Edge AI Solutions

Empowering AI at the Edge with Software and Services





www.advantech.com



Explore Our Solutions Built to Accelerate your Business

The range of use cases for Edge AI is highly diverse, necessitating varied AI computing capabilities and industrial designs. Advantech collaborates with multiple partners to develop edge AI solutions that integrate hardware and software, facilitating the accelerated deployment of AI across industries such as manufacturing, healthcare, retail, and smart cities. Our offerings harness CPU, VPU, or GPU technologies from eco-partners, delivering multifaceted AI solutions that efficiently scale AI applications while optimizing budget and resource utilization.



Comprehensive Offerings

Ranging from AI acceleration modules, inference systems to servers.



Industrial Grade

All products include 5 to 7 years longevity support with strict revision management.



Edge AI Design-in Services

Including advanced thermal solutions, embedded software service, and peripheral integration.



Full Software Support

Including Edge AI SDK, DeviceOn, Robotic Suite, and Ubuntu.

intel AMDA HAILO KINArA AXELERA

AI Acceleration Modules >>













Edge Al Inference Systems >>

AIR-000/100/300 Series 13~275 TOPS (INT8)

Diversity Drives Edge Al

Edge AI applications span a wide range of industries, including manufacturing, smart cities, retail, and healthcare. With the vast potential of AI across such a diverse spectrum, a variety of hardware is necessary to efficiently handle different AI workloads with distinct computing power optimizations and combinations.

Advantech offers comprehensive edge AI solutions, including AI acceleration modules, inference systems, and servers that utilize a variety of AI technologies from eco-partners. This provides customers with a broad selection to swiftly deploy AI at scale across the entire range of workloads, all while optimizing budget and resource utilization.



EAI Series » Al Acceleration Modules

Advantech offers the EAI series, which features low-power plug-in AI modules and GPU cards designed to expedite real-time inference on edge platforms. This enhances the scalability and affordability of AI applications at the edge.



Key Features



AI Acceleration at the Edge

Varied AI add-on solutions for different computing power and inferencing requirements



Superior Image Processing and Analysis

Optimize imaging quality and enhance visual processing flexibility to improve AI-assisted vision analysis



Quick Integration for Industrial Usage

Open standard form factors with optimized thermal solutions for a wide temperature range



Easy-to-use AI Development Toolkits

Including inference benchmark evaluation, runtime SDK integration, and consulting services



AIR-000/100/300 Series » Edge Al Inference Systems

Advantech AIR-000/100/300 series edge AI inference systems offer a variety of CPU options and integrate AI capabilities seamlessly. They feature a compact design and low power consumption, making them ideal for applications in autonomous robotics and automated optical inspection fields.

Key Features



Scalable Edge AI Performance

Comprehensive range of CPUs, including Intel and NVIDIA Jetson, with built-in AI capabilities



Ready-to-deploy AI System

Pre-installed with Ubuntu OS and software toolkits such as Edge AI SDK, DeviceOn, and Robotic Suite



Vertical I/O Support for Industrial Applications

 Rich I/Os enable connectivity for multi-function devices and video capture interfaces



Easy Installation for Harsh Environments

Compact and rugged design supports a wide range of operating temperatures and voltage inputs





Performance/ Watt

AIR-500 Series » Edge Al Servers

Advantech's AIR-500 series AI servers provide extreme AI power with high-performance CPUs and GPUs. They are designed for graphics-intensive workloads, on-premise training applications in medical, automation, and LLM training fields.

VIDIA CERTIFIED

Key Features



Unmatched Acceleration for Al, and Data Analytics

High-performance CPU with definable GPU configuration



AI-Ready Enterprise Platforms

Can be configured with licensed NVIDIA AI Enterprise & Intel Geti Platform



0

Robust and Reliable Designs

Effective thermal management and anti-vibration bracket design



NVIDIA Certified Edge AI Server Ensures GPU performance and AI SDK

Ensures GPU performance and AI SDK compatibility.



Edge Al Servers – 6

Data Bandwidth

End-to-End Edge Al Solutions

Versatile, Accelerated Deployment at Every Phase

The aim is to make Edge AI adoption easier, Advantech offers end-to-end software solutions to address Edge AI implementation challenges across different stages. Initially, it helps enterprises gauge computational needs through its Edge AI SDK, aiding in resource assessment like CPU vs. GPU usage. During model training, Advantech's integration with Intel[®] GETi[™] and NVIDIA AI Enterprise tools expedites development, providing value-added benefits. This integration extends Intel[®] GETi[™]'s capabilities to cover both model development and deployment, leveraging NVIDIA's ecosystem for further resources and collaboration. For deployment, DeviceOn's OTA Update and Container Management features streamline the process, enabling continuous data collection for model retraining, ensuring accuracy and validity. Aligning Software Solutions Across Phases of AI Implementation.



