

#### TechTalks:

# The Evolution of Edge Computing and AI

Steen Graham, General Manager, IoT Ecosystem & Channels, Intel Corporation Dr. Jau Huang, Chairman & CEO, CyberLink Stephen Huang, AVP of Embedded-IoT Group, Advantech



# Paradigm Shift: Edge Computing and Edge AI





### **Edge Computing and Intelligence Solutions**







AloT Solutions & Marketplace

Multi-Cloud Interoperability & Application Services





Recognition















Embedded Computers Edge Intellige

Edge Intelligence Servers

**Digital Signage Players** 

**Al Inference Systems** 





























Camera

ADVANTECH



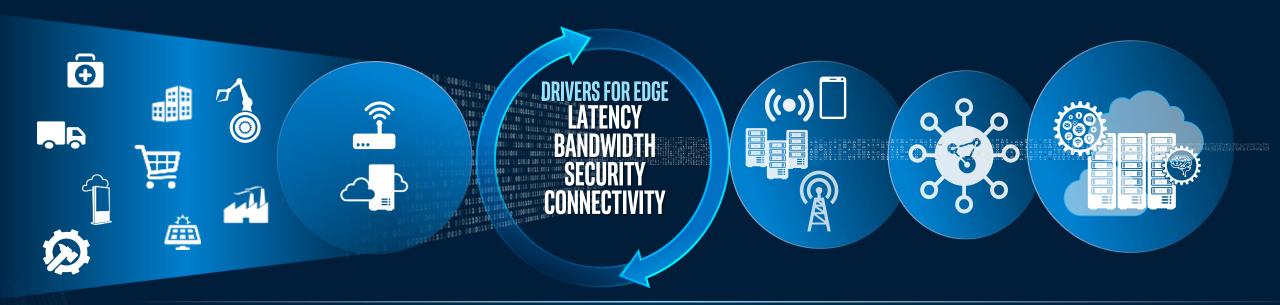
# EDGE COMPUTING EVOLUTION - PANEL

2020 ADVANTECH EMBEDDED IOT WORLD PARTNER CONFERENCE

STEEN GRAHAM 葛雲龍

GENERAL MANAGER, IOT ECOSYSTEM & CHANNELS
INTERNET OF THINGS GROUP
INTEL CORPORATION

# PARADIGM SHIFT TO EDGE COMPUTING





**ON-PREMISE EDGE** 



**NETWORK HUB OR** 

NETWORK

**CLOUD DATA** 



of enterprise-generated data will be created and processed outside the a traditional centralized data center or cloud by 20251

share of AI tasks taking place on edge devices (vs. cloud) in 2023<sup>2</sup>

growth in devices with edge AI capabilities by 2023<sup>2</sup>

Gartner; https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders

# DEEP LEARNING WITH OPENVINOT TOOLKIT

# **OpenVINO**













- ✓ NETWORK LOAD OPTIMIZATION
- ✓ COMMAND LINE DEPLOYMENT MANAGER
- ✓ CUSTOM LAYERS
- **✓ PARALLEL & HETEROGENOUS INFERENCING**
- **V** SUPPORT 10<sup>TH</sup> GEN INTEL® CORE™ PROCESSORS

