PocketGo – Portable Universal Data Collection Platform

Product Overview

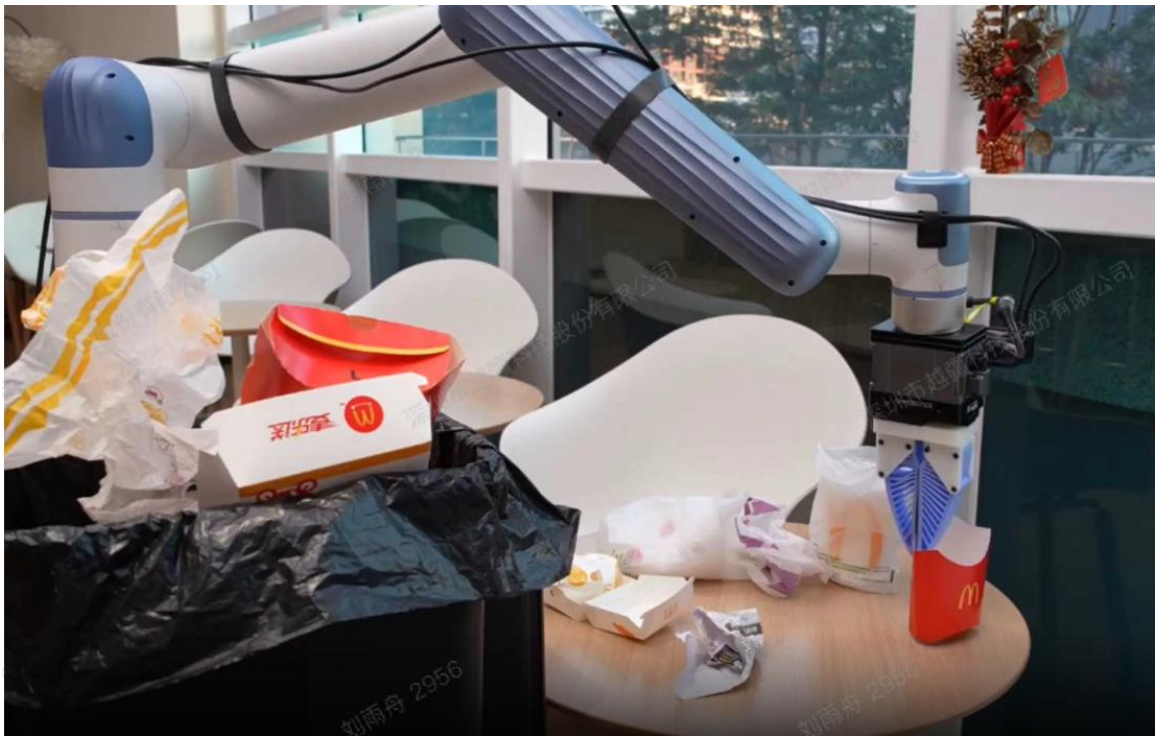
Amid the global advancement toward Artificial General Intelligence (AGI), *Embodied AI*—as a core pathway to realizing AGI—is unlocking revolutionary possibilities in robotic cognition. Built on the foundation of Behavior Cloning, PocketGo utilizes imitation learning based on human demonstrations and integrates industrial-grade data acquisition kits, high-precision motion trajectory capture, and multimodal data fusion algorithms. This enables robots to evolve full-stack skills, from basic pick-and-place tasks to intricate, fine-grained operations—paving the way for a breakthrough in how intelligent agents interact with the physical world. With a modular architecture that ensures rapid integration with mainstream robotic systems, PocketGo offers cross-scenario solutions spanning domestic, commercial, and industrial applications. It significantly lowers the entry barrier to embodied AI research and accelerates the transformation of AGI technology from the lab to real-world deployment.

Handheld and end-effector kit combinations, low-cost, high-efficiency tools for effective data collection

Exceptional compatibility, fully adaptable with the entire Nova and CR series robotic arms

Application Scenarios

- **Commercial: Automated cleaning, warehouse sorting, fresh produce handling, retail restocking. *Example*: Table cleaning in restaurants.**



- **Household Services: Kitchen management, environment tidying, elderly care. Example: Daily household chores.**

- **Industrial: Precision assembly, flexible production lines, automated sorting. *Example*: Item sorting in industrial environments.**

