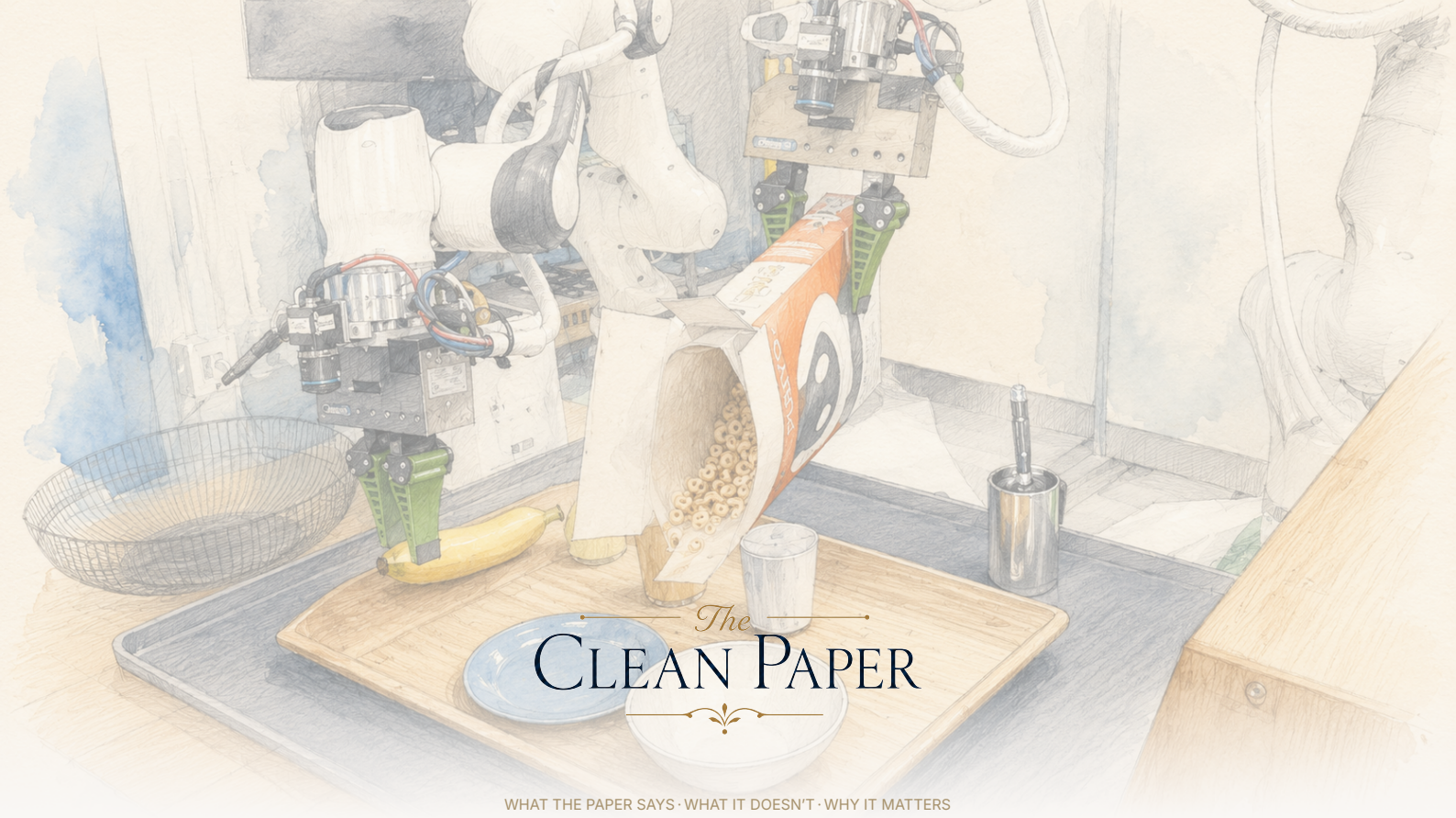


A funksionojnë vërtet më mirë “foundation models” e mëdha për robotë? Përgjigjja e kujdesshme: modestisht po — dhe shumica e studimeve mezi mund ta matin

