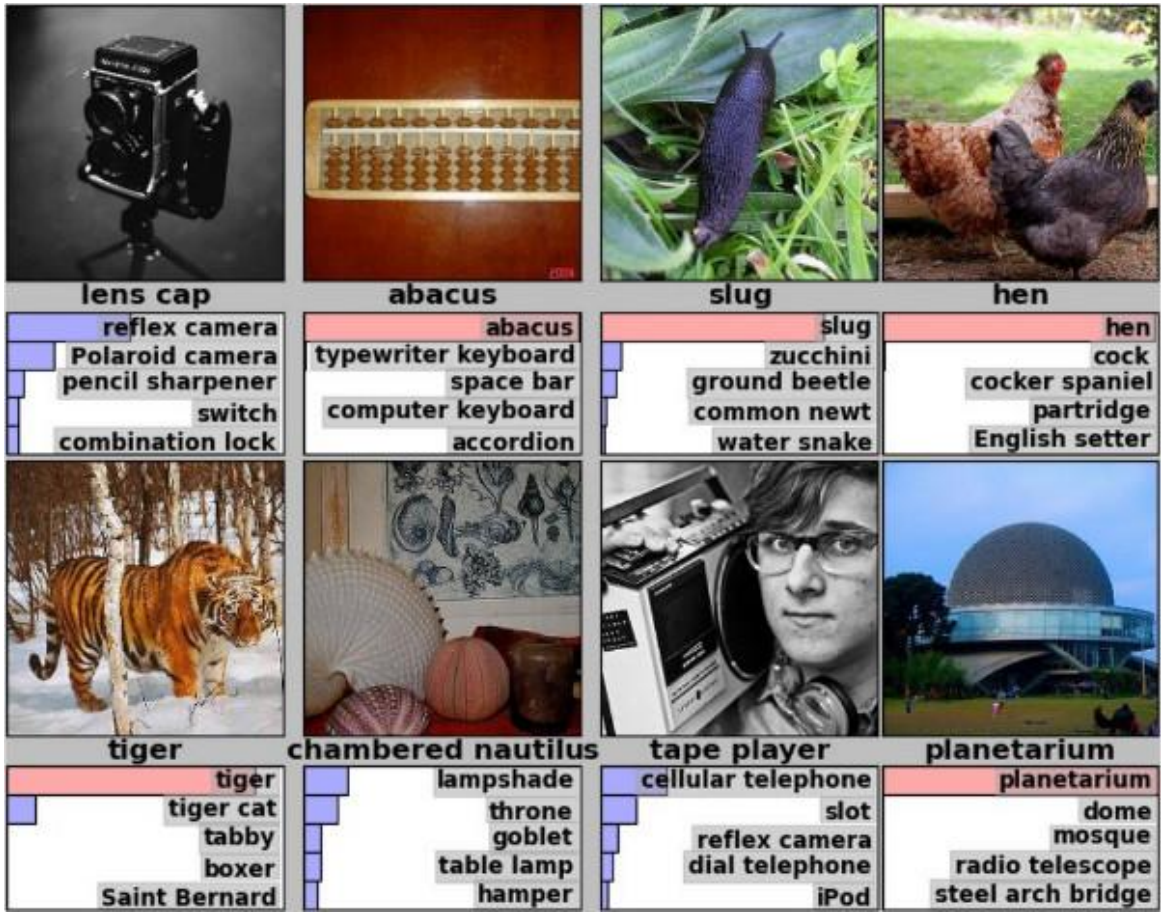
- 1980년대 산업계에 전문가 시스템이 도입되며 본격적으로 확산
- **퍼지(Fuzzy)이론**을 통해 다중 값 논리방법을 이용하는 방법이 주로 활용
- 1975년에 영국의 런던대학 에브라힘 맘다니 교수가 증기기관 제어 적용에 성공

전자 산업 분야	비디오 캠코더, 세탁기, 온수기, 에어컨, TV, 진공 청소기
중공업 분야	보일러-터빈, 아아크 용접기, 화력발전소
제어계측 분야	서보시스템, 하수처리 및 관리, 잠수정제어
의료기기 분야	알레르기 진단 장비, 혈압 측정기
로봇산업 분야	순응 제어기, 모빌 로봇, 정밀 부품 조립, 능동제어
교통제어 분야	엘리베이터 제어, 자기 부양 장치
패턴인식 분야	한글인식, 펜 컴퓨터, 칼라 복사기, CRT 제조
전력산업 분야	부하 변동 주파수 제어, 전력 손실 복원

인공지능

3 세대 - 딥러닝

-2000년대에 이르러 힌튼교수의 Deep Belief Network를 기반으로 심층신경망(딥러닝) 기술이 실용화 가능성
- Deep-CNN(Convolution Neural Network: 합성곱신경망, 이미지 인식/분류 특화모델)은 이미지 인식 성능 평가에서 2011년에는 26%인식 오류율을 보였으나, 2015년4년만에 3.5%로 개선하는 괄목할 성과
-실증화된 인공지능(AI)기술은 CNN, RNN(Recurrent Neural Network : 음성과 문자분야에 강한 신경망)으로 발전



인공지능

- 이세돌과 알파고의 바둑 대결

인공지능

LG 전자의 로봇킹

-구글과 페이스북은 딥러닝을 활용한 얼굴 인식 기술에서 99.96%와 97.25%의 정확도
-아마존은 2014년 인공지능을 활용하여 결제예측배송 특허
-영국 유니버시티 칼리지런던(UCL), 셰필드대, 미국 펜실베니아주립대의 공동 연구의 결과물로 만들어진 인공지능 판사는 79% 정확도로 재판의 결과를 예측
-IBM AI 로스는 파산관리변호사로 공식선임
-미국종양학회의 IBM왓슨의 대장암/직장암 진단 정확도는 이미 90%가 넘는 수준
-테슬라의 AutoPilot은 인간 개입을 배제한 자율주행이 가능한 수준
-멘스(Siemens)의 스마트팩토리는 매일 5,000만건으로 제조공정의 75%를 자동으로 작업을 지시

