ENHANCING ENDPOINT INTELLIGENCE
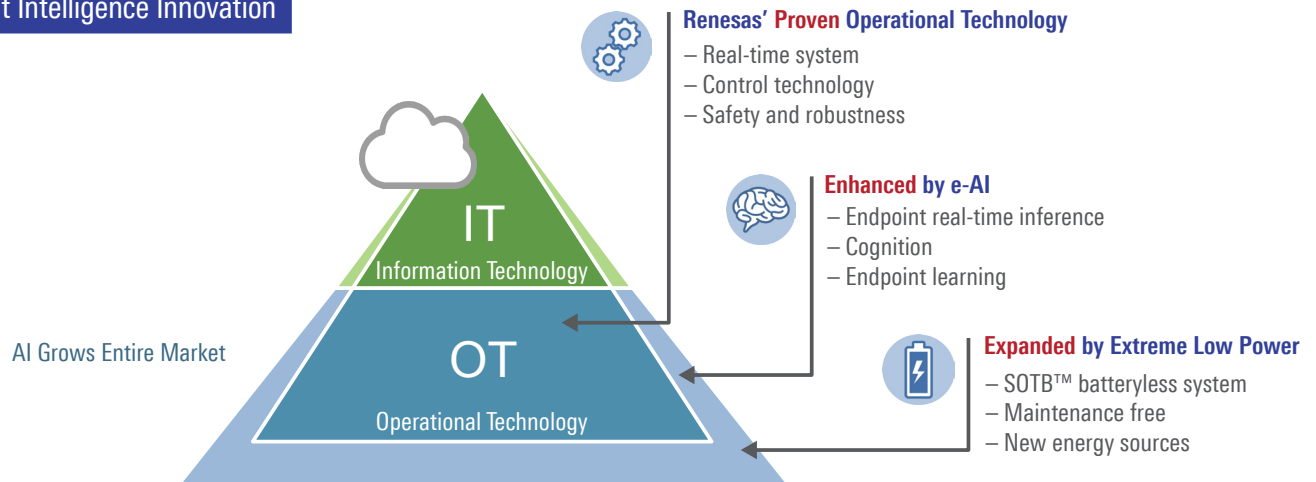
With Embedded Artificial Intelligence (e-AI) from Renesas
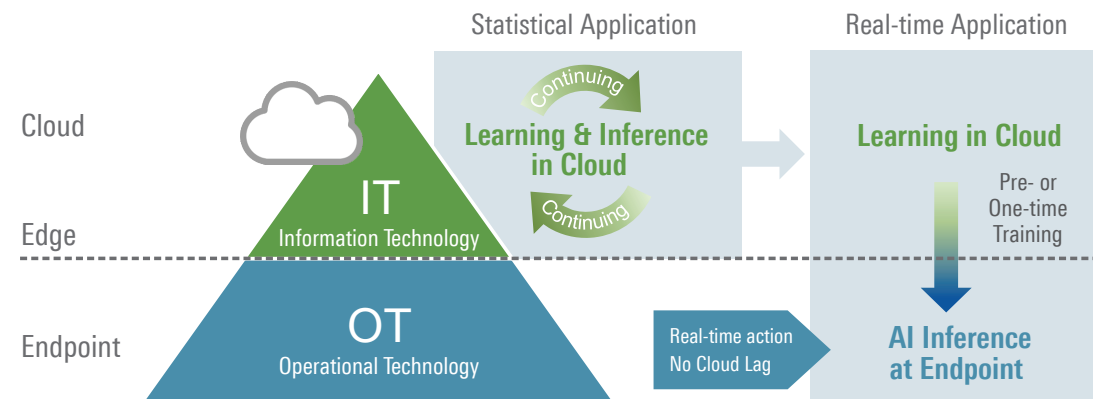
Real-time Intelligence without Cloud Lag

Artificial Intelligence is rapidly driving growth in the information technology (IT) and operational technology (OT) domains. For years, Renesas has been a leader in OT endpoint applications with microprocessor and microcontroller solutions. Leveraging that experience, Renesas' e-AI solutions are enhancing OT-based systems and products that we use around us every day by placing AI where it matters the most — at the endpoint — while decoupling dependency on the Cloud for real-time decisions and real-time action. Additionally, Renesas will expand e-AI application possibilities with the use of its exclusive extreme low-power process technology, Silicon On Thin Buried Oxide or SOTB™, to enable batteryless solutions powered only by harvested ambient energy. Think of the possibilities.

