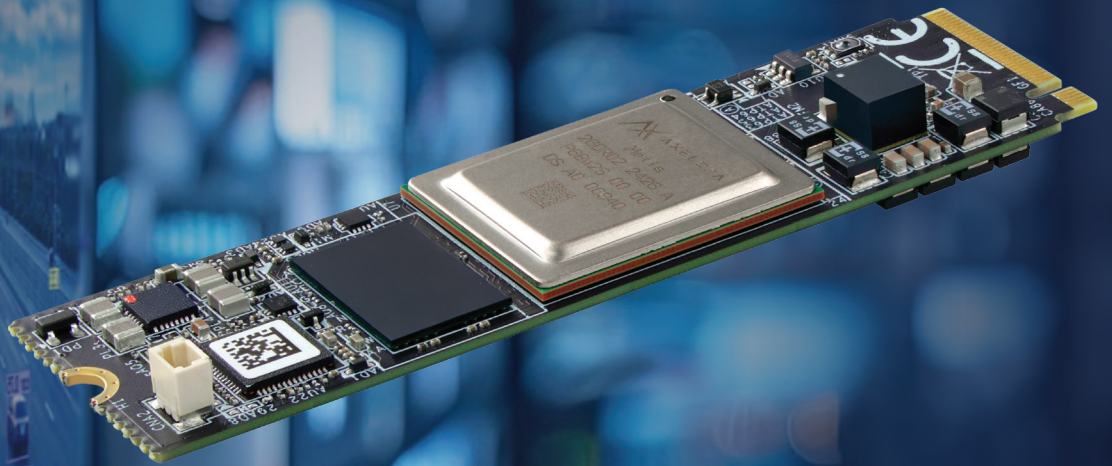


Axelera Embedded 110m

Optimized for constrained environments.
Engineered for embedded systems.



Security



Industry 4.0



Retail



Mobility



Logistics



Robotics



Medical



Hospitality



Utilities



Agritech



Drones



& More



Axelera® Embedded 110m card, powered by Metis AIPU, is perfect for constrained devices that require high performance AI inference.

Key Features

- High performance **Axelera Embedded 110m card** feature **Metis AIPU** to enable state-of-the-art AI inference in small footprint devices.
- A single board can run inference on multiple cameras as well as support many independent parallel neural networks.
- A wide range of end-to-end AI pipelines and models are available out of the box.
- Hassle free evaluation and software integration thanks to **Voyager® SDK**.
- Uncompromised prediction accuracy thanks to advanced quantization tools.

Key Technical Specifications

Form Factor M.2 2280 M-key

Host Interface PCIe Gen3 x4 – 4 GB/s bidirectional

AIPU 1x Metis AIPU

AIPU Memory 1 GB DRAM

Peak INT8 TOPS 214

Operating Temperature¹ -20 to +70°C

Thermal Solution Optional standalone active cooling

Typical Application Power 3.5-9 W

Security Features Secure Boot, Root of Trust

1. Extended temperature variant [-40 to +85°C] available upon request.

Proven in Key Markets

Companies in multiple market segments have already adopted Axelera Embedded-based AI acceleration for different applications such as:



Surveillance: provide multi-camera AI processing capabilities in small form factors for easy deployment in shops, building sites, factories, venues.



Robotics: enable advanced cooperation between human and robots thanks to high resolution advanced gesture recognition, pose estimation and human intention decoding.



Medical: improve accuracy of real time diagnostics, monitoring, imaging and analysis.



Agritech: enhance precision in agricultural practices with crop analysis, smart irrigation/pest control, automated harvesting.

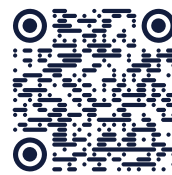


Drones: deliver real-time video processing for navigation, object detection and surveillance in lightweight, battery-powered UAV systems.

Performance Benchmarks

See how Axelera Embedded 110m card compares against other AI accelerators across real-world neural networks.

View the latest results:



axelera.ai/axelera-products/embedded#110m

Axelera® technology integrates seamlessly with host CPUs based on both x86 and ARM architectures. Our team actively tests different systems from vendors making it easy for embedded developers to prototype AI applications.