NVIDIA AI

Jetson Nano

Jetson 모듈은 소형 폼 팩터 형태로 뛰어난 성능 및 에너지 효율성을 제공하여, 결과적으로 최신 AI, 딥 러닝 및 추론을 위한 최첨단 임베디드 시스템

NVIDIA Jetson 모듈 사양 비교

	Jetson Nano	Jetson TX2 Series			Jetson AGX Xavier Series	
		TX2 4GB	TX2	TX2i	Jetson AGX Xavier 8GB	Jetson AGX Xavier
GPU	NVIDIA Maxwell™ architecture with 128 NVIDIA CUDA® cores	NVIDIA Pascal™ architecture with 256 NVIDIA CUDA cores			NVIDIA Volta™ architecture with 384 NVIDIA CUDA cores and 48 Tensor cores	NVIDIA Volta™ architecture with 512 NVIDIA CUDA cores and 64 Tensor cores
	0.5 TFLOPs (FP16)	1.3 TFLOPS (FP16)			5.5 TFLOPS (FP16) 11.1 TOPS (INT8)	11 TFLOPS (FP16) 22 TOPS (INT8)
CPU	Quad-core ARM® Cortex®-A57 MPCore processor	Dual-core Denver 2 64-bit CPU and quad-core ARM A57 complex			6-core ARM v8.2 64-bit CPU, 8MB L2 + 4MB L3	8-core ARM v8.2 64-bit CPU, 8MB L2 + 4MB L3
DLA	-	-			4.1 TFLOPS (FP16) 8.2 TOPS (INT8)	5 TFLOPS (FP16) 10 TOPS (INT8)
Memory	4 GB 64-bit LPDDR4	4 GB 128-bit LPDDR4	8 GB 128-bit LPDDR4		8 GB 256-bit LPDDR4x 1333MHz - 85GB/s	16 GB 256-bit LPDDR4x 2133MHz - 137GB/s
Storage	16 GB eMMC 5.1	16GB eMMC 5.1	32GB eMMC 5.1		32GB eMMC 5.1	
Video Encode	250MP/sec 1x 4K @ 30 (HEVC) 2x 1080p @ 60 (HEVC) 4x 1080p @ 30 (HEVC)	500MP/sec 1x 4K @ 60 (HEVC) 3x 4K @ 30 (HEVC) 4x 1080p @ 60 (HEVC) 8x 1080p @ 30 (HEVC)			2x464MP/sec 2x 4K @ 30 (HEVC) 6x 1080p @ 60 (HEVC) 14x 1080p @ 30 (HEVC)	2x1000MP/sec 4x 4K @ 60 (HEVC) 8x 4K @ 30 (HEVC) 16x 1080p @ 60 (HEVC) 32x 1080p @ 30 (HEVC)
Video Decode	500MP/sec 1x 4K @ 60 (HEVC) 2x 4K @ 30 (HEVC) 4x 1080p @ 60 (HEVC) 8x 1080p @ 30 (HEVC)	1000MP/sec 2x 4K @ 60 (HEVC) 4x 4K @ 30 (HEVC) 7x 1080p @ 60 (HEVC) 14x 1080p @ 30 (HEVC)			2x690MP/sec 2x 4K @ 60 (HEVC) 4x 4K @ 30 (HEVC) 12x 1080p @ 60 (HEVC) 24x 1080p @ 30 (HEVC) 16x 1080p @ 30 (H.264)	2x1500MP/sec 2x 8K @ 30 (HEVC) 6x 4K @ 60 (HEVC) 12x 4K @ 30 (HEVC) 26x 1080p @ 60 (HEVC) 52x 1080p @ 30 (HEVC) 30x 1080p @ 30 (H.264)
Camera	12 lanes (3x4 or 4x2) MIPI CSI-2, DPHY 1.1 (1.5 Gbps)	12 lanes MIPI CSI-2, D-PHY 1.2 (30 Gbps)			16 lanes MIPI CSI-2, 8 lanes SLVS-EC D-PHY (40 Gbps), C-PHY (64 Gbps)	16 lanes MIPI CSI-2, 8 lanes SLVS-EC D-PHY (40 Gbps), C-PHY (109 Gbps)
Power	5W | 10W	7.5W | 15W		10W | 20W	10W | 20W	10W | 15W | 30W
Connectivity	Wi-Fi requires external chip	Wi-Fi requires external chip	Wi-Fi onboard	Wi-Fi requires external chip	Wi-Fi requires external chip	
	10/100/1000 BASE-T Ethernet	10/100/1000 BASE-T Ethernet			10/100/1000 RGMII Gigabit Ethernet	
Mechanical	69.6 mm x 45 mm 260-pin edge connector	87 mm x 50 mm 400-pin connector			100 mm x 87 mm 699-pin connector	

Jetson Nano

NVIDIA Jetson Family Specifications				
	Xavier NX (15W)	Xavier NX (10W)	AGX Xavier	Jetson Nano
CPU	4x/6x Carmel @ 1.4GHz or 2x Carmel @ 1.9GHz	4x/ Carmel @ 1.2GHz or 2x Carmel @ 1.5GHz	8x Carmel @ 2.26GHz	4x Cortex-A57 @ 1.43GHz
GPU	Volta, 384 Cores @ 1100MHz	Volta, 384 Cores @ 800MHz	Volta, 512 Cores @ 1377MHz	Maxwell, 128 Cores @ 920MHz
Accelerators	2x NVDLA		2x NVDLA	N/A
Memory	8GB LPDDR4X, 128-bit bus (51.2 GB/sec)		16GB LPDDR4X, 256-bit bus (137 GB/sec)	4GB LPDDR4, 64-bit bus (25.6 GB/sec)
Storage	8GB eMMC		32GB eMMC	16GB eMMC
AI Perf.	21 TOPS	14 TOPS	32 TOPS	N/A
Dimensions	45mm x 70mm		100mm x 87mm	45mm x 70mm
TDP	15W	10W	30W	10W
Price	$399		$999	$129

TOPS : Tera operations per second

TFLOPS : Tera Floating point operations per second

Jetson Nano

JETSON NANO	JETSON TX2 SERIES (TX2, TX2 4GB, TX2i*)	JETSON XAVIER NX	JETSON AGX XAVIER SERIES (AGX Xavier 8GB, AGX Xavier)
			