Versatile, Accelerated Deployment at Every PhaseToolkit

Streamlines the AI Implementation Process

Through comprehensive software deployment and integration with hardware design, service capabilities, and ecosystem partners. This saves time and reduces costs associated with AI development, training, and deployment, making Edge AI adoption more accessible for enterprises.

Efficient Resource Utilization

This approach optimizes computational resources like CPU and GPU usage for Al applications, minimizing waste and maximizing performance in Edge AI deployments. This leads to cost savings, improved productivity, and enhanced competitiveness.

Enhanced Scalability and Flexibility

Advantech's recommended solutions for each stage empower enterprises to adjust to evolving business demands and technological progressions. Through the integration of development toolkits from chip partners and service providers, Advantech delivers a flexible approach that caters to various customer needs.

0

Rapid AI SDK Development Toolkit

Creating an AI-ready development environment can often feel like a formidable challenge. However, Advantech Edge AI SDK, an exceptional toolkit designed to seamlessly support leading AI platforms and their associated inference SDKs, empowers developers to effortlessly harness the boundless potential of AI inference.

Inference Benchmark Tool

No-code GUI for rapid AI inference assessment

Inference Runtime SDK

Compatible installation on Advantech edge devices

TensorRT OpenVINO ROCm HailoRT elQ Neural Processing SDK NeuroPilot

Inference Deployment Platform

Edge

Managing AI containers remotely with scalable updates

Global Customer Support

Inference computing and edge-to-cloud consulting services

AloT Device Management and Edge Orchestration

Leveraging Advantech's 40 years of embedded expertise, DeviceOn, an edge orchestration platform, remotely manages over 10,000 AloT devices across x86/RISC, Windows/Linux/ Android, and in private or public clouds.

DeviceOn

Securely Automated Device Onboard

Auto Device Registration, TPM certification, PKI architecture , and Device Security, enables fast and secure connection establishment

Remote Device Monitor and Control

Power On/Off, device remote control with scheduling rules, protection management, and Windows Lockdown

Software OTA Updates & Container Mgmt.

Remotely deliver and update software, firmware, drivers, and configuration at scale.

Edge Integrated Security

Ransomware protection, application whitelisting, system backup, and bare-metal recovery

NVIDIA AI Enterprise Production-Grade AI Development Platform

The platform provides best-in-class development tools, frameworks, and pretrained models for AI practitioners, as well as reliable management and orchestration for IT professionals to ensure performance, high availability, and security.

Comprehensive AI Devt. Suite

- Data preparation and training
- · AI inference library and services
- ** Powered by NVIDIA RAPIDS, TAO Toolkit, TensorRT, Triton Inference Server

Lifecycle Management & Security Updates

- · 2 production branches per year with 9-month lifetime
- Monthly CVE patches and bug fixes

Expert Support and Services

- Advantech Tier 1 services for AI dev infra: OS, drivers, and AI SDKs
- · Direct access to Nvidia AI experts

Advanced Software Offerings

- Cross-regional infra: Microsoft Azure
- Remote deployment: Advantech DeviceOn
- IoT security: Trellix, IEC62443 aligned solutions

Annual Subscription

Standard support included

• 1-/3-/5-year subscription license

- Perpetual License
 - 5-year standard support included Free Trial
 Extendable support on a yearly basis

Notes

- Per GPU Licensing
- Optional Upgrade for Business Critical Support
- Discounted Pricing for Advantech NCS S/W Bundle

Embedded Ubuntu Services

Ubuntu Desktop is a secure OS for Edge AI development, offering extensive software repositories and support for popular programming languages and frameworks.

🗘 Ubuntu

User-Friendly Interface

Simplifies navigation for developers and streamlines task execution.

Seamless Integration with Cloud

Enables developers to easily deploy and scale edge AI applications.

Various Tool Options

Supports a wide range of programming languages and frameworks such as Python, TensorFlow, and PyTorch.

Extensive Software Repositories

Allows developers to easily install and manage Al development tools and libraries.

Intel GeTi -Computer Vision Al Platform

Intel's new software platform for building computer vision models in a fraction of the time and with less data. The platform eases laborious data labeling, model training and optimization tasks across the AI model development process, empowering teams to produce custom AI models at scale.

intel ceti

Smart Annotations

Expedite data annotation and easily segment images with professional drawing features like a pencil, polygon tool and OpenCV GrabCut.