Endpoint Intelligence Innovation

Renesas' Proven Operational Technology
— Real-time system
— Control technology
— Safety and robustness

Enhanced by e-AI
— Endpoint real-time inference
— Cognition
— Endpoint learning

Expanded by Extreme Low Power
— SOTB™ batteryless system
— Maintenance free
— New energy sources

IT
Information Technology

OT
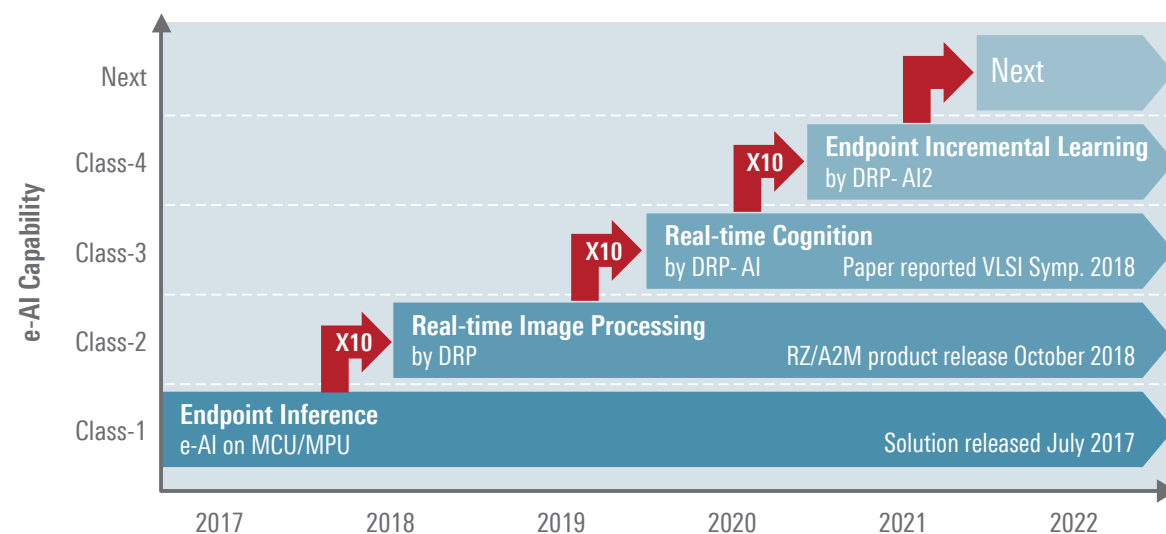Operational Technology

AI Grows Entire Market

e-AI: Local Real-time AI by Inference

- Traditional statistical AI applications execute completely in the Cloud
- Real-time applications cannot tolerate cloud lag at the endpoint
- e-AI takes immediate action locally through inference from cloud-trained AI neural networks

Statistical Application

Real-time Application

Cloud

Continuing
Learning & Inference in Cloud
Continuing

Learning in Cloud
Pre- or One-time Training

IT
Information Technology

Edge

OT
Operational Technology

Endpoint

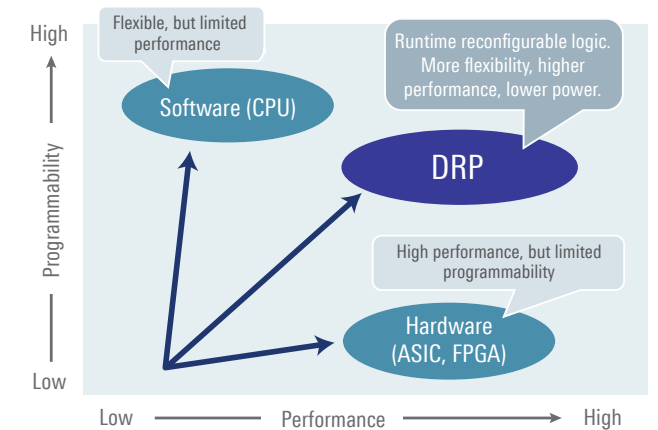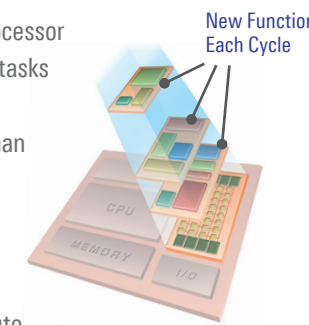Real-time action No Cloud Lag

