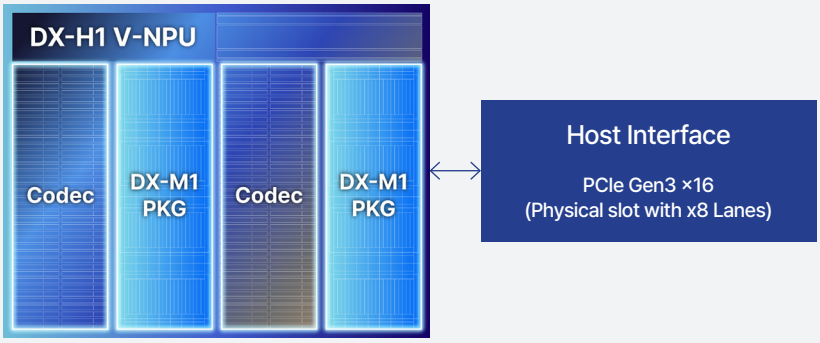
DX-H1 V-NPU

Redefining AI Power:
Massive,sustainable intelligence at 30W

Type: AI Accelerator + Codec **Performance: 50 TOPS / 30W**

Functional Block Diagram



Specifications

Features	Details
AI Performance	50 TOPS (INT8)
Codec Performance	Video Decoding: H.264/265 64CH (1080P @ 30fps)
	Video Encoding: H.264/265 32CH (1080P @ 30fps)
Host Interface	PCIe Gen3 ×16 Physical slot with x8 Lanes
Memory	VPU : 4GB x 4 LPDDR5, NPU : 2GB x 4 LPDDR5
Power Consumption	30W
Thermal Solution	Active Cooling (Integrated Fan & Heatsink Module)
Form Factor	Low Profile PCIe Card
Dimensions	166.9 mm × 68.10 mm × 18.33 mm (without Braket)
OS Support	Windows 11/10
	Debian-based Linux (Ubuntu 24.04/22.04/20.04 LTS), Docker
	Yocto Linux
	Android
AI Frameworks	Ultralytics, TensorFlow, PyTorch, ONNX, Keras
System Support	x86, ARM Based Architecture

Overview

Integrated decoding, encoding, and transcoding turn the DX-H1 V-NPU into a complete foundation for real-time, large-scale monitoring.

What once required rooms full of expensive, power-hungry servers can now be placed into the hands of any organization.

The barrier to advanced AI falls away, and what remains is possibility, accessible, sustainable, and inevitable.