>40 PRE-TRAINED MODELS



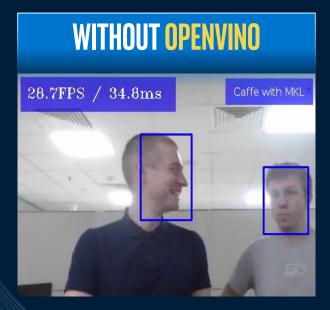




FACTORY YIELD IMPROVEMENT

# OPENVINO™ DELIVERING PERFORMANCE & SCALE

SAME HARDWARE, BETTER SOFTWARE











1STREAM

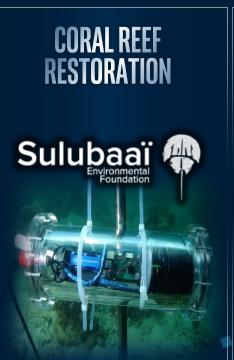
16 STREAMS

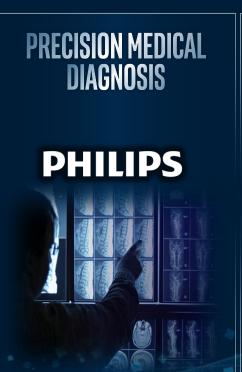
49 STREAMS



# OPENVINOT TOOLKIT DELIVERING AI IN PRODUCTION



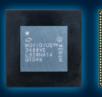








# NEXT GEN MOVIDIUS™ VPU





### **BUILD FOR EDGE AI**

**DEEP LEARNING | VISION | MEDIA** 

- ✓ LARGER MEMORY BANDWIDTH
- ✓ MULTI-NETWORK EXECUTION
- ✓ POWER EFFICIENCY WITH LOW COST
- ACCELRATED WITH OpenVINO

### HIGH EFFICIENCY

**GROUNDBREAKING PERFROMANCE** 

4X NVIDIA 1.25X ASCEN 310 THROUGHPUT (#IMAGES PER SECOND)

ON PAR NVIDIA @ 1/5TH POWER



The above is preliminary performance data based on pre-production components. For more complete information about performance and benchmark results, visit <u>www.intel.com/benchmarks</u>. See backup for configuration details. Comparison of Frames Per Second utilizing Resnet-50, Batch 1.

# ACCELERATE EDGE AI DEVELOPMENT TO DEPLOYMENT

**EDGE AI SOLUTIONS** 









**DEVELOPER OFFERINGS** 





**CLOUD-EDGE DEVELOPER WORKFLOW INTEGRATION** 





ONNX 🗘 OpenVINO





ADVANTECH















**VEGA-330** With 2x Intel® Movidius™ Mvriad™ X VPU

**VEGA-340** With 4x Intel® Movidius™ Myriad™ X VPU

# FUELING THE EDGE BUILDOUT





# DIVERSE PRODUCT PORTFOLIO

WITH SEAMLESS SOFTWARE SCALABILITY



ROBUST DEVELOPER TOOLS

FOR CHOICE OF EDGE USE CASES

VAST ECOSYSTEM SCALE

FOR DIVERSE NEEDS AT THE EDGE



### DISCLAIMER

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

For more information go to <u>www.intel.com/benchmarks</u>. Performance results are based on testing as of Oct 31, 2019 and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure.

Product	Intel Keem Bay VPU	NVIDIA Jetson TX2	Huawei Atlas 200 (Ascend 310)	NVIDIA Xavier AGX
Testing as of	10/31/2019	10/30/19	8/25/19	10/22/19
Precision	INT8	FP16	INT8	INT8
Batch Size	1	1	1	1
Sparsity	50% weight sparsity	N/A	N/A	N/A
Product Type	Keem Bay EA CRB Dev kit (preproduction)	Jetson Developer kit	Atlas 200 Developer kit	Jetson Developer kit
Mode	N/A	nvpmodel 0 Fixed Freq	N/A	nvpmodel 0 Fixed Freq
Memory	4GB	8GB	8GB	16GB
Processor	ARM* A53 x 4	ARM*v8 Processor rev 3 (v8l) × 4	ARM* A53 x 8	ARM*v8 Processor rev 0 (v8l) × 2
Graphics	N/A	NVIDIA Tegra X2 (nvgpu)/integrated	N/A	NVIDIA Tegra Xavier (nvgpu)/integrated
OS	Ubuntu 18.04 Kernel 1.18 (64-bit) on Host Yocto Linux 5.3.0 RC8 on KMB	Ubuntu 18.04 LTS (64-bit)	Ubuntu 16.04	Ubuntu 18.04 LTS (64-bit)
Hard Disk	N/A	32GB	32GB	32GB
Software	Performance demo firmware	JetPack: 4.2.2	MindSpore Studio, DDK B883	JetPack: 4.2.1
Listed TDP	N/A	10W	20W	30W

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. Check with your system manufacturer or retailer or learn more at www.intel.com.

Intel, the Intel logo, Xeon™ and Movidius™ are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries. \*Other names and brands may be claimed as the property of others.

© Intel Corporation.