AI Inference at Endpoint

e-AI Capability Advancements

- Renesas is evolving e-AI. Classes 1 through 4, and beyond, increase capability incrementally at each step while keeping similar power consumption
- Exclusive Dynamically Reconfigurable Processor (DRP) technology and architecture accelerate image processing, object recognition, AI, and cognitive decision making

- Each evolution step represents 10 times the previous computing power due to DRP (see below) advancement
- Class 4 represents capability of incremental learning without connection to the Cloud to solve complex graphical problems and process multi-sensor inputs for robotics

e-AI Capability

Next — **Next**

Class-4 — X10 — **Endpoint Incremental Learning** by DRP- AI2

Class-3 — X10 — **Real-time Cognition** by DRP- AI   Paper reported VLSI Symp. 2018

Class-2 — X10 — **Real-time Image Processing** by DRP   RZ/A2M product release October 2018

Class-1 — **Endpoint Inference** e-AI on MCU/MPU   Solution released July 2017

2017   2018   2019   2020   2021   2022

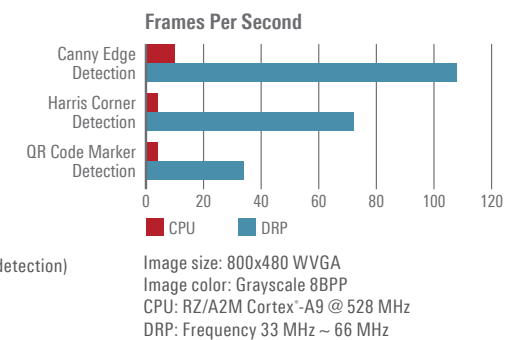## Dynamically Reconfigurable Processor (DRP)

- DRP – A Wired Logic Architecture
  - Multi-application, massively parallel processor
  - Offloads main processor for specialized tasks
- Extreme Efficiency
  - Higher performance and lower power than use of CPU, GP-GPU, DSP, or FPGAs
  - Reduced memory requirements and memory access
- Flexibility
  - Run-time reconfigurable logic can execute different tasks as needed on each processor cycle
  - Continuous new functions available to deployed products extend product life
- Acceleration
  - Image processing: edge detection, gray level, feature extraction, and more
  - Next: AI acceleration

New Function Each Cycle

High / Low Programmability

Flexible, but limited performance
Software (CPU)

Runtime reconfigurable logic. More flexibility, higher performance, lower power.
DRP

High performance, but limited programmability
Hardware (ASIC, FPGA)

Low — Performance — High

Example of DRP vs CPU Performance

Process	Execution Time (ms)	
	DRP	CPU
Canny Edge Detection	9.3	138.3*
Harris Corner Detection	13.8	294.1*
QR Marker Detection	31.3	223.0**

* CPU: Using OpenCV (cv::medianBlur+cv::Canny)
** QR Marker detection: ZBar (cv::medianBlur+Zbar detection)

Frames Per Second

Canny Edge Detection
Harris Corner Detection
QR Code Marker Detection

0   20   40   60   80   100   120

■ CPU ■ DRP

Image size: 800x480 WVGA
Image color: Grayscale 8BPP
CPU: RZ/A2M Cortex®-A9 @ 528 MHz
DRP: Frequency 33 MHz ~ 66 MHz