Toyota Research Institute trajnoi “large behavior models” — politika robotike të paratrajnuara mbi ~1,700 orë manipulimi të larmishëm — dhe i testoi kundrejt politikave një-detyre nga zero me rigorozitet të pazakontë: prova të verbra, të randomizuara, me kampion të madh (~1,800 reale, 47,000+ simulime) dhe statistikë reale. Pas fine-tuning-ut për detyrë, modelet e mëdha performuan më mirë mesatarisht, kërkuan rreth 3–5× më pak të dhëna dhe ishin më robustë kur kushtet ndryshonin; performanca u rrit gradualisht me më shumë pretraining. Por pa fine-tuning nuk i mundën në mënyrë konsistente modelet një-detyre, disa efekte ishin tepër të vogla për t’u parë pa kampione të mëdha, dhe një zgjedhje e zakonshme normalizimi pati më shumë rëndësi se arkitektura. Është mbështetje e matur për drejtimin e robot foundation models — jo robot i përgjithshëm, jo generalist zero-shot dhe jo “kërcim emergjent” — plus një paralajmërim se shumë robotikë mund të jetë duke matur zhurmë.

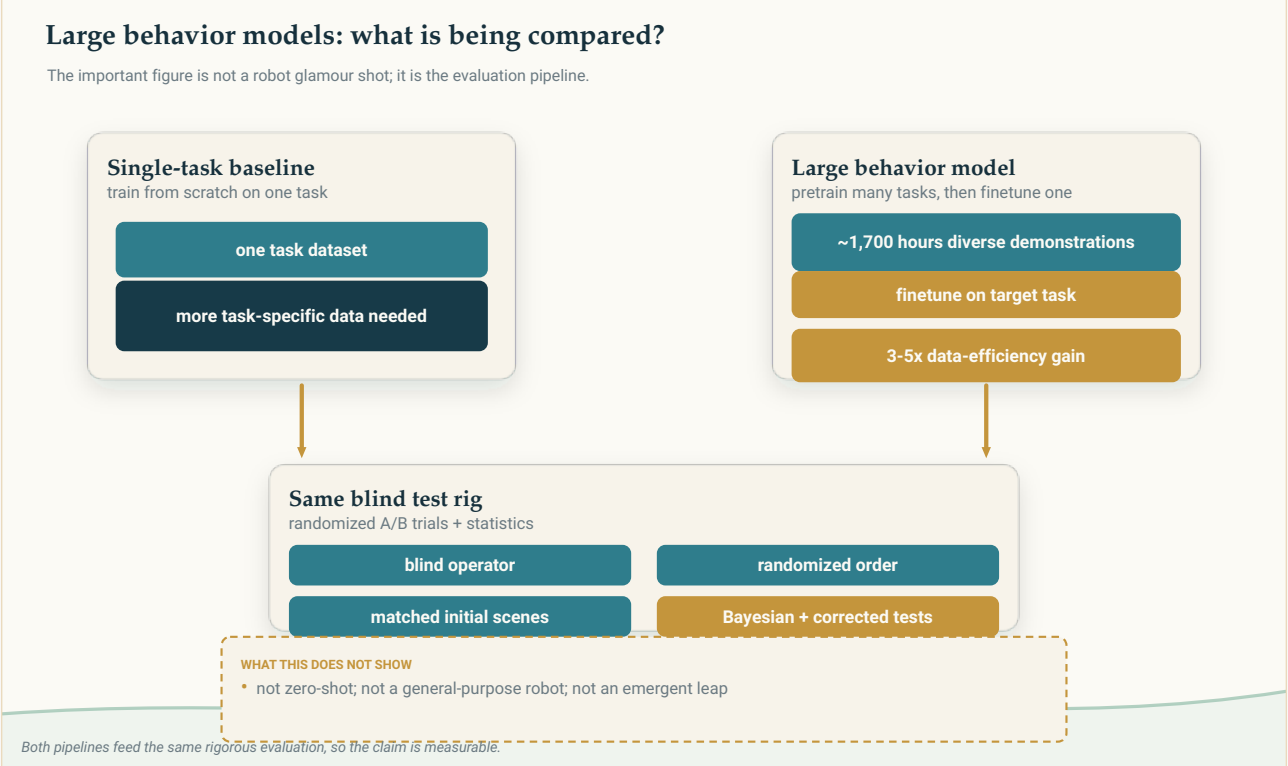
Written by Lucio Vaglio · Cross-checked by Laura Nesso · Edited by Michele Renda · Figures by Laura Nesso and Nova Vela · AI-assisted, human-reviewed