VOYAGER TOOLCHAIN

Voyager Wingman Builds it for You

- **Describe the pipeline, get a running code.** Voyager Wingman builds the pipeline, compiles it, runs it on the card and benchmarks the result, drawing on curated set of workflows, tools and knowledge for Axelera hardware.
- **Need to port your existing project?** Voyager Wingman adapts your existing inference stack, including for non-Axelera hardware, and builds your Axelera pipeline, often from a single prompt.

Voyager SDK Gives You Direct Control

- **Fast end-to-end integration.** A high-level pipeline framework builds optimized AI applications with pre-optimized models, datasets, models and business logic in very few lines of code.
- **Measure before you commit.** Built-in tools and turnkey Model Zoo networks let you evaluate performance, accuracy and power consumption on the card in minutes.
- **Direct hardware control.** For teams running their own pipelines, low-level APIs give direct access to the inference hardware.

A Simple, Feature-Rich SDK

160+ pre-optimized models across 16 task families, with 24 end-to-end reference pipelines as templates.

- **Detection** Ultralytics YOLO, YOLOX, YOLO-NAS, SSD-MobileNet, RetinaFace
- **Classification** ResNet, EfficientNet, MobileNet, DenseNet, RegNet, SqueezeNet, Inception, ResNeXt
- **Segmentation and Pose** Ultralytics YOLO for instance, semantic, pose and oriented bounding boxes, plus U-Net
- **Language models** Phi-3-mini, Llama 3.2, Llama 3.1
- **Specialist** Re-identification, depth estimation, license plate recognition, super-resolution, face recognition

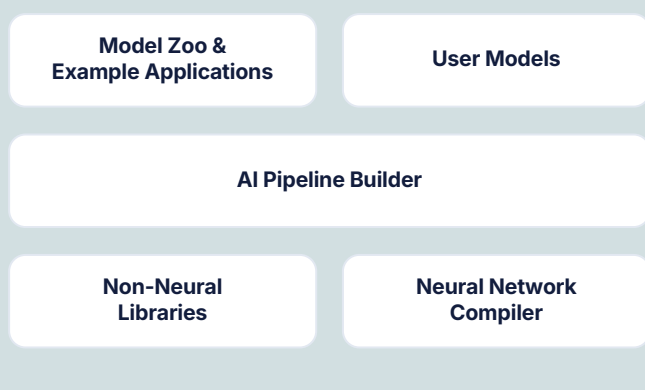
Bring Your Own Models

Pair your own weights with a standard architecture and deploy them through the same path every Model Zoo model uses. Voyager Wingman works alongside it, turning a plain-language description into the compile, run and validate loop on your card.

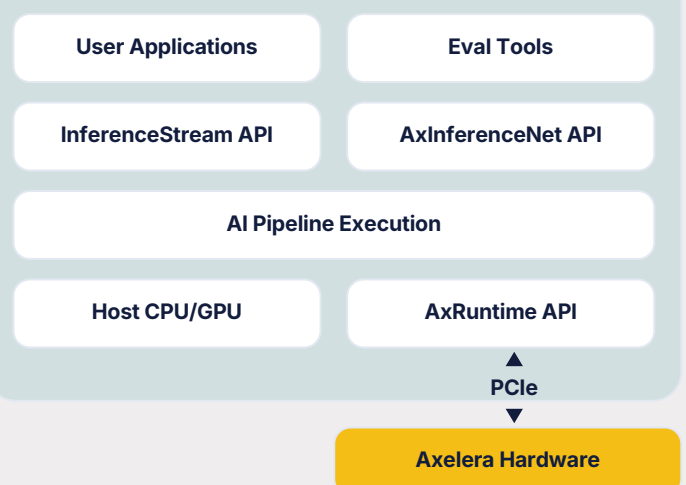
Compiler and Tooling

PyTorch and ONNX models compile with automatic quantization and graph optimization. Full pre- and post-processing libraries included, with accuracy and performance evaluation built in.

Develop



Run



To order the Axelera Embedded 110m, visit store.axelera.ai/products. Part numbers are listed in the product datasheet.