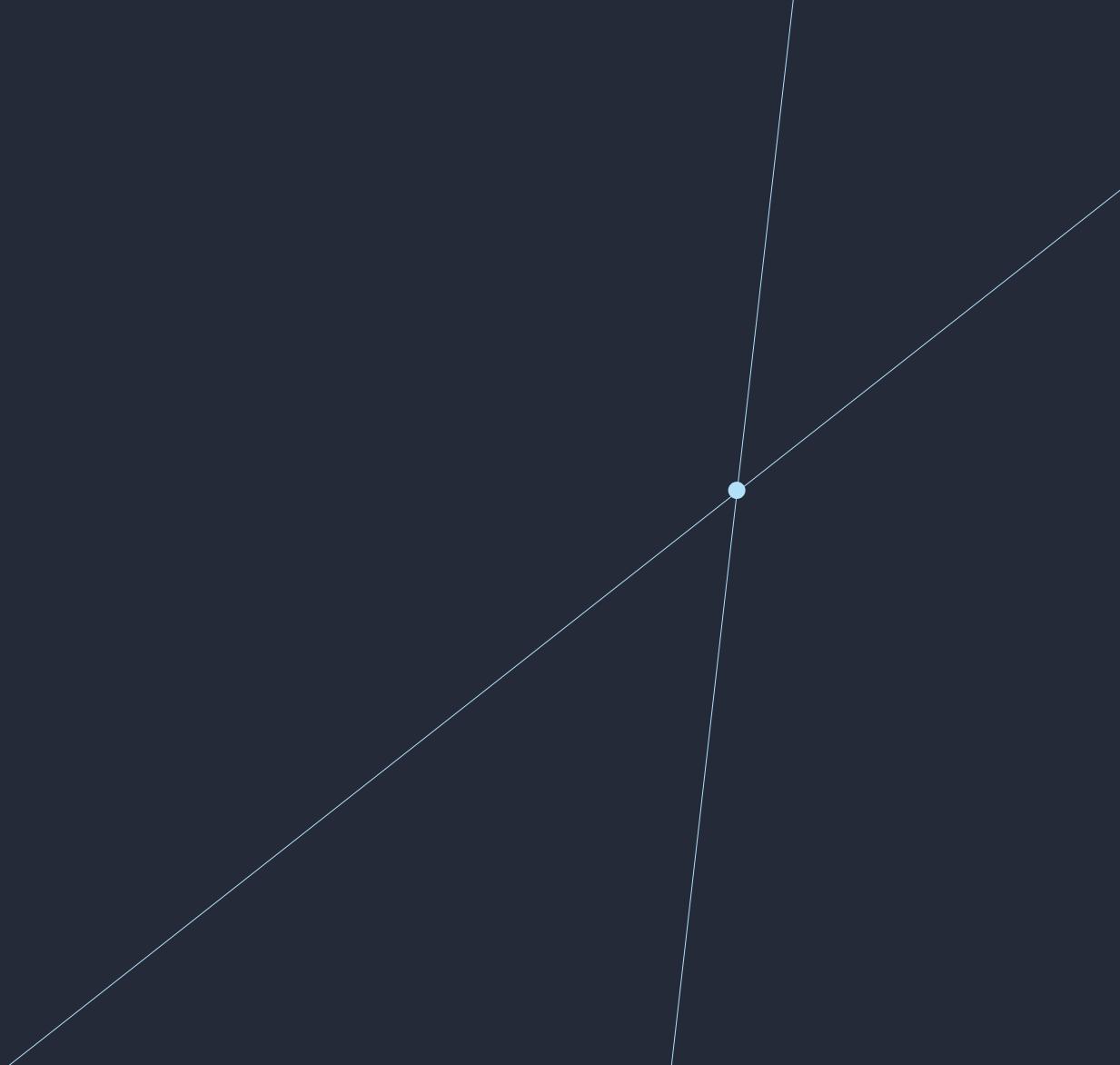




Hardware-aware AI Model Optimization Platform and  
On-device AI Solutions



# We Bring the Benefits of AI Everywhere

Nota AI bridges the gap between AI technology and real-world application, pioneering the democratization of AI by making it more accessible to everyone, everywhere.

