SmartFly info CRL-G116U-P3DF

SmartFly info PCIe Gen3 AI Accelerator Card (CRL-G116U-P3DF) User Manual

Model: CRL-G116U-P3DF | Brand: SmartFly info

1. Introduction

This manual provides essential information for the installation, operation, and maintenance of your SmartFly info PCIe Gen3 AI Accelerator Card, model CRL-G116U-P3DF. This card is designed for Edge AI Inference, leveraging Google Coral Edge TPU M.2 modules to deliver powerful AI inference capabilities. It is compatible with standard PCI Express Gen 3 x16 slots and features an optimized thermal design for stable performance under high loads.



Figure 1: Front view of the AI Accelerator Card, showcasing its dual cooling fans and PCIe interface.



> *Worldwide Spending on Artificial Intelligence Systems Will Be Nearly **$98 Billion** in 2023, According to New IDC Spending Guide*

> *AI to create **$13 TRILLION** additional economic activity by 2030. Machines to impact jobs – in a good way **58M NEW JOBS** by 2022, According to Tractica*

ASUS IoT AI Accelerator Card

ASUS IoT AI accelerator CRL-G18U-P3DF is a full height, half length, PCIe Gen3 AI acceleration card based on Google Coral Edge TPU processor. It allows users to build intelligent devices that are efficient, private and ideal for edge AI inference.

Powerful AI Inference Capability

- Support up to 8 x Google Edge TPU M.2 modules

Easy-to-Use Pre-trained AI Models

- Google TensorFlow Lite pre-trained ML models can be easily compiled and run on this model

Easy Installation, Common Expansion Slot

- Compatible general PCI Express Gen 3 x16 slot

Stable At High-Loading

- Optimized thermal design with high quality Copper heatsink and twin turbofans



Product Overview – CRL-G18U-P3DF

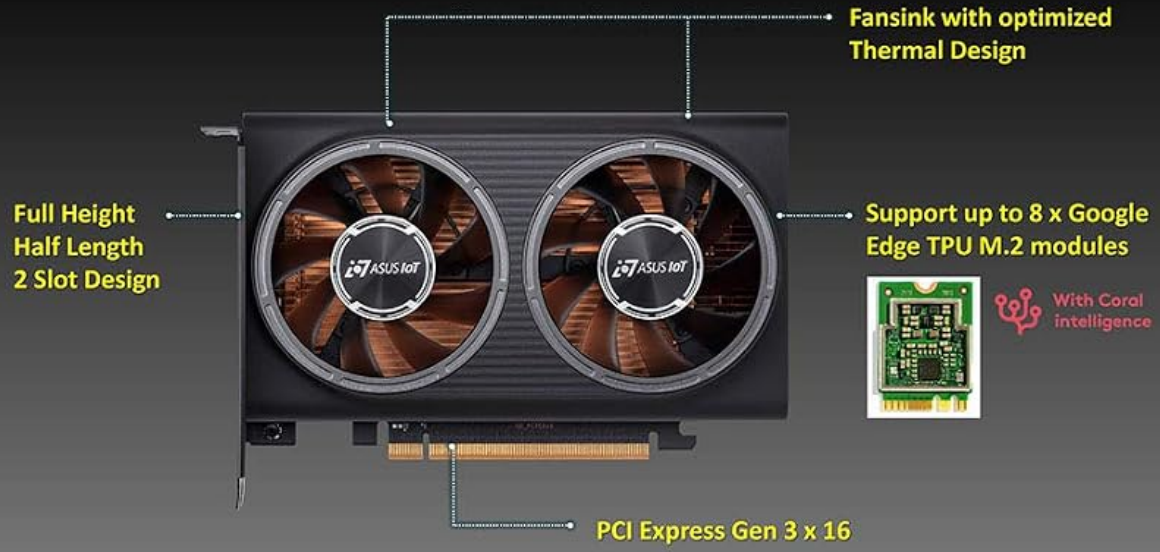


Figure 2: Product overview detailing the card's features, including support for Google Edge TPU modules and its robust thermal solution.

2. Safety Information

Please read and follow these safety guidelines to prevent damage to the device or injury to yourself:

- Always disconnect power from your computer before installing or removing any internal components.
- Handle the card by its edges to avoid touching sensitive components.
- Ensure proper grounding to prevent electrostatic discharge (ESD). Use an anti-static wrist strap if available.
- Install the card in a well-ventilated computer case to prevent overheating.
- Do not expose the card to moisture or extreme temperatures.
- Keep out of reach of children.

