

# MiTAC Industrial Edge AI Solutions

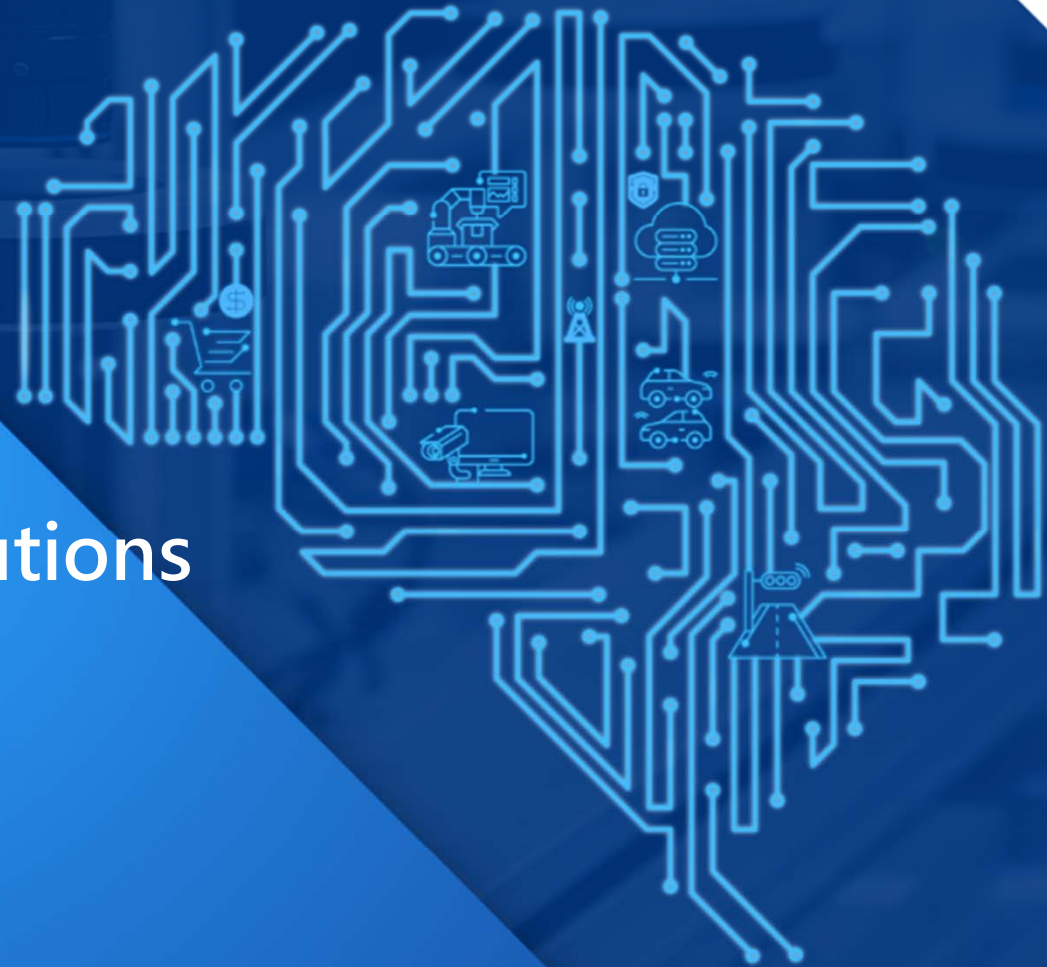
V1.1

REPORTER

MiTAC

DATE

2022/11/7



# MiTAC Industrial Edge AI Solution



LOW  LATENCY	HIGH  SECURITY	LOW  BANDWIDTH	HIGH  RELIABILITY
--------------------	----------------------	----------------------	-------------------------

Edge Inferencing Workstation / Machine Learning Workstation

Edge Inferencing System

MZ1 + 2\*NVIDIA GPU Card  
or 2\*Ainstein-H8 (Inference)

MX1 + 1\*NVIDIA GPU Card  
or 1\*Ainstein-H8 (Inference)

MX1 + VPU

MP1 + VPU

MB1 + VPU

NV1 + VPU  
Smart NVR

SBC + VPU

Traffic Management & Tolling	Traffic Monitoring	Intersection Safety	Public Health Monitoring
Autonomous Delivery	Quality Inspection	Factory Safety	Retail Automated Checkout
Smart Building	Advanced Driver Assistance (ADAS)	Front Facing Perception	Access Control

Your success is our ambition.