Interactive Model Training

Build models faster as active learning selects the most informative data samples to achieve working models with fewer data

Starter, Annual License

Business, Annual License

- 1 name user, 1 instanceRecommended for POC
- Basic support
- 3 named users for small team, 1 instance
- 2 concurrent model trainings
 - Basic support

Hyperparameter Optimization

Refining hyperparameters is critical to the model's learning process. With built-in optimization, the Intel Geti platform makes a data scientist's job easier.

Production-Ready Models

Output deep learning models in TensorFlow or PyTorch formats or as an optimized model for OpenVINO toolkit to run on Intel[®] architecture CPUs, GPUs and VPUs.

Starter, Annual License

• 10 named users for improved, 1 instance

Basic support

• Up to 4 concurrent model trainings

30-Day Free Trial

- Drive proof of value
 - Incur no cost for using the software and generating models
 - Enables unlimited usage

Windows Embedded and Windows IoT

Windows IoT supports Edge AI development with various development tools and integrates popular AI frameworks. It seamlessly connects with Azure IoT services, providing an easy-to-use interface and interaction with AI models.

Windows IoT

Integration with Azure Services

Allows developers to easily install and manage Al development tools and libraries.

Wide-Range Development Tools and Frameworks

Supports popular AI development frameworks like TensorFlow and ONNX for deploying AI models on edge devices.

Easy Updates and Maintenance

Simplifies the process of updating and maintaining security patches, new features, and AI model updates

Scalability

Facilitates easy scaling of edge AI solutions from small to large deployments.

Success Stories



AI-Enhanced Diagnosis Workstation

Advanced ICT with AI-enhanced diagnostics boosts efficiency in clinical decision-making and medical operations. By allowing personnel to focus on clinical tasks requiring human judgment, the technology reduces error rates. The AIR-510, equipped with a powerful Intel processor and extendable graphics performance, provides performance, reliability, and expandability for medical visualization and data processing applications.

SOLUTIONS

NVIDIA-certified system ensures compatibility with graphic card performance, NVIDIA medical AI models, and NVAIE

BENEFITS

- 10 years of longevity and safety compliance to IEC 62368-1 (CB & UL Certified)
- · One-stop shopping for complete hardware systems, Windows IoT or Ubuntu operation systems, and NVAIE or Geti solution software licenses





Enhance Factory Operator Safety with AI Technology

Workplace safety is crucial for employee protection. Our customer sought a vision Al solution to enhance operational safety, efficiency, and productivity. Advantech and CyberLink implemented AI facial recognition for access control and securing restricted areas. Another ISV partner added PPE and idle time detection for improved safety and management

SOLUTIONS

Inference- AIR-020, Software- CyberLink FaceMe, OS-Ubuntu 20.04, SDK- NVIDIA JetPack

BENEFITS

- · Leverage solutions developed by ISV partners
- · Compact AI system designed for critical working environments
- · Ready OS and JetPack enable guick AI development



AI-Powered Surveillance System for Unmanned Self-Service Gas Stations

The Al-driven automatic refueling monitoring system addresses labor shortages at gas stations. The AIR-030, edge AI system with NVIDIA Jetson AGX Orin, monitors the refueling process from nozzle removal to return, using AI image recognition to allow refueling only when the posture is normal. Refueling is automatically halted if suspicious behavior is detected



SOLUTIONS

Inference- AIR-030 with NVIDIA Jetson AGX Orin, OS- Linux Ubuntu 20.04, SDK- Edge AI SDK

BENEFITS

- Ready-to-use AI inference system with Ubuntu 20.04 built-in and JetPack 5.1.2 SDK
- Al computing power analyzes imaging to monitor and control refueling
- Rich I/O supports high resolution camera connection and industrial protocol communication



Turkey Deploys AI-Assisted Traffic Surveillance and Analysis

Transportation plays a pivotal role in cultivating a content urban populace. Enhancing traffic efficiency to alleviate road congestion is essential for convenient transportation. Advantech's EAI-3100 graphics card and ISSD's VIERO-AI Analysis system provide real-time traffic monitoring. This AI-accelerated analysis tracks vehicle count, density, and speed to enhance urban traffic flow and safety.

