White paper

Technical Application Note — nanoONYX-SDI + Hailo-8R for AI processing

March 2026

This application note presents an analysis of the benefits of the nanoONYX-SDI + Hailo-8R integration. It describes the role of the Hailo-8R accelerator, the advanced Hailo software capabilities, the test results, the performance characteristics, and the complete workflow for model updates via HEF.

Note:
The HEF model (Hailo Executable Format) is the executable file format used by AI accelerators developed by Hailo, including the Hailo-8R AI Accelerator. It contains an AI model that has been compiled and optimized to run directly on the Hailo chip**.**

Technical Description — nanoONYX-SDI (English)

The nanoONYX-SDI is an ultra-compact, rugged embedded computer designed for demanding applications operating in harsh environments. It combines a high-performance x86 architecture, advanced modularity, and native support for professional HD-SDI video streams, while maintaining mechanical and thermal robustness compliant with military and avionics constraints.

Thanks to its modular design, strong processing capabilities, and support for embedded AI through extension cards such as the Hailo-8R, the nanoONYX-SDI is an ideal platform for real-time vision, mission-critical, and control applications.

The nanoONYX-SDI is a compact, rugged, high-performance embedded computer designed for operation in harsh and demanding environments. With its powerful x86 architecture, advanced modularity and native support for HD-SDI video acquisition, it is an ideal platform for mission-critical vision, surveillance, and real-time processing applications.

Key Strengths

- High-performance x86 architecture in a rugged compact chassis.
- Native real-time HD-SDI video handling.
- Advanced modularity with multiple extension slots.

Main Technical Features

- Intel® Core™ i7-1185GRE processor (or rugged equivalent).
- Up to 16 GB DDR4 (32 GB DDR5 in future versions).
- Up to 1 TB SSD storage.
- Up to 2 real-time HD-SDI inputs.
- 4 expansion slots (Mini PCIe, M.2 depending on model).

- Compatible with AI accelerators such as Hailo-8R (13 TOPS).
- Operating temperature from −40 °C to +55 °C without external airflow (up to +71 °C depending on cTDP).
- MIL-STD-810, MIL-STD-461, DO-160 qualified.
- Interfaces: Ethernet 10/100/1000, USB, Serial, DVI-D, GPIO.
- Operating systems: Linux and Windows.

Detailed Role of the Hailo-8R in the Test

Within the nanoONYX-SDI architecture, the Hailo-8R accelerator handles 100% of AI inference workloads (object detection, segmentation, classification). The Intel® CPU remains dedicated to mission logic, HD-SDI acquisition, and system supervision.

