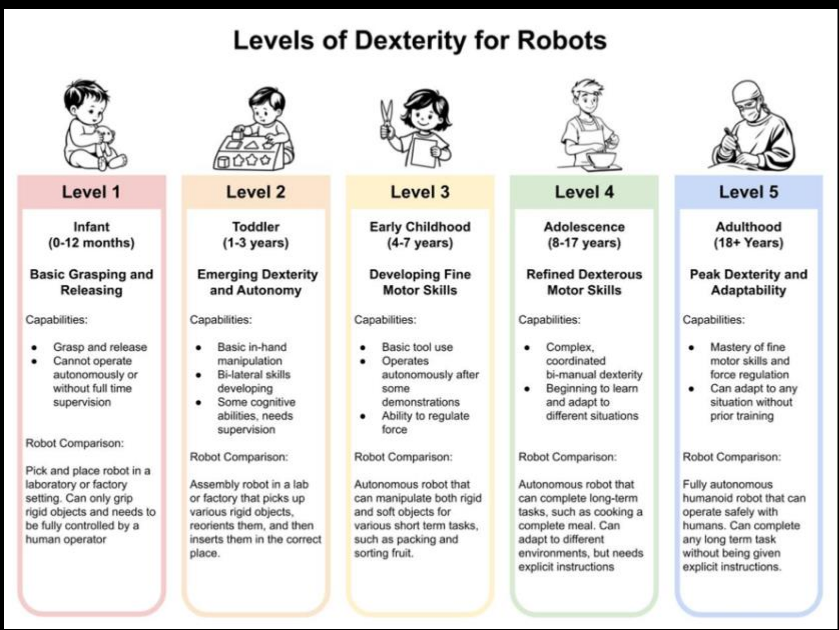
The Future of Humanoid Manipulation

Introduction

Humanoids are set to make a significant impact on the world in the next five years. One of the most crucial aspects of their development is manipulation. Recent advancements in AI make solving humanoid hand dexterity challenges feasible within this decade. This is where the true value of humanoids lies, as 98% of human jobs rely on hand usage. However, current humanoid hand development remains weak.

Levels of Dexterity for Robots

To better define and track progress in robotic hand dexterity, we classify manipulation capabilities into five levels:



1. **Level 1: Infant (0-12 months) - Basic Grasping and Releasing**
   - Capabilities: Simple grasp and release, unable to operate autonomously.
   - Robot Comparison: Pick-and-place robots requiring full human control.
2. **Level 2: Toddler (1-3 years) - Emerging Dexterity and Autonomy**
   - Capabilities: Basic in-hand manipulation, early bi-lateral skills, minimal cognitive ability.
   - Robot Comparison: Assembly robots handling rigid objects with some reorientation.
3. **Level 3: Early Childhood (4-7 years) - Developing Fine Motor Skills**
   - Capabilities: Basic tool use, some autonomous operation, early force regulation.
   - Robot Comparison: Autonomous robots sorting rigid and soft objects.
4. **Level 4: Adolescence (8-17 years) - Refined Dexterous Motor Skills**
   - Capabilities: Complex bi-manual dexterity, adapting to different situations.
   - Robot Comparison: Robots capable of long-term tasks, requiring explicit instructions.
5. **Level 5: Adulthood (18+ years) - Peak Dexterity and Adaptability**
   - Capabilities: Mastery of fine motor skills, real-time adaptability without prior training.
   - Robot Comparison: Fully autonomous humanoids capable of safely operating with humans and executing long-term tasks without explicit guidance.

To push humanoid dexterity to higher levels, the following are our approaches.

1. Hardware

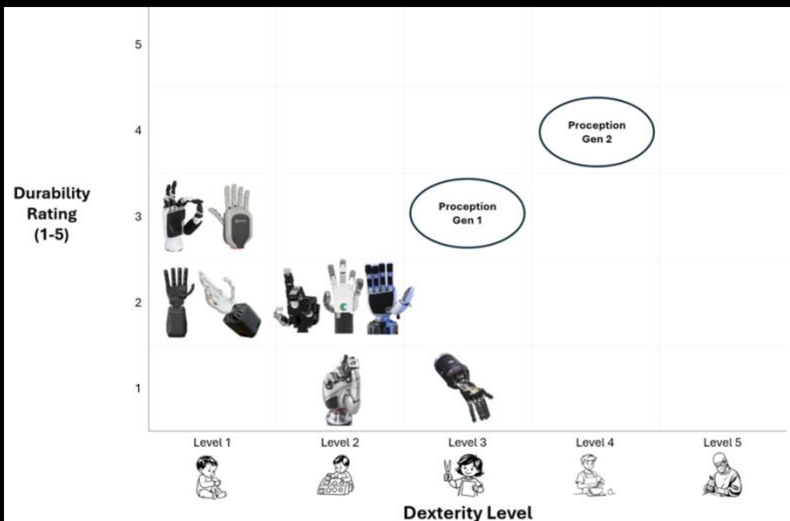
Hand is the foundation of humanoid manipulation. Current robotic hands struggle to balance durability, strength, form factor and dexterity. Many existing designs focus on one aspect while sacrificing others, leading to suboptimal performance in real-world applications. Our benchmarking research reveals key limitations in the current market:

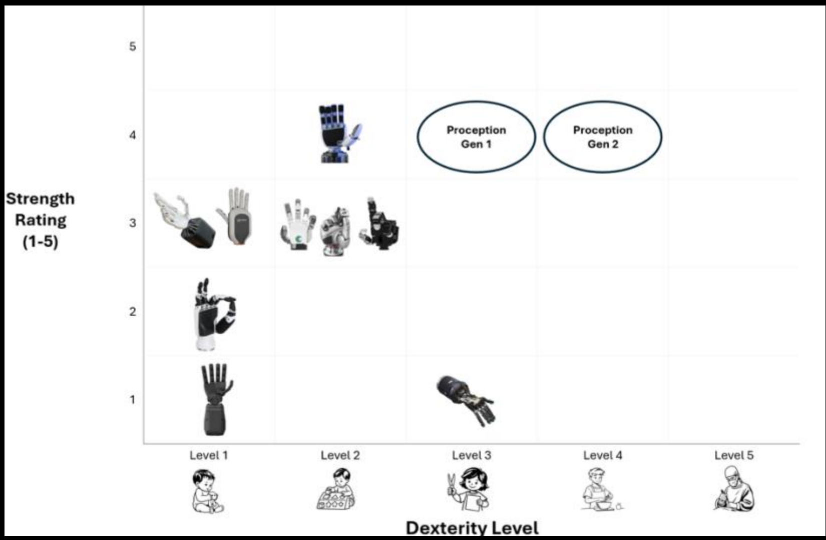
- **Durability**: Most robotic hands are fragile and unable to withstand prolonged real-world usage.
- **Strength**: Many hands lack sufficient grip force to handle heavy or complex objects.
- **Form factor**: Most hands are either too bulky to use human tools, or too different from human hand form (gripper/4 fingers, etc.)
- **Dexterity**: Few solutions achieve fine motor skills comparable to human hands.

Why We Are Developing Our Own Hand Hardware

Given these limitations, we have chosen to develop our own robotic hand hardware to solve the manipulation problem effectively. Our approach prioritizes:

- **A balance of durability, strength, form factor, and dexterity** to create a versatile, high-performance hand.
- **Advanced material and actuator selection** for improved robustness and lifespan.
- **Custom tendon-driven mechanisms** to enable natural hand movements and force control.





2. Data

Our key philosophy is that data collection is one of the biggest bottlenecks in robotic learning. Traditional teleoperation methods are extremely inefficient in terms of scaling data and consuming capital.

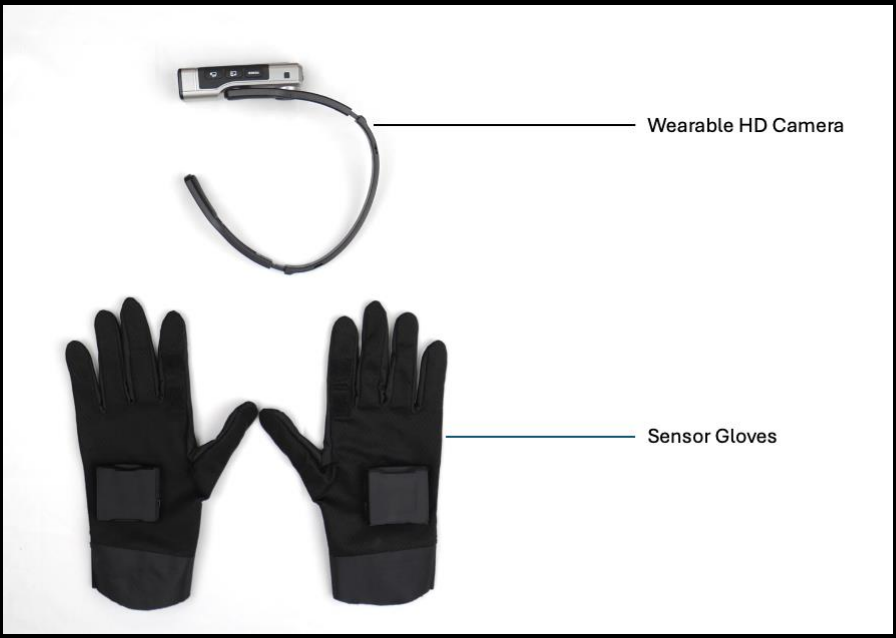
The most common practice for data collection in this industry is tele-operation. It's heavily constrained:

- **Most companies have very few humanoids (mostly less than 100)**
- **Humanoids are expensive to build and operate (for data collection purposes)**
- **Tele-operating humanoids are difficult and more than 100 times slower to collect data than human.**

Overall, we believe tele-operation based data collection is a new form of brute-force and it's extremely inefficient for capital and time.

