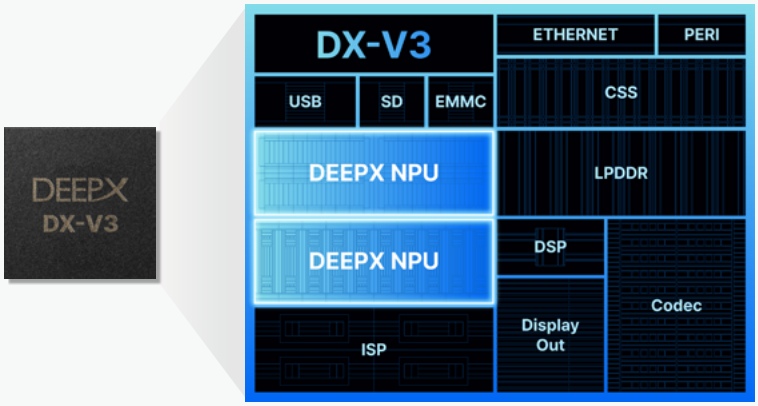
DX-V3: AI Enabler for Every Vision System

The DX-V3 is a high-performance AI SoC for diverse vision systems like 3D sensing, AI NVR, and autonomous robotics. This all-in-one SoC platform, optimized for compact, real-time edge vision AI, accelerates intelligent vision solution development by integrating advanced AI processing, image/video processing, and versatile connectivity.

"Integrating powerful AI vision processing and essential core features into a single chip, DEEPX drives innovative edge AI solutions for diverse intelligent systems."

DX-V3 Chip Functional Block Diagram



Key Features

Category	Details
AI Performance	• Up to 13 TOPS (INT8) for real-time AI inference
Powerful CPU	• Quad-core Cortex-A53 CPU @ 1.2GHz • Full Linux OS support
Advanced ISP & Video Processing	• Integrated 12MP ISP with 3A (AE/AF/AWB), and noise reduction • Supports H.264/H.265 encoding/decoding up to 4K@60fps
Multi-Channel Video Input/Output	• Captures video from 4x MIPI CSI-2 and legacy BT.656/1120 input • Output via MIPI DSI, CSI, and parallel interface for display or streaming
High-Speed I/O & Storage	• Includes USB 3.0, Gigabit Ethernet, and SD/eMMC interface
Edge AI Integration	• All-in-one SoC optimized for compact, real-time edge AI vision
DSP	• 75 GFLOPS (SLAM/Radar support)

"DX-V3 Engineering Sample (ES) scheduled for release in Q4 2025."

DEEPX

DX-V3 Specifications

> Type: AI Vision SoC
> Video Codec: H.265/264 4K @60P
> Memory: x16 2ch LPDDR4X/LPDDR5
> DSP: 75 GFLOPS (SLAM/Radar Support)
> Process: Based on TSMC 12nm process

Support DXNN® SDK

DXNN® SDK is a comprehensive SW development environment for deploying AI on DEEPX NPUs. It integrates tools for compiling, optimizing, simulating, and inferring the latest AI models, such as YOLO, ViT, and VLMs. And it provides an optimized, ready-to-use environment as the DX-All Suite package to support fast and efficient AI development.

Target Applications

- 3D Sensing & Stereo Cameras
- AI NVR (Network Video Recorder)
- AI CCTV
- Robotics
- Automotive
- Video Conferencing Cameras
- CMS (Camera Monitoring System)
- Autonomous Robotic Platforms
- Drones
- AR/VR
- SBC (Single Board Computer)
- ADAS/AD

DEEPX HQ
5F, 20 Pangyoyeok-ro 241beon-gil, Seongnam-si, Gyeonggi-do, South Korea

USA
1735 Technology Drive, Suite 740 San Jose, CA 95110

Taiwan
No. 66-10, Yucheng St., Nangang Dist., Taipei City 115012, Taiwan (R.O.C.)

Europe
Tel-Aviv Israel

Buy Now! - DX TechBridge Program
Contact Sales - sales@deepx.ai

DXNN® SDK

DXNN® (DEEPX Neural Network) SDK streamlines AI deployment on DEEPX NPUs by integrating version-aligned tools for compilation, optimization, simulation, and inference. For **efficient development**, it's offered as the **DX-AS (All Suite), a fully integrated and optimized package.**

How It Works: 4-Step AI Deployment with DXNN® SDK

Step 1

Prepare Your ONNX Model

Convert a pre-trained AI model from frameworks such as TensorFlow, PyTorch, or scikit-learn into the ONNX format, which enables model interoperability.

AI Model Traning

TensorFlow | TensorFlow Lite | PyTorch | Keras | XGBoost | Darknet

↓

User Must Convert to ONNX Before Input

Step 2

Convert Model for DX-NPU

Use the DX-COM compiler to convert the ONNX model into a format optimized for DEEPX NPU hardware, ensuring efficient execution on the target NPU.

User Must Convert to ONNX Before Input

↓

DX-COM™

↓

DEEPX AI Model

Step 3

Real-time Operating DX-RT

Deploy the optimized model using the DX-RT runtime API. Input data is fed into the model, and inference results are produced in real-time on-device.

DEEPX AI Model | Input Tensor | Output Tensor

↓ ↓ ↑

DX-RT™ API

↕

DEEPX AI Products

Step 4

Run Your Application

Apply NPU-generated AI inference results to applications like Image & Video Analytics, Object Detection, Robotics and Automation, and Smart Homes and Factories.

DEEPX AI Products Applications

↓

Image & Video Analytics | Object Detection

Robotics and Automation | Smart Homes and Factories

DXNN® Full Stack Architecture

DXNN® Full Stack Architecture streamlines AI model deployment onto DEEPX products using its two-stage AI Model Compile and Runtime Environments.

1) AI Model Compile Environment

AI Model Input Sources

TensorFlow | TensorFlow Lite | PyTorch | Keras | XGBoost

↓

ONNX User Model
User Must Convert to ONNX Before Input

↓

AI Model Compiler

DX-COM®	DX-TRON	DX-SIM
NPU Model Compiler	Graphical Model Viewer	Model Execution Simulator

↓

DX-Model Zoo
DEEPX Pre-compiled Models

or

Compiled User Model
Optimized Binary for DEEPX NPU

→

2) AI Model Runtime Environment

AI Model Input Sources

DX-APP — SDK Usage Examples | User App — Your Custom Application

AI App Dev Tool

DX-STREAM™

gstreamer

Ecosystem

aws AWS IoT Greengrass | WNDRVR | NetworkOptix

CVEDIA | VCA | Baidu 百度

DX-RT™

C++-based DX-RT™ API | Python-based DX-RT™ API

Core Runtime Library

Device Drivers (Windows or Linux)

↕ PCIe

DEEPX AI Products

DX-M1™	DX-M1M	DX-H1	DX-V3	DX-M2™
AI Accelerator	AI Accelerator	AI Accelerator	AI Vision SoC	GenAI Accelerator

DEEPX HQ
5F, 20 Pangyoyeok-ro 241beon-gil, Seongnam-si, Gyeonggi-do, South Korea

USA
1735 Technology Drive, Suite 740 San Jose, CA 95110

Taiwan
No. 66-10, Yucheng St., Nangang Dist., Taipei City 115012, Taiwan (R.O.C.)

Europe
Tel-Aviv Israel

Buy Now! - DX TechBridge Program

Contact Sales - sales@deepx.ai