

The gap between human intelligence and artificial intelligence

WONG Limsoon



NUS
National University
of Singapore

National University of Singapore

What does mastering a subject mean

Expertise is not just knowing the rules; it is knowing:

Why the rules work

When the rules fail, and

When the rules no longer apply

AI is often strongest only at recognizing patterns similar to what it has seen before

“AI does not know the answer today, but surely it will know tomorrow”

AI currently learns *after* humanity discovers something

Human experts can discover it *before* it is known

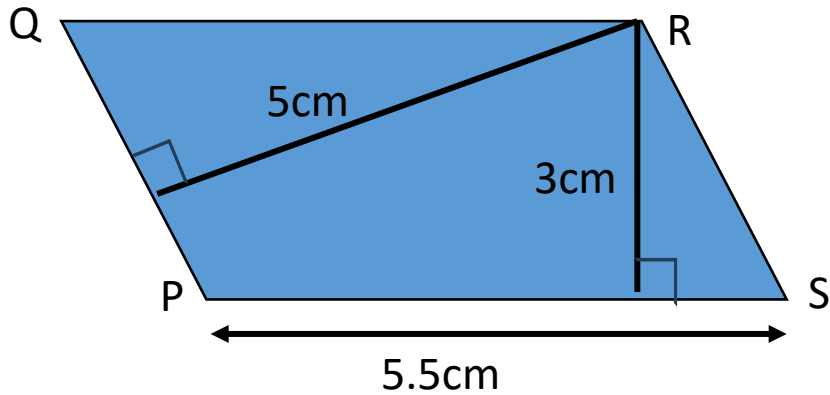
This is the difference between *absorbing answers* and *figuring them out*

If AI truly knew how to figure out new answers by itself, there would be no need to wait for someone to first discover and teach those answers

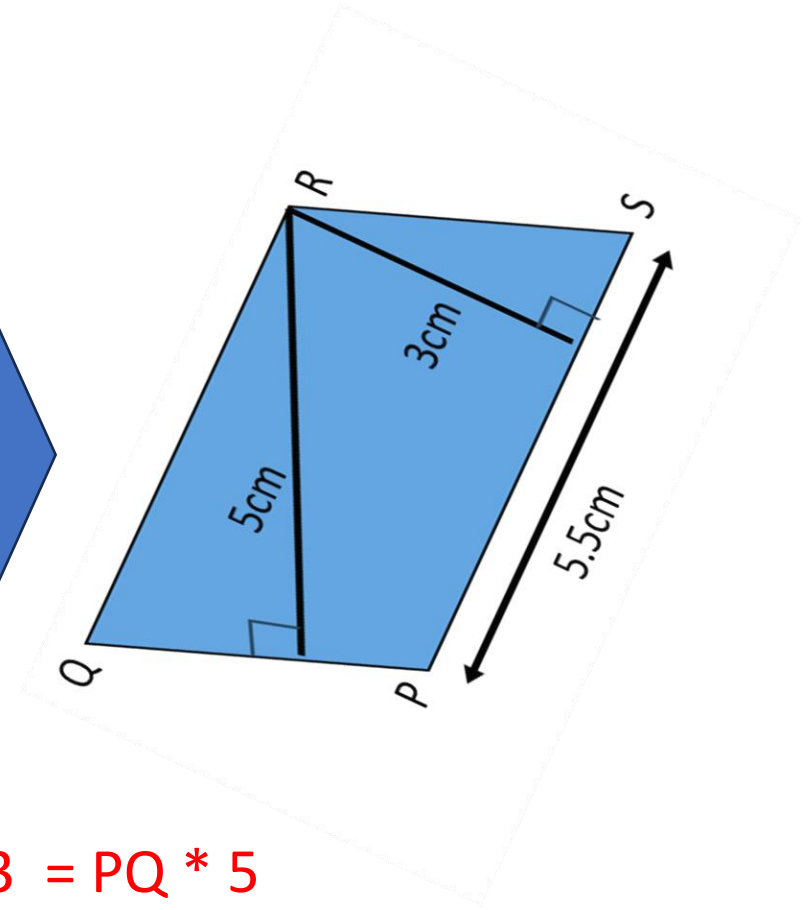
The delay itself reveals the gap

Story 1

My son's homework



I am given a parallelogram PQRS. The base PS = 5.5cm. The height from PS to the opposite side QR is 3cm. A perpendicular line from PQ to the opposite corner R is 5cm. What is the perimeter of PQRS?