Key Features

- Portable and Mobile Design: Lightweight and compact with a quick-clip waist attachment for convenient carrying. Ready for data collection with virtually no setup required. Capable of collecting up to 4,000 datasets per day.
- High-Performance Action Camera: Equipped with a wide-angle fisheye GoPro camera for an expansive field of view. Combined with a built-in IMU sensor, it enables robust motion tracking even during fast movements.
- Multimodal Data Fusion Algorithm: Utilizes visual-inertial OBR-SLAM3 monocular SLAM technology to achieve global 6D pose tracking by fusing camera and IMU data. Effectively addresses scale ambiguity caused by motion blur or texture loss in traditional handheld devices, enabling high-precision motion trajectory capture and faithful learning path reproduction.
- Strong Generalization Capability: Able to generalize across hundreds of zero-shot tasks with only a small amount of sample data, and adaptable to new environments and objects.
- High Anti-Interference Performance: Integrated multi-spectral fusion technology ensures 94% trajectory capture integrity in low-light conditions, making it suitable for complex lighting environments like factories and warehouses.
- Universal and Efficient: Collected data and trained AI models are universally compatible across robotic platforms, supporting the entire Nova and CR series collaborative arms. This reduces data collection costs and improves training efficiency by up to 3 times compared to traditional

methods.

- Wireless and Convenient Recording: Supports wireless control for easy and precise data recording, ensuring efficient and accurate data acquisition.
- Versatile End-Effector Gripper Design: The soft grippers effectively conform to the shape of the object being grasped, provide cushioning against end-effector collisions, and possess self-recovery capabilities. In contrast, the rigid grippers offer high rigidity and resistance to deformation, making them suitable for scenarios involving heavy loads or requiring high precision.

Teaching Resources

- **Development of Multimodal Perception Systems**
  - Hardware Components: High-precision IMU + Wide-angle Camera
  - Teaching Experiments:
    - Implementation of 6D Trajectory Reconstruction Algorithms
    - Spatiotemporal Alignment Error Compensation Experiment
    - Multispectral Fusion Practice in Low-light Environments

- **Theory and Algorithmic System of Imitation Learning**
  - Principles of Behavior Cloning (BC) / Inverse Reinforcement Learning (IRL)
  - Diffusion Policy / Meta-learning Generalization Mechanisms
  - Comparison of UMI Framework vs. Traditional DAgger Algorithm

- Utilization Rate Analysis: Single-modal Demonstrations vs. Multimodal Demonstrations

- **Innovative Topics in Transfer Learning**
  - Few-shot Multi-task Strategy Distillation (<10 demonstrations per task)
  - Cross-modal Imitation from Human Video to Robotic Actions
  - Demonstration Data Augmentation Based on Physical Priors

Configuration List

Part Number	Portable Universal Data Collection Platform	Core Components	Quantity	Remarks
1200000321	Handheld Gripper Kit	Handheld Gripper*1, GoPro*1	1	Standard
1200000322	End Gripper Kit	Robotic Arm End Gripper*1, GoPro*1	1	Standard
____	Software Platform	Detailed User Manual, Secondary Development SDK	1	Standard
1100000208	Mini AI Reasoning Computer	AI Inference Host*1	1	Optional
1200000394	Mobile Stand	Mobile Vertical Robotic Arm Stand	1	Standard
1200000162	Nova 5	High-Precision 6-Axis Collaborative Robotic Arm	1	Order Separately

Core Specifications

- Handheld Gripper Kit:
  - Opening Range: 73.5 ± 1.5 mm
  - Effective Gripping Depth: 100 mm
  - Maximum Gripping Weight: ≤ 2 kg
  - Net Weight: 0.93 kg
  - Soft Fingertip × 1 (pre-installed on gripper end)
  - Rigid Fingertip × 1
  - Operating Environment: Room temperature 16–28° C

- ○ Camera
  - ■ Field of View (FOV): 155°
  - ■ Image Sensor: 1/1.9" CMOS Sensor
  - ■ Effective Pixels: 24.69 MP (dynamic mode), 27.13 MP (still mode)
  - ■ Video Resolutions:
    - ● 5.3K (8:7) @ 30/25/24 fps
    - ● 4K (9:16) @ 60/50/30/25 fps
    - ● 4K (8:7) @ 60/50/30/25/24 fps
    - ● 5.3K (16:9) @ 60/50/30/25/24 fps
  - ■ Stabilization System: HyperSmooth 6.0 Advanced Stabilization