Paper: *A Careful Examination of Large Behavior Models for Multitask Dexterous Manipulation*, Toyota Research Institute Large Behavior Model Team — J. Barreiros, A. Beaulieu, et al.; senior authors incl. R. Ambrus, B. Burchfiel, S. Feng, H. Kress-Gazit (Cornell), R. Tedrake, Science Robotics (2026); preprint arXiv:2507.05331 · DOI: 10.1126/scirobotics.aea6201

Një punim robotike ku tema e vërtetë është ndershmëria

Robotika është në mes të etheve për "foundation models". Ideja, e huazuar nga AI-ja e gjuhës dhe imazhit, është joshëse: në vend që një robot të trajnohet për një detyrë në një kohë, trajno një **"large behavior model" (LBM)** të madh mbi një grumbull të gjerë demonstrimesh dhe merr një sistem me aftësi më të përgjithshme e që përshtatet shpejt. Entuziazmi — dhe investimi — është i madh. Titulli shkruhet vetë: *truri i robotit me përdorim të përgjithshëm ka ardhur*.

Ky punim nga Toyota Research Institute është interesant pikërisht sepse refuzon ta shkruajë atë titull. Kontributi kryesor nuk është një robot më spektakolar. Është një vështrim i fortë ndaj një pyetjeje mashtruesisht të thjeshtë — *a funksionon vërtet më mirë qasja me model të madh, dhe si do ta dinim?* — me një kujdes statistikor që, sipas vetë autorëve, është i pazakontë në fushë. Përgjigjja është një "po, por" vërtet i dobishëm, dhe një paralajmërim se shumë punime robotike mund të jenë duke matur zhurmë.



Të dyja qasjet futen në të njëjtën vlerësim të verbër dhe të randomizuar. Pretendimi nuk është se foundation models të robotëve janë generalistë zero-shot, por se pretraining-u mund të përmirësojë efikasitetin e të dhënave dhe robustësinë kur matet me kujdes.

Çfarë bënë autorët

Ata ndërtuan LBM në një kuptim konkret: **politika visuomotore të bazuara në diffusion** — një Diffusion Transformer që lexon imazhe nga kamerat, një udhëzim të shkurtër gjuhësor dhe pozicionet e nyjeve të robotit, dhe nxjerr seri të shkurtra komandash motorike në 10 Hz. Modelet u **paratrajnuan mbi afërsisht 1,700 orë** demonstrime robotike — mbi 500 detyra të ndryshme të mbledhura brenda laboratorit plus dataset-e publike — dhe pastaj u **fine-tune-uan** për detyra individuale. Krahasimi gjatë gjithë punimit është me një **politikë një-detyre të trajnuar nga zero** vetëm mbi të dhënat e asaj detyre.

Por zemra e punimit është *vlerësimi*, që autorët e trajtojnë si rezultat kryesor. Për të shmangur vetëmashtimin përdorën:

- **A/B testing të verbër dhe të randomizuar** në botën reale — operatori nuk e dinte cilën politikë po testonte dhe rendi randomizohej.
- **Kushte fillestare të kontrolluara dhe të përsëritshme** — operatorët e përputhnin skenën me një overlay imazhi para çdo prove.
- **Shumë prova** — 50 rollout-e reale për detyrë, politikë dhe kusht; 200 për detyrë në simulim. Në total rreth **1,800 rollout-e reale të verbra dhe mbi 47,000 rollout-e simulimi**.
- **Statistikë të mirëfilltë** — vlerësime Bayesian-e të probabilitetit të suksesit dhe teste hipotezash dyshe me korrigjim për krahasime të shumëfishta, jo thjesht shikim të error bars. Madje bënë QA në një të katërtën e provave të vlerësuara nga njerëzit për të matur gabimin e vlerësimit.