SOLUTIONS

Inference-EAI-3100 Intel Arc A370M Graphics card, Software-ISSD VIERO-AI system, SDK-Intel OpenVINO toolkits

BENEFITS

- Ready to run sophisticated AI inference models, e.g., Yolov8.
- Inference latency reduced by 50% compared to previous Movidius solution.
- Construction fee costs reduced by over 33%



Success Stories — 12



HAILO



Model		AIR-030	AIR-150	AIR-310
CPU/Platform		NVIDIA Jetson AGX Orin 32G/64G	13th Gen Intel® Core™ i3/i5 processor	14th Gen Intel [®] Core™ i3/i5/i7/i9 processor
Al Performance		up to 275 TOPS	Bundled with Hailo-8 AI module, up to 26 TOPS	Compatible with Intel Arc A370M/ Quadro® A2000, up to 60W
Memory	Technology	LPDDR5	DDR5 5200 MHz	2 x DDR5 5600MHz
	Max. Capacity	32GB/64GB	Up to 64 GB	Up to 64 GB
Display		1 x HDMI 2.0, 3840 x 2160@60Hz	2x HDMI 2.0, 4096x2160@60Hz	1 x HDMI 2.0, 4096x2160@60Hz 1 x DP++, DP 1.4a, up to 4096 x 2304@60 Hz
Ethernet	Speed/Controller	3 x 2.5GbE, Intel I225-LM	1x GbE, Intel I219-LM 1 x 2.5 GbE, Intel I226-LM	1 x GbE, Intel I219-LM 2 x 2.5 GbE, Intel I226-LM
	PoE	LAN1 & 2 optional, by adding MIOe-PSE	-	-
1/0 Ports	USB 3.0 / USB 2.0	4 x USB 3.2 2 x USB 2.0 (internal)	3 x USB 3.2 1 x USB 2.0	4 x USB 3.2
	COM	4 x RS232/422/485	2 x RS-232/422/485 1 x RS-485	2 x RS232/422/485
., 0 . 0.10	DIO	16-bit	8-bit	16-bit
	CANBus	2	2	2
	Audio	Line-out	Line-out/Mic-in (switch)	Line-out/Mic-in (switch)
	eMMC	64GB	-	-
Storogo	2.5" SATA/ SATA Slim	-	1x SATA Slim	1x 2.5" SATA
Storage	M.2	1 x M.2 B-Key 2280/3052 (PCIe x2, USB 3.0)	1 x M.2 M-Key 2280 (PCIe Gen4 x4, SATA)	1 x M.2 M-Key 2280 (PCIe Gen3 x4, SATA)
	SD Card	1 x SD 3.0 slot	-	-
Expansion	M.2	1 x M.2 E-Key 2230	1 x M.2 E-Key 2230 1 x M.2 B-Key 3042 (default w/ Hailo module)	1 x M.2 E-Key 2230
	PCI Express	1 x PCIe x16 (Optional, signal: PCIe x 8)	-	-
	GPU Card	-	-	1 x MXM3.1 Type A up to 60W
	MIPI	2	-	-
Others	Trusted Platform Module	on-board TPM 2.0	on-board TPM 2.0	on-board TPM 2.0
Power	Power Input	9-36V	12-24V	12-24V
Operating System		Built-in Linux Ubuntu 20.04 (JetPack 5.1.2)	Windows 11/10 IoT/Ubuntu 22.04	Windows 11/10 IoT/Ubuntu 22.04
Operating temp. (with 0.7 m/s air flow)		-10~55°C (MODE_50W)	-20~60°C	-20~55°C (w/o MXM GPU) 0~50°C (w/ MXM GPU)
Dimensions (W x H x D)		200 x 220 x 74 mm	156 x 112 x 60 mm	215 x 220 x 55 mm

Note: "-" : means Not Applicable (N/A)