0.5 TFLOPS (FP16) 5-10 W 45 mm x 70 mm $129	1.3 TFLOPS (FP16) 7.5-15 W* 50 mm x 87 mm Starting at $249	6 TFLOPS (FP16) | 21 TOPS (INT8) 10-15 W 45 mm x 70 mm $399	20-32 TOPS (INT8) 5.5-11 TFLOPS (FP16) 10-30 W 100 mm x 87 mm Starting at $599

Jetson Nano

JETSON XAVIER NX

Xavier Performance, Less Power, Smaller Size, Lower Price

Smallest Jetson form factor

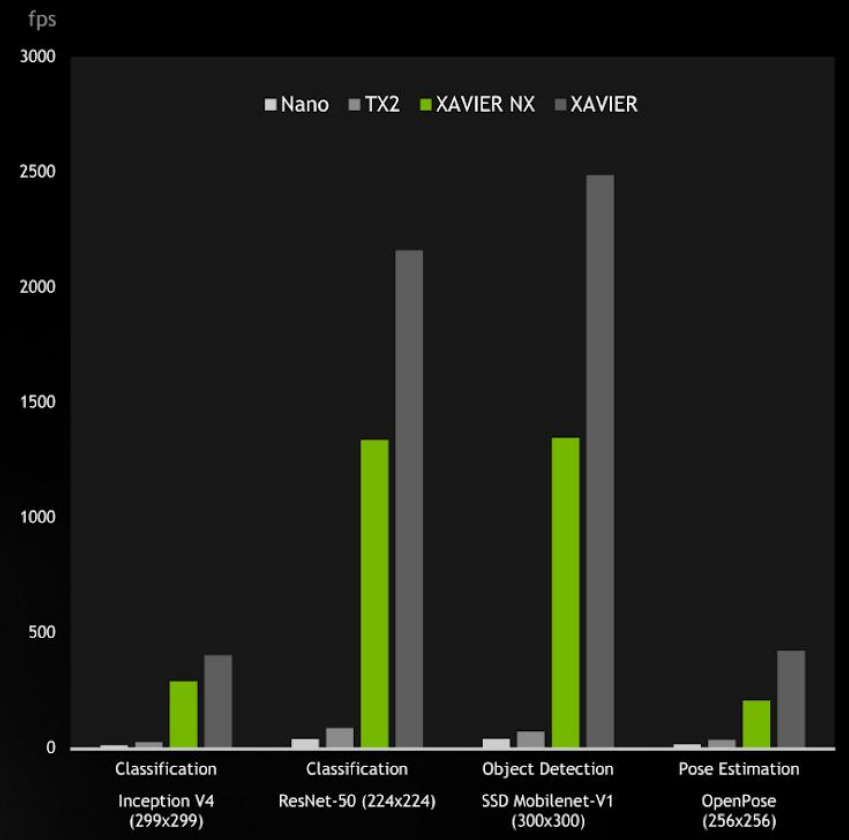
Pin compatible with Jetson Nano

Up to 15x higher perf than Jetson TX2, smaller size, same power

Start development today on Jetson AGX Xavier Developer Kit

One software architecture

$399



Intel

What's Inside the OpenVINO™ toolkit

Intel® Deep Learning Deployment Toolkit

Model Optimizer
Convert & Optimize

1R

Inference Engine
Optimized Inference

Code Samples & 10 Pre-trained Models

IR = Intermediate Representation file

Traditional Computer Vision Tools & Libraries

Optimized Libraries

OpenCV* | OpenVX* | Photography Vision

Code Samples

For Intel® CPU & CPU with integrated graphics

Increase Media/Video/Graphics Performance

Intel® Media SDK
Open Source version

OpenCL™
Drivers & Runtimes

For CPU with integrated graphics

Optimize Intel® FPGA

FPGA RunTime Environmen
(from Intel® FPGA SDK for OpenCL™)

Bitstreams

For CPU with integrated graphics

OS Support CentOS* 7.4 (64 bit) Ubuntu* 16.04.3 LTS (64 bit) Microsoft Windows* 10 (64 bit) Yocto Project* version Poky Jethro v2.0.3 (64 bit)