- ● End-Effector Gripper Kit
  - ○ Standard Opening Range: 73.5 ± 1.5 mm
  - ○ Object Size Range: 5 mm–75 mm
  - ○ Effective Gripping Depth: 100 mm
  - ○ Maximum Gripping Weight: ≤ 2 kg
  - ○ Net Weight: 1.58 kg
  - ○ Soft Fingertip × 1 (pre-installed on gripper end):
  - ○ Hard Fingertip × 1
  - ○ Soft materials: effectively absorb end-collision impact, conform to the shape of the object being grasped, and support self-recovery
  - ○ Hard materials: high rigidity, not easily deformed
  - ○ Operating Environment: Room temperature 16–28° C
  - ○ Camera:
    - ■ Field of View (FOV): 155°

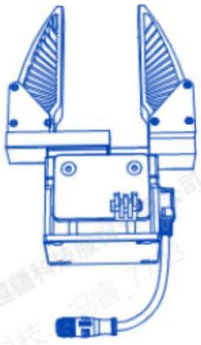
- Image Sensor: 1/1.9" CMOS Sensor
- Effective Pixels: 24.69 MP (dynamic mode), 27.13 MP (still mode)
- Video Resolutions:
  - 5.3K (8:7) @ 30/25/24 fps
  - 5.3K (16:9) @ 60/50/30/25/24 fps
  - 4K (8:7) @ 60/50/30/25/24 fps
  - 4K (9:16) @ 60/50/30/25 fps
  - 4K (16:9) @ 120/100/60/50/30/25/24 fps
- Stabilization System: HyperSmooth 6.0 Advanced Stabilization

Shipping List

- **Handheld Gripper Kit**

DOBOT

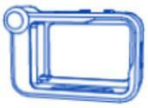
便携式通用数据采集平台 PocketGo
末端夹爪套件发货清单



末端夹爪
End Gripper

Gopro hero 12 相机
Gopro hero 12 camera

Gopro hero 12 广角镜头
Gopro hero 12 wide-angle lens



GoPro hero 12 媒体组件
Gopro hero 12 media mod

视频采集卡采集盒
Video capture box

Type-C usb 数据线
Type-C USB cable

相机线 Micro HDMI转HDMI
Camera cable Micro HDMI to HDMI

电动夹爪延长线
Electric gripper extension cable

快速使用说明书
Quick Start User Manual

机械臂与电动夹爪转接安装板
Cobot and electric gripper
adapter mounting plate

硬夹爪

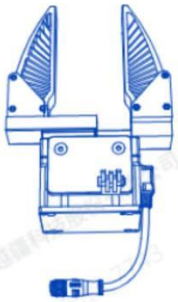
配件/Accessories

圆柱头内六角螺丝 Cylinder head hexagon screw	内六角沉头螺丝 Hexagonal countersunk screw	六角螺母 Hexagonal nut	内六角扳手 Internal hexagonal wrench	防滑胶带 Anti slip tape	魔术贴束线带 Velcro binding tape
M3*35×2 M4*12×6 M4*16×4 M5*20×2	M6*8×6	M3×2 M5×2	2.5mm×1 3mm×1 4mm×1	2.54cm*1m×1	2cm*3m×1
					

- **End-Effector Gripper Kit**

DOBOT

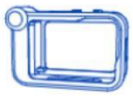
便携式通用数据采集平台 PocketGo
末端夹爪套件发货清单



末端夹爪
End Gripper

Gopro hero 12 相机
Gopro hero 12 camera

Gopro hero 12 广角镜头
Gopro hero 12 wide-angle lens



GoPro hero 12 媒体组件
Gopro hero 12 media mod

视频采集卡采集盒
Video capture box

Type-C usb 数据线
Type-C USB cable

相机线 Micro HDMI转HDMI
Camera cable Micro HDMI to HDMI

电动夹爪延长线
Electric gripper extension cable

快速使用说明书
Quick Start User Manual

机械臂与电动夹爪转接安装板
Cobot and electric gripper adapter mounting plate

硬夹爪

配件/Accessories

圆柱头内六角螺丝 Cylinder head hexagon screw	内六角沉头螺丝 Hexagonal countersunk screw	六角螺母 Hexagonal nut	内六角扳手 Internal hexagonal wrench	防滑胶带 Anti slip tape	魔术贴束线带 Velcro binding tape
M3*35×2 M4*12×6 M4*16×4 M5*20×2	M6*8×6	M3×2 M5×2	2.5mm×1 3mm×1 4mm×1	2.54cm*1m×1	2cm*3m×1