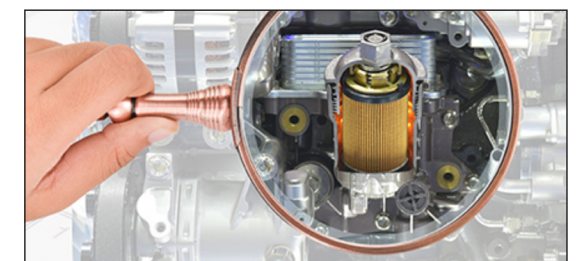**Renesas RZ/A2M Microprocessor with DRP – A hybrid approach to e-AI**

- e-AI Class-2 solution
- 528 MHz Arm® Cortex®-A9 CPU for AI Inference
- 66 MHz DRP for Image Processing
- 4 MB Internal SRAM
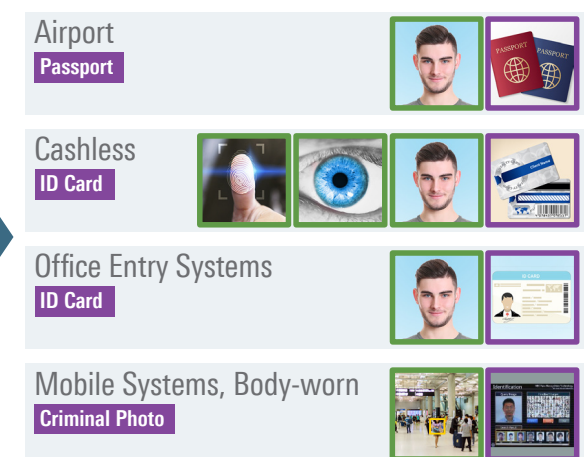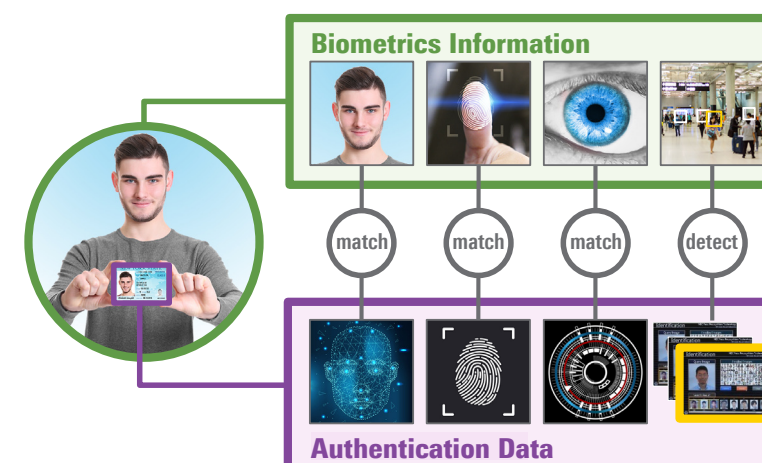- MIPI Camera Interface

Learn more:
https://www.renesas.com/RZA2M

e-AI Use Cases

Class 1: e-AI Failure Prediction for Motors

- Detects previously invisible faults in real time by minutely analyzing oscillation waveforms from motors through current, vibration, or sound
- Predicts failure before it occurs to enable early warning
- Improves service quality, avoids downtime, and reduces maintenance costs

Classes 2 and 3: e-AI Multimodal Biometrics Authentication by Image Recognition

Biometrics Information

match   match   match   detect

Authentication Data

Airport — **Passport**

Cashless — **ID Card**

Office Entry Systems — **ID Card**

Mobile Systems, Body-worn — **Criminal Photo**

e-AI Deployed at Renesas Semiconductor Factory

Smart Factory moves from Preventive Maintenance to Predictive Maintenance

- Successfully detected defective wafers using e-AI, same as human experts could do
- Reduced false alarms from 50 incidents per month to ZERO
- Anomaly detection rate improved by 6x
- Reduced engineering resources required to respond
- Eliminated requirement to set statistical thresholds

Renesas installed over 150 AI units into one of its own semiconductor factories, with 3,000 more AI units on the way

Renesas Naka Wafer
Fabrication Factory

Add-on AI Units

Learn more about Renesas e-AI solutions at:

https://www.renesas.com/us/en/solutions/key-technology/e-ai.html