3. Package Contents

Verify that all components are present in your package:

- 1 x AI Accelerator PCIe Card (Model CRL-G18U-P3DF, which is the base card that supports additional Edge TPU modules, up to 16x for the CRL-G116U-P3DF configuration).
- Documentation (this user manual).

4. Specifications

The following table details the technical specifications of the SmartFly info PCIe Gen3 AI Accelerator Card:

Feature	Description
Model Name	PCIe Gen3 AI Accelerator (CRL-G116U-P3DF)
Graphics Coprocessor	Google Coral Edge TPU
Brand	SmartFly info
Video Output Interface	Multiple (HDMI, DisplayPort, DVI or similar) - *Note: This refers to potential system integration, the card itself is an accelerator.*
Graphics Processor Manufacturer	Google
Video RAM Type	Unknown
Recommended Product Applications	Professional
Included Components	AI Accelerator PCIe Card with 8 x Google Edge TPU (CRL-G18U-P3DF)
Compatible Devices	Devices with PCIe Gen3 slot
Graphics Card Interface	PCI-Express x16
Power Connector	1 x 6-pin 12V External Power

Feature	Description
Operating System Support	Ubuntu 18.04, Debian 10
Operating Temperature	-0°C to 55°C
Non-Operating Temperature	-40°C to 85°C
Relative Humidity	0% to 95%
Dimensions (Width x Height x Depth)	42.1 mm x 186.3 mm x 186.3 mm
Weight (CRL-G116U-P3DF)	516 g

AI ACCELERATOR CARD
DATASHEET

CRL-G18U-P3DF / CRL-G116U-P3DF

Full height, half length, PCIe Gen3 AI accelerator card based on Google® Coral Edge TPU processor, enabling AI-based real-time decision process at edge



Features

- Support up to 8 & 16 x Google® Coral Edge TPU M.2 modules
- Support TensorFlow Lite machine learning framework
- Compatible with PCI Express 3.0 x16 expansion slot
- Optimized thermal design with twin tuborfans

With Coral intelligence TensorFlowLite

Specifications

Main Chip	Core	Google® Coral Edge TPU Processor
PCIe Interface	Technology	PCI Express 3.0 x16
Software	Supported Framework	TensorFlow Lite
	Precision	INT8
	Performance	CRL-G18U-P3DF: 32 TOPS CRL-G116U-P3DF: 64 TOPS
Thermal Solution	FAN design	Active Fan
Power	Power Connector	1 x 6-pin 12V External Power
	Power Consumption	CRL-G18U-P3DF: 36 W CRL-G116U-P3DF: 52 W
Operating System	Linux	Ubuntu 18.04, Debian 10 *For latest OS support list, please check https://iot.asus.com/
Environment	Operating Temperature	0~55° C
	Non-Operating Temperature	-40~85° C
	Relative Humidity	0%~85%
Dimension	Width	42.1 mm
	Height	126.3 mm
	Depth	186.3 mm
Weight	Weight	CRL-G18U-P3DF: 516 g CRL-G116U-P3DF: 530 g