[video pending]

Krahasim krah për krah i modeleve duke përgatitur një tavolinë mëngjesi: majtas baseline një-detyre, djathtas LBM. Të dy videot luhen me shpejtësi 1×. Kjo është një detyrë e vlerësuar, jo provë autonomie me përdorim të përgjithshëm.

Kjo makineri matëse është vetë pika. I gjithë punimi argumenton se pa të nuk mund ta dallosh përmirësimin real nga fati.

Çfarë gjetën

Modelet e mëdha të fine-tune-uara i mundin modelet një-detyre të nisura nga zero — mesatarisht. Të mbledhura nëpër detyra, një LBM i paratrajnuar dhe pastaj i fine-tune-uar për një detyrë e tejkaloi në mënyrë të besueshme një politikë të trajnuar nga zero mbi të njëjtën detyrë, si në simulim, ashtu edhe në botën reale, me ndarje statistikisht domethënëse. Në nivel detyre, LBM-ja e fine-tune-uar ishte statistikisht po aq e mirë ose më e mirë pothuajse gjithmonë (3/3 detyra reale, 15/16 në simulim).

Fitorja më e madhe dhe më e qartë është efikasiteti i të dhënave. Një LBM i fine-tune-uar arriti performancë të barabartë me modelin nga zero me rreth **3–5× më pak të dhëna specifike për detyrën**. Në një detyrë reale, shtrimi i tavolinës së mëngjesit, një LBM i fine-tune-uar me vetëm **15%** të demonstrimeve mundi një politikë nga zero të trajnuar me **100%**.

Pretraining-u ndihmon më shumë kur kushtet ndryshojnë. Kur mjedisi i testit u ndryshua qëllimisht nga kushtet e trajnimit ("distribution shift"), avantazhi i LBM-së së fine-tune-uar *u rrit*. Në një grup simulimesh, ajo mundi statistikisht baseline-in në 3 nga 16 detyra në kushte normale, por 10 nga 16 nën distribution shift. Meqë përdorimet reale gjithmonë largohen disi nga kushtet e trajnimit, kjo robustësi mund të jetë gjetja më praktike.

Më shumë të dhëna pretraining ndihmuan në mënyrë të lëmuar. Performanca u rrit vazhdimisht me shtimin e të dhënave — **pa kërcim të papritur ose "emergjencë" aftësish** në shkallët e testuara. E dobishme, e parashikueshme, jo dramatike.

Por historia e generalistit pa fine-tuning nuk qëndroi. Një LBM i paratrajnuar i përdorur *zero-shot*, pa fine-tuning për detyrën, **nuk** i mundi në mënyrë të qëndrueshme politikat një-detyre. Një rrjet i vetëm mund të kryente shumë detyra, por ëndrra "vetëm thuaji çfarë të bëjë" nuk u realizua këtu; autorët ia atribuojnë një pjesë kufizimit të enkoderit të vogël gjuhësor.

Dhe fitimet ishin aq të vogla sa mund të humbnin — ose të shpikeshin — lehtë. Shumë efekte u bënë të dukshme vetëm me kampione më të mëdha se zakonisht dhe teste të kujdesshme. Autorët thonë shprehimisht se, duke parë madhësinë e efekteve dhe zhurmën, **ka rrezik domethënës që shumë punime robotike po matin zhurmë statistikore**. Gjetën edhe se një zgjedhje e zakonshme — si **normalizohen** të dhënat — ndikonte rezultatet më shumë se ndryshimet arkitekturore, dhe një **bug normalizimi** në pretraining doli në pah vetëm *pas* përfundimit të vlerësimeve.

Çfarë do të thotë me gjasë

Leximi i mbrojtshëm: pretraining-u në shkallë të madhe mbi të dhëna të larmishme robotike është përbërës realisht i vlefshëm — kërkon më pak të dhëna për një detyrë të re dhe e bën politikën më të fortë kur bota nuk përputhet me trajnimin. Kjo mbështet drejtimin ku fusha po investon. Por fitimet janë *modeste dhe të kushtëzuara* — kryesisht pas fine-tuning-ut dhe më të qarta në agregat e nën stres — jo ardhja e një roboti të përgjithshëm plug-and-play.