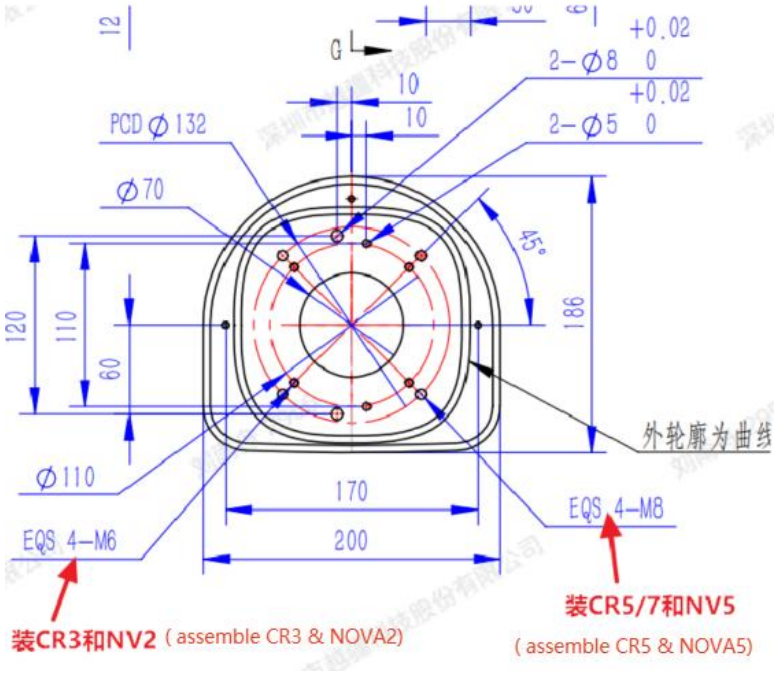
- **Mini AI Reasoning Computer (Standard Configuration)**

Item	Model	Qua	Remark

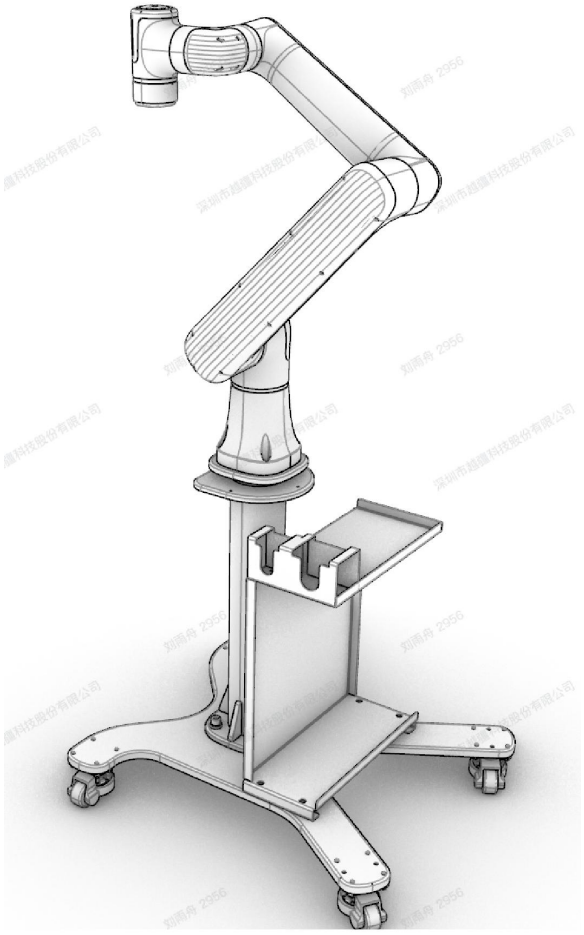
		ntity	
Processor	Intel Core i9 series	1	
TDP	35 W	-	*Based on actual power consumption
Graphics Card	NVIDIA RTX Series: Computational Power Specifications (Minimum Delivery Standards): CUDA Core Count ≥ 4608 cores Memory Capacity ≥ 11GB GDDR6 Memory Bus Width ≥ 192-bit	1	*Due to fluctuations in host machine prices and market conditions, the actual configuration and model may vary. The actual product shall prevail. However, the host machine's performance is sufficient to support edge-side model inference and complete the kit's functions.
Memory	32GB DDR5 per stick	2	64GB total
Disk Storage	2T SSD	1	
SIze	168×168×139mm	-	*Subject to actual conditions
Image		-	

● Mobile Stand

○Size: 680 × 600 × 664 mm (L × W × H)
○Equipped with four swivel casters
○Compatible with: Nova 5, Nova 2, CR3, CR5, CR7



Note: CR10 and higher models cannot be installed due to insufficient mounting space.



