

DX-M1ML Chip

Real-time, on-device intelligence will be universal, affordable, and transformative



Type: AI Accelerator

Performance: 13 TOPS / 3W (Typical)

Overview

Experience the future of on-device AI with the **DEEPX DX-M1ML**. This AI processor delivers a stunning **13 TOPS** while maintaining an ultra-low power consumption of just **3W(Typical)**. We prioritize Inferences Per Second (IPS) per Watt, so every operation delivers maximum, real-world value.

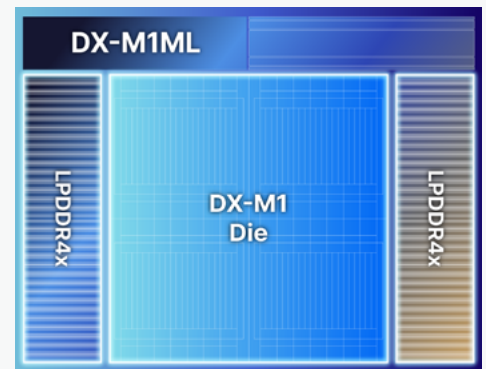
The **DX-M1ML** integrates effortlessly with your host system, accelerating diverse applications like **computer vision and audio analysis**.

We believe that powerful AI should be accessible to all. Simply use your models from popular frameworks like Ultralytics Yolo 26, TensorFlow, PyTorch or ONNX, and our tools will automatically optimize them for the DX-M1ML. This allows you to focus on innovation, not integration.

Specifications

Features	Details
AI Performance	13 TOPS (INT8)
Host Interface	PCIe Gen3 x4 (Supports Gen 1/2/3 & x1/x2/x4)
Memory	Integrated 512MB LPDDR4x (4266 MT/s)
Power Consumption	3W (Typical)
Operating Temperature	0 ~ 70°C (Commercial)
	-40 ~ 85°C (Industrial)
Package	FCL-BGA, 21 × 21mm 396-ball, 1.0 mm Pitch
OS Support	Windows 11/10
	Debian-based Linux (Ubuntu 24.04/22.04/20.04 LTS)
	Yocto Project
	Docker
AI Frameworks	Ultralytics, TensorFlow, PyTorch, ONNX, Keras
System Support	x86, ARM Based Architecture

Functional Block Diagram



Target Applications

- AI CCTV
- ADAS
- Industrial Robot Arms
- AI Dashcams
- Autonomous Mobile Robots (AMRs)
- Drones
- Edge Servers
- AIoT Devices
- Remote Patient Monitoring (RPM)

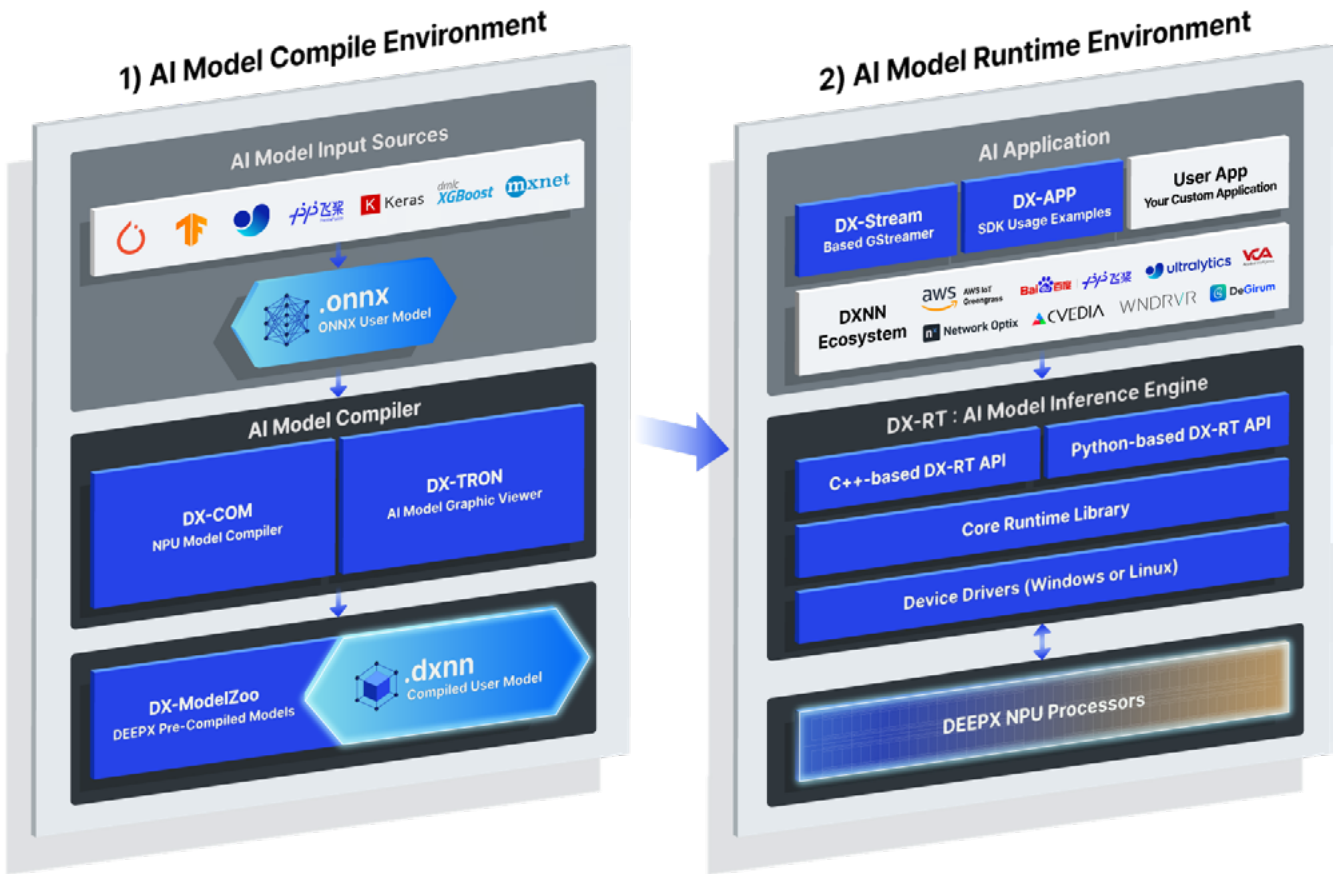
DEEPX

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

Third Party Space
 User Space
 DEEPX SDK Space



1) AI Model Compile Environment

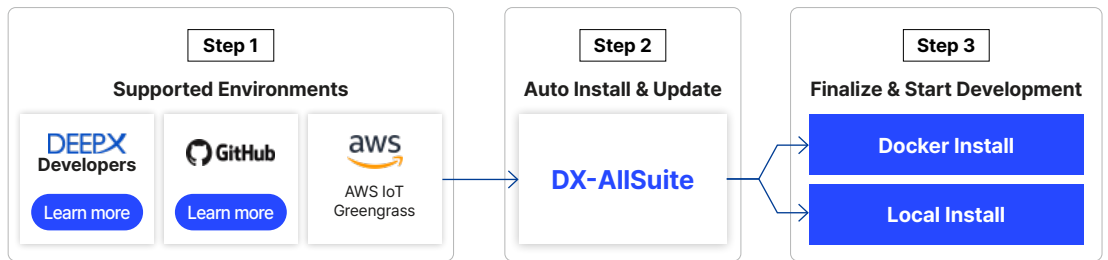
Optimizes models from various sources into DEEPX-specific .dxnn formats via the ONNX framework.

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.



DEEPX HQ
3F, 20 Pangyoeyeok-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