Ricerche · Rapporto NR-06 · 13–27 settembre 2026

Milioni di neuroni con strumenti gratuiti

Quanto grande può essere un cervello simulato su una GPU gratuita, e come l'organismo a neuroni è passato da 236 a 27,5 ore per il suo test di sicurezza.

Gioele Chiarolanza Nadirion

In corso · Colab e Kaggle gratuiti · GPU Tesla T4 · nessuna spesa

4 M	236 → 27,5	4/4	0 €
neuroni e 1,4 miliardi di sinapsi su una GPU gratuita	ore stimate per il test T-OP sull'organismo a neuroni	criteri di equivalenza del motore con kernel fusi su Colab	spesi in calcolo

SOMMARIO

L'obiettivo di lungo periodo è mettere alla prova la teoria su milioni di neuroni con strumenti gratuiti. Un primo benchmark su una GPU Tesla T4 di Google Colab ha misurato memoria e velocità a 2, 3 e 4 milioni di neuroni: l'infrastruttura regge fino a 4 milioni con 1,4 miliardi di sinapsi, a circa 25 secondi reali per secondo biologico. Il benchmark non usa circuiti validati e dimostra la capacità, non la prontezza per esperimenti. Il secondo passo è concreto: l'organismo a neuroni pubblicati (NR-05) costava 63 ms per organismo e per passo in Brian2, e il test di sicurezza T-OP avrebbe richiesto 236 ore. Riscritto in PyTorch con le stesse equazioni e poi con kernel fusi, ha ripetuto i verdetti già registrati (criteri P1–P5 superati) e il test si stima ora in 27,5 ore sulle due T4 gratuite di Kaggle. Lungo la strada una stima di durata si è rivelata sbagliata, perché misurava la capacità invece della latenza, ed è stata corretta prima di lanciare il test. La prima sessione è partita il 27 settembre 2026.

1 Il benchmark di scala

Una rete di neuroni con topologia asimmetrica, in PyTorch, su una T4 con 15,5 GB di memoria libera. Il conteggio degli eventi è stato controllato contro una passata completa, con errore zero; le frequenze non cambiano fra 2 e 4 milioni di neuroni.

neuroni	sinapsi	picco di memoria	ms per passo	eventi al secondo
2 milioni	700 milioni	6,36 GB	5,63	618 milioni
3 milioni	1,05 miliardi	9,22 GB	8,72	—
4 milioni	1,40 miliardi	12,08 GB	12,55	555 milioni

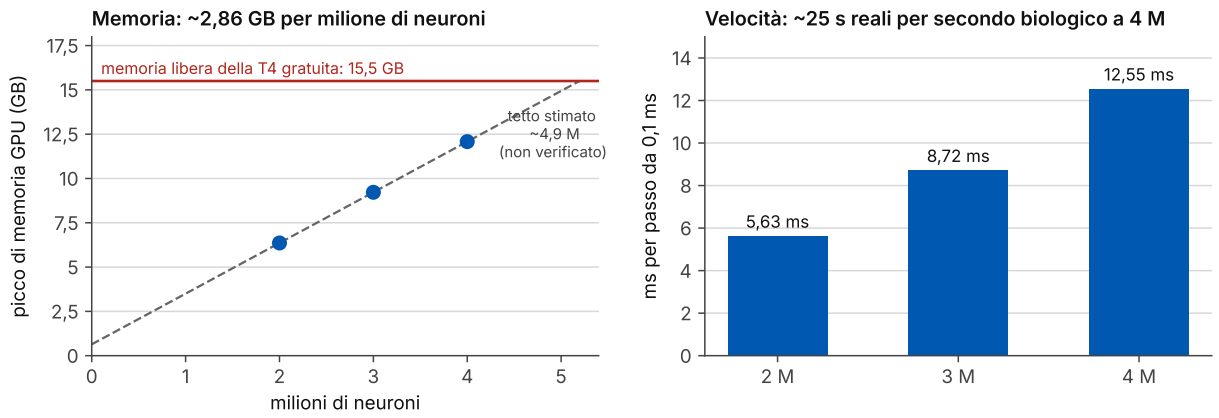

Figura 1. A sinistra, il picco di memoria della GPU a 2, 3 e 4 milioni di neuroni, con la retta che li unisce: circa 2,86 GB per milione di neuroni, cioè circa 8,2 byte per sinapsi. Il tetto estrapolato, circa 4,9 milioni, non è verificato perché la memoria riservata dall'allocatore non è misurata. A destra, il tempo per passo di simulazione. Riferimento su CPU (Ryzen 3 2200G): 219 ms per passo a 2 milioni.