AI Acceleration Modules & GPU Cards



Edge Al Se	rvers			
	Model	AIR-510	AIR-520	AIR-530
CPU/Platform		14th Gen Intel® Core™ i3/i5/i7/i9 processor	AMD EPYC 7003 series	Arm Cortex-A78AE / IGX Orin Platform
Al Performance		NVIDIA Certified with RTX 6000 Ada (Graphics card up-to 350W)	NVIDIA Certified with 2x RTX A5000 (Graphics card up-to total 700W)	NVIDIA Certified with RTX 6000 Ada (Planning)
Memory	Technology	DDR5 5600 MHz up to 192GB (48GB per DIMM)	DDR4 3200MHz up to 768GB (128GB per DIMM)	LPDDR5 4300 MHz, 64GB
	Socket	4 x UDIMM	6 x RDIMM	on-board
Display		1 x HDMI 2.0, 4096x2160@60Hz 2 x DP 1.2,up to 4096 x 2304 @ 60Hz	1 x VGA via Aspeed AST2500 BMC	1 x DP 1.4a
Ethernet	Speed/Controller	1 x GbE, Intel I219-LM 1 x 2.5 GbE, Intel I226V	1 x Gbe from BMC 2 x 2.5 GbE, Intel I226V 2 x 10 GbE, Intel X550	2 x GbE from BMC or Orin 2 x 100 GbE QSFP28 from ConnectX-7
I/O Ports	USB 3.0 / USB 2.0	7 x USB 3.2 4 x USB 2.0	3 x USB 3.2	4 x USB 3.2
	СОМ	6 (5 x RS-232, 1 x RS-232/422/485 with auto flow control), optional	1 x RS-232	4 (TBD)
	DIO	1x 8-bit DIO	1x 8-bit DIO	1 (TBD)
	eMMC	-	-	64GB
Storage	2.5" SATA	8 x 2.5" SSD (4 x 2.5"+ 4 x 3.5" by project)	6 x 2.5" SSD (4 x 2.5"+ 4 x 3.5" by project)	4 x 2.5" SSD (4 x 3.5" by project)
	M.2	1 x M-Key 2280 (PCIe Gen4 x4)	1 x M-Key 2280 (PCIe Gen4 x4)	1 x M-Key 2280 (PCIe Gen4 x4)
Expansions	M.2	-	-	1 x B-Key 2280 (PCIe x1 from BMC) 1 x M.2 E-Key 2230
	PCI Express	1 x PCle x4 (Gen3) 1 x PCle x4 (Gen4) 1 x PCle x16 (Gen5)	4 x PCle x16(Gen4)	1 x PCle x8 (Gen4) 1 x PCle x16 (Gen4)
Others	IPMI	-	Aspeed AST2500 BMC IPMI 2.0 with virtual media over LAN and KVM-over-LAN support	Aspeed AST2600 BMC
Power	Power Input	100-240V AC	100-240V AC	100-240V AC
Operating System		Windows 10 IoT / Ubuntu 22.04	Windows Server 2019 / Ubuntu 22.04	IGX SW (OS & NIVIDA AI stack)
Operating temp. (with 0.7 m/s air flow)		0 ~ 40°C	0 ~ 40°C	0 ~ 40°C
Dimensions (W x H x D)		380 x 176 x 467 mm	380 x 176 x 467 mm	380 x 176 x 467 mm

Note: "-" : means Not Applicable (N/A)







Model	EAI-2100	EAI-3100	EAI-3101		
GPU	Intel [®] Arc [™] A370M	Intel [®] Arc [™] A370M	Intel [®] Arc [™] A380E		
Xe-cores	8	8	8		
Intel® XMX Engines	128	128	128		
Graphics Clock	1550 MHz	1550 MHz	2000 MHz		
FP32 Performance	4.198 TFLOPS	4.198 TFLOPS	5.018 TFLOPS		
Memory Capacity	GDDR6 4GB	GDDR6 4GB	GDDR6 6GB		
Memory Speed	14 Gbps	14 Gbps	15.5 Gbps		
Signal Interface	MXM 3.1 PCI Express 4.0 x8 with 3.0 Backwards Compatibility	PCI Express 4.0 x16, PCIe x8 signal	PCI Express 4.0 x16, PCIe x8 signal		
Dimensions (L x H x W)	82 x 70 mm	168 x 110 x 34.8 mm, Dual slot	152 x 110 x 34.8 mm, Dual slot		
Form Factor	MXM Type A 3.1	PCI Express 4.0 x16	PCI Express 4.0 x16		
Power Consumption	Up to 40W	Up to 60W	Up to 85W		
Thermal Solution	Fansink (optional), Heat spreader (optional)	Single active fan	Single active fan		
Operating Temperature	0~60°C (depends on thermal solution)	0~60°C (depends on thermal solution)	0~60°C (depends on thermal solution)		
Power Connector	Reserved 4-pin power connector for fansink	8-pin power input (optional)	8-pin power input		
Power Range	9-20V input (Fan control: 12V only)	DC 12V input	DC 12V input		
Display I/O	Reserved 4 x DP 1.4a	2 x DP 1.4a 2 x HDMI 2.0b	4 x DP 1.4a		
DP Max. Resolution	7680 x 4320@60Hz	7680 x 4320@60Hz	7680 x 4320@60Hz		
HDMI Max. Resolution	-	4096 x 2160@60Hz	-		
Supported Graphics Engine	DirectX12, OpenGL 4.6, OpenCL 3.0 HW Encoding: H.264/H.265(HEVC)/AV1 HW Decoding: H.264/H.265(HEVC)/AV1/VP9				
Operating system	Windows 11, 10 64-bit Ubuntu 22.04 LTS				