

مرشد الهجدة: روبوت بشري للاستقبال يتكلم العربية

Murshid Al-Hajda: an Arabic-speaking reception humanoid — تقرير تقني بأسلوب الأوراق البحثية، مسودة للمراجعة الداخلية.

مركز أبحاث الهجدة · روبوت الهجدة (HUM-001) · <https://lab.alhajda.com/research/humanoid/paper> · 2026-10-29

Abstract

We present the design of «Murshid Al-Hajda», an Arabic-speaking reception and guide humanoid built on the open-source Berkeley Humanoid Lite platform. The robot pairs a 22-actuator, 3D-printed body with a three-layer control stack: Mohsen, an Arabic language model, for dialogue and task planning; a skills layer based on open vision-language-action models (OpenVLA, $\pi 0/\pi 0.5$, GR00T N1.7) trained from teleoperated demonstrations with LeRobot; and reinforcement-learning locomotion policies, all bounded by an independent safety supervisor. We describe a staged job roadmap (reception, in-building inventory and delivery, light assembly, elder-care assistance) with explicit hardware, safety limits and readiness levels, a bill of materials of about USD 5,540 per unit (USD 6,647 with 20% contingency) and a five-unit pilot batch at about USD 7,412 per unit. In simulation we reproduce stable walking of the open Unitree G1 model with a published policy, and report an honest CPU-only baseline for training our own policy with Augmented Random Search, together with the GPU budget needed to train a deployable policy with MuJoCo Playground. No hardware has been built yet; this is a draft for internal review.

الملخص

نقدم تصميم «مرشد الهجدة»، روبوتًا بشريًا للاستقبال والإرشاد يتكلم العربية، مبنياً على منصة Berkeley Humanoid Lite المفتوحة. يجمع الروبوت جسداً مطبوعاً مزوداً بـ 22 مشغلاً مع مكس تحكم من ثلاث طبقات: «محسن» للحوار والتخطيط بالعربية، وطبقة مهارات مبنية على نماذج رؤية-لغة-فعل مفتوحة تُدرَّب من عروض بالتحكم عن بعد، وسياسات مشي بالتعلم المعزز، يحدّها جميعاً مشرف سلامة مستقل. نعرض خارطة وظائف متدرجة بعتها وحدود سلامتها وجاهزيتها الصادقة، وقائمة مواد بنحو 5,540 دولاراً للوحدة، ونتائج محاكاة أولية. لم نبني أي عتاد بعد؛ هذه مسودة للمراجعة الداخلية.

1. المقدمة

بيعت نحو 7,000 روبوت بشري كامل الحجم عالمياً في 2025، ومعظمها للبحث، والأسعار تنخفض بسرعة³¹. في المقابل تبني الشركات الكبرى للإنجليزية أولاً، بينما تجرّب المملكة روبوتات إرشاد صوتية في مطار المدينة والحرمين^{29,28}. نسال: ما أصغر روبوت بشري مفيد يستطيع فريق صغير في السعودية بناءه، وكيف يتوسع لوظائف أخرى دون إعادة تصميم؟

مساهمات هذا التقرير: (1) مواصفة «مرشد الهجدة» على عتاد مفتوح تجارياً؛ (2) بنية «جسد واحد ووظائف كثيرة» تربط نموذجاً لغوياً عربياً بنماذج VLA مفتوحة وسياسات حركة؛ (3) خارطة وظائف بمستويات جاهزية صريحة؛ (4) خطة إنتاج وتكلفة لنموذج ودفعة من 5 وحدات؛ (5) نتائج محاكاة قابلة لإعادة الإنتاج وقياس صادق لحدود الحوسبة.

2. أعمال ذات صلة

- روبوتات مفتوحة: Berkeley Humanoid Lite بأقل من 5,000 دولار وعلب تروس دائرية مطبوعة ونقل مباشر من المحاكاة للواقع.⁷
- تعلم الحركة: MuJoCo Playground يدرّب مشي Berkeley Humanoid في أقل من 15 دقيقة على بطاقتي RTX 4090¹⁴، ووصفة منشورة تدرب مشيًا قابلاً للنقل في 15 دقيقة على بطاقة واحدة¹⁶، وARS يثبت أن سياسات خطية بالبحث العشوائي تنافس على مهام MuJoCo¹⁷.
- نماذج VLA: OpenVLA¹ و $\pi^0/\pi^{0.5}$ ² وGR00T N1.7 للروبوتات البشرية³، ومنظومة LeRobot للبيانات والتدريب⁴.
- النشر الصناعي: Digit في اللوجستيات²³، وFigure في BMW²⁴، وUBTech في Airbus²⁵؛ والمنزل ما زال صعبًا (56%) نجاح²⁶.