Block Diagram

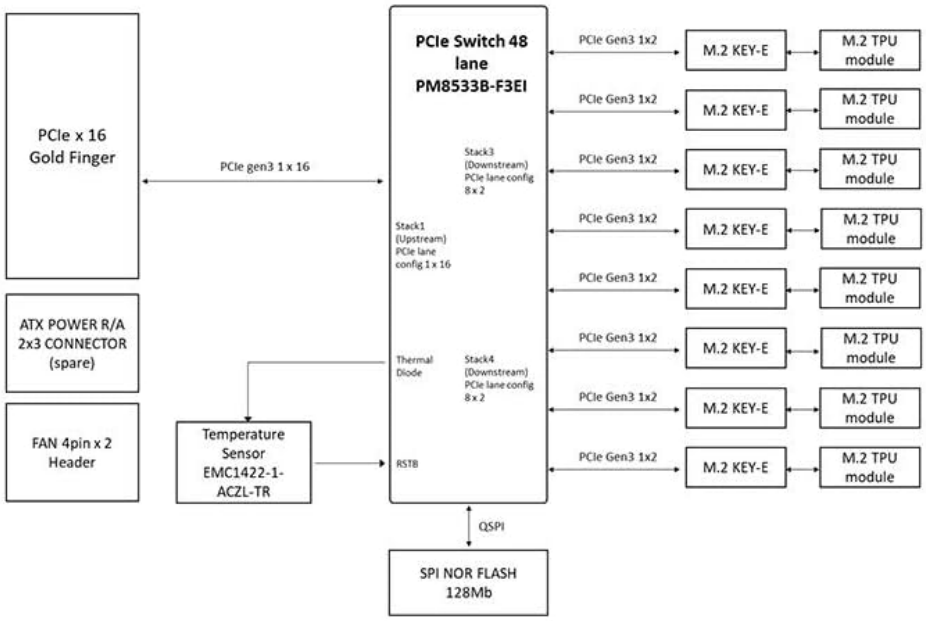


Figure 3: Official datasheet providing detailed specifications and a block diagram of the AI Accelerator Card.

Detail Specification



ubuntu® Debian TensorFlowLite

AI With Coral intelligence

Main Chip	NPU	Google® Edge TPU Processor
PCIe Interface	Technology	PCI Express 3.0 x16
Software	Supported Framework	TensorFlow lite
	Precision	INT8
	Performance	INT8, 32TOPS
Thermal Solution	FAN design	Active Fan
Power	Power Connector	1 x 6-pin 12V External Power
	Power Consumption	36W (Default)
Operating System	Linux	Ubuntu 18.04, Debian 10 *For latest OS support list, please check https://iot.ASUS IoT.com/products/AI-accelerator/AI-Accelerator-PCIe-Card/
Environment	Operating Temperature	0~55° C
	Non-Operating Temperature	-40~85° C
	Relative Humidity	0%~85%
Dimension	Width	42.1mm
	Height	126.3mm
	Depth	186.3mm
Weight	Weight	516g

Advance Thermal Design

Perfect combination for the powerful plug-and-play experience



- Optimized thermal design with twin Axial-tech turbofans which propel cool air over a Copper heatsink in order to dispatch the most heat from the cores, guarantee the best performance at high workload.
- One-piece cover effectively solidify the structure of the card

Comprehensive Google Coral Tool Set

Simplify your programming, speed up time to market

Tool	Purpose
Edge TPU Runtime on Mac OS/Windows	Edge TPU runtime for macOS and Windows, also the USB and PCIe drivers required on Windows
Edge TPU Runtime on Linux	All Coral software tools and libraries as Debian packages
PyCoral API	Python API for Python programmer
Libcoral API	C++ API for C++ programmer
Pre-compiled Models	Pre-compiled ML models from Google
Edge TPU Compiler	Compiles TensorFlow Lite ML model for Edge TPU
Profiled-based Partitioner	Segments a big ML model for multiple Edge TPUs

Figure 4: Detailed specifications and an illustration of the advanced thermal design, including twin Axial-tech turbofans and a copper heatsink.

5. Setup

5.1. Hardware Installation

1. **Prepare your system:** Power off your computer and disconnect the power cable. Open the computer case.
2. **Locate a PCIe slot:** Identify an available PCI Express Gen 3 x16 slot on your motherboard.
3. **Insert the card:** Carefully align the AI Accelerator Card with the PCIe slot and press down firmly until it is securely seated. Ensure the retention clip (if present) locks into place.
4. **Connect power:** Connect a 6-pin 12V external power cable from your power supply unit (PSU) to the power connector on the AI Accelerator Card.
5. **Secure the card:** Fasten the card to the computer case with the appropriate screw.
6. **Close the case:** Replace the computer case cover and reconnect the power cable.