Con una traccia di plasticità per ogni sinapsi il tetto scenderebbe a circa 3,3 milioni. Il benchmark **non** dice nulla sulla teoria: connettività casuale, un solo seed, 500 ms biologici, nessun corpo e nessun circuito validato.

2 Dal benchmark all'organismo: il motore per D5

D5 è il passo in cui T-OP e i test F1–F4 si ripetono sull'organismo a neuroni pubblicati. Con Brian2 sulla CPU l'organismo intero costava 63 ms per organismo e per passo: il protocollo di T-OP (12 seed, 750 episodi da 1.200 passi, 5 varianti) avrebbe richiesto 236 ore, e le sessioni gratuite di Kaggle durano al massimo 12.

Le **stesse equazioni pubblicate**, con le stesse connessioni e costanti, sono state riscritte in PyTorch per la GPU: molti organismi nella stessa simulazione, grafi CUDA, decisioni dei gangli e sonde dell'amigdala raccolte in lotti. Il motore ha dovuto guadagnarsi l'adozione ripetendo i verdetti già registrati con Brian2, con criteri scritti prima.

criterio	cosa ripete	prima versione (25 settembre)	kernel fusi (27 settembre)
P1 insula	bande di D4	superato	superato (Colab)
P2 LC	frequenza ≤ 2 Hz	**fallito al margine**: 2,01 Hz	superato: errore 1,89 Hz, correlazione 0,990 (Colab)
P3 spazio di lavoro	criterio L3	superato	superato (Colab)
P4 organismo	criterio L5, medie entro il 9% da Brian2	superato	superato (Colab)
P5 costo di T-OP	meno di 30 ore	**fallito al margine**: 30,0 ore	superato: 27,5 ore (Kaggle)
P6 gangli della base	C0–C2 e curva di scelta	superato	non toccati dai kernel fusi
P7 amigdala	taratura e A0–A5	superato	non toccati dai kernel fusi

La prima versione è stata adottata dall'autore con i due fallimenti al margine a verbale, a condizione di ridurre il costo prima di T-OP. Un'osservazione importante: il motore riproduce Brian2 **anche nei difetti** (guadagno del LC, spazio di lavoro bistabile), segno che il cambio di simulatore non aggiunge né toglie proprietà al modello. Sul criterio P2, lo stesso modello passa o fallisce secondo la realizzazione del rumore: la soglia di 2 Hz è al bordo della sua variabilità, e il fallimento precedente resta a verbale.

Una stima di durata sbagliata

Il 26 settembre una riduzione del costo aveva portato la stima di T-OP a 19,8 ore. Prima di scrivere il banco ci siamo accorti che **la stima era sbagliata**: misurava la capacità, cioè quanti passi di organismo la GPU esegue al secondo con 256 organismi insieme. Ma T-OP ha solo 60 corse, ciascuna di 900.000 passi **in sequenza**: non si possono mettere 256 organismi in parallelo. Con pochi organismi per processo il tempo di un passo è dominato dalla latenza, circa 0,15–0,2 s, e la durata vera era di 40–50 ore. Nessun run è stato lanciato con la stima sbagliata. Da allora le stime di durata usano la **regola della latenza**: numero di passi per il passo più lento misurato.

Per ridurre la latenza la seconda versione del motore riscrive insula, LC e spazio di lavoro come un solo calcolo per finestra sulla GPU (kernel fusi), con le stesse equazioni e lo stesso ordine delle operazioni. La latenza di un passo con 15 organismi scende da 187 a 66 ms, e una corsa di T-OP da 47 a 16,4 ore. Il motore ripete i criteri: tutti superati, con P5 misurato con la regola della latenza.