3. حالة الاستخدام والمتطلبات

- البيئة: مداخل جامعات ومعارض ومتاحف وصالات خدمة؛ أرض مستوية؛ حشود متوسطة.
- الوظائف: ترحيب، إجابة عربية من قاعدة معرفة المكان، إشارة للاتجاه، مرافقة قصيرة، تحويل لموظف.
- المعلم: يمشي 10 أمتار ويحجب 20 سؤالاً عربيًا صحيحًا؛ وقوف ساعة؛ إيقاف طارئ أقل من 100 مللي ثانية.
- القيود: لا تخزين لصور الوجوه أو التسجيلات، ولوحة واضحة بأن الزائر يتحدث مع آلة.

4. تصميم العتاد

نعتمد Berkeley Humanoid Lite (نحو 1 م، 16 كجم): 10 مشغلات M6C12 و12 مشغل 5010، كلها بعلب تروس دائرية مطبوعة بـPLA ومتحكمات B-G431B-ESC1 ومشغلات AS5600، وحاسوب صغير وبطارية LiPo 6S وحساس BNO085، بقائمة مواد رسمية 4,350.59 دولارًا^{7,6}. نضيف للمرشد Jetson Orin NX 16GB وكاميرتين وسماعة¹² ومصفوفة ميكروفونات وزر إيقاف طارئ (تقدير)، فتصبح الوحدة 5,540 دولارًا.

5. البرمجيات: بنية «جسد واحد، وظائف كثيرة»

- الطبقة 3: محسن (نموذج لغوي عربي) مع Whisper للتعرف على الكلام³⁰؛ يخطط ويستدعي المهارات بالاسم.
- الطبقة 2: مكتبة مهارات؛ كل مهارة = سياسة + شروط + كاشف نجاح + غلاف سلامة + اختبارات. المهارات تُتعلم من عروض بالتحكم عن بعد بصيغة LeRobot⁴ وتُكثّر في المحاكاة، ثم تُضبط سياسات صغيرة أو نماذج VLA بـLoRA^{2,4}.
- الطبقة 1: سياسة مشي RL بتردد 50 هرتز ومتحكم PD في المشغل¹⁴.
- المشرف: دائرة وبرنامج مستقلان بحدود سرعة وقوة ومنطقة، يتجاوزان كل الطبقات.
- خارطة الوظائف: استقبال (نبدأ) ← جرد وتوصيل داخلي ← تجميع خفيف ← مساعدة كبار السن (بحث فقط، ISO 13482²¹).