MiTAC Confidential – not for public disclosure



# About HAILO-8



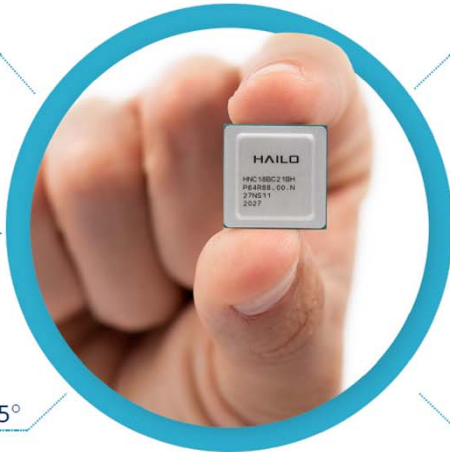
**High Performance**  
26 TOPS  
Efficient AI architecture



**Comprehensive SW Tools**  
Mature dataflow compiler  
Efficient RT library



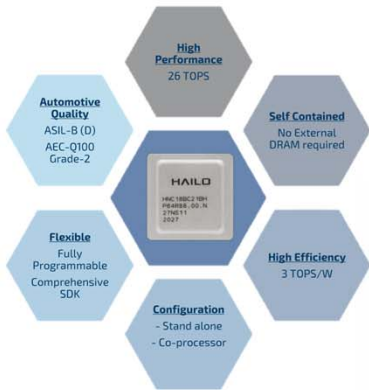
**Industrial & Automotive Grades**  
Industrial: -40° up to 85°  
Automotive: -40° up to 105°