$$\begin{aligned}PS * 3 &= PQ * 5 \\5.5 * 3/5 &= PQ \\3.3 &= PQ \\Perimeter &= 2 * 5.5 + 2 * 3.3 = 17.6\end{aligned}$$

Did the AI figure out that rotation preserves area by itself, or someone told that to the AI?

Current AI can learn the rules (if told what the rules are), but is not good at inferring the rules

ARC-AGI-3: A New Challenge for Frontier Agentic Intelligence

ARC Prize Foundation

We introduce ARC-AGI-3, an interactive benchmark for studying agentic intelligence through novel, abstract, turn-based environments in which agents must explore, infer goals, build internal models of environment dynamics, and plan effective action sequences without explicit instructions. Like its predecessors ARC-AGI-1 and 2, ARC-AGI-3 focuses entirely on evaluating fluid adaptive efficiency on novel tasks, while avoiding language and external knowledge. ARC-AGI-3 environments only leverage Core Knowledge priors and are difficulty-calibrated via extensive testing with human test-takers. Our testing shows humans can solve 100% of the environments, in contrast to frontier AI systems which, as of March 2026, score below 1%. In this paper, we present the benchmark design, its efficiency-based scoring framework grounded in human action baselines, and the methodology used to construct, validate, and calibrate the environments.

Story 2

A video-processing algorithm: Eigenbackground

Oliver et al., *TPAMI*, 2000; <https://doi.org/10.1109/34.868684>

“In order to reduce the dimensionality of the space, in PCA only M eigenvectors (eigenbackgrounds) are kept, corresponding to the M largest eigenvalues ... Note that moving objects, because they don't appear in the same location in the N sample images and they are typically small ... the portions of an image containing a moving object cannot be well described by this eigenspace model ... whereas the static portions of the image can be accurately described as a sum of the various eigenbasis vectors ... the eigenspace provides a robust model of the probability distribution function of the background but not for the moving objects”

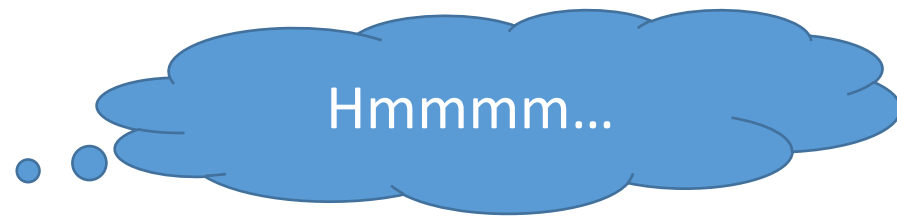


Fig. 1 The result of the Eigenbackground (Fig. 9 of [19]) in background modeling. From top to bottom, the original images, ground truth and the results of foreground detection using background models created by the Eigenbackground method (courtesy from [19])

What AI says... hmm?

Deepseek: For a very static background (e.g., hallway, office) with small, fast-moving objects that never repeat at the same pixel locations during training, the quoted logic works

Manus: Yes, the idea is brilliant. It uses reconstruction error as a detection tool: by modeling only the persistent background, anything the system cannot explain is identified as motion. It is an elegant, foundational intuition in computer vision

ChatGPT: The idea is fundamentally correct and was historically influential for background subtraction