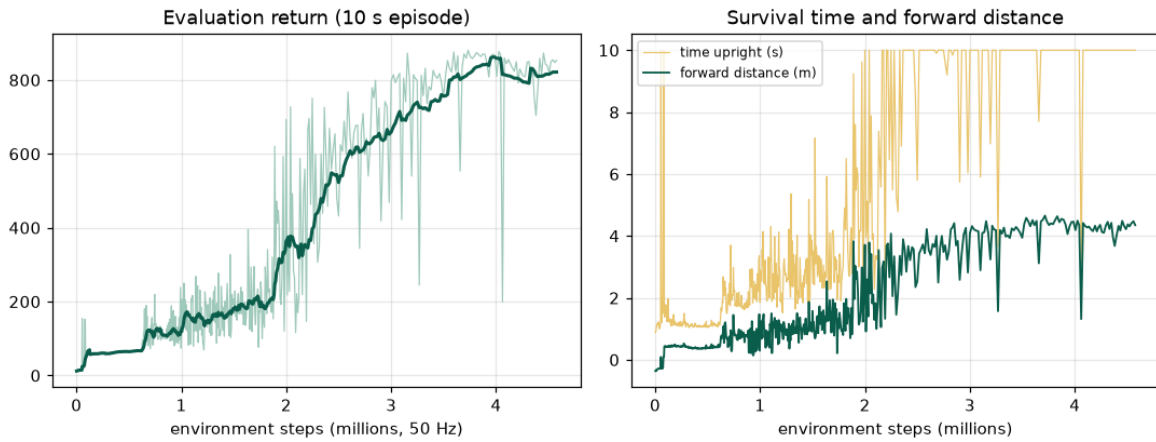
6. المحاكاة والنتائج

- إعادة إنتاج مشي منشور: نموذج Unitree G1 مع سياسته المدربة المفتوحة يقف ويمشي ويستدير ويتوقف في MuJoCo دون سقوط؛ متوسط السرعة 0.477 م/ث مقابل أمر 0.5، وأدنى ارتفاع للحوض 0.76 م²⁰. السياسة ليست من تدريبنا.

- تدريب سياستنا على المعالج: نموذج Berkeley Humanoid (12 مفصلاً، 16 كجم) بإعدادات MuJoCo Playground نفسها ($kp=35$ ، خطوة 2 مللي ثانية، تحكم 50 هرتز)¹⁵، وخوارزمية ARS بسياسة خطية وساعة مشية¹⁷، بمكافأة تتبع سرعة 0.5 م/ث.
- السياسة الخطية بعد 610 تكرارًا (4,575,731 خطوة تحكم ≈ 25.4 ساعة من الزمن المحاكى) على 2 نواتي معالج:
- متوسط البقاء واقفاً 16.26 ثانية من 20، و 7 من 10 حلقات اكتملت دون سقوط، ومتوسط التقدم للأمام 7.103 م بسرعة متوسطة 0.371 م/ث (الأمر 0.5 م/ث).
- بصراحة: هذه أول سياسة مشي من تدريبنا نحن، لكنها ليست جاهزة للواقع: سياسة خطية بخطوات قصيرة، تنجرف جانبياً (نحو 1.4 م في 20 ثانية في الحلقة المعروضة)، ودُرِّبَت بلا ضوضاء حساسات ولا دفعات ولا عشوائية في الكتل والاحتكاك — وكلها شروط ضرورية للنقل من المحاكاة للواقع. وفي تجربتين سابقتين وجد التدريب «ثغرات» في المكافأة (الوقوف بلا حركة، ثم الاهتزاز لخداع قراءة السرعة اللحظية) فأصلحنا المكافأة؛ سجلات التجارب الثلاث محفوظة.
- قياس الحوسبة: MuJoCo Playground على هذا المعالج (بلا بطاقة رسومية) أعطى نحو ≈ 71 خطوة/ثانية، بينما وصفة Playground الرسمية تحتاج 150,000,000 خطوة بـ 8192 بيئة متوازية¹⁵ — أي أسابيع على المعالج. MuJoCo العادي أسرع بكثير على المعالج ($\approx 5,850$ خطوة تحكم/ثانية لكل نواة) لكنه بلا توازٍ ضخم.
- ما يلزم: ساعة إلى ثلاث ساعات بطاقة رسومية لتدريب سياسة قابلة للنقل بوصفة Playground (المصدر: أقل من 15 دقيقة على بطاقتي RTX 4090¹⁴؛ تقديرنا لبطاقة أقدم أبداً بعدة أضعاف). الخيارات المجانية: Kaggle يعطي بطاقة P100 بحصة نحو 30 ساعة أسبوعياً¹⁸، و Colab المجاني يعطي بطاقات غير مضمونة لجلسات حتى 12 ساعة¹⁹. لا نحتاج حوسبة مدفوعة.

(الفيديو على الصفحة: lab.alhajda.com/research/humanoid/paper/)

Berkeley Humanoid, ARS linear policy, CPU only (2 cores), Al-Hajda Lab, 2026-10-29



منحنى تعلم سياستنا على المعالج (ARS): مكافأة التقييم والمسافة المقطوعة لكل تكرار.

7. الإنتاج والتكلفة

- الوحدة: 5,540 دولارًا، و 6,647 دولارًا ($\approx 24,930$ ريال) مع احتياط 20%، قبل الشحن والجمارك والضريبة^{12,6}.
- دفعة 5 وحدات مع طابعتين¹¹ وعدة ومنصة اختبار وغيار 10% واحتياط 15%: $\approx 7,412$ دولارًا للوحدة.
- الجدول: نحو 10 أسابيع من وصول القطع للنموذج الأول؛ والطباعة هي عنق الزجاجة في الدفعة.