**Power Efficiency**  
Typical Power Consumption: 2.5W

**Single Chip Solution**  
No External DRAM required

**Scalable & Flexible**  
Multi-streams  
Multi-model  
Multi-chip



Single-VPU Modules      4/8-VPU PCIe x16 Card

**HAILO**  
Empowering Intelligence

Delivering Data-Center Class Performance at the Edge



Industrial 4.0



Smart City



Smart Retail



ITS



Automotive



Smart Home

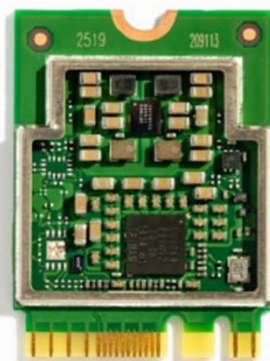
# Unprecedented AI Performance

Highest Performance AI module on the Market



Intel Myriad X

87 FPS  
35 FPS/W



Google Edge TPU

385 FPS  
275 FPS/W



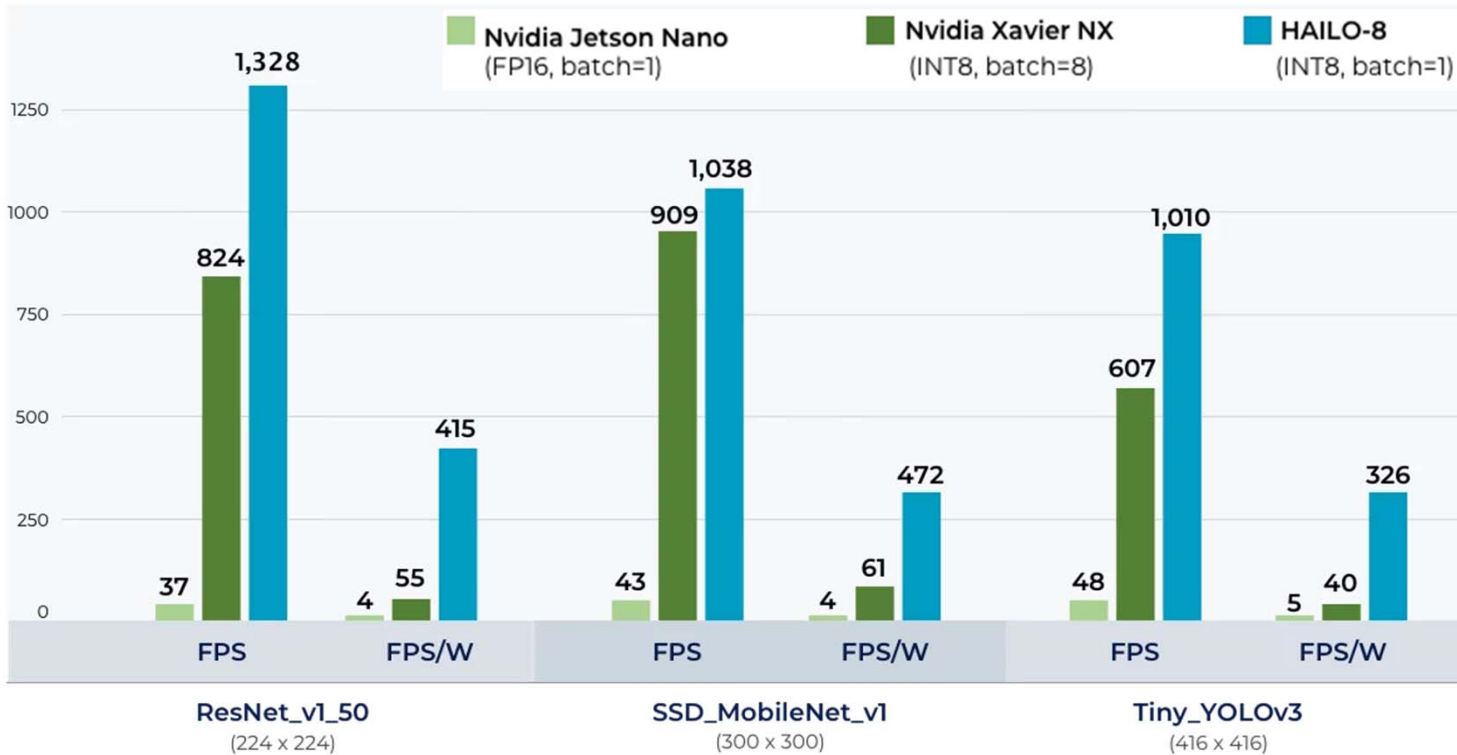
Hailo-8™

2,613 FPS  
1,267 FPS/W

- Hailo-8™ delivers better throughput: **x30** better than Myriad X and **x6** than Edge TPU
- Power Efficiency is **x30** better than Myriad X and **x4** than Edge TPU
- Provides the scalability to run advanced video analytics NN models in high-resolution & high-frame rate

# Unprecedented AI Performance

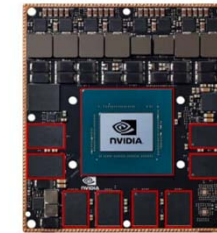
Higher Performance and As Much As x8 the Power Efficiency of nVidia's Best Edge Device



**X15 Better**  
Area Efficiency

NVIDIA AGX Xavier

Hailo-8™



General Purpose GPU  
+ External Memory

Dedicated AI Chip  
No External Memory

**X20 Better**  
Power Efficiency

ResNet-50 Benchmark

Device	Total Power [Watt]	Total Power Efficiency [TOPS/W]
Hailo-8™	1.7	2.8
Nvidia Xavier AGX	32	0.14

Conditions:

- TOPS (8-bit): Xavier 32 TOPS, Hailo-8 26 TOPS
- 224x224 image resolution feed @ 656 FPS
- 8-bit precision
- Batch size = 1

• **Xavier NX results are using batch=8** (while Hailo-8 and Jetson Nano are using batch=1) and that **Jetson Nano is limited to FP16** (while Hailo-8 and Xavier NX are INT8). Nvidia results for batch=1 and INT8, respectively, are expected to be lower.  
 • **Hailo-8** figures are based on SDK version 3.12.0 (November 2021), measured at room temperature on a single Hailo-8 device through PCIe interface on a [Hailo evaluation board](#) (system host: Intel® Core™ i5-9400 CPU @ 2.90GHz)  
 • FPS & power figures for Nvidia Jetson Nano and Xavier NX are sourced from <https://developer.nvidia.com/embedded/jetson-benchmarks> and [https://github.com/NVIDIA-AI-IOT/jetson\\_benchmarks/tree/master/benchmark\\_csv](https://github.com/NVIDIA-AI-IOT/jetson_benchmarks/tree/master/benchmark_csv), retrieved 12/07/21