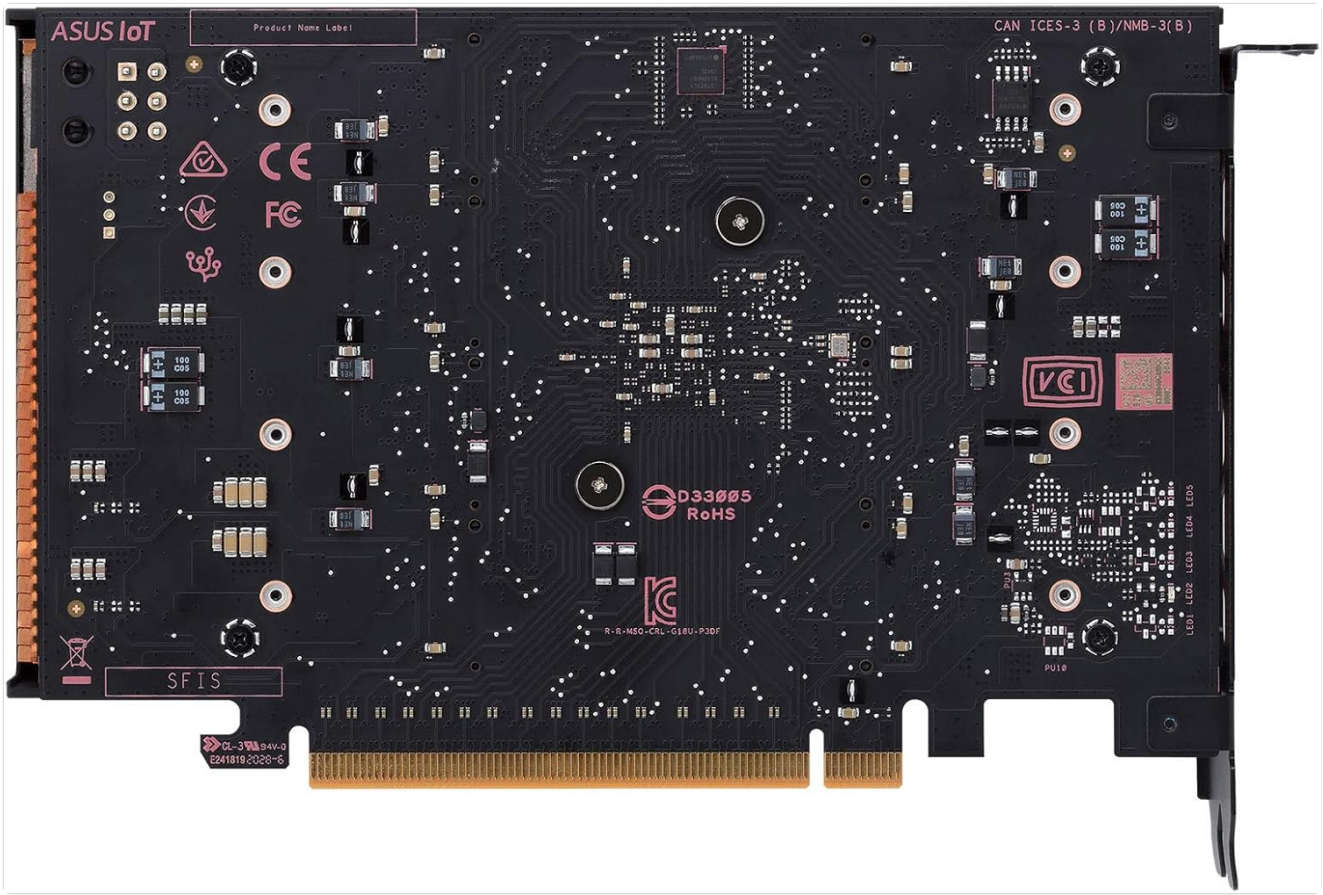


Figure 5: Rear view of the AI Accelerator Card, illustrating the PCIe connector for proper installation into a motherboard slot.

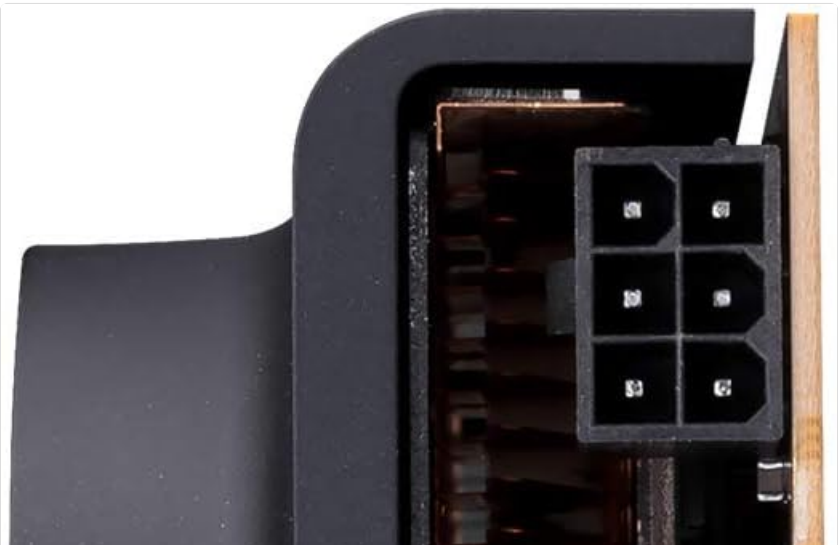




Figure 6: Side view of the AI Accelerator Card, showing the 6-pin external power connector that must be connected to the PSU.