## Product Value

### AI Model Optimization

Maximized Productivity

### On-device AI Solutions

High Scalability Across Industries

## Product Portfolio

Hardware-aware AI Model Optimization Platform and On-device AI Solutions

### AI Model Optimization Platform

Model Development ➞ Model Optimization ➞ Model Testing ➞ Deployment

### NetsPresso.

Task-specific  
lightweight model

Device-tailored  
optimized model

Validated performance  
on real devices

### On-device AI Solutions



Nota Vision Agent



ITS Solution



DMS & FR Solution



Industrial Safety Solution



Surveillance Solution

## Global Acclaims

**24+**

				
CB Insights AI 100: The most promising AI Startups 2025			CB Insights LLM Ops Market Map Global Leader in Hardware-aware AI Optimization 2023	
MarketsandMarkets Industry-leading Edge AI Software Vendor 2023			tinyML Awards 2023 Best Product of the Year 2023	
Plug and Play Tech Center Startup Award 2022			Embedded Tech Convention USA 2022 Excellence in Engineering 2022	

## Academic Achievement

**40+**

	CVPR 2025 Workshop Efficient LLaMA-3.2-Vision by Trimming Cross-attended Visual Features		COLING 2025 Where do LLMs Encode the Knowledge to Assess the Ambiguity?
	AI City Challenge CVPR 2024 2nd Place, Track 4: Road Object Detection in Fish-Eye Cameras		ECCV 2024 SNP: Structured Neuron-level Pruning to Preserve Attention Scores

## Clients & Partners

**101+**

Semiconductor and Hardware Manufacturing	    
IT Services and Telecommunications	   
Construction, Manufacturing, Distribution, Smart Cities	    

## Milestone

	2015-2017	2018-2020	2021-2023	2024-Present
	<ul style="list-style-type: none"> <li>Nota Inc. established</li> <li>Seed funding (USD 0.2M)</li> </ul>	<ul style="list-style-type: none"> <li>Pre-series A funding (USD 1.2M)</li> <li>Series A funding (USD 6.6M)</li> <li>Europe Subsidiary established in Berlin, Germany</li> </ul>	<ul style="list-style-type: none"> <li>Series B funding (USD 14.6M)</li> <li>Nota America Inc. established in Sunnyvale, US</li> <li>NetsPresso® Platform launched</li> </ul>	<ul style="list-style-type: none"> <li>Series C funding (USD 19.9M)</li> <li>Garnered 'A, A' in Tech Evaluation</li> <li>Selected for National AI Foundation Model Project</li> </ul>

# NetsPresso®

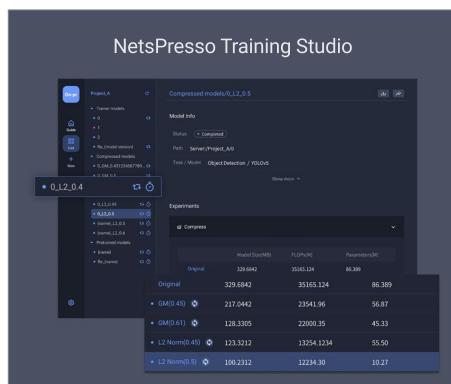
Hardware-aware AI Optimization Platform Tailored for Any Model, Any Device

Workflow	Model Development	Model Optimization	Model Testing
NetsPresso Core Engine	Model Zoo Trainer Compressor	Graph Optimizer Quantizer IR Converter	Simulator Profiler Device Farm
Available on	CLI	 PyNetsPresso	
	GUI	NetsPresso Training Studio	NetsPresso Optimization Studio NetsPresso Benchmark Studio
Customer Value	Seamless Model-to-Device Compatibility		High-Performance Optimization Real-Device Validation that Minimizes Trial and Error

Workflow 01

## Model Development

Easily create task-specific lightweight models with high performance



### Core Modules

Model Zoo

Trainer

Compressor

### Key Features

Task-specific, device-compatible pre-trained models

Easy-to-use advanced compression techniques, such as structured pruning

Secure training in a local environment with zero risk of data leakage

Available on	PyNetsPresso	NetsPresso Training Studio
--------------	--------------	----------------------------

## Workflow 02

# Model Optimization

Newly Featured

Optimize AI models for size, speed, and compatibility, all tailored to your target device

The screenshot shows the NetsPresso Optimization Studio interface. On the right, a neural network diagram is displayed with nodes for Input, Conv, Relu, MaxPool, and Conv. The input layer is labeled 'x'. The first convolutional layer has dimensions 'W @ 64x64x56+56' and 'B @ 64'. The Relu layer follows. The MaxPool layer has dimensions '64x64x56+56'. The second convolutional layer has dimensions 'W @ 64x3x1x7x7' and 'B @ 64'. Below the diagram, there is a note: 'Scroll To Top'.