# HAILO-8 VPU Measured Benchmarks

Model	Type	Input Resolution	FPS	Total Power [W]	FPS/W
ResNet-v1-50	Classification	224x224	1,328	3.1	428
MobileNet-v2-1.0	Classification	224x224	2,613	2.1	1,267
MobileNet_v3 <sup>4</sup>	Classification	224x224	3,468	1.8	1,878
RegNetx_800mf	Classification	224x224	2,447	2.0	1,232
EfficientNet-M	Classification	240x240	889	3.2	278
SSD-MobileNet-v1	Object Detection	300x300	1,038	2.3	452
Tiny-YOLOv3	Object Detection	416x416	1,010	3.2	315
YOLOv3 <sup>5</sup>	Object Detection	608x608	60	4.2	14
YOLOv4 <sup>5</sup>	Object Detection	512x512	72	3.1	23
YOLOv5m	Object Detection	640x640	218	4.3	50

**Notes:**

1. Based on Dataflow compiler version 3.14.0 (Q1 2022)
2. Measurements are done in room temperature through PCIe interface on Hailo-8 evaluation board
3. System host: Intel(R) Core(TM) i5-9400 CPU @ 2.90GHz
4. MobileNet-V3 - the benchmarked model flavor is Mobilenet V3 Large Minimalistic
5. Performance figures are gives for processing 8 simultaneous streams

