

# Hailo-8R Mini PCIe Al Acceleration Module

Deliver server class performance to edge devices



## **Key Features & Benefits**

Provides 13 Tera-Operations Per Second (TOPS)

Best-in-class power efficiency

Enabling real-time, low latency, and high-efficiency Al inferencing on edge devices Highest cost-efficiency (TOPS/\$) compared with existing solutions

Scalable, enabling simultaneous processing of multi-streams & multi-models

Robust software suite supports state-of-the-art deep learning models & applications out-of-the-box Supporting extended temperature range of -40°C to 85°C

Fast time-to-market using a standard form factor module

# **Technical Specifications**

Form factor

PCI Express Mini Card Full-Mini F1

Dimensions

30×50.95 mm

Interface

PCIe Gen 3.0, 1-lane

Supported OS

Linux & Windows

Supported AI Frameworks TensorFlow, TensorFlow Lite, Keras, PyTorch & ONNX Supported host architecture X86 or ARM based

Certification CE. FCC

## **Seamless Integration**

Compatible with the PCI Express Mini (mPCle) form factor, the module can be plugged into an existing edge device with an mPCle Full-Mini socket to process deep learning inferencing in real time and with low power consumption.

## Featuring the Hailo-8R AI Processor

Hailo-8R unique, powerful, and scalable structure-driven dataflow architecture takes advantage of the core properties of neural networks. It enables edge devices to run deep learning applications at full scale more efficiently, effectively, and substantially than traditional solutions, while significantly lowering costs.

### **Comprehensive Software Suite**

Al software suite which seamlessly integrates with existing deep learning development frameworks to allow smooth and easy integration in existing development ecosystems. The Hailo Al software suite includes:

#### **Model Build Environment Runtime Environment Model Build Computer Host Processor** Machine Learning Applications Frameworks **TAPPAS TAPPAS** 作 Application A set of full application examples TensorFlow TensorFlowlite examples, implementing pipeline elements and K $\dot{O}$ ⇎ pre-trained AI tasks PyTorch ONNX 7 Hailo HailoRT HailoRT User OS IP Model Production-grade, Models Stack 700 precompiled, runtime software for the host ♠ PCIe processor Hailo Dataflow Hailo-8 Device Compiler Hailo-8 Firmware ■ Hailo SW component Other SW component

### Model Zoo

Over 100 of common & state-of-the-art pre-trained and retrainable models in TensorFlow & ONNX

### Hailo Dataflow Compiler

A software toolset for optimization & translation of trained models to Hailo-8 format

# Hailo Ecosystem

Hailo collaborates with leading partners to seamlessly integrate their products with ours, creating easy-to-use, dependable, and highquality solutions that adhere to the industry's most rigorous standards. Our extensive network of trusted partners covers various regions and market segments, delivering innovative solutions to our customers.

To view our extensive network click here

## **System Usage**

mPCIe module connected to various types of host processors via PCIe.



### Part Number

Hailo-8R mPCle module Extended Temperature: HMP1RB1C2GAE

www.hailo.ai

Click here for more information & product inquiry