5.2. Software Installation

The AI Accelerator Card requires specific software to function correctly. It is compatible with Linux-based operating systems.

1. **Operating System:** Ensure your system is running a supported operating system such as Ubuntu 18.04 or Debian 10. For the latest OS support list, please refer to the official ASUS IoT website.
2. **TensorFlow Lite:** The card is designed to run pre-trained ML models from Google TensorFlow Lite. Install the necessary TensorFlow Lite runtime and development tools.
3. **Google Coral Tool Set:** Utilize the comprehensive Google Coral Tool Set for compiling and deploying models. This includes tools for Edge TPU runtime, Python API, and model compilers.

6. Operating Instructions

The SmartFly info PCIe Gen3 AI Accelerator Card is optimized for Edge AI inference. Once installed and configured with the appropriate software, you can deploy and run your machine learning models.

- **Model Deployment:** Compile your pre-trained ML models (e.g., from Google TensorFlow Lite) for the Edge TPU using the provided tools.
- **Application Integration:** Integrate the compiled models into your applications. The card supports various frameworks and can be used with visual programming tools like ASUS IoT Node-RED AI Nodes for wiring AI applications.
- **Inference Execution:** Execute inference tasks on the card. The Edge TPU is designed for high-efficiency, low-power inference at the edge.
- **Monitoring:** Monitor the performance of your AI applications, including inference time, CPU, and RAM usage, through relevant system tools or dedicated dashboards.