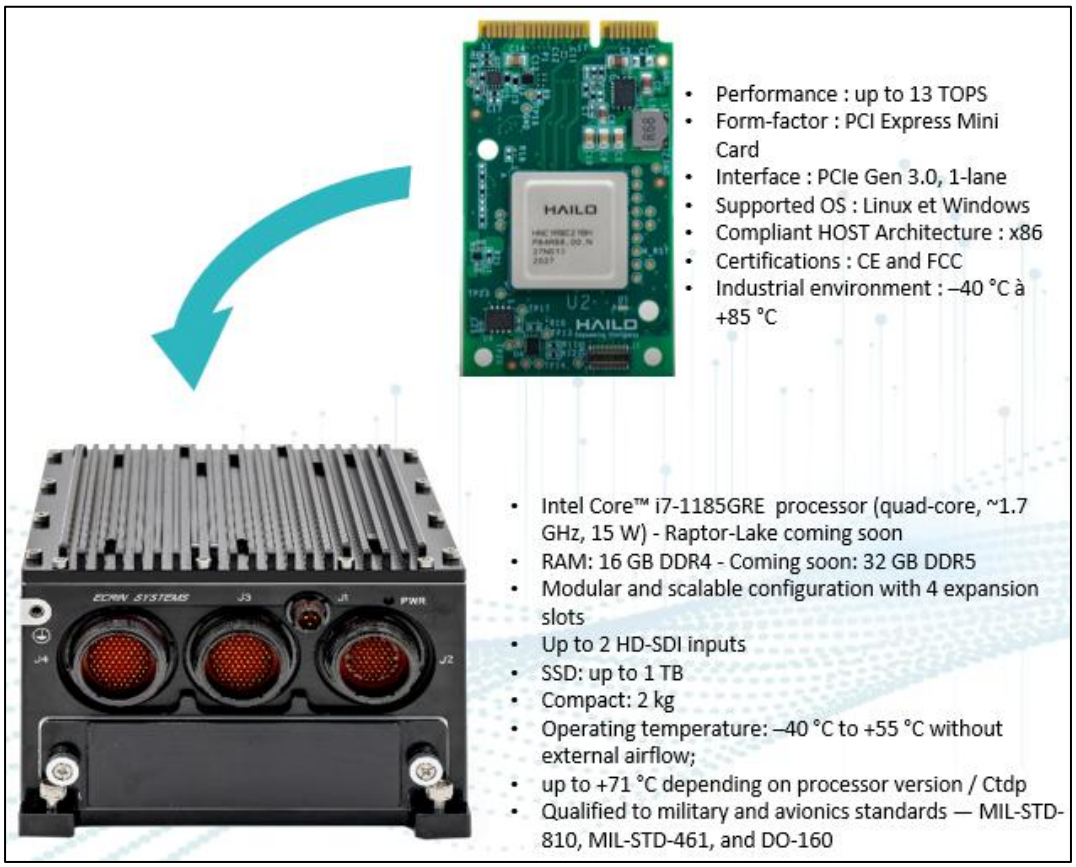

fig 1. The system consists in the integration of the Hailo-8R miniPCIe card in the nanoONYX-SDI

The test results showed that CPU load remains unchanged with or without AI processing, confirming full AI offload to the Hailo-8R.

Objective:

Demonstrate that real-time AI capabilities can be added to an existing rugged system without impacting power consumption, thermal budget, or software architecture, and without degrading HD-SDI video functions.

Most standard processors, whether CPUs or general-purpose GPUs, are not particularly efficient when handling intensive AI workloads. The main reason is that traditional processors are designed for sequential or general-purpose computations rather than the highly parallel operations required by modern AI algorithms. Tasks such as matrix multiplications, tensor operations, and deep neural network inference require massive parallelism and high memory bandwidth—capabilities that typical processors cannot fully exploit. As a result, AI training or inference on conventional processors is often slow, energy-inefficient, and generates substantial heat, requiring heavy cooling solutions.

This is where the Hailo-8R demonstrates its true relevance. Unlike general-purpose processors, the Hailo-8R is architected specifically for AI workloads. It features highly parallel processing units optimized for tensor computations, low-latency memory access, and specialized accelerators for common AI operations. Its design is also optimized for energy efficiency and thermal management, allowing it to maintain high performance while minimizing heat generation. This reduces the need for bulky cooling systems, lowers power consumption, and improves system reliability.

In short, the Hailo-8R delivers superior AI performance, efficient energy use, and optimized thermal dissipation, making it a highly relevant choice for organizations looking to deploy AI at scale.

Test Setup and Implementation Environment

The following equipment and system configuration were used to perform the functional benchmark and validate the integration of the AI accelerator within the nanoONYX-SDI platform.

Video Input Source

The video input stream is provided by an HD-SDI camera configured with the following parameters:

- Resolution: 720p
- Frame rate: 30 fps

The camera delivers a continuous HD-SDI signal to the system through the frame grabber interface.

Video Output Display

The processed video stream is displayed on a digital monitor using a DVI-D interface, allowing real-time visualization of the video pipeline and AI processing results.

Processing Platform: nanoONYX-SDI

The processing platform used for the evaluation is the nanoONYX-SDI rugged computing system, equipped with a COM-Express TGU9A2 module featuring an Intel® Core™ i7-1185GRE processor.

The system includes several extension slots used for the functional setup.