Kuptimi më i qetë dhe ndoshta më i rëndësishëm është metodologjik. Punimi është, në efekt, një vizore matjeje: tregon sa prova duhen realisht për një pretendim të besueshëm për një politikë robotike dhe nënkupton se shumë entuziazëm i publikuar mbështetet në tepër pak të dhëna. Fusha ka më shumë nevojë për këtë korrigjim sesa për një model tjetër.

Çfarë nuk provon kjo

- **Nuk është robot me përdorim të përgjithshëm.** Fitoret demonstrohen për një arkitekturë specifike diffusion të fine-tune-uar për secilën detyrë, në mjedise të kontrolluara dhe nga demonstrime me teleoperim — jo robot autonom që kryen punë arbitrare me komandë.
- **Nuk validon përdorimin zero-shot.** Pa fine-tuning, modeli i madh nuk i mundi në mënyrë konsistente baseline-et një-detyre.

- **Nuk është provë e një "kërcimi emergjent".** Shkallëzimi i përmirësoi gjërat gradualisht; nuk ka diskontinuitet që mbështet rrëfimin "dhe papritur u bë i aftë".
- Numrat janë **relativë dhe laboratorikë**. Normat absolute të suksesit u rregulluan qëllimisht rreth 50% për ta bërë krahasimin të ndjeshëm; nuk matin besueshmërinë në botën reale dhe puna është një arkitekturë nga një laborator.
- Nuk shpjegon **pse** një politikë e caktuar fiton ose dështon; disa detyra ku LBM ishte më e keqe raportohen, por nuk shpjegohen.
- Nuk thotë asgjë për **sigurinë, autonominë ose deployment-in** jashtë rig-ut të vlerësimit.

Sa e fortë është prova?

Për pretendimet krahasuese qendrore — LBM-të e fine-tune-uara i mundin baseline-et nga zero në agregat, duan disa herë më pak të dhëna dhe janë më robustë ndaj distribution shift — provat janë të forta **dhe** jashtëzakonisht të kontrolluara: të verbra, të randomizuara, me kampion të madh, testim statistikor dhe QA të scoring-ut. Është rast i rrallë ku metodologjia është mjaft e mirë që përfundimet kryesore të merren pothuajse sipas vlerës nominale.

Caveat-et e ndershme i ngre vetë punimi. Error bars kapin rastësinë e **vlerësimit**, por jo rastësinë e **trajnimit** — trajnimi i të njëjtit model dy herë mund të prodhojë politika të ndryshme dhe ky variacion nuk është në statistika. Detyrat reale kishin 50 prova secila, mjaft për efekte mesatare por jo domosdoshmërisht për të vogla. Kushtëzimi gjuhësor përdorte enkoder modest, ndaj pretendimet për "vetëm thuaji robotit" mund të ndryshojnë në sisteme më të mëdha. Dhe ka deklarimin e sinqertë të bug-ut post-hoc të normalizimit. Asnjë nuk e fundos rezultatin; janë pikërisht lloji i gjërave që punimi thotë se fusha shpesh i anashkalon.

Një shënim burimor në të njëjtin frymë: ky shpjegim bazohet në preprint-in e autorëve. Nuk arritëm të marrim versionin e botuar në revistë, ndaj nuk kemi kontrolluar ndryshime mes preprint-it dhe tekstit final.

Pse ka rëndësi

"Robot foundation models" është shprehje e krijuar për teprim dhe ky studim mund të keqlexohet në të dy drejtimet — si triumf "funksionon!" ose si hedhje poshtë "është hype". Leximi i saktë është më i dobishëm: pretraining-u në të dhëna të larmishme jep përfitime reale, të matshme, por të moderuara — kryesisht më pak të dhëna për detyrë dhe më shumë robustësi — dhe përmirësimi me shkallën është i parashikueshëm.