**Model ID:** inavn\_ib083cb17369548  
**Inference Time:** 392.17 ms  
**Peak Memory:** 0-60  
**Compute Units:** 50 50 150  
**Latency:** 50 50 150 ms  
**Framework:** ONNX + TFLite  
**Input Specs:** float32 [1, 3, 128, 128]  
**Created Time:** 2024.11.07 16:48  
**Target Hardware:** NPU-UXI-03  
**Runtime:** Python 3.12, TensorFlow 2.8.1  
**Updated Time:** 2024.11.07 18:57

**Total 0 Models**

Layer ID	Layer Name	Operator Type	Static Input Size	Latency	Precision	Compute Unit	Sensitivity Rank	Target Precision	Latency
01	Conv_JermeC...	FusedConv	133,156	19 ms	FP_32	CPU	19	✓ int8	138,156 (+3.5%)
02	Bnrm_JermeC...	FusedConv	133,156	19 ms	NPU	NPU	19	✓ int8	138,156 (+3.5%)
03	Conv_JermeC...	FusedConv	68,156	19 ms	GPU	GPU	19	✓ int8	138,156 (+3.5%)
04	Conv_JermeC...	FusedConv	68,156	19 ms	CPU	CPU	19	✓ int8	138,156 (+3.5%)
05	Conv_JermeC...	FusedConv	68,156	19 ms	NPU	NPU	19	✓ int8	138,156 (+3.5%)
06	Conv_JermeC...	FusedConv	68,156	19 ms	GPU	GPU	19	✓ int8	138,156 (+3.5%)
07	Conv_JermeC...	FusedConv	68,156	19 ms	CPU	CPU	19	✓ int8	138,156 (+3.5%)
08	Conv_JermeC...	FusedConv	68,156	19 ms	NPU	NPU	19	✓ int8	138,156 (+3.5%)
09	Conv_JermeC...	FusedConv	68,156	19 ms	GPU	GPU	19	✓ int8	138,156 (+3.5%)
10	Dnrm_JermeC...	FusedConv	68,156	19 ms	CPU	CPU	19	✓ int8	138,156 (+3.5%)
11	Flrn_JermeC...	FusedConv	68,156	19 ms	GPU	GPU	19	✓ int8	138,156 (+3.5%)
12	Conv_JermeC...	FusedConv	68,156	19 ms	CPU	CPU	19	✓ int8	138,156 (+3.5%)

## Key Features

Computational optimization ensuring both compatibility and performance

Enhanced speed and memory efficiency with minimal accuracy loss

Automatic model conversion for effortless SDK integration

## Core Modules

Graph Optimizer

Quantizer

IR Converter

## Available on

PyNetsPresso

NetsPresso Optimization Studio

## Workflow 03

# Model Testing

Quickly and accurately validate performance on real devices

The screenshot shows the NetsPresso Benchmark Studio interface. It features a dashboard with four main sections: 'Convert', 'Benchmark', 'Simulator', and 'Profiler'. The 'Convert' section includes options for 'ONNX to TensorFlow', 'ONNX to TensorFlow Lite', 'ONNX to OpenVINO', and 'TensorFlow Keras to TensorFlow Lite'. The 'Benchmark' section lists various target devices: Andante Hilda Vision (Arm Mali-G71), Renesas R8D1 (Arm Cortex-A75/A76), All Ensemble DevKit E7 (Arm Cortex-A75/A76/Mali-G71), NVIDIA Jetson AGX Orin, NVIDIA Jetson Nano, NVIDIA AWD-T4, NVIDIA Jetson Xavier, NVIDIA Jetson Nano, NVIDIA Jetson AGX Xavier, NVIDIA Jetson TX2, Raspberry Pi 5 (Arm Cortex-A75), Raspberry Pi 4B, and Raspberry Pi ZeroZW. There are also buttons for 'Check the supported model range' and 'Request to add the device you want'.