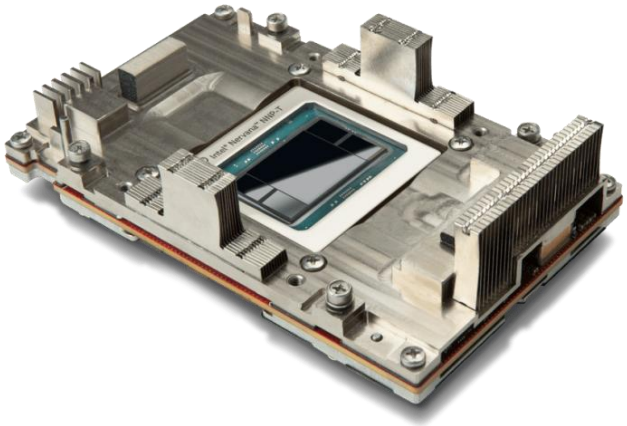
Intel

Movidius
- 인텔 NUC 사용
- 개발 kit인 '인텔 오픈비노(Intel OpenVINO)'
- 1W에서 1 TOPS

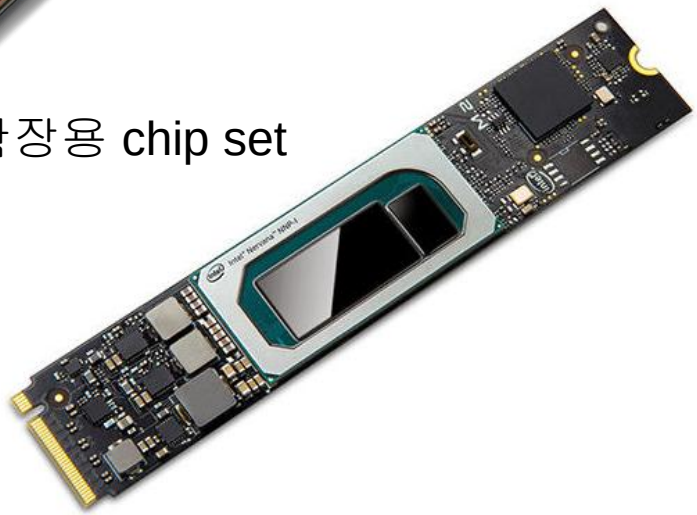
Intel

< 기계 학습용 >

1. 너바나 ‘NNP-T1000’ 신경망 프로세서 : GPU를 겨냥



2. 너바나 NNP-I1000 : 확장용 chip set



Jetson Nano AI 플랫폼

Jetson Nano

OPEN FRAMEWORK SUPPORT

MACHINE LEARNING