# MiTAC Einstein-H8 Highlights

## Multi-Processor (4 or 8 VPU) PCIe x16 Visual AI Card

Active/Passive Cooling

Industrial Grade  
-40°C ~ 60°C / 85°C

Low Power  
Typical > 20W / 30W

Scalable & Flexibility

Number of VPU  
4 / 8

High Performance  
104TOPS / 208TOPS

ML Framework Support  
ONNX/TensorFlow/pyTorch

Comprehensive AI SW  
Suit Support

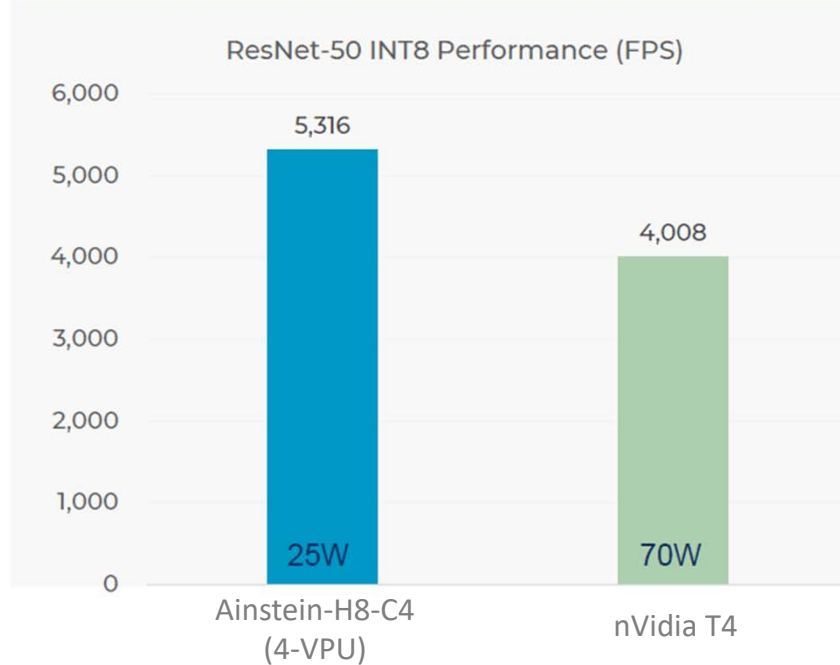
**AINSTEIN-H8-C4** 4 x Hailo-8™ AI Core

**AINSTEIN-H8-C8** 8 x Hailo-8™ AI Core

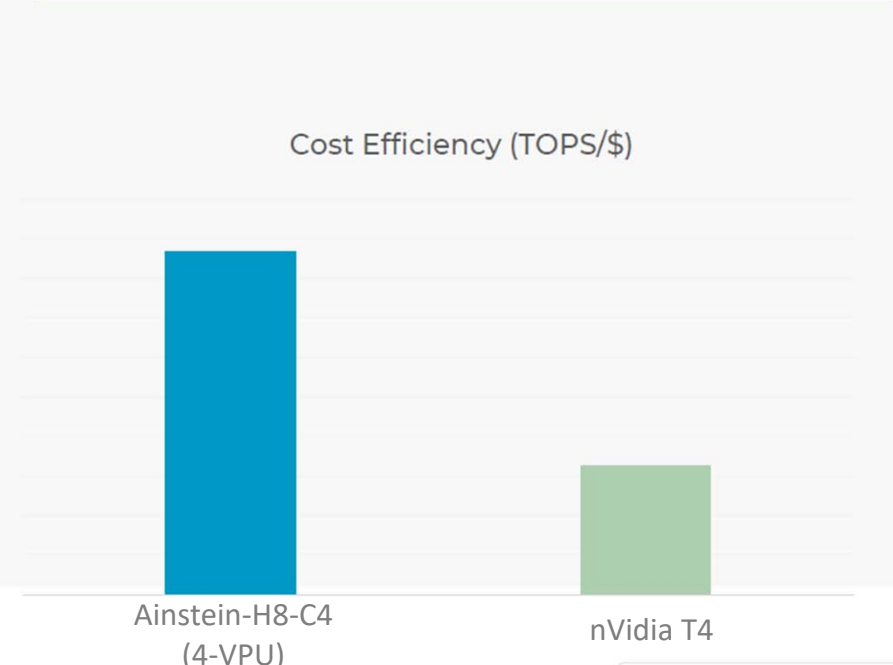
# Ainstein-H8-C4 Compares to NVIDIA T4

## Cost and Power Efficiency

Get more performance at **1/3** of the power

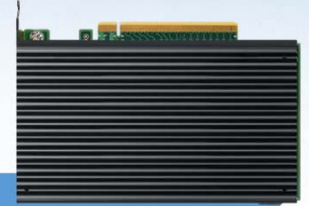


Get **X3** more performance for every \$ spent





# MiTAC Einstein-H8 Spec (1/2)



ITEM	SPEC DESCRIPTION
PCIe I/F	PCIe Gen3 x16
AI Performance	104 TOPs, 5,300 FPS of ResNet, and 900 FPS of YOLO v5m 208 TOPs, 10,600 FPS of ResNet, and 1,800 FPS of YOLO v5m
AI Processor	4 x Hailo-8™ AI Processor (Default SKU) 8 x Hailo-8™ AI Processor (Optional SKU)
ML Frameworks Support	ONNX; TensorFlow; pyTorgc
Runtime Support	GStreamer; ONNX Runtime; Native C/C++
OS Compatibility	Linux ; Windows
Power Consumption	4 x Hailo-8™: Max. 42W; Typ. 20W / 8 x Hailo-8™: Max. 74W; Typ. 30W
Operating Temperature	-40°C ~ 60°C / 85°C (depends on SKU)
Storage Temperature	-40°C ~ 85°C
Operating Humidity	10% ~ 95% R/H (Non-condensing)
Dimension	200mm x 111mm x 36mm (including heatsink fan)