My Socratic ladder

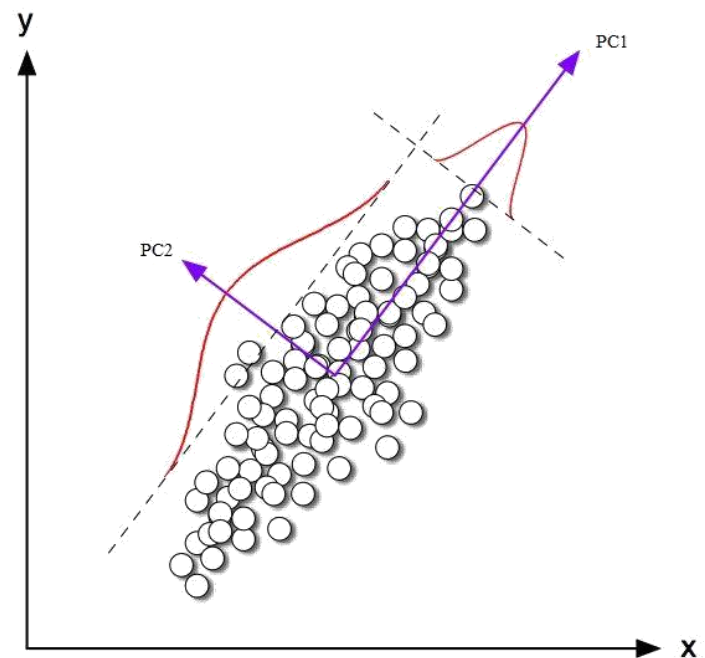
What is PCA?

In a surveillance video, does foreground or background change more?

Which PCs encode foreground?

Which PCs encode background?

What do you now think of the eigenbackground paper?



“In order to reduce the dimensionality of the space, in PCA only M eigenvectors (eigenbackgrounds) are kept, corresponding to the M largest eigenvalues ... Note that moving objects, because they don't appear in the same location in the N sample images and they are typically small ... the portions of an image containing a moving object cannot be well described by this eigenspace model ... whereas the static portions of the image can be accurately described as a sum of the various eigenbasis vectors ... the eigenspace provides a robust model of the probability distribution function of the background but not for the moving objects”

Current AI can apply rules, but is blind to when the rules break and is bad at fixing when things get broken



Fig. 6 Modeled Background and detected Foregrounds using 10 strongest eigenvectors

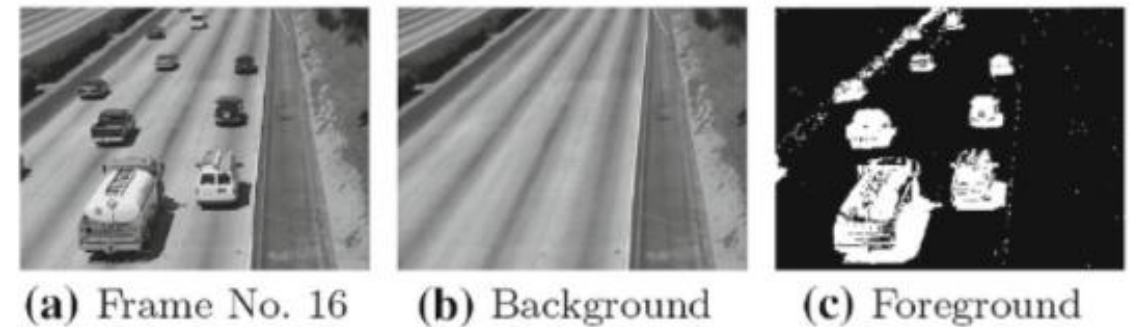


Fig. 7 Modeled Background and detected Foregrounds using 10 weakest eigenvectors

J Math Imaging and Vis, 2022; <https://doi.org/10.1007/s10851-022-01080-4>

Story 3

Brain vs ChatGPT, Gemini, ...

Why is AI so inefficient?