Caffe

Caffe2

Keras

mxnet

PYTORCH

TensorFlow

ROBOTICS / IOT

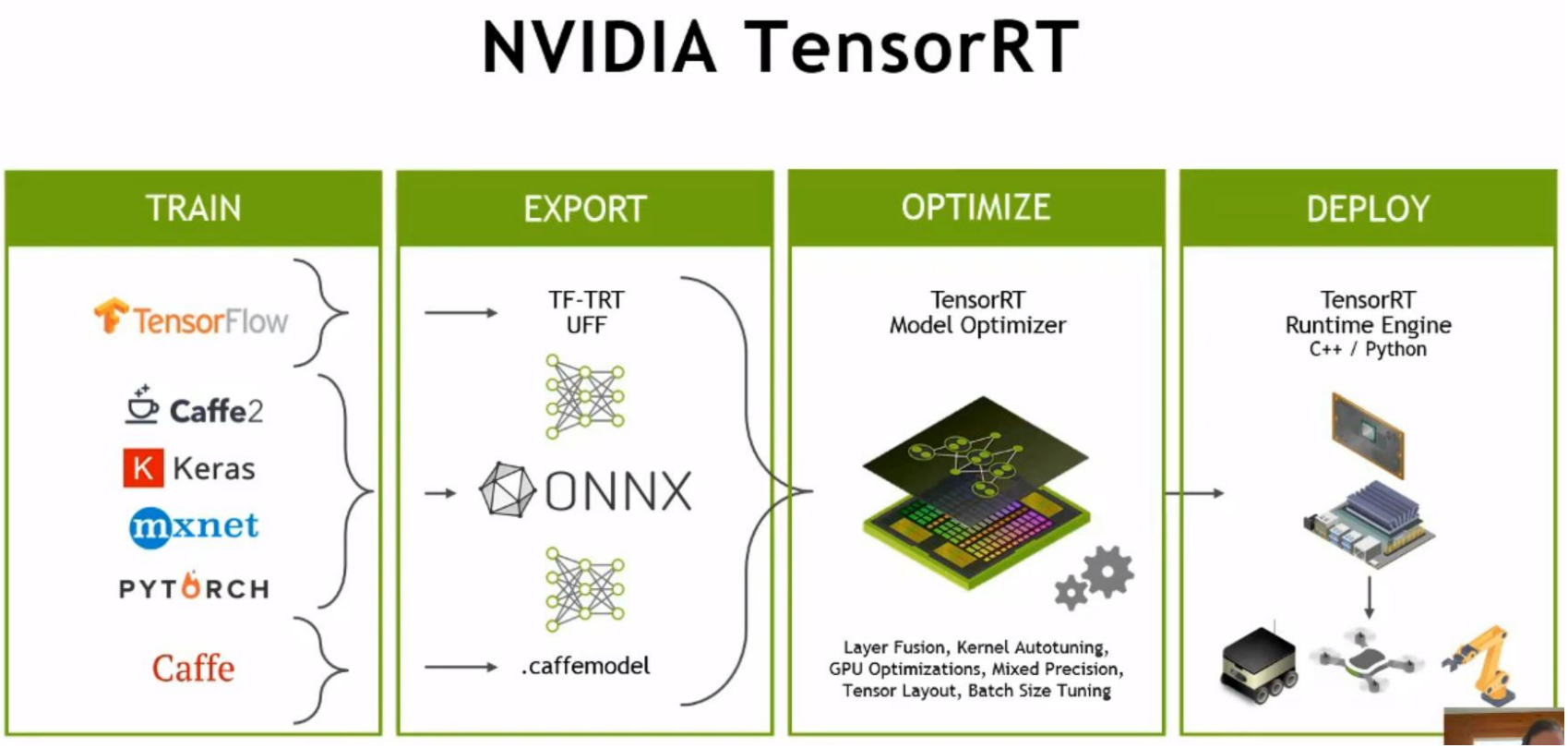
AWS Greengrass

docker

MPI

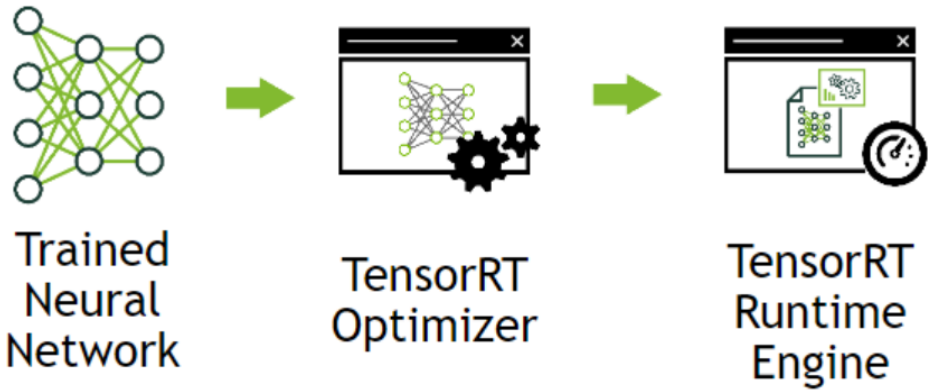
ROS

Jetson Nano

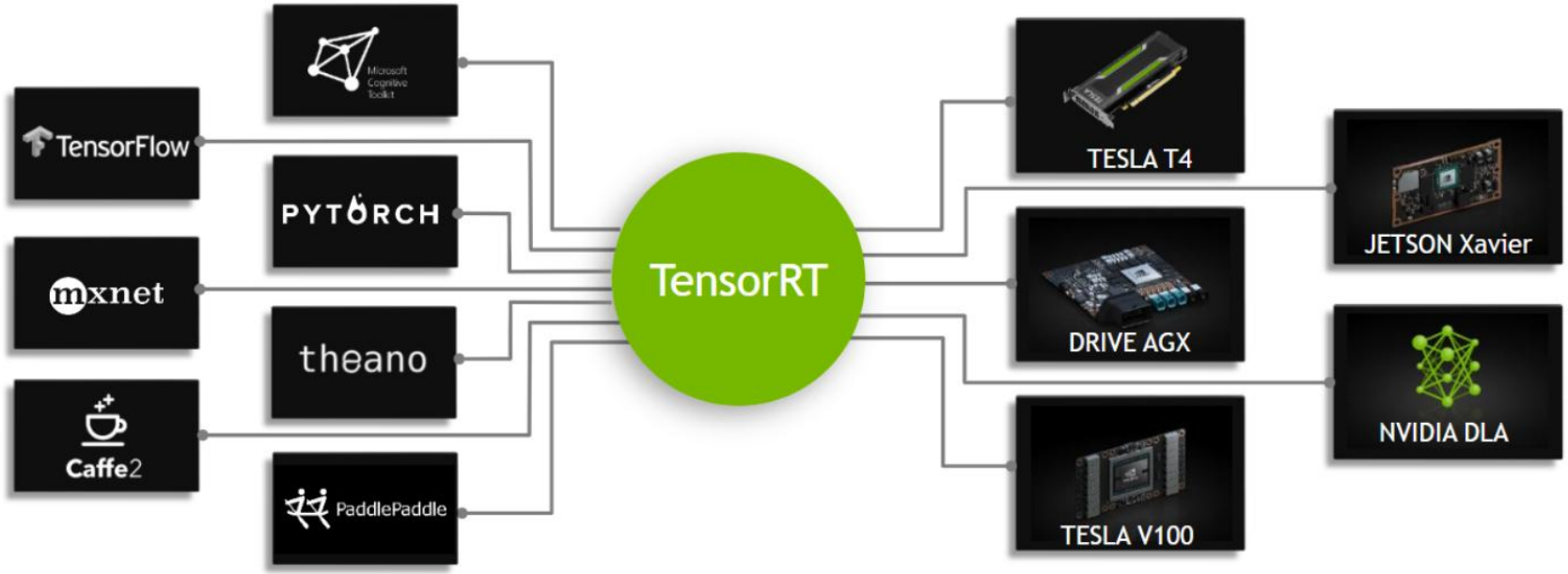


ONNX(open neural network exchange : 페이스북과 마이크로소프트가 만든 개방형 포맷으로 다른 프레임워크에서 모델을 교환할 수 있도록 만든것