8. القيود (بصراحة)

- لم نبني أي عتاد؛ كل أرقام الأداء الحقيقي في هذا التقرير منقولة من مصادرها لا من تجاربنا.

- سياسة المشي التي تمشي جيدًا في عرضنا ليست من تدريبنا؛ وسياستنا المدربة على المعالج ما زالت بعيدة عن مشي قابل للنقل للواقع.
- نموذج المحاكاة هو Berkeley Humanoid (الأخ الأكبر) لا Lite نفسه؛ نحتاج نموذج MJCF للـ Lite قبل النقل.
- لم نختبر نماذج VLA على عتادنا، ولم نقس دقة محسن في أسئلة الأماكن بالصوت في بيئة صاخبة.
- أسعار أمريكية من مصدر واحد؛ زمن الطباعة والتجميع للدفعات تقديرات.
- لا يوجد معيار سلامة منشور للروبوتات المتوازنة ديناميكيًا²²؛ معاييرنا مقترحة.

9. العمل التالي

1. تدريب سياسة المشي بوصف Playground على بطاقة مجانية (Kaggle) ونقلها لنموذج Lite.
2. مكتبة مهارات v0 للمرشد (6 مهارات كلام وإيماءة) واختبار محسن صوتيًا.
3. بعد موافقة إلياس: بناء مشغل واحد واختباره، ثم النموذج الأول.
4. تجربة ميدانية لأسبوعين بإشراف، وقياس نسبة الإجابات الصحيحة والأعطال لكل ساعة.

10. السلامة والأخلاقيات

- لوحة واضحة: «تتحدث مع روبوت»، ولا تخزين لصور الوجوه أو الأصوات.
- مشرف بشري وحاجز وزر إيقاف طارئ في كل تجربة.
- لا مساعدة جسدية لأي شخص قبل تحقق قانوني وتنظيمي.
- الالتزام بأطر سدايا للذكاء الاصطناعي وحماية البيانات³¹.

حالة هذا التقرير

مسودة لمراجعة John. لم يُرسل لأي مجلة أو مؤتمر أو أرشيف. النشر على arXiv يحتاج موافقة إلياس وحسابًا (وقد يحتاج تزكية في قسم CS.RO)، ويقترح تحويله للإنجليزية كاملاً مع الإبقاء على نسخة عربية.

المصادر

1. OpenVLA: An Open-Source Vision-Language-Action Model (7B, 970k episodes, 64×A100 TRI و Berkeley و Stanford وآخرون: 2024-06-13) — <https://openvla.github.io/>
2. openpi — π_0 / π_0 -FAST / $\pi_0.5$ open weights, Apache-2.0, GPU requirements — <https://github.com/Physical-Intelligence/openpi> (2026-10-29)
3. Isaac GR00T N1.7 — open VLA for humanoids (3B, ~28 Hz on Jetson AGX Thor, SONIC whole-body control) — NVIDIA: <https://github.com/NVIDIA/Isaac-GR00T> (2026-10-29)
4. LeRobot — datasets, teleoperation, ACT/Diffusion/ π_0 /GR00T/SmolVLA policies (Apache-2.0) — Hugging Face: <https://github.com/huggingface/lerobot> (2026-10-29)
5. Berkeley Humanoid Lite BOM (Google Sheet: actuators \$188.69 / \$136.36, robot total: UC Berkeley Hybrid Robotics \$4,350.59, Amazon/Digikey links) — https://docs.google.com/spreadsheets/d/1AQEHcH_nPkXYfor2-h7bwNIUMmsePtAm53epnsWgZXc/edit?usp=sharing (2026-10-29)
6. Demonstrating Berkeley Humanoid Lite (arXiv:2504.17249) — under \$5,000, cycloidal 3D-printed gearboxes, zero-shot sim-to-real — <https://arxiv.org/abs/2504.17249> (2025-04-24)
7. Bambu Lab P1S — \$579 typical price (\$509 on sale), 256×256×256 mm — <https://us.store.bambulab.com/products/p1s> (2026-10-29)
8. ToddlerBot BOM — Jetson Orin NX 16GB \$955, cameras, speaker — <https://motari.io/r/toddlerbot> (2026-10-01)
9. MuJoCo Playground (Berkeley Humanoid < 15 min, G1 < 30 min on 2× RTX 4090) — Google DeepMind وآخرون (RSS 2025) — <https://arxiv.org/abs/2502.08844> (2025-02-12)