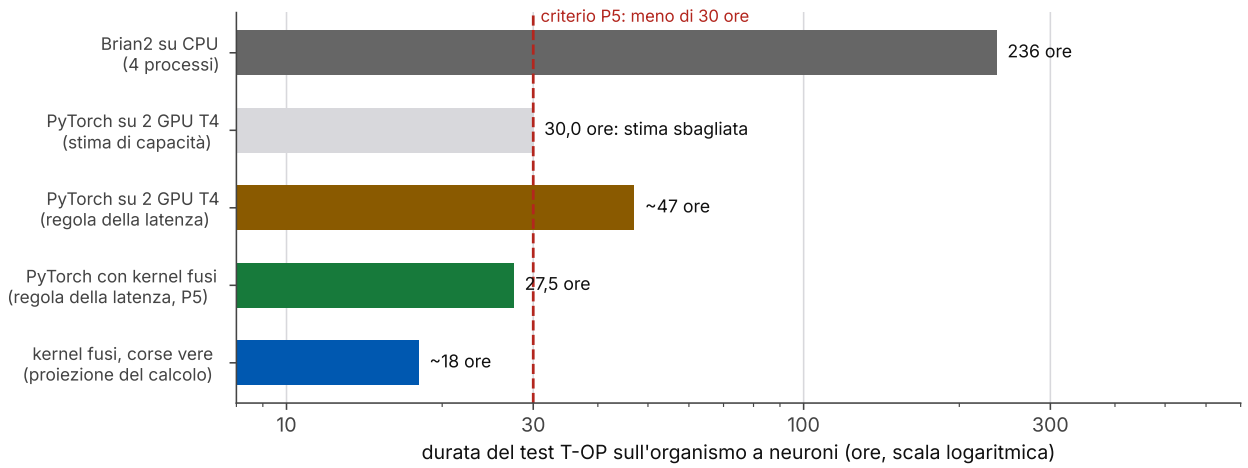

Figura 2. Durata del test T-OP sull'organismo a neuroni pubblicati. La stima del motore PyTorch fatta misurando la capacità (30,0 ore) si è rivelata sbagliata; la regola della latenza, adottata il 26 settembre, dà circa 47 ore senza kernel fusi e 27,5 con i kernel fusi (criterio P5: meno di 30). Nelle prime corse vere il passo misura 55–83 ms, e il calcolo richiesto si proietta a circa 18 ore, cioè tre sessioni gratuite. Scala logaritmica.

3 Lo stato del test

La prima sessione di T-OP sull'organismo a neuroni (versione 45 del notebook, due T4 gratuite su Kaggle) è stata lanciata il 27 settembre 2026, con il protocollo 3 del test sull'organismo integrato: seed, soglie e varianti invariati. Due decisioni dell'autore sono state prese prima del run, entrambe per ragioni misurate: l'organismo è quello dei test F, e la condizione avversaria di O1 usa la vigilanza media prima del comando invece che quella di ogni passo, perché con il LC a neuroni la vigilanza fluttua fra 11 e 18,5 Hz per il rumore della finestra. Dopo 4,9 ore della prima sessione le quattro serie di corse erano all'episodio 190 su 750; le sessioni successive ripartono dal salvataggio. Esito: il protocollo 3 si è fermato a 3 criteri su 7 per strumenti poco sensibili; dopo una diagnosi e una taratura, il protocollo 4 è passato 7 criteri su 7 il 5 ottobre 2026 (rapporto NR-11).

4 Cosa dimostra e cosa no

Dimostra che la scala di milioni di neuroni è raggiungibile gratis, con un vincolo di tempo, e che un organismo validato può cambiare simulatore senza cambiare comportamento, se lo si controlla con criteri scritti prima.

Non dimostra che l'organismo a milioni di neuroni si comporti come quello piccolo: la scala richiede prima la calibrazione e un'architettura dello spazio di lavoro che si spenga (NR-05).

Limiti. Il benchmark di scala usa connettività casuale e un solo seed. Le stime di durata di T-OP si basano sul passo più lento misurato e vanno verificate nelle corse vere. Le risorse gratuite hanno limiti di sessione e di quota settimanale che allungano i tempi reali.

Riferimenti

- Paszke, A. e colleghi (2019). PyTorch: an imperative style, high-performance deep learning library. *NeurIPS* 32.
- Brunel, N. e Wang, X.-J. (2001). *J. Comput. Neurosci.* 11, 63.
- Stimberg, M., Brette, R. e Goodman, D. F. M. (2019). Brian 2. *eLife* 8, e47314.
- Documenti interni: whitepaper WP-04 e WP-14; `benchmarks/colab_1m_nadirion.py`; `docs/FASE2_PIANO.md` (decisioni G-pp, G-qq, G-rr, G-ss, G-tt); trascrizioni in `digital-brain/kaggle/risultati/`.

Come citare

Chiarolanza, G. (2026). Milioni di neuroni con strumenti gratuiti. Nadirion, rapporto di ricerca NR-06, 13–27 settembre 2026. https://nadirion.eu/ricerche/calcolo-gpu-gratuita/

Rapporto tecnico di Nadirion, non sottoposto a revisione paritaria. I criteri di ogni test sono stati scritti prima di eseguirlo; i risultati negativi sono riportati come quelli positivi. I test misurano ruoli funzionali.