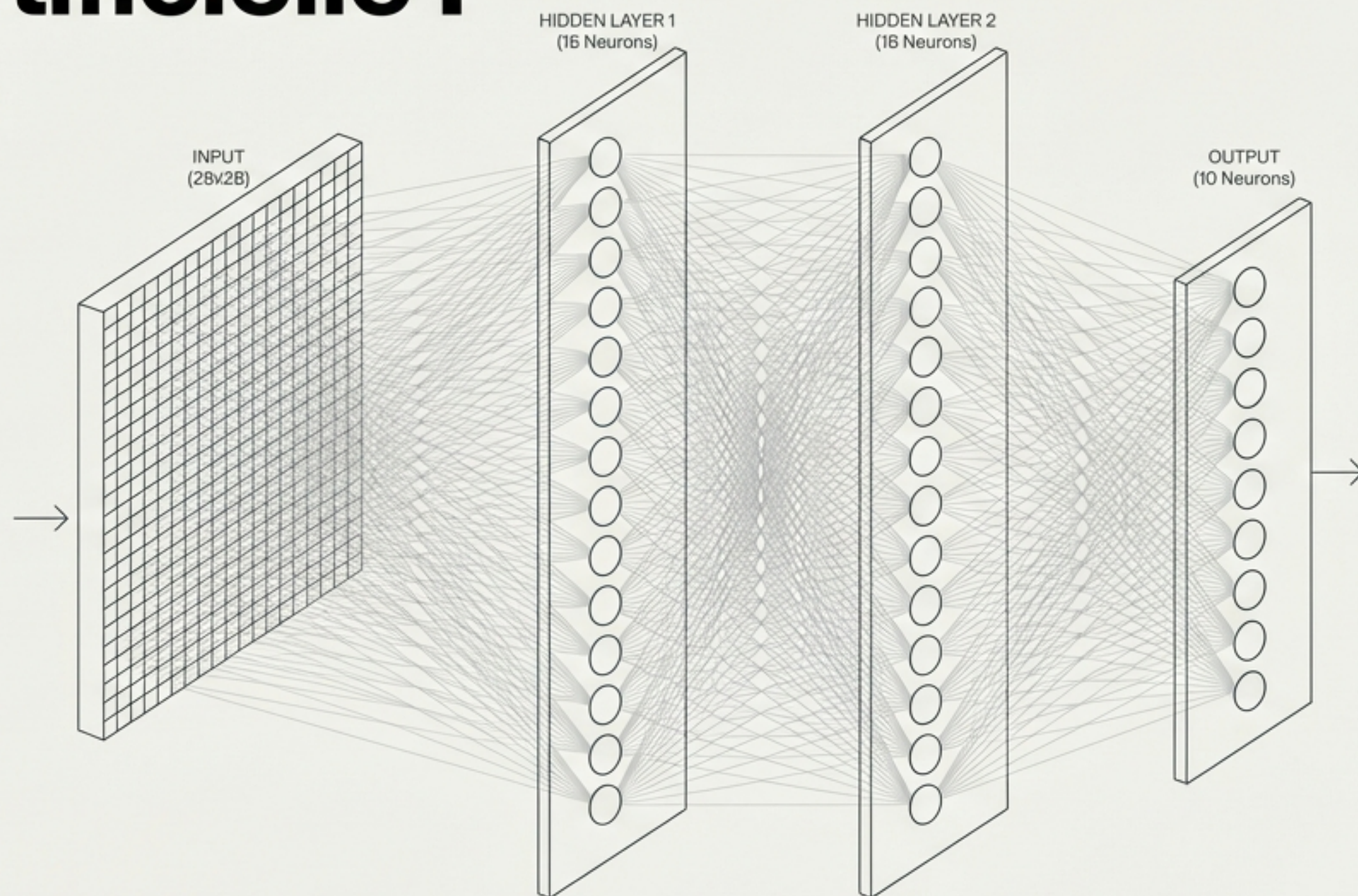


Anatomie d'une Intelligence Artificielle :