Jetson Nano

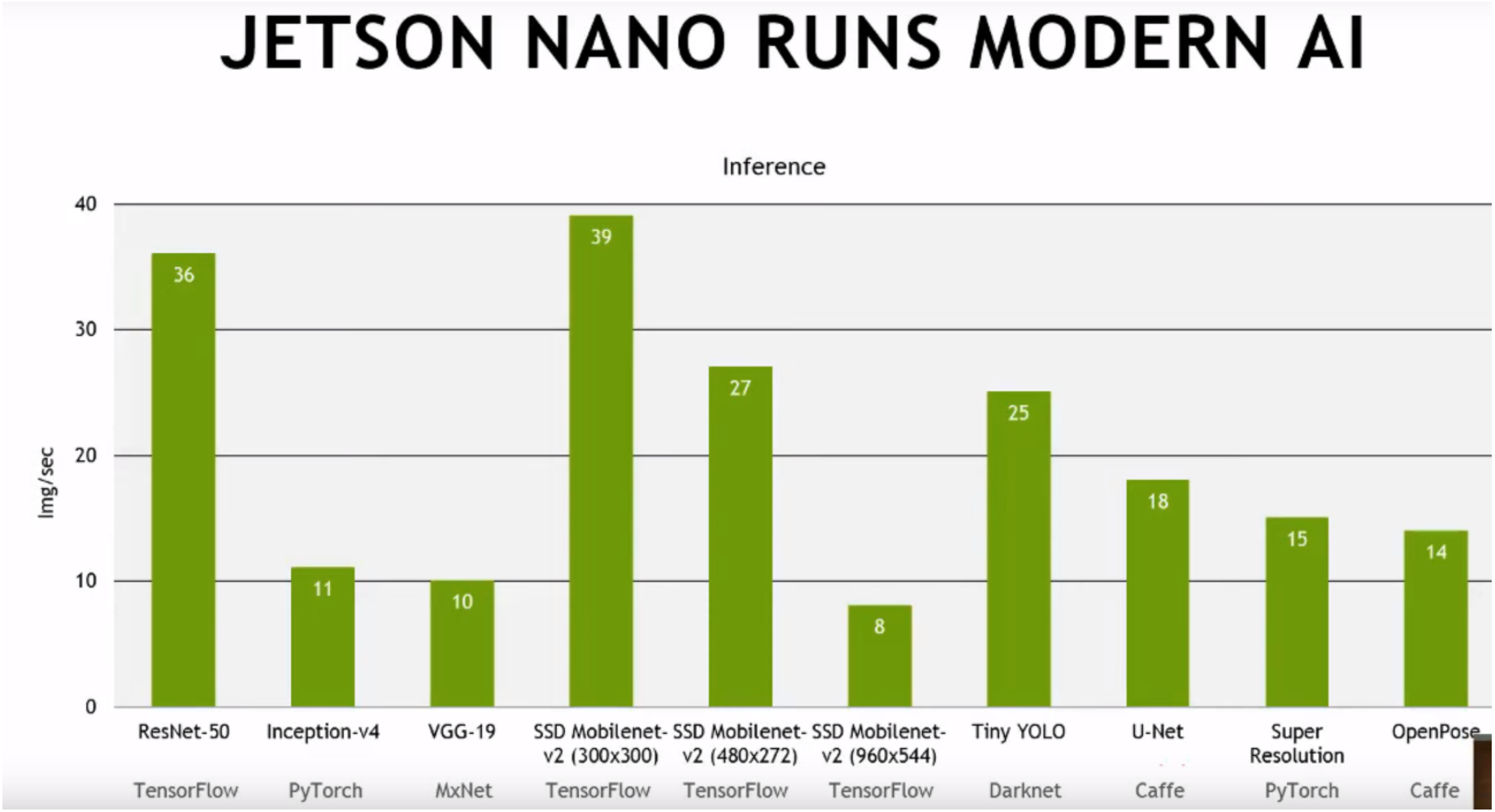


TensorRT™
- 딥러닝 추론 엔진
- TensorFlow, Caffe, PyTorch, MXNet, etc에 동작되도록 디자인됨



cuDNN : CUDA Deep Neural Network Library

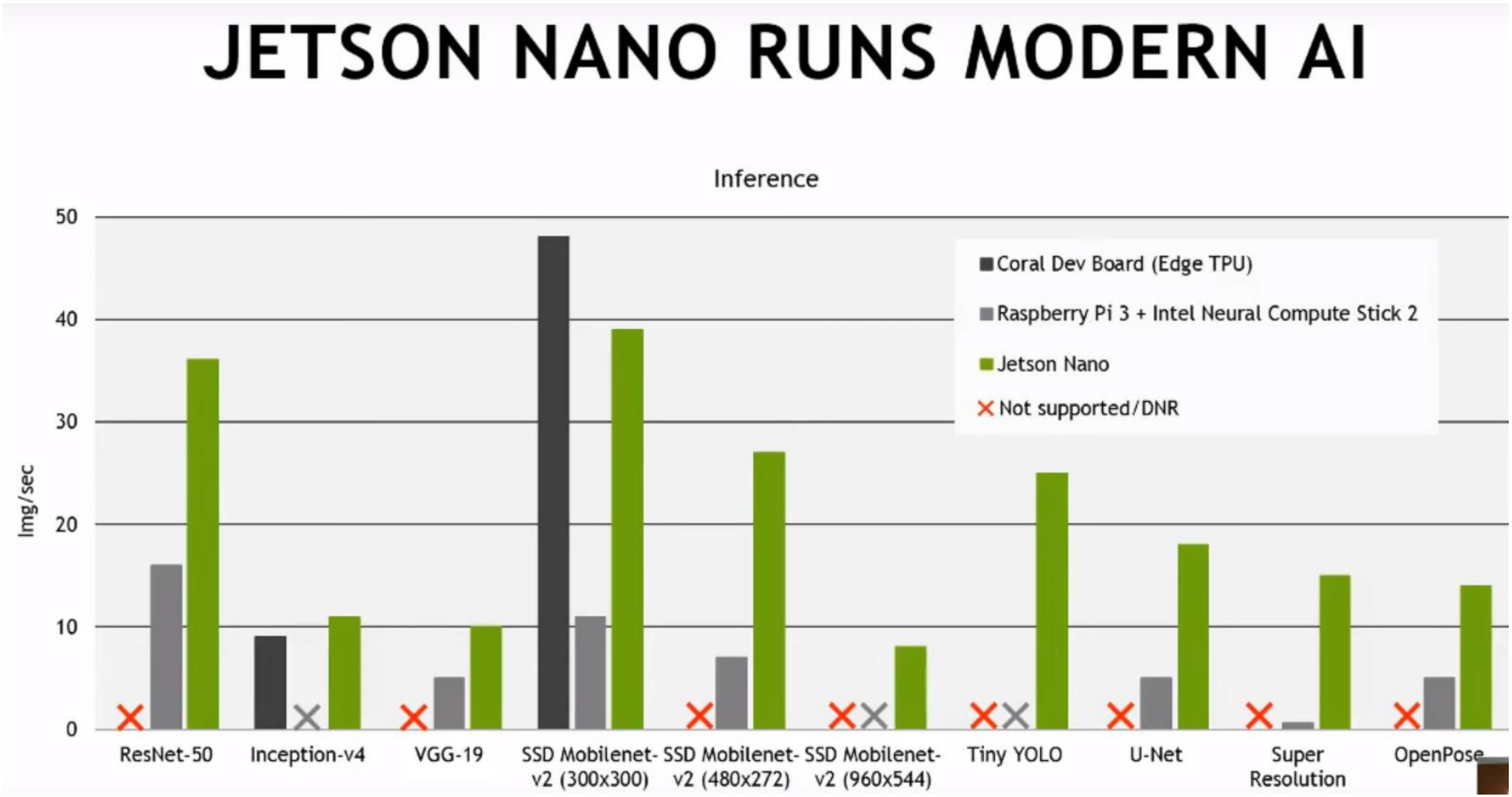
Jetson Nano



OpenPose : Caffe와 OpenCV를 기반으로 구성된 손, 얼굴 포함 몸의 움직임을 추적해주는 API

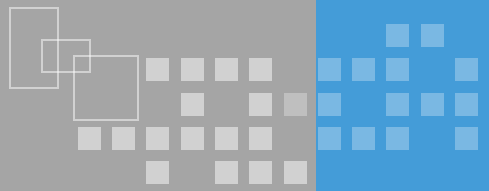
SSD(single shot Multibox Detector) Mobilenet V2

Jetson Nano


JETSON NANO RUNS MODERN AI
Inference
Img/sec
50
40
30
20
10
0
Coral Dev Board (Edge TPU)
Raspberry Pi 3 + Intel Neural Compute Stick 2
Jetson Nano
Not supported/DNR
ResNet-50
Inception-v4
VGG-19
SSD Mobilenet-v2 (300x300)
SSD Mobilenet-v2 (480x272)
SSD Mobilenet-v2 (960x544)
Tiny YOLO
U-Net
Super Resolution
OpenPose