# MiTAC Ainstein-H8 Spec (2/2)



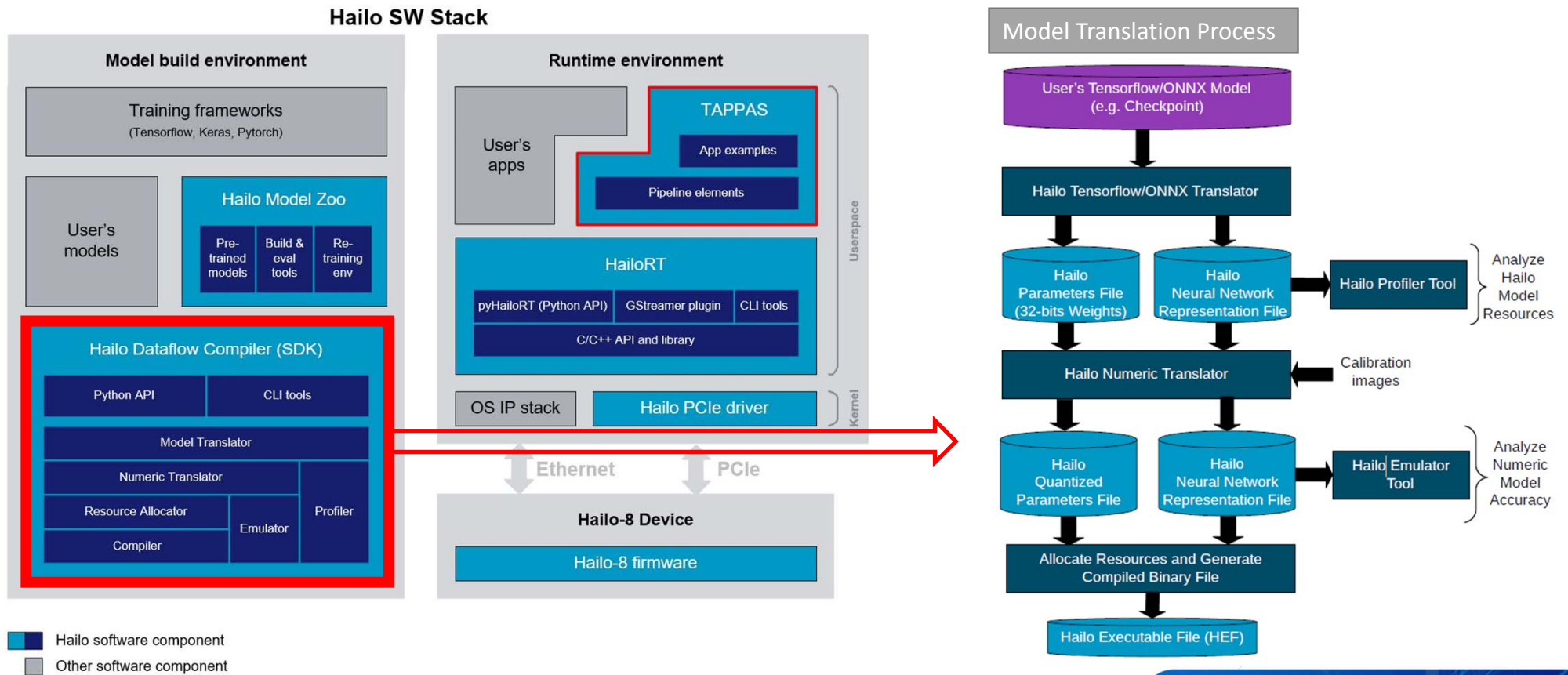
ITEM	SPEC DESCRIPTION
Internal Slot	2 x 2-pin FAN header (1 for embedded fan and 1 reserved for customer use)
Thermal	Heatsink with embedded system fan
Accessories	Full height PCIe bracket with venting holes
Certification	CE, FCC

## Ordering Information

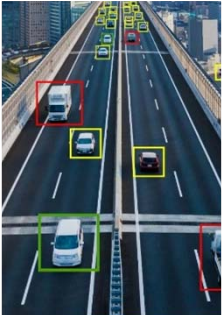
Model Number	AI Processor	Thermal Solution	Operating Temperature
AINSTEIN-H8-C4-PC	4 x Hailo-8™ AI Core	Passive Heatsink Cooling	-40 ~ 60°C
AINSTEIN-H8-C8-PC	8 x Hailo-8™ AI Core	Passive Heatsink Cooling	-40 ~ 50°C
AINSTEIN-H8-C4-AF	4 x Hailo-8™ AI Core	Active Fan with Heatsink	-40 ~ 85°C
AINSTEIN-H8-C8-AF	8 x Hailo-8™ AI Core	Active Fan with Heatsink	-40 ~ 75°C