## Core Modules

Simulator

Profiler

Device Farm

## Key Features

Validated real-world device performance

Precise bottleneck analysis for optimized performance

Accelerate development by minimizing model tuning iterations

## Available on

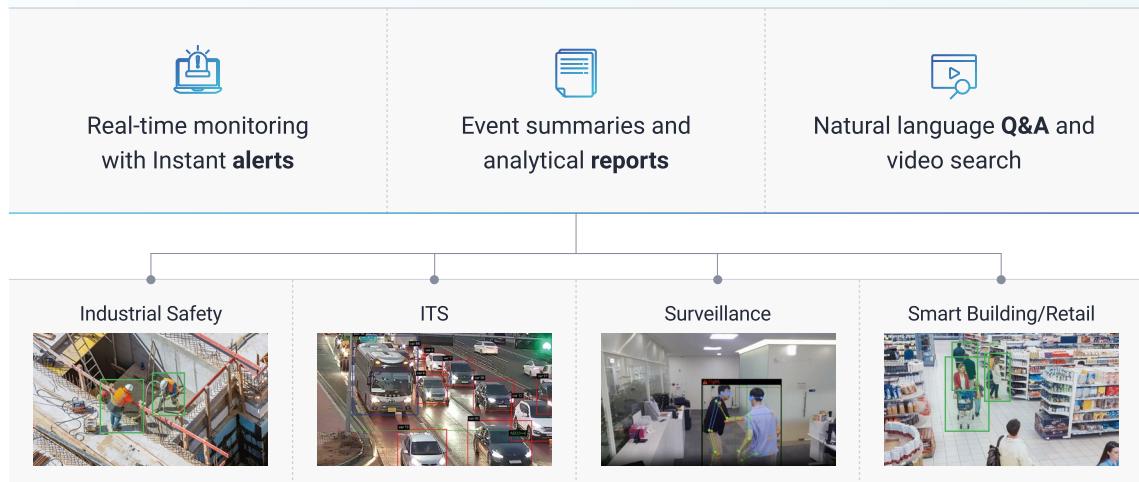
PyNetsPresso

NetsPresso Benchmark Studio

# Nota Vision Agent (NVA)

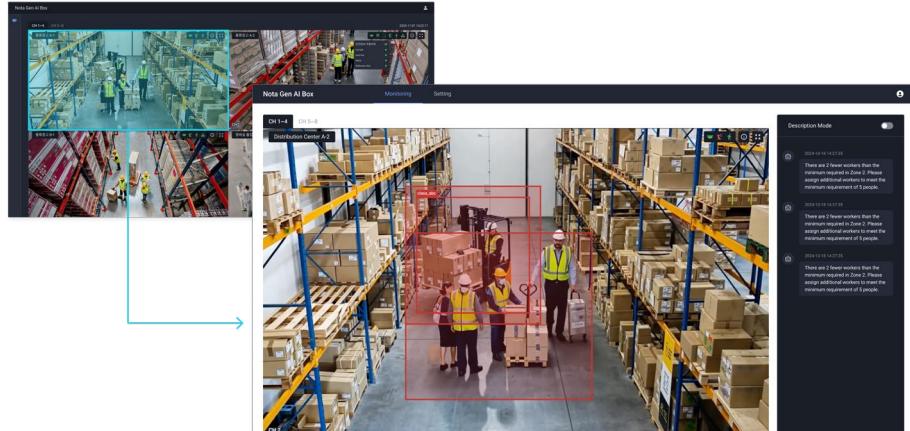
A Next-gen AI Agent Transforming Visual Data into Actionable Insights

A tailor-made agent for any industry, any situation



## Customer Value

Nota Vision Agent (NVA) is an advanced AI solution powered by Vision-Language Models (VLM), revolutionizing visual data analysis and utilization.



### Enhanced Safety

Proactive accident prevention and SOP compliance

### Maximized Productivity

Real-time risk detection and automated reporting

### Simplified Operation

Seamless integration with existing systems without equipment replacement

## Distinctive Features

Legacy Video Management Solution	Nota Vision Agent
Limited capability to analyze sequential actions, resulting in low accuracy for event detection in complex environments	» Real-time sequential action tracking with high accuracy in complex environments
Manual video searches are time-consuming, and metadata-based searches are often inaccurate	» Intuitive natural language-based video search and AI-generated summaries improve operational efficiency
Manual report generation after events occur, consuming significant time and effort	» Automatically generates event timelines and provides detailed AI-driven analytics reports

## Use Cases

NVA delivers lightweight, high-speed agents flexibly adaptable to diverse industrial environments, enhancing safety and dramatically improving operational efficiency.