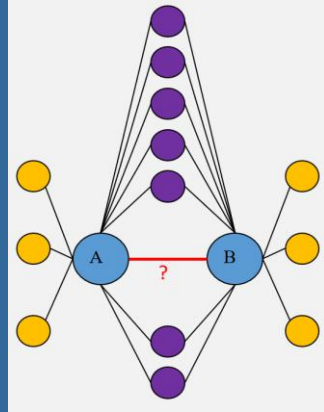
Does it need to be so inefficient?

Feature	LLM	Brain
Within-layer connectivity	Dense	Mostly sparse
Between-layer connectivity	Dense (layer-to-layer)	Highly sparse, variable distance
Global connectivity	Only via attention (soft/dynamic)	Extremely recurrent, local + long-range connections
Sparsity	Mostly dense per layer	Overall extremely sparse

Task	Approx Energy
Human brain answering "1 + 1"	~1–20 J
ChatGPT answering "1 + 1"	~300–3000 J
Ratio	~100× to 3000× more energy

So even a trivial question costs **two to three orders of magnitude more energy** for a frontier LLM than for a human brain.

Insight



If two proteins (A, B) share many interaction partners, they are likely to directly interact

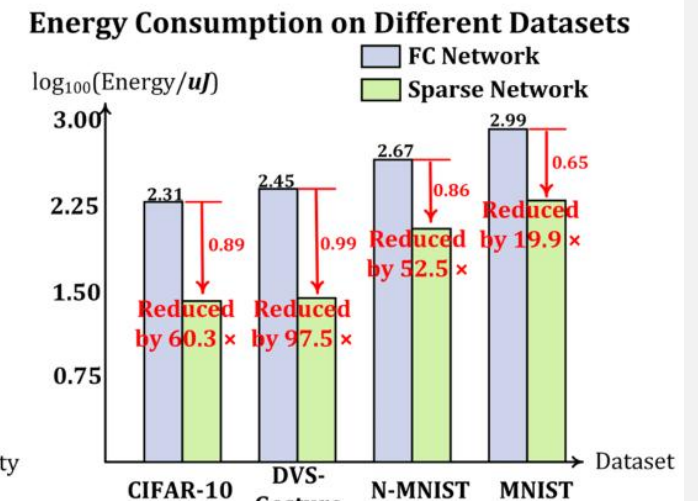
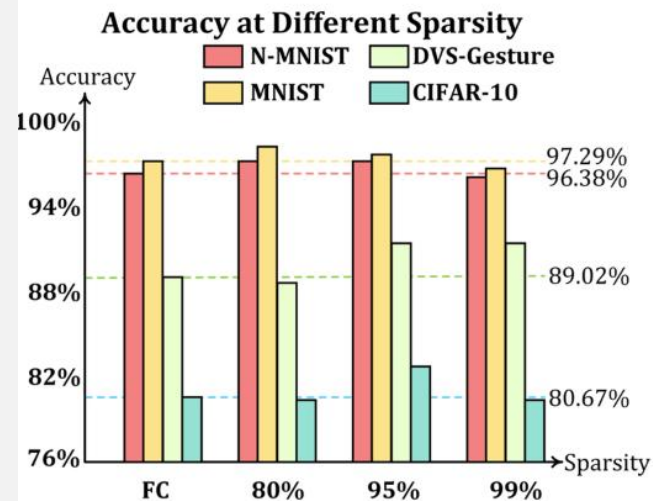
If a shared partner interact with many other proteins, then its interactions with A & B may be “incidental”

Cannistraci’s centrality-hierarchy-redundancy (CHR) hypothesis on brain network



Ultra-sparse Cannistraci-Hebb neural networks

train → prune → CHR rewire → train → prune → CHR rewire → ...



Hua, et al. Cannistraci-Hebb training on ultra-sparse spiking neural networks, ICLR, 2026

The purpose of computing is insight, not numbers

--Richard Hamming, 1962

Human intelligence

Insight + logic \Rightarrow elegant solution

I hope AI will get there eventually

But it probably won't get there with current architecture & learning paradigms

Human brain possesses the capacity for insight

Researchers should study and mirror its structural architecture to build the next generation of truly intelligent AI systems