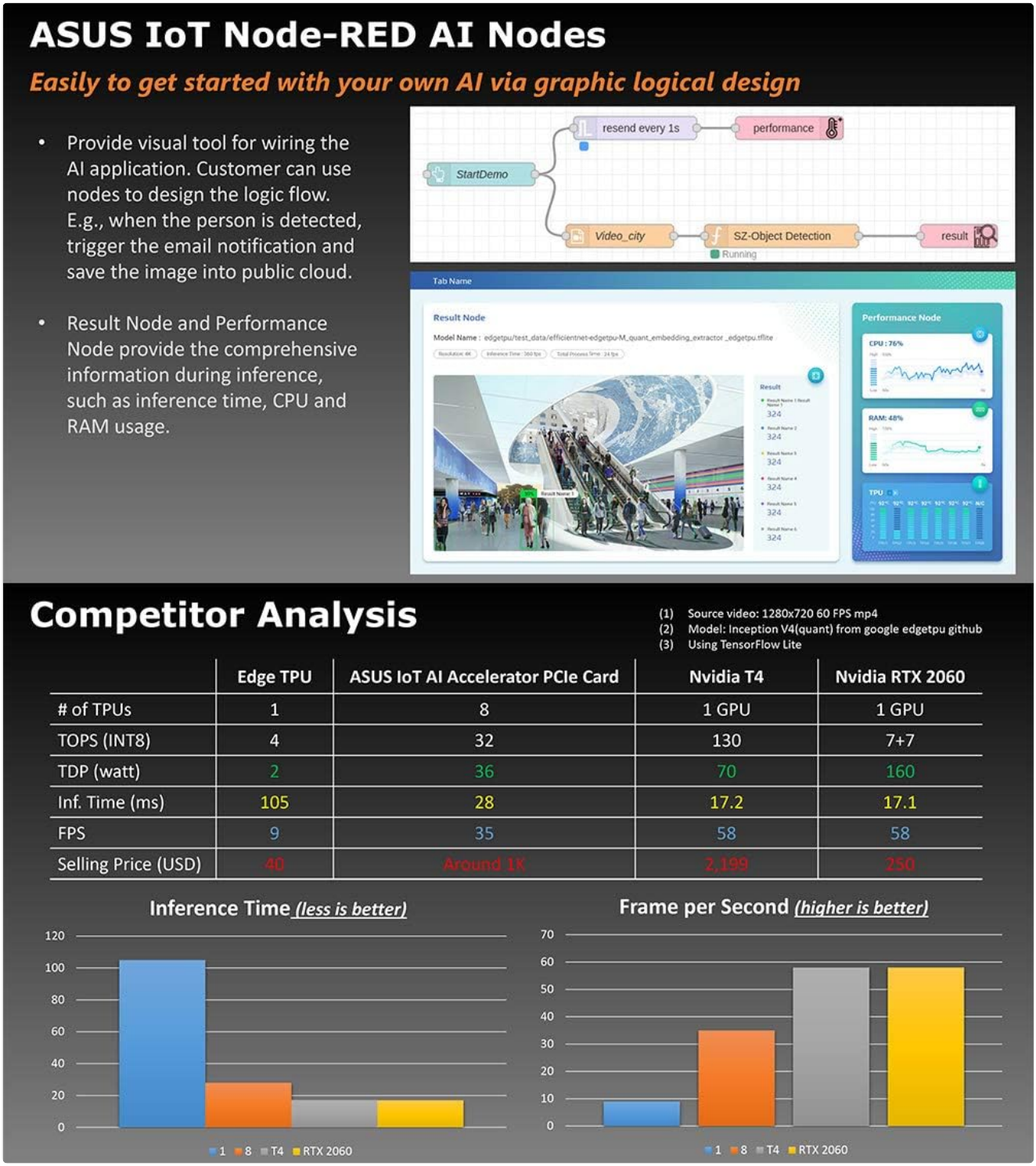


Figure 7: An example of using ASUS IoT Node-RED AI Nodes for visual design of AI applications and monitoring inference results and performance metrics.

7. Maintenance

To ensure the longevity and optimal performance of your AI Accelerator Card, follow these maintenance guidelines:

- **Cleaning:** Periodically clean the card and its fans to prevent dust buildup, which can impede airflow and lead to overheating. Use compressed air to gently remove dust. Ensure the system is powered off and unplugged before cleaning.
- **Ventilation:** Ensure your computer case has adequate airflow. Do not obstruct the card's fans or heatsink.
- **Temperature Monitoring:** Monitor the operating temperature of your system, especially under heavy AI inference loads, to ensure it remains within the specified operating temperature range (0°C to 55°C).

8. Troubleshooting

If you encounter issues with your AI Accelerator Card, consider the following basic troubleshooting steps:

- **No Detection:** Ensure the card is fully seated in the PCIe slot and the 6-pin power connector is securely attached. Check your system's BIOS/UEFI settings to confirm PCIe slot detection.
- **Performance Issues:** Verify that the correct drivers and TensorFlow Lite runtime are installed. Check for any background processes consuming excessive system resources. Ensure adequate cooling.
- **Software Errors:** Consult the documentation for TensorFlow Lite and Google Coral tools for specific error messages. Ensure your operating system is up-to-date.
- **System Instability:** If your system crashes or becomes unstable, ensure your power supply unit (PSU) can provide sufficient power to all components, including the AI Accelerator Card.

If problems persist, refer to the support section for further assistance.

9. Warranty Information

Warranty information for the SmartFly info PCIe Gen3 AI Accelerator Card is not explicitly specified in the provided product details. Please refer to your purchase documentation or contact the seller/manufacturer directly for warranty terms and conditions.

10. Support

For technical assistance, additional resources, or to report issues, please use the following contact information:

- **Wiki/Documentation:** For comprehensive guides and information, visit the official product wiki: https://youyeetoo.com/blog/ai-accelerator-pcie-card-94
- **Technical Support Email:** For technical issues after purchase, you can contact the support team at: monica@smartfire.cn
- **Seller Contact:** Alternatively, you can click on "SmartFly Tech" on the product page and ask a question directly through the Amazon platform.

Ask a question about this manual

Ask about setup, troubleshooting, compatibility, parts, safety, or missing instructions. Manuals+ will review the question and use this page's manual context to help answer it.

Name

Anonymous

Email

you@example.com

After a successful submission, we securely hash the supplied account or form email before sending it to Microsoft Advertising for conversion attribution. Privacy details.

Question

Example: How do I reset this device or fix this error?

Details

Model number, symptoms, what you tried, or the section of the manual you are using.

Submit question