Case 1 Industrial Safety	Case 2 ITS	Case 3 Surveillance
<p> <b>Factory Safety Supervisor</b></p> <p>Please check whether proper safety equipment is worn and whether the work is being carried out in pairs when using a ladder.</p> <p>The worker is not wearing a helmet and is working alone on the ladder. Immediate corrective action is required.</p>  <p>Stream: Camera_046</p>	<p> <b>Traffic Accident Investigator</b></p> <p>Please explain the cause of the accident at the intersection this morning, and how it was handled.</p> <p>Here's the footage from the incident at 7:23 AM. I'll prepare a detailed report, including the cause and follow-up actions.</p>  <p>Stream: Camera_047</p>	<p> <b>Municipal Surveillance Officer</b></p> <p>Please find the person who dumped trash here yesterday around 10PM.</p> <p>The person who dumped trash at 10:23 PM yesterday has been identified on video. I'll show the footage.</p>  <p>Stream: Camera_048</p>

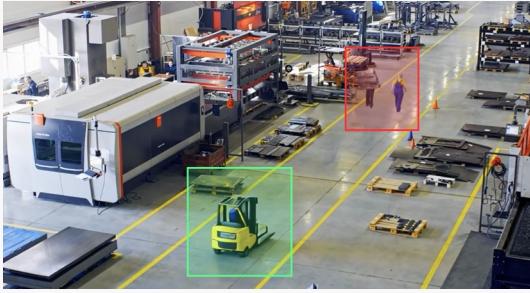


Easily adaptable across diverse industries to solve complex challenges beyond traditional visual data analytics.

# Industrial Safety Solution

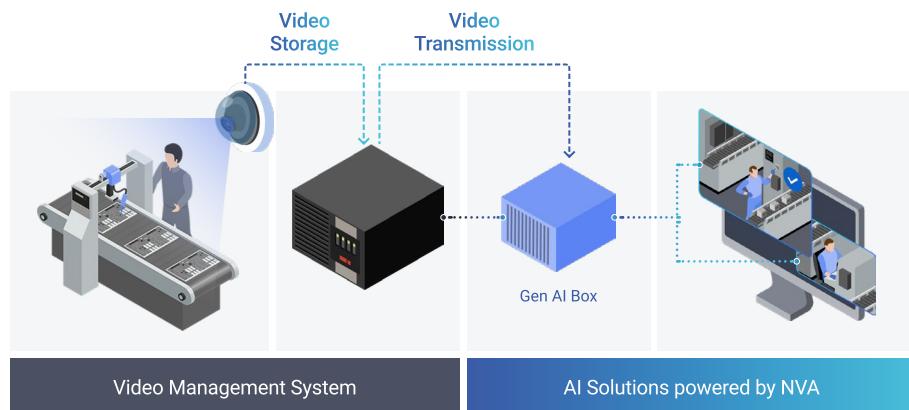
Powered by NVA

A Generative AI Solution that Identifies and Understands On-site Hazards

Worker Safety Management	Work Environment Safety Management
<ul style="list-style-type: none"> <li>Detection of personal protective equipment (PPE) usage</li> <li>Detection of abnormal worker behavior</li> <li>Detection of worker location and access to hazardous areas</li> </ul> 	<ul style="list-style-type: none"> <li>Detection of workplace hazards</li> <li>Heavy equipment access control</li> <li>Detection of obstacles in the workplace</li> <li>Monitoring of specific areas</li> </ul> 

Customer Value	Enhanced Workplace Safety	Improved Operational Efficiency	Maximized Productivity
	<b>Enhanced Workplace Safety</b> Strengthens preventive measures to minimize accidents	<b>Improved Operational Efficiency</b> Streamlines operations by improving hazard detection and response	<b>Maximized Productivity</b> Optimizes workforce allocation with automated monitoring

**Pipeline** Plug-and-play 'Gen AI Box' with your current video management system and start using it immediately on-site – with no need for replacements or additional purchases



# Surveillance Solution

Powered by NVA

A Video Surveillance Solution Protecting Public Safety and the Environment

<b>Public Safety Management</b>	<b>Disaster Prevention</b>
Detection of safety incidents such as fights, and unauthorized intrusions	
<b>Crowd Safety Management</b>	<b>Illegal Activity Detection</b>
Detecting crowd density at public events such as parks, festivals, and concerts	