## DISCLAIMER

#### Results from Slide-4: Testing by Intel as of August 7<sup>th</sup>, 2019

#### Core™i7: (for all scenarios)

Platform: Intel(R) Core(TM) i7-8700T CPU @ 2.40GHz / 6 cores x 2 Threads, HT ON, Turbo ON, Total Memory 64GB DDR4-2400MHz. Model Name: Z370M-DS3H-CF. BIOS Version: F11. Ubuntu 16.04.6 LTS with kernel 4.15.0-55-generic.

#### Caffe\* with MKL

Public distribution of Caffe with Intel® MKL optimizations enabled, for more information visit http://caffe.berkeleyvision.org MKL - Math Library for Intel®-Based Systems for more information: <a href="https://software.intel.com/en-us/mkl">https://software.intel.com/en-us/mkl</a>

#### OpenVINO (Scenarios Core™i7 + OpenVINO™)

Intel® Open Visual Inference & Neural Network Optimization software toolkit. For more information: <a href="https://software.intel.com/en-us/openvino-toolkit">https://software.intel.com/en-us/openvino-toolkit</a> OpenVINO™ Toolkit R2'2019 for Linux. Topology: face-detection-retail-0004/INT8. Scenarios (Core™i7+OpenVINO™, Core™i7+iGPU+OpenVINO™) Precision: mixed FP32+INT8. Scenario (Core™i7+iGPU+HDDL R8+OpenVINO™) Precision: FP16.

#### HDDL R8 (Scenarios Core™i7 + HDDL R8 + OpenVINO™)

Intel® Vision Accelerator Design with Intel® Movidius™ VPU PCIe card (HDDL-R8).

# **Edge AI Solutions**

**Vehicle Robot Behavior License Plate Facial** Path People **Applications** Recognition Classification **Vision** Recognition **Detection Planning Counting On-Premise Training Cloud Training Training & Modeling Enterprise Services**  Unified Al Platform intel Deep Learning Suite Azure **Training**  Model Management **WISE-PaaS/AFS** ON DIDIA. **Machine Learning** • Cross Clouds Support kubernetes docker Kubeflow **Training Server** mxnet Caffe ONNX **Supported Frameworks** \*TensorFlow **Model Deployment Device Management** SDK Inference WISE-PaaS/ DeviceOn **OpenVINO** Device Monitor & Update Deployment 3rd Party Alert & **Pretrained** Management Wizard Control Model Management Action Docker Container(Ubuntu, Win10) **Solutions Acceleration Modules Inference Systems** NVIDIA M.2+Ekey 4K Multi-display **Din-Rail High Performance** PClex4 Rugged AIR-200 **MiniPCle Facial Recognition** VEGA-320 VEGA-330 VEGA-340 **AIR-100** AIR-101 AIR-300

## **Edge AI Acceleration Modules**



**Imagine Accelerators** 

**CPUs** 

**Neural Compute Engine** 

**Vision Accelerators** 

#### **VEGA-320**



M.2 2230, One Intel Movidius X VPU on-board

#### **VEGA-330**



miniPCIe, One/Two Intel Movidius X VPUs on-board

#### **VEGA-340**



PCIe x 4, Four/Eight Intel Movidius X VPUs on-board

#### **VEGA-341**



PCIe x 4, Intel
Next Gen Movidius VPUs on-board





# **Edge Inference Systems**

Enabling Real-time Intelligence



#### **AIR-100**

- ✓ Multi-4K display
- ✓ Cost sensitive



#### **AIR-101**

- ✓ Compact design
- ✓ Low power



#### **AIR-200**

- ✓ Extended temp.
- ✓ Multi-channel processing



#### **AIR-300**

- ✓ High performance CPU
- ✓ High density image processing

Powered by:



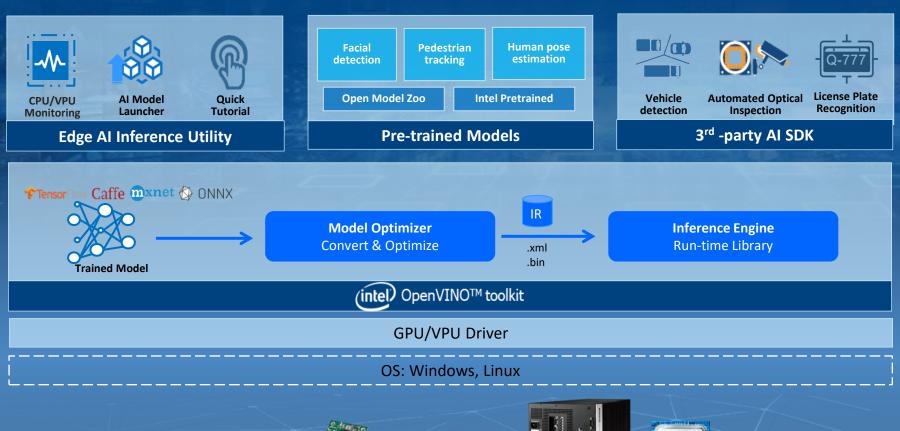








## **Edge AI Suite with OpenVINOR3**



#### **OpenVino R3 New Features:**

- Model Loading Optimization
- Command Line Deployment
- Open Model Zoo Optimizations



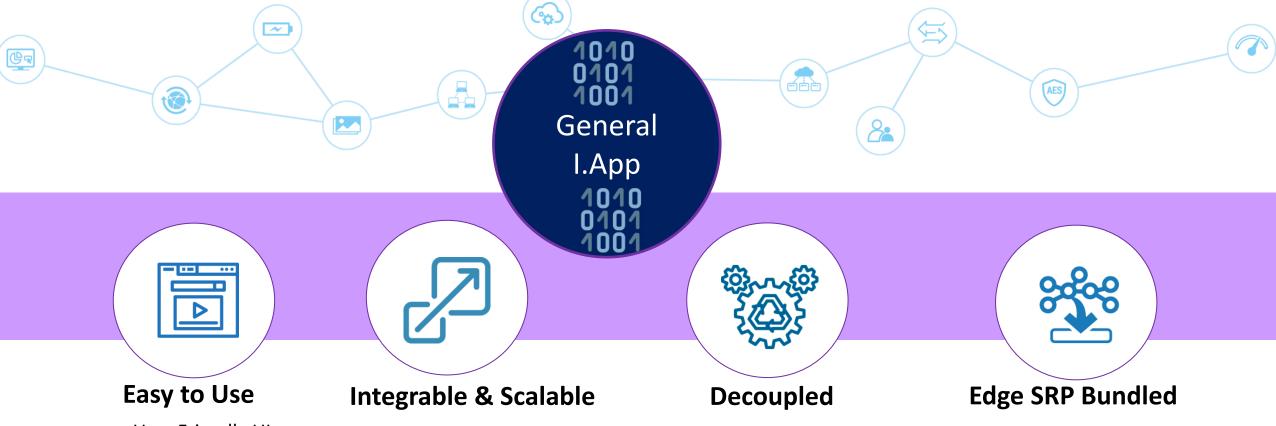
**Advantech Edge AI Platforms** 

### **EIoT Edge SRP & I.App**

**Manufacturing DFSI Smart City DFSI Domain DFSI WISE-PaaS Marketplace FaceView E21 OTA ePaper General LAPP** DeviceOn DeviceOn DeviceOn DeviceOn Software, firmware & Facial recognition • EPD device auto-discover Protocol conversion **OTA** configuration updates Customer behavior analytics E2I **ePaper FaceView**  Fast transmission Edge intelligence Advantech BIOS updates VIP/Watchlist mgt · Device association Real-time virtualization **Azure AKS** Compute Resource Management **IoT Device Connecting** Data & Storages Services IoT - PaaS Azure Automated ML X CLOUD FOUNDRY **kubernetes L**RabbitMQ mongoDB Azure Digital Twin WISE-PaaS/EnSaaS **kubernetes** laaS **WISE-STACK/ Data Service Server Public Cloud Microsoft Azure Private Cloud** 3<sup>rd</sup> Party Al.Inference **OpenVINO SDK Edge Al Suite** DeviceOn **McAfee Security Acronis Backup Edge.SRP** (HW+SW) **Edge Intelligence Servers Edge AI AIR Series** 



# **Data Driven General I.App**



- User Friendly UI
- Easy to Connect & Configure
- Easy for DFSI to Integrate
- Large-scale Duplication