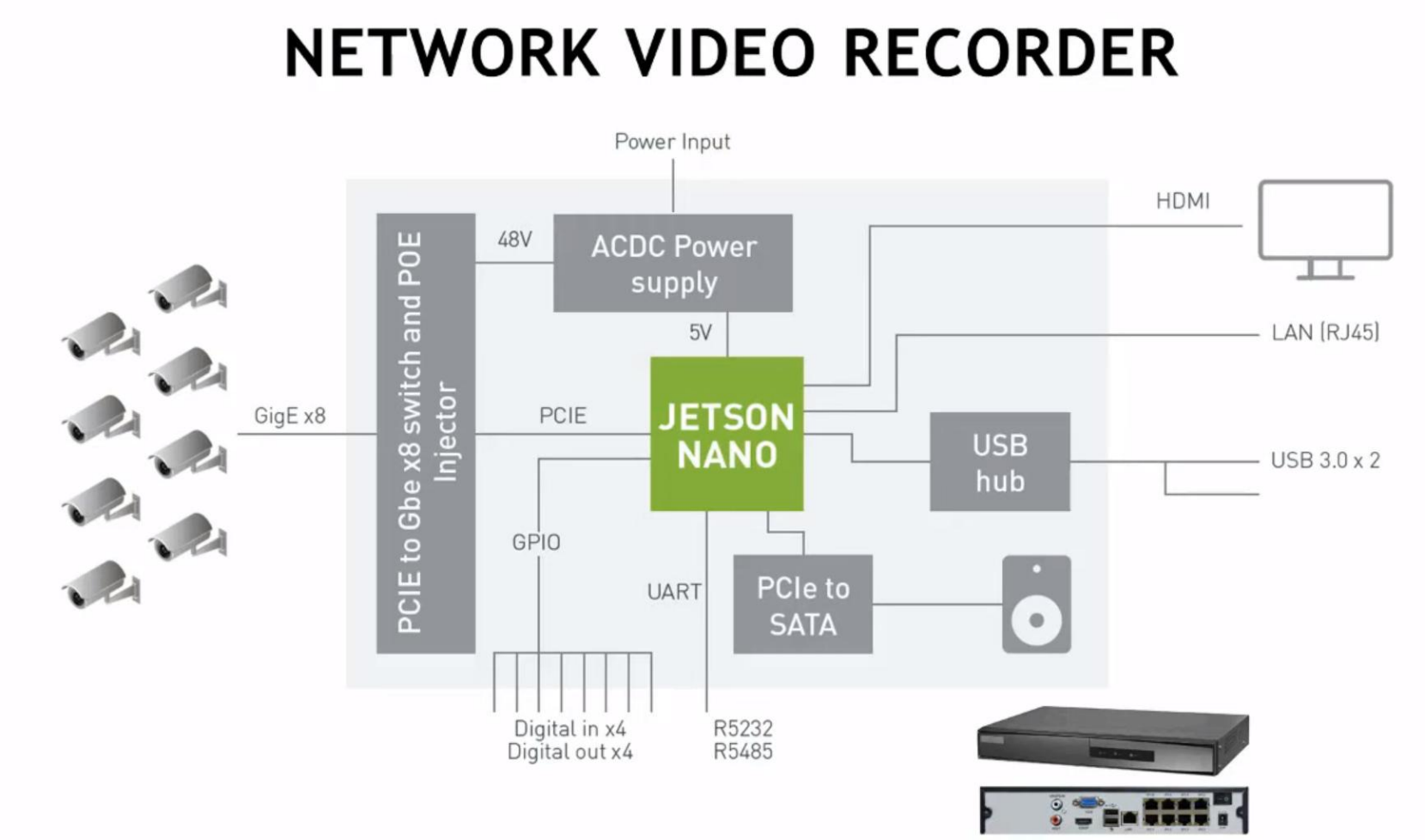
Jetson Nano

Model	Application	Framework	NVIDIA Jetson Nano	Raspberry Pi 3	Raspberry Pi 3 + Intel Neural Compute Stick 2	Google Edge TPU Dev Board
ResNet-50 (224×224)	Classification	TensorFlow	36 FPS	1.4 FPS	16 FPS	DNR
MobileNet-v2 (300×300)	Classification	TensorFlow	64 FPS	2.5 FPS	30 FPS	130 FPS
SSD ResNet-18 (960×544)	Object Detection	TensorFlow	5 FPS	DNR	DNR	DNR
SSD ResNet-18 (480×272)	Object Detection	TensorFlow	16 FPS	DNR	DNR	DNR
SSD ResNet-18 (300×300)	Object Detection	TensorFlow	18 FPS	DNR	DNR	DNR
SSD Mobilenet-V2 (960×544)	Object Detection	TensorFlow	8 FPS	DNR	1.8 FPS	DNR
SSD Mobilenet-V2 (480×272)	Object Detection	TensorFlow	27 FPS	DNR	7 FPS	DNR
SSD Mobilenet-V2 (300×300)	Object Detection	TensorFlow	39 FPS	1 FPS	11 FPS	48 FPS
Inception V4 (299×299)	Classification	PyTorch	11 FPS	DNR	DNR	9 FPS
Tiny YOLO V3 (416×416)	Object Detection	Darknet	25 FPS	0.5 FPS	DNR	DNR
OpenPose (256×256)	Pose Estimation	Caffe	14 FPS	DNR	5 FPS	DNR
VGG-19 (224×224)	Classification	MXNet	10 FPS	0.5 FPS	5 FPS	DNR
Super Resolution (481×321)	Image Processing	PyTorch	15 FPS	DNR	0.6 FPS	DNR
Unet (1x512x512)	Segmentation	Caffe	18 FPS	DNR	5 FPS	DNR



Jetson Nano 응용

Jetson Nano


NETWORK VIDEO RECORDER
Power Input
ACDC Power supply
48V
5V
PCIE to Gbe x8 switch and POE Injector
GigE x8
PCIE
JETSON NANO
HDMI
LAN (RJ45)
USB hub
USB 3.0 x 2
GPIO
UART
PCIe to SATA
Digital in x4
Digital out x4
R5232
R5485