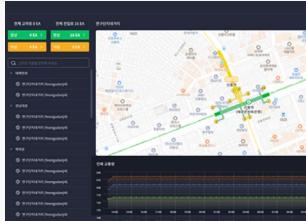
<b>Customer Value</b>	<b>Enhanced Public Safety</b>	<b>Improved Crisis Response Capabilities</b>	<b>Optimized Operational Efficiency</b>
	Enables real-time crime detection through CCTV surveillance	Enables rapid response to mitigate damages	Improves operational readiness through an integrated control system

<b>Success Cases</b>	 <b>DAEJEON</b>  <b>경찰청</b> Conducted a field demonstration for crowd safety monitoring in collaboration with Daejeon Metropolitan	 <b>GUNPO</b> Implemented an illegal waste dumping detection project in partnership with Gunpo City	 <b>Korea Forest Service</b>  <b>Incheon Metropolitan City</b> Carried out a field demonstration for natural disaster detection in cooperation with the Korea Forest Service and Incheon Metropolitan City
----------------------	---	---	---

# ITS Solution

Powered by NVA

An Intelligent Transportation System Creating Smarter, Safer Roads for Everyone

<h2>Traffic Monitoring</h2> <ul style="list-style-type: none"><li>Intersection traffic analysis</li><li>Level of Service (LOS)</li></ul> 	<h2>VRU Safety</h2> <ul style="list-style-type: none"><li>Road safety assessment</li><li>Safe crossing solution</li></ul> 
<h2>Incident Management</h2> <ul style="list-style-type: none"><li>Automatic incident detection</li><li>Instant alerts and analysis with VLM</li></ul> 	<h2>Smart Parking</h2> <ul style="list-style-type: none"><li>Parking space management</li><li>Autonomous parking assistant</li></ul> 

<h2>Customer Value</h2>	<h3>Optimized Traffic Flow</h3> <p>Analyzes real-time traffic data to reduce congestion and improve flow</p>	<h3>Enhanced Road Safety</h3> <p>Identifies hazards instantly to help prevent road accidents</p>	<h3>Cost-Efficient Deployment</h3> <p>Seamlessly integrates with existing CCTV infrastructure</p>
-------------------------	--	--	---

## Success Cases

2025

Signed Contract with Dubai RTA for Generative AI-based ITS Solution



2024

Achieved highest rating in smart intersection performance assessment by Korea's Ministry of Land, Infrastructure and Transport

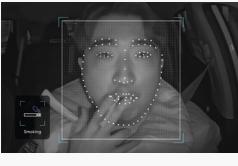
2023

Completed AI Smart Parking pilot project in Milton Keynes, UK



# DMS & FR Solution

Enhancing In-Cabin Safety and Security with Real-Time Precision Video Monitoring

DMS (Driver Monitoring Systems)		FR (Facial Recognition)	
<ul style="list-style-type: none"><li>Real-time detection of drowsiness, distraction, smoking, and phone usage</li><li>Evaluates stress and fatigue levels by analyzing driving patterns, sudden acceleration, and braking behaviors</li></ul>		<ul style="list-style-type: none"><li>Driver Authentication &amp; Shift Detection</li><li>Applies facial recognition to secure restricted areas and control equipment access</li><li>Restricts unauthorized access to enhance security</li></ul>	
<p>Drowsiness</p> 		<p>Distraction</p> 	
<p>Smoking</p> 		<p>Phone Usage</p> 	
<p>Customer Value</p>		<p>Enhanced Driving Safety</p> <p>Tracks driver fatigue and distractions to prevent accidents</p>	
<p>Efficient Security Management</p> <p>Provides secure identity checks with facial recognition</p>		<p>Optimized for Low-Power Devices</p> <p>Designed to perform reliably on low-power devices</p>	
<p>Customer Benefit</p>		<p>High performance on low-spec edge devices</p>	
<p>Compatible with a wide range of cameras</p>		<p>Flexible installation at various locations</p>	
<p>Robust performance in various environments</p>			



**Email** contact@nota.ai

**Website** nota.ai

**Offices**

HQ: Daejeon, South Korea

R&D and Business Center: Seoul, South Korea

US Subsidiary: Sunnyvale, US

Europe Subsidiary: Berlin, Germany



Nota AI® is the trademark of Nota Inc.

© 2025, Nota Inc.