Target Applications

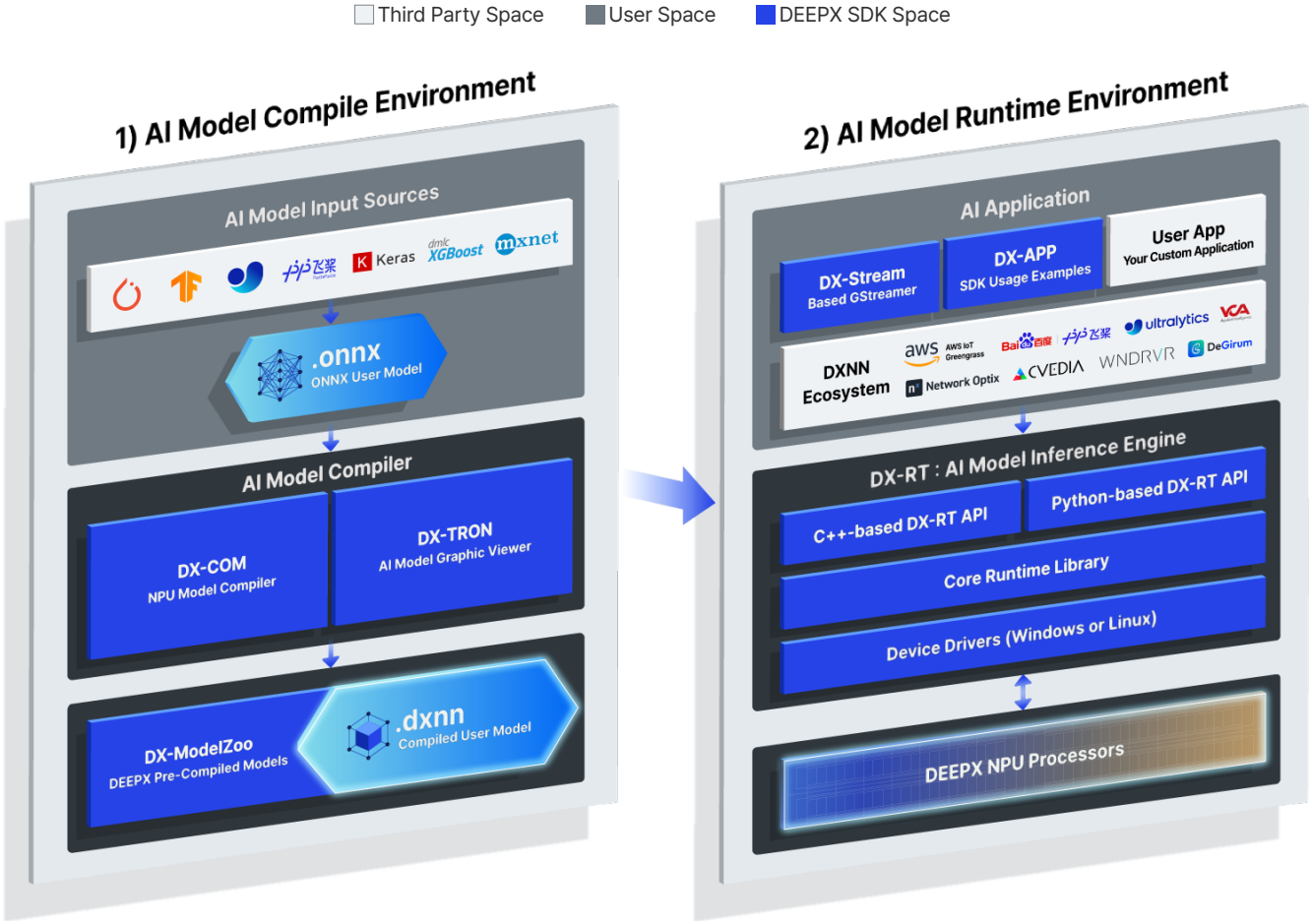
- Security
- Smart Buildings
- Smart Factory
- Smart Cities
- Content Delivery Network (Smart Search)
- Content Delivery Network
- Logistics & Warehouses
- Smart Transportation
- Smart Retail

DEEPX

DXNN® - DEEPX SW Development Kit

DXNN® (DEEPX Neural Network) SDK streamlines the AI deployment pipeline on DEEPX NPUs. By integrating essential tools for compilation, optimization, simulation, and inference, it ensures high development efficiency. Experience a ready-to-use development environment with DX-AS (All Suite), a fully integrated and version-aligned package designed for rapid scaling.

DXNN SDK Full Stack Architecture



1) AI Model Compile Environment

Optimizes models from various sources into DEEPX-specific .dxnn formats via the ONNX framework.

2) AI Model Runtime Environment

Deploys optimized models onto hardware via DX-Stream and DX-RT engines using C++ or Python APIs.

DX-AllSuite

DX-AllSuite: The Single Package for Your Complete DXNN Environment. Simplify setup, installation, and updates across local machines and Docker.

Step 1
Supported Environments

- DEEPX Developers — Learn more
- GitHub — Learn more
- aws AWS IoT Greengrass

Step 2
Auto Install & Update

DX-AllSuite

Step 3
Finalize & Start Development

- Docker Install
- Local Install

DEEPX HQ
3F, 20 Pangyoyeok-ro 241beon-gil, Seongnam-si, Gyeonggi-do, South Korea

USA
1735 Technology Drive Suite 740. San Jose, CA U.S.A

China
No 0910, Yangguang Yuehai Building, Nanshan District, Shenzhen, Guangdong, China

Taiwan
Nanjing E. Rd., Zhongshan Dist., Taipei City 104, Taiwan (R.O.C.)

Shop Now!
Contact Sales
sales@deepx.ai