# HAILO's Comprehensive AI Software Suite

Seamlessly Integrates with Existing ML Development Frameworks



# MiTAC Edge AI Application Scenarios (1)



## Intelligent Transportation System

- Autonomous Driving
- Mobile NVR
- Intelligent Passenger Information System

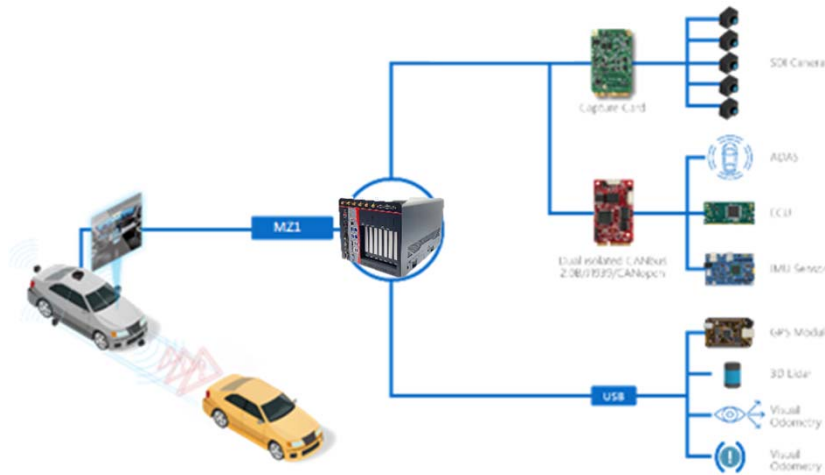


## Smart Manufacturing

- Robotic Vision
- AI Inspection
- Workspace / Worker Safety

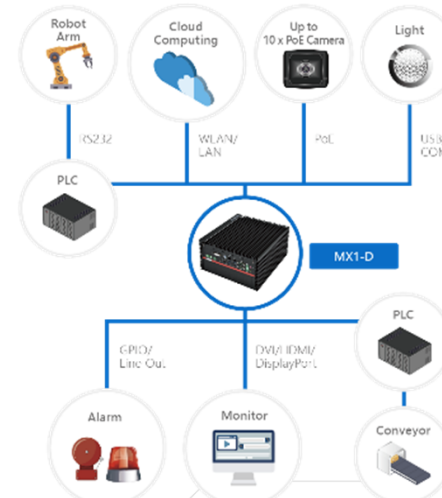
### Autonomous Driving Control Unit

Solutions: MZ1 + 2x Ainstein-H8 Card



### Robotic Vision + Control

Solutions: MX1-10FEP-D + Ainstein-H8 Card



# MiTAC Edge AI Application Scenarios (2)



## Smart City

- Traffic Monitoring & Control
- Public Safety
- Infrastructure & Services Management

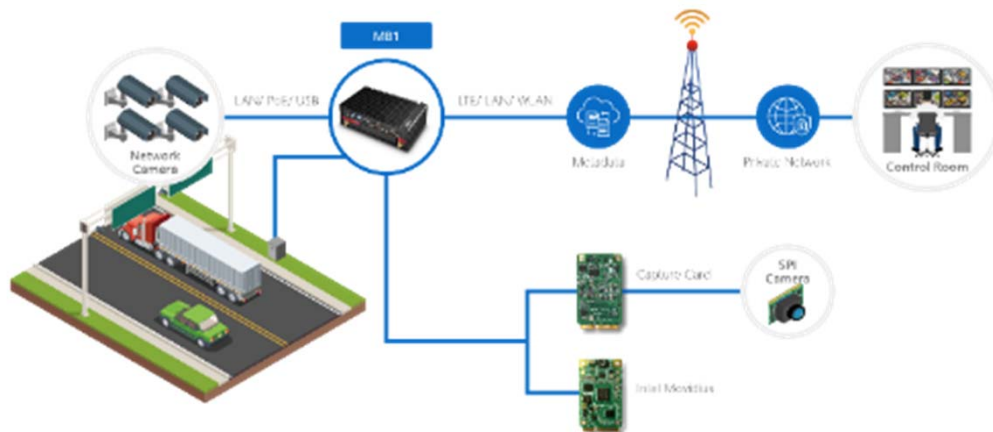


## Smart Retail

- Smart Analytics
- Self-Checkout
- Automated Store

### Traffic Flow Monitoring & Analytics

Solutions: MB1-10AP + HAILO VPU Module



### NVR + AI Analytics

Solutions: NV1 + HAILO VPU Module



# Thank you

YOUR SUCCESS IS OUR AMBITION.