Extension slots configuration:

- Slot #1: miniPCIe HD-SDI frame grabber card
- Slot #2: CAN Bus interface card
- Slot #3: miniPCIe Hailo-8 AI accelerator card
- Slot #4: Not used

This configuration enables real-time acquisition of the HD-SDI video stream, AI inference processing through the Hailo accelerator, and visualization of the processed video output.

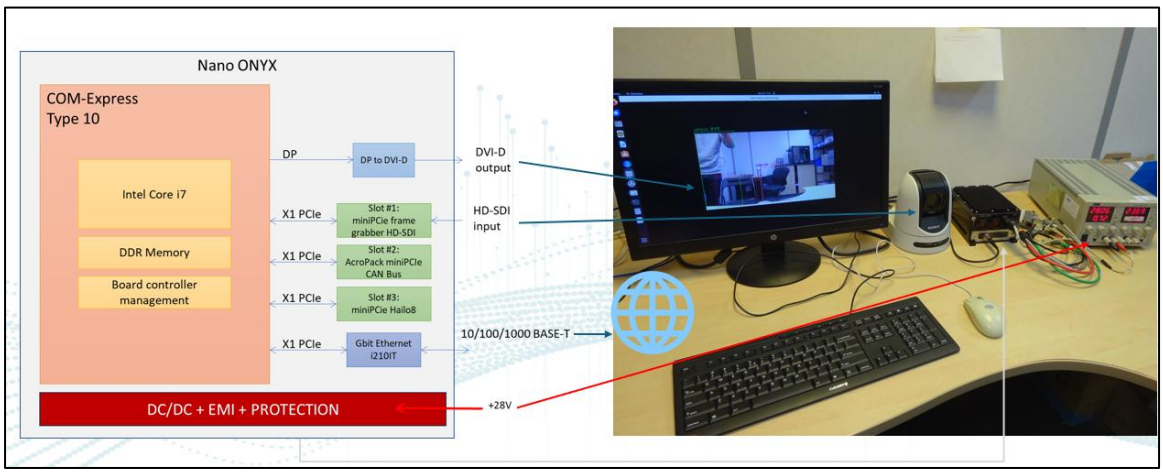

fig 2. Test set-up : camera, nanoONYX, Screen

Launching the Hailo application

cd /home/ecrin/TEST/hailo-apps-infra/
source setup_env.sh
hailo-detect-simple -i /dev/video0 --arch hailo8l

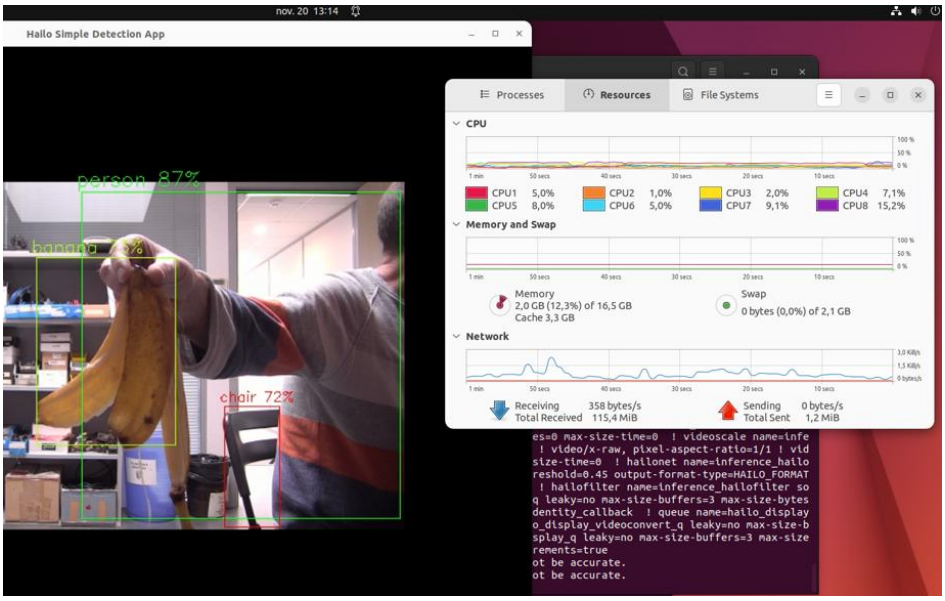

Fig.3 Display of the video capture and processing, as well as CPU load.