Jetson Nano

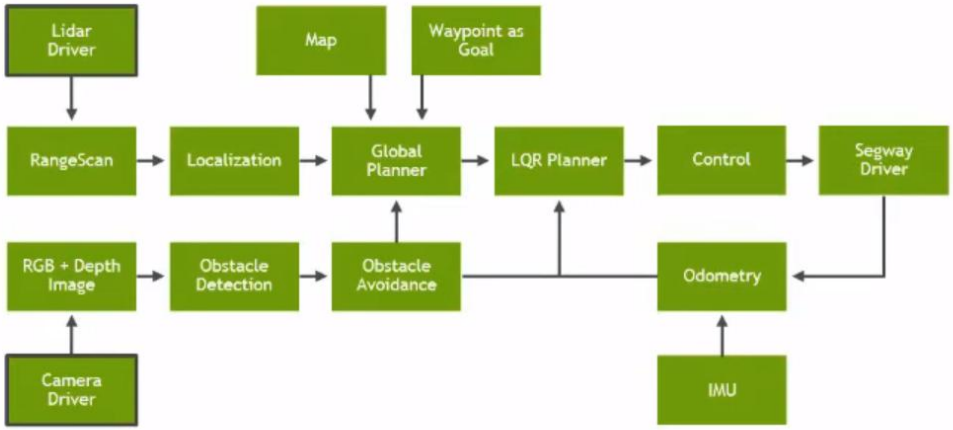
ISAAC(아이작) 플랫폼 : NVIDIA의 로봇 애플리케이션 플랫폼

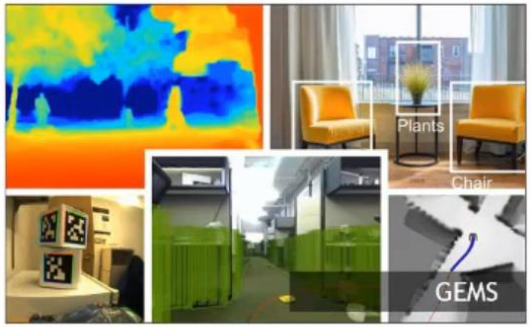
Jetson Nano

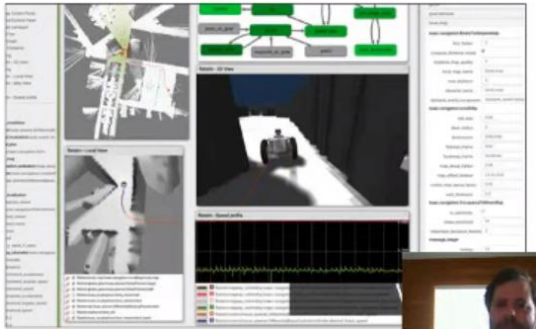
ISAAC ROBOTS

NVIDIA Carter

NVIDIA Kaya







Jetson Nano

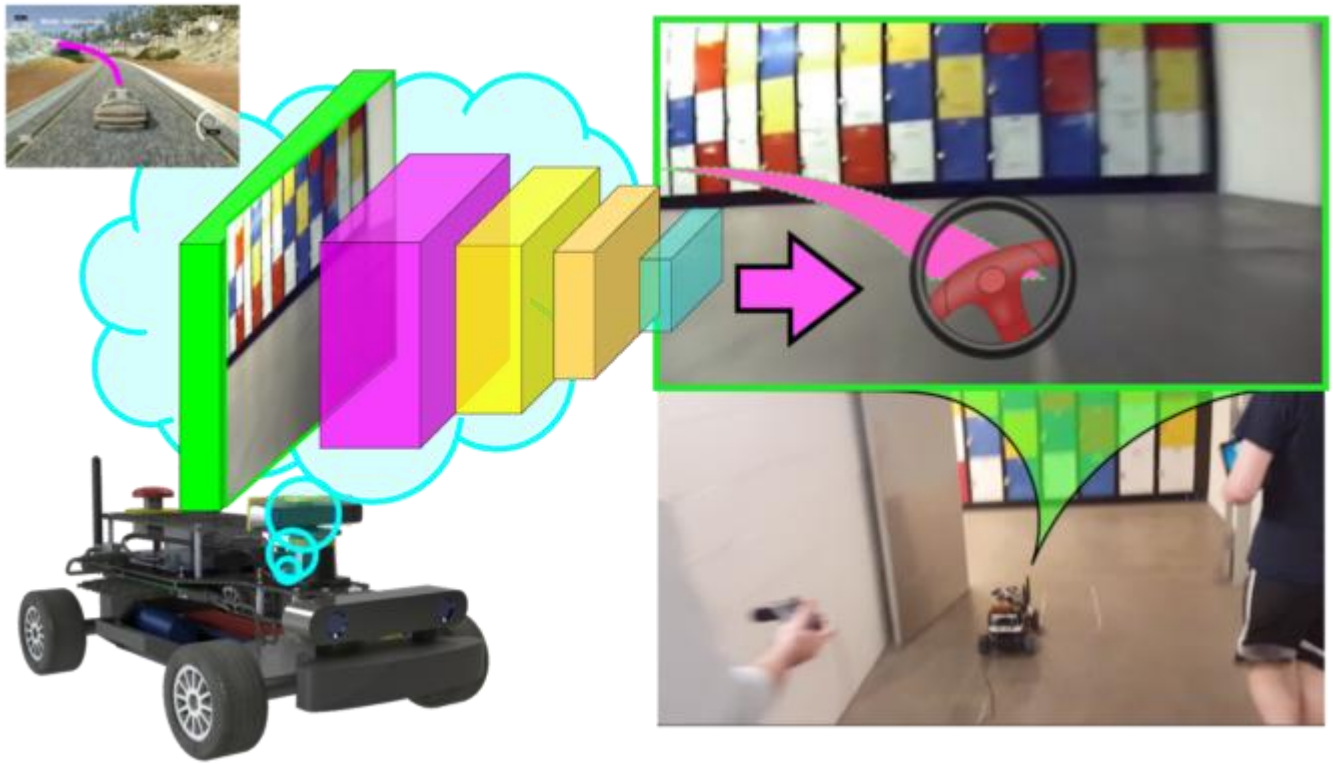
- Jetson Nano 개발 kit으로 만들 수 있는 것

1.) 자율 기계 & 로봇공학
2.) 지능형 비디오 분석 & 스마트 시티
3.) 인공지능 사물인터넷(AIoT)

Jetson Nano

미국 매사추세츠공대(MIT) 학생
-엔비디아의 젯슨 AGX 자비에(Xavier)와 쿼드로(Quadro) RTX 기반 데이터 사이언스 워크스테이션
-학생들은 자동차에 텐서플로우(TensorFlow) 뉴럴 네트워크를 학습시켜 자율주행하는 법을 학습

Jetson Nano



THANK YOU

Suggestions Questions