Arsyeja më e thellë pse ka rëndësi është se punimi e kthen rigorozitetin ndaj vetë fushës. Duke treguar se efektet reale janë aq të vogla sa zhduken me vlerësim të dobët, dhe se një zgjedhje e mërzitshme si normalizimi mund të ketë më shumë peshë se një arkitekturë e zgjuar, ai argumenton se shumë "progres" në robot learning ka nevojë për matje më të forta përpara se të besohet. Një punim që e shpenzon kredibilitetin duke ruajtur kufirin mes rezultatit dhe dëshirës bën diçka më të rrallë e më

të vlefshme se një vend i parë në leaderboard.

Përmbledhje e shkurtër

Studiuesit e Toyota Research Institute trajnuan "large behavior models" — politika robotike diffusion të paratrajnuara mbi rreth 1,700 orë manipulimi të larmishëm — dhe i krahasuan me politika një-detyre të trajnuara nga zero me një protokoll jashtëzakonisht rigoroz: i verbër, i randomizuar, me kampion të madh (≈1,800 prova reale dhe 47,000+ simulime) dhe statistikë të mirëfilltë. Pas fine-tuning-ut për detyrë, modelet e mëdha performuan më mirë në agregat, kërkuan rreth **3–5× më pak të dhëna specifike** dhe ishin **më robustë kur kushtet ndryshonin**, ndërsa performanca u rrit gradualisht me më shumë pretraining. Por **pa fine-tuning** nuk i mundën në mënyrë konsistente modelet një-detyre, disa efekte ishin aq të vogla sa u panë vetëm me kampionet e mëdha, dhe një zgjedhje e zakonshme normalizimi pati më shumë ndikim se arkitektura. Është mbështetje e fortë e matur për drejtimin e foundation models të robotëve — **jo** robot i përgjithshëm, jo generalist zero-shot dhe jo "kërcim emergjent" — plus një paralajmërim të mprehtë se shumë robotikë mund të jetë duke matur zhurmë.

Kontroll pa teprime

Çfarë tregon punimi: Me vlerësim rigoroz të verbër dhe me fuqi statistikore (≈1,800 rollout-e reale + 47,000+ simulime), politikat diffusion të paratrajnuara shumëdetyrëshe dhe pastaj të fine-tune-uara i tejkalojnë në agregat politikat një-detyre nga zero, arrijnë performancë të barabartë me ~3–5× më pak të dhëna specifike dhe janë më robustë nën distribution shift; performanca rritet gradualisht me të dhënat e pretraining-ut.

Çfarë është e besueshme, por jo e provuar: Që përfitimet transferohen te vision-language-action models shumë më të mëdha; që scaling-u i lëmuar vazhdon përtej intervalit të të dhënave të testuara.

Çfarë nuk tregon: Robot të përgjithshëm ose zero-shot; kërcim emergjent aftësish; shpjegime për dështimet specifike; besueshmëri absolute në botën reale; apo diçka mbi sigurinë dhe deployment-in autonom.

Kufizimet kryesore: Statistikat kapin rastësinë e vlerësimit, jo të trajnimit; 50 prova reale për detyrë mund të humbin efekte të vogla; një arkitekturë dhe një laborator; enkoder gjuhësor modest; bug normalizimi i gjetur pas vlerësimeve; analiza bazohet në preprint.

Sa besim duhet të ketë një lexues i përgjithshëm? Të lartë se pretraining-u shumëdetyrësh + fine-tuning jep përfitime reale të moderuara, sidomos efikasitet dhe robustësi, dhe se këto u matën jashtëzakonisht mirë. Të lartë se **nuk** është robot i përgjithshëm, zero-shot apo kërcim emergjent. Mesatar për sa larg shkallëzohen fitimet. Dhe paralajmërimi i autorëve duhet marrë seriozisht: efektet janë mjaft të vogla që studimet me fuqi të ulët mund të raportojnë zhurmë.

Burimet

arXiv:2507.05331
https://arxiv.org/abs/2507.05331

Peer-reviewed version (not retrieved) — Science Robotics, 10.1126/scirobotics.aea6201
https://doi.org/10.1126/scirobotics.aea6201

Project page
https://toyotaresearchinstitute.github.io/lbm1/