Our unique hand design includes wearable skin sensors, which enabled a much more efficient data collection scheme.

Data will be collected by human operators only, without constraints of humanoids. Operators will wear HD cameras on their head and sensor glove as they do their jobs. Their hand pose, tactile feedback, and their field of view videos will be logged.



Our data strategy can be summarized as follows

- **Developing scalable data collection methodologies** that minimize the need for expensive human teleoperation.
- **Leveraging simulation environments** to generate high-quality training data efficiently.
- **Integrating real-world sensor feedback** to continuously refine and improve manipulation performance.
- **Avoiding brute-force teleoperation**, which relies on expensive robots and multiple shifts of operators, making it unsustainable for large-scale data collection.

3. AI

The core of our AI system revolves around seamlessly integrating human tactile intuition with robotic dexterity. Our approach leverages a dual-glove system—one pair for a human operator and another for a robotic counterpart—each embedded with high-fidelity tactile and position sensors. This setup enables real-time data acquisition of human tactile interactions, forming the basis for an advanced learning model that optimizes the robotic hand's actuation.

AI components are deeply embedded within our development roadmap, structured into four key areas:

- **Foundational Lower Cortex Model - Tactile-Actuation Learning Loop**

  By capturing rich tactile data from human demonstrations, our AI models learn the complexities of touch, grip force modulation, and dynamic adaptation to various objects and surfaces. This enables the generalization of learned behaviors to previously unseen objects, forming a foundational model for dexterous in-hand manipulation.

  - The collected data is used to train neural networks that map sensor inputs to precise actuator responses. This ensures the high DOF tendon-driven robotic hand achieves high precision with human-like dexterity
  - Reinforcement Learning (RL) techniques further refine the control policies by minimizing errors and enhancing adaptability to novel interaction
  - Additional vision data, combined with large language models (LLMs) allow for scenario labeling, task identification, and contextual reasoning. The key characteristics of each interaction are tokenized and leveraged as modulation parameters to adjust grasping behavior dynamically.

- **Digital Twin - Closing Gap between Simulation and Real Hardware**

  - A critical aspect of our AI framework involves leveraging real-world tactile data to bridge the disparity between simulated models and physical hardware performance. This includes modeling hardware variability to refine the distribution of disturbed parameters.
  - By continuously integrating feedback from real-world interactions, our self-improving models refine simulation parameters and control strategies, ensuring greater alignment with physical execution.
  - This approach enhances predictive accuracy and improves robotic performance, making the system more robust to variations in hardware build.

- **AI Agent – Profession-specific Dexterity Fine-Tuning**
  - By collecting the operational data from specific professional fields, we refine our foundational model while also enabling fine-tuned applications for specialized roles for our customers.
  - We develop personalized AI agents that tailor robotic dexterity to industry-specific requirements, ensuring optimal adaptation for various use cases, from surgical precision to industrial manipulation
  - This approach enables a seamless transition from generalized learning to highly customized robotic behavior for end-user applications
- **AI-driven Generative Design**
  - Beyond optimizing control policies, we leverage AI-driven generative design techniques to accelerate the permutation of the robotic hand design to enhance the physical structure and performance.
  - This involves using the digital twin model and foundational control policy to expedite the design space exploration in simulation, reducing the physical design iteration and shortens the design-to-deployment timeline.

By integrating advanced AI methodologies across learning, simulation, professional adaptation, and generative design, our system pioneers a new paradigm in human-inspired robotic dexterity. Our approach not only enhances real-world robotic interaction but also accelerates the evolution of robotic hands toward unprecedented levels of adaptability and efficiency.

Conclusion

Achieving Level 5 dexterity is critical for humanoids to reach their full potential. The recent breakthroughs in AI and robotics provide a strong foundation for rapid advancements in this field. With focused development, humanoids will revolutionize industries by performing complex tasks that currently require human hands. Our proprietary hardware, data, and AI-driven approach will be key enablers in bridging the gap between existing robotic hands and the dexterity required for true humanoid autonomy.