- Function Decoupled
- Available on MarketPlace

- Optimized for Edge SRP
- AloT Applications Oriented



# FaceView I.App Empowering AI Facial Recognition



- 98.5% accuracy rate
- 10E-6 error rate



Real-time Identification & Notification

- Face identity & labeling
- VPU/GPU acceleration





 Dashboard for decision makers



- Easy enrollment
- VIP/Watchlist management



Corporate



Retail



Hospitality



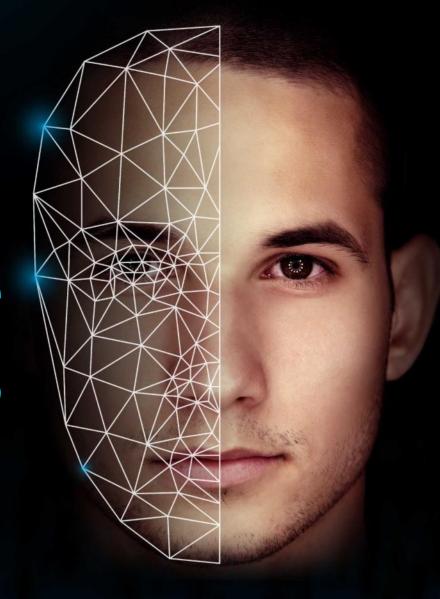
**Bank/Government** 

# Al-driven IoT at the Edge

**Dr. Jau Huang**Chairman and CEO



Al Facial Recognition by **CyberLink** 



# About CyberLink World's Top Multimedia & Al Software Company

- A 20-year public company since 2000
- Market leader in multimedia; over 3 billion software shipped to date
- From #1 media player, PowerDVD to cutting-edge

Al Facial Recognition solutions



### PowerDVD, used to be on 80% of all PC





### PowerDirector 18

(10 PC Mag Editors' Choices in a row, from v9 to v18)









































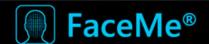


The World's Top Cross-Platform AI Facial Recognition Engine

99.5% ACCURACY RATE

OPTIMIZED FOR EDGE IoT & Server

CROSS PLATFORM
Windows | Linux
iOS | Android



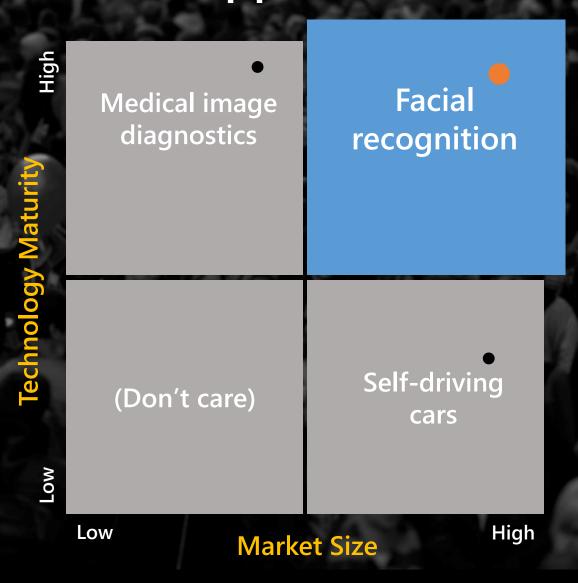
# Top Ranking by NIST (WILD, 1E-4)

Vendor Ranking	Company	FNMR	Country/ Region
1	Hikvision	2.71%	China
2	N-Tech Lab only 0.27%	2.75%	Russia
3	Imperial College London	2.76%	UK
4	DeepGlint	2.78%	China
5	Ever Al	2.78%	US
6	Vocord	2.80%	Russia
7	Toshiba	2.82%	Japan
8	VisionLabs	2.85%	Netherlands
9	Camvi Technologies	2.88%	US
10	Anke Investments	2.88%	China
11	Panasonic R+D Center Singapore	2.95%	Japan
12	CyberLink FaceMe	2.98%	Taiwan
13	Innovatrics	3.01%	Sloak
14	Hengrui Al Technology Ltd	3.03%	China
15	Neurotechnology	3.03%	Lithuania
16	KanKan Ai	3.04%	China
17	Vigilant Solutions	3.06%	US
18	Saffe Ltd	3.08%	Germany
19	Idemia	3.09%	France
20	Tech5 SA	3.10%	South Africa

Source: NIST FVRT 1:1 2019.07.03 Top 20 vendor ranking, each vendor may submit up to 2 algorithms.