- mujoco_playground — BerkeleyHumanoidJoystick env, PPO config 150M steps / 8192 envs :Google DeepMind .10
(Apache-2.0) — https://github.com/google-deepmind/mujoco_playground (2026-10-29)
- Learning Sim-to-Real Humanoid Locomotion in 15 Minutes (one RTX 4090) — <https://arxiv.org/abs/2512.01996> :arXiv .11
(2025-12-01)
- Simple random search provides a competitive approach to reinforcement learning (ARS, :Recht, Guy, Mania .12
arXiv:1803.07055) — <https://arxiv.org/abs/1803.07055> (2018-03-19)
- Efficient GPU Usage Tips — free P100, ~30 h/week quota — <https://www.kaggle.com/docs/efficient-gpu-usage> :Kaggle .13
(2026-10-29)
- FAQ — free tier GPUs not guaranteed, notebooks up to 12 h — <https://research.google.com/colaboratory/faq.html> :Google Colab .14
(2026-10-29)
- unitree_rl_gym (BSD-3) — G1 model + pretrained policy — https://github.com/unitreerobotics/unitree_rl_gym :Unitree Robotics .15
(2025-07-25)
- ISO 13482:2014 — Safety requirements for personal care robots (to be replaced by ISO/FDIS 13482) — <https://www.iso.org/standard/53820.html> :ISO .16
(2026-10-29)
- ISO/CD 25785-1 — Safety for dynamically stable industrial mobile robots (draft) — <https://www.iso.org/standard/91469.html> :ISO .17
(2026-07-08)
- Agility Unveils Digit 5 (65,000+ hours, 100,000+ totes at GXO) — <https://www.agilityrobotics.com/content/agility-unveils-digit-5-humanoid-robot-built-for-cooperatively-safe-work-at-scale> :Agility Robotics .18
(2026-09-15)
- F.03 Arrives at BMW (logistics parts sequencing, Helix 02) — <https://www.figure.ai/news/f-03-at-bmw> :Figure AI .19
(2026-06-30)
- Airbus Buys Robots From UBTech (Walker S2 pilot) — <https://www.forbes.com/sites/johnkoetsier/2026/01/19/humanoid-robots-building-airplanes-airbus-buys-robots-from-ubtech/> :Forbes .20
(2026-01-19)
- Figure Helix 2.5: home-task success 9% → 56% in unseen homes — <https://www.forbes.com/sites/johnkoetsier/2026/09/21/figures-billion-dollar-physical-ai-bet-delivered-a-6x-jump-in-robot-chore-success/> :Forbes .21
(2026-09-21)
22. صحيفة مال: — <https://maaal.com/news/details/%D8%B1%D8%A6%D9%8A%D8%B3-%D8%A7%D9%84%D9%87%D9%8A%D8%A6%D8%A9-%D8%A7%D9%84%D8%B9%D8%A7%D9%85%D8%A9-%D9%84%D9%84%D8%B7%D9%8A%D8%B1%D8%A7/> — تدشين التشغيل التجريبي للمساعد الذكي «سند» في مطار المدينة
(2026-05-09)
- Saudi Arabia deploys AI-powered multilingual robot for pilgrims at two holy mosques — <https://gulfnews.com/world/gulf/saudi/saudi-arabia-deploys-ai-powered-multilingual-robot-for-pilgrims-at-two-holy-mosques-1.500540029> :Gulf News .23
(2026-05-14)
- Robust Speech Recognition via Large-Scale Weak Supervision (Whisper) — <https://arxiv.org/abs/2212.04356> :Radford وآخرون) OpenAI .24
(2022-12-06)
25. مركز أبحاث الهجرة: — <https://lab.alhajda.com/research/humanoid/> — روبوت الهجرة — المرحلة 1: بحث وتصميم ومحاكاة (55 مصدرًا)
(2026-10-27)