Software Platform

1. Data Collection Module

- **Function Description:**

  ○ Supports dynamic recording of task videos from the handheld data acquisition terminal, including map-building videos, gripper calibration videos, and task collection videos.

  ○ Provides QR code scanning for quick access to installation and calibration guides.

  ○ Supports standardized data procedures, including video naming conventions, path management, SLAM trajectory generation, and quality assessment.

2. SLAM Localization and Trajectory Generation Module

- **Core Tools:**

  ○ **ORB-SLAM3 Integration:** Enables high-precision localization and trajectory computation based on environmental feature points.

  ○ **Map Construction:** Initializes scene maps using desktop visual calibration boards; supports dynamic environment updates.

- **Trajectory Correction:** Detects trajectory breaks and feature point losses; offers data cleaning suggestions.

- **Application Scenarios:** Dynamic position correction during data collection; autonomous inference path planning.

3. Calibration Tools Module

- **Supported Features:**
  - **Gripper Calibration:** Calibrates gripper opening and force parameters using continuous open-close motion videos.
  - **Camera Calibration:** Configures video capture card and host interfaces to ensure camera media component synchronization.
  - **Environmental Calibration:** Initializes SLAM scenes using visual calibration boards.
- **Calibration Process:** Scan QR codes to access instructional videos and record standardized calibration data.

4. Data Processing and Quality Assessment Module

- **Functional Tools:**
  - **Data Standardization:** Renames video files per protocol and generates Zarr compressed packages.
  - **Trajectory Visualization:** Visualize trajectory maps in the `batch/logs/trajectory` directory.
  - **Quality Evaluation:** Analyzes SLAM success rates and identifies low-quality data (e.g., insufficient features, lighting interference).
- **Output:** Supports deletion of abnormal trajectory files to enhance the quality of training datasets.

5. Deep Learning Training and Inference Module

- **Core Features:**
  - **Model Training:** Supports training of diffusion models using the PyTorch framework; accelerated by CUDA 11.8 and cuDNN.
  - **Autonomous Inference:** Performs robotic tasks such as grasping and placing using pretrained models (requires >2GB GPU memory).
  - **Model Deployment:** Provides sample models (e.g., `clean_desk_ckpt`) with secondary development APIs.
- **Training Environment:** Ubuntu 20.04/23.10, Anaconda virtual environment (Python 3.9), Docker image with ORB-SLAM3.

6. Robotic Arm Control Module

- **Control Commands:**
  - **Motion control:** point-to-point movement, speed ratio adjustment, homing calibration, suction/gripper toggle.
  - **Remote mode:** Real-time communication between robotic arm and host via IP address configuration.

- **Integrated Tools:**
  ○ **DobotStudioPro:** Records initial joint positions and sets up a virtual control panel (axis coordinates, angle, and speed control).
  ○ **Script Control:** Supports Python scripts (e.g., `04_0_init_joints.py`) for automated initialization and inference.

7. Communication & Interface Module

- **Communication Protocols:**
  ○**TCP/IP:** Remote control of the robotic arm via host connection**.**
  ○**Docker Interface:** Loads preset images (e.g., `chicheng/openicciatest`) to support both offline and online deployment.
- **Hardware Interfaces:** Type-C charging port, USB 3.0 host port, video capture card connection.

8. System Configuration & Development Support Module

- **Environment Configuration:**
  ○**Operating System:** Ubuntu 20.04 / 23.10
  ○**Dependencies:** CUDA 11.8, cuDNN, Anaconda, Docker, ExifTool
- Development Support:
  ○ **Python demo scripts:** OpenCV-based scripts for data processing, model training, and inference
  ○ **Halcon Integration:** Supports Qt-based visual algorithm extensions (e.g., defect detection, image stitching)

9. Safety and Exception Handling Module

- **Safety Guidelines:**
  ○**Operation Warnings:** During SLAM, restrict camera rotation to ±90° (RZ axis); avoid plain-color backgrounds to ensure sufficient environmental texture.
  ○**GPU Memory Management:** Training requires ⩾11GB GPU memory (RTX 4090 recommended); inference requires ⩾2GB.
- **Exception Handling:**
  ○ **Q&A Support:** Troubleshooting for SLAM failures, poor data quality, insufficient memory, etc.
  ○ **Log Recording:** Stores detailed runtime logs in the `batch/logs` directory.

Note:
All modules are built on the PocketGo hardware platform, including handheld and robotic arm-mounted terminals, enabling a fully closed-loop data workflow — from environment calibration, data acquisition, and model training to autonomous robotic inference**.**