# Killer Apps for IoT





### Al on Video Goes to the Edge

(Bandwidth & Computation, Cost/Performance, advantages)

Edge-IoT



Billions
Detection & recognition

**Edge-Workstations** 



Gateways
Data & analytics

**Cloud Server** 

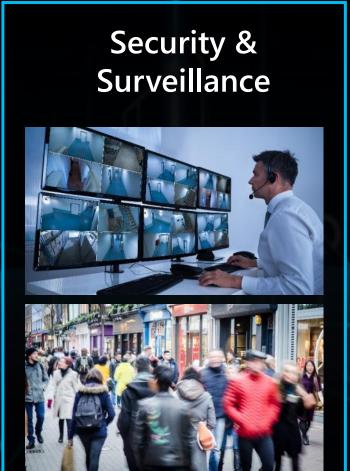


Data center Al training & inference

# Market Demand for Facial Recognition

(a few examples)







# Smart Kiosk Using Facial Recognition

- Customer demographics, e.g., gender, age, emotions
- Personalized offering
- Authentication

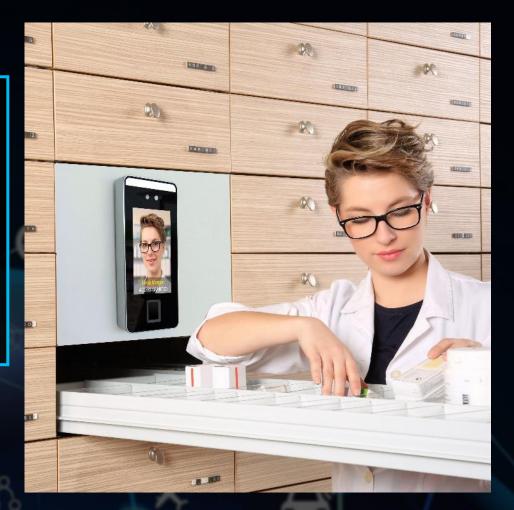






# Medicine Cabinet Access with Facial Recognition

- Fingerprint does not work for 2 reasons
- Light-weight AI engine (inference engine) required



## FaceMe® - Built for AloT

**Optimized for Edge Computing** 

**Edge Device** 















**Embedded System Mobile Device** 

**POS** 

**Kiosk** 

**Digital Signage Door Security** 

**VMS** 

OS











**Inference Engine** 









**CPU** 















## FaceMe® - Optimized for both Cloud & AloT

False Acceptance Rate (FAR) 1E-6

	UH Model	VH Model	H Model
True Acceptance Rate (TAR)	98.89%	97.99%	94.98%
Execution time (on Celeron)	68.0 ms	6.9 ms	4.7 ms
Deep Learning Model	250MB	17MB	4MB

- VH Model runs nearly 10X faster & 15X smaller by only sacrificing 0.9% accuracy.
- H Model can even run on Snapdragon 410 chip (very low computation power).
- Both VH and H Models are ideal for AloT devices.



### An Example on Hardware Optimization

FaceMe on Intel® Movidius™ VPU

Intel Movidius VPU (Vision Processing Chip) aims to run Al models in small factor devices with very low power







Intel next-gen Movidius VPU, codenamed Keem Bay

20X Performance Speedup on "Face Recognition"



### FaceMe | Face Recognition w/Ultra High Model **Joyce Joyce** [25~30] Female [30~35] Female 49 visits, 0:35:52 49 visits, 0:37:56 Movidius VPU Celeron CPU 3868.7ms 195.1ms Time required for Face Recognition (Smaller is faster)

# Conclusion: FaceMe is Designed for ...





# **Best Partnership in IoT Ecosystem**







**Dr. Jau Huang**Chairman and CEO

For more details CyberLink.com/FaceMe





#### **Acceleration Modules**

#### **Inference Systems**

#### **Solutions**



**VEGA-320** 



**VEGA-330** 

M.2, 1 VPU MiniPCle, 1/2 VPUs





**VEGA-340** PClex4, 4/8 VPUs



**AIR-100** Quad 4K Displays



**AIR-101** Din Rail



**AIR-200** Ruggedized



**AIR-300** High performance



**Facial Recognition** 



# Co-Creating the Future of the IoT World