Full Workflow for Model Updates (HEF)

The workflow for updating models on the nanoONYX-SDI consists of three main steps:

Model preparation, either BYOD (retraining networks from the Model Zoo) or BYOM (importing ONNX or TFLite models).

Compilation, using the Dataflow Compiler (on a PC x86) to perform optimization and quantization, resulting in the generation of a HEF file.

Deployment, transfer the HEF file to the nanoONYX-SDI, then load and execute it using HailoRT. Note that there is no on-device learning or autonomous model updates.

Why the nanoONYX-SDI Cannot Compile Models

The Dataflow Compiler requires a full development workstation (x86 architecture, Python environment, and heavy dependencies). Model compilation involves intensive analysis, optimization, and quantization, which is unsuitable for an embedded production system. Best practice: compile the model on a PC, and execute it on the nanoONYX-SDI.

Results and Conclusion

The tests performed with the nanoONYX and the Hailo-8R card demonstrate that the object recognition application is fully operational. The CPU load remains unchanged whether the AI processing is enabled or not, confirming that the inference workload is entirely offloaded to the Hailo-8R accelerator. The total power consumption of the nanoONYX-SDI platform was measured at 21.5 W, while the Hailo-8R module itself consumes approximately 2.5 W. Real-

time performance is preserved, as the HD-SDI video stream is processed without any impact on the mission logic or overall system responsiveness.

	CPU Intel	Hailo-8R
Processor	i7-1185GRE – Quad Core	Hailo
Consumption w/o AI processing	19,6W	X
Consumption with AI processing	19,6W	Very low: 2,5W
nanoONYX overall power consumption	21,5W	
Processor load with AI processing	No change	n/a

The nanoONYX-SDI + Hailo-8R combination validates a powerful, energy-efficient approach to embedded edge AI: real-time inference, free CPU resources, controlled power budget, and industrial ruggedness. This architecture is immediately applicable to defense, avionics, security, and Industry 4.0 systems.

Glossary

Term	Definition / Description
HEF	Hailo Executable Format - Compiled and optimized model file for Hailo AI accelerators, ready for deployment on nanoONYX-SDI.
BYOD	Bring Your Own Data - Retraining pre-existing Model Zoo networks using your own dataset.
BYOM	Bring Your Own Model -Importing an existing AI model (ONNX/TFLite) to run on Hailo hardware.
Model Zoo	Library of pre-trained AI models available for reuse or retraining.
ONNX	Open Neural Network Exchange - Open format for AI models enabling interoperability between different frameworks (e.g., PyTorch, TensorFlow).
TFLite	TensorFlow Lite - Lightweight version of TensorFlow for deployment on embedded and mobile devices.
Dataflow Compiler	Software tool to optimize, quantize, and generate HEF files from AI models, requires x86 workstation.
HailoRT	Runtime environment for loading and executing HEF models on Hailo accelerators.
Frame grabber	Hardware that captures video frames from a video stream for processing.
DVI-D	Digital Video Interface used for displaying video output on monitors.
CPU load	Percentage of processor usage, used to measure system performance impact.
TOPS (Tera Operations Per Second)	Measure of AI accelerator performance (Hailo-8R: 13 TOPS).
MIL-STD-810 / MIL-STD-461 / DO-160	Military and avionics standards for environmental robustness, electromagnetic compatibility, and durability.

Term	Definition / Description
Tensor operations	Mathematical operations on multidimensional arrays used in deep learning computations.
Quantization	Optimization process converting model weights/activations to lower-precision numbers to reduce memory and computation.

Contact us

For more information about ECRIN Systems and how our solutions can benefit your applications visit our website https://ecrin.com/en

Visit

Contact info
+33 (0)4 76 92 20 01
sales@ecrin.com
ECRIN Systems
143 Rue Louis Neel
38920 Crolles
FRANCE