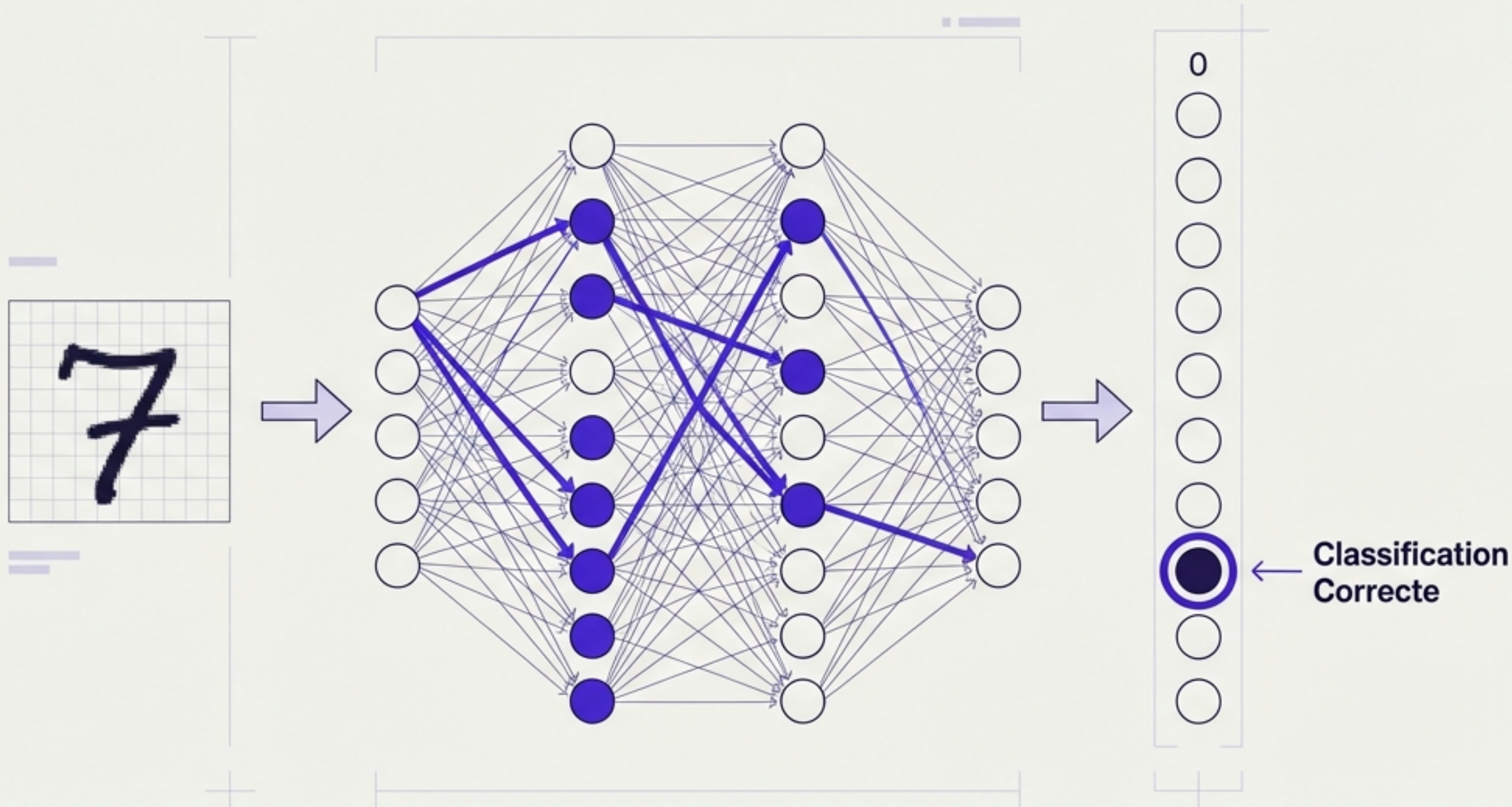
Ce que notre réseau voit vraiment



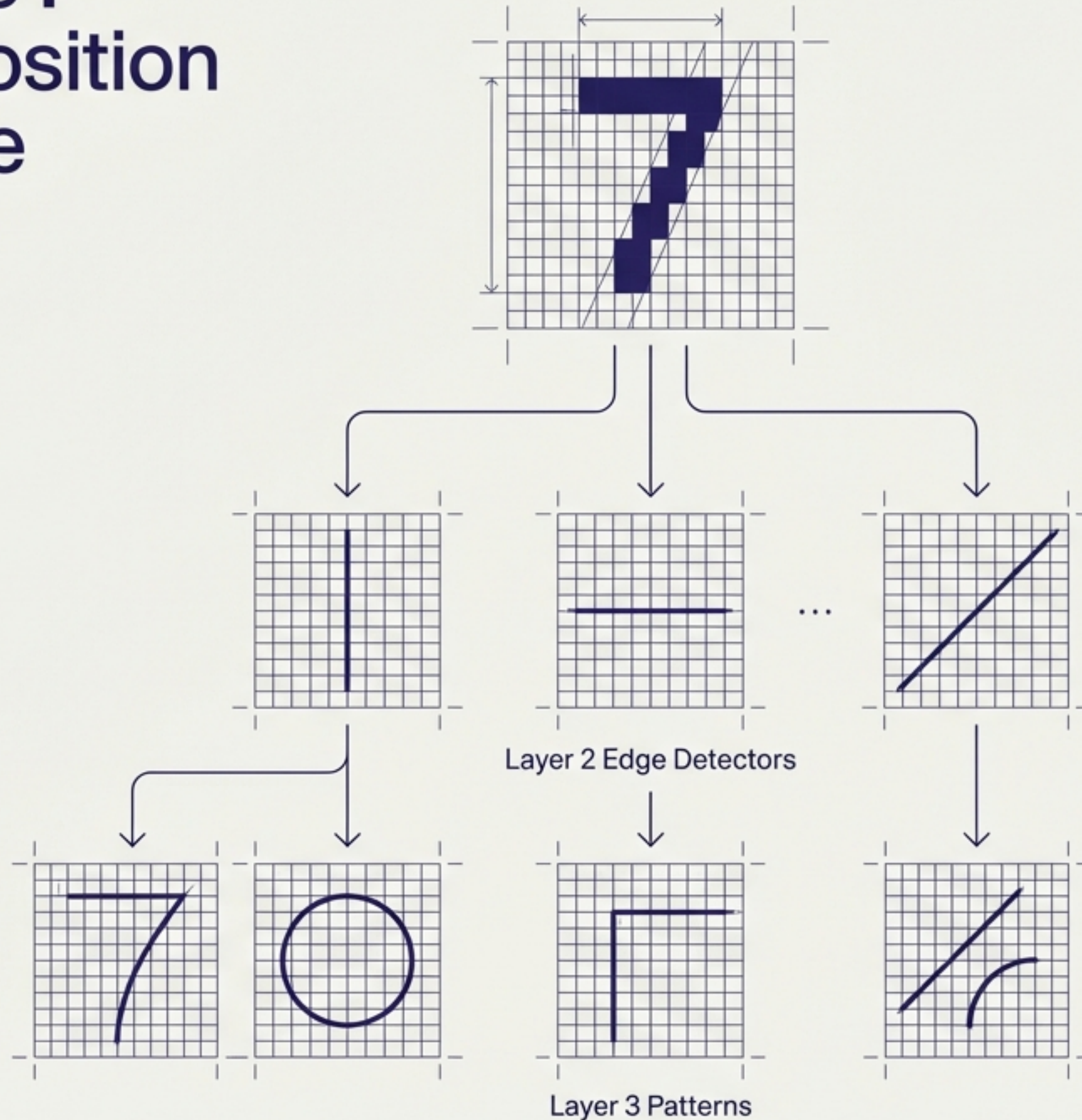
Analyse du fonctionnement interne et des limites du perceptron multicouche.

Un taux de réussite de 96 %

Avec deux couches cachées de 16 neurones, le réseau classe correctement 96 % des nouvelles images. C'est un résultat 'pas mal' pour une structure simple, bien que les humains et les réseaux modernes atteignent > 99,75 %. Compte tenu de la complexité de la tâche, cette performance sans instruction explicite est remarquable.



L'Hypothèse : La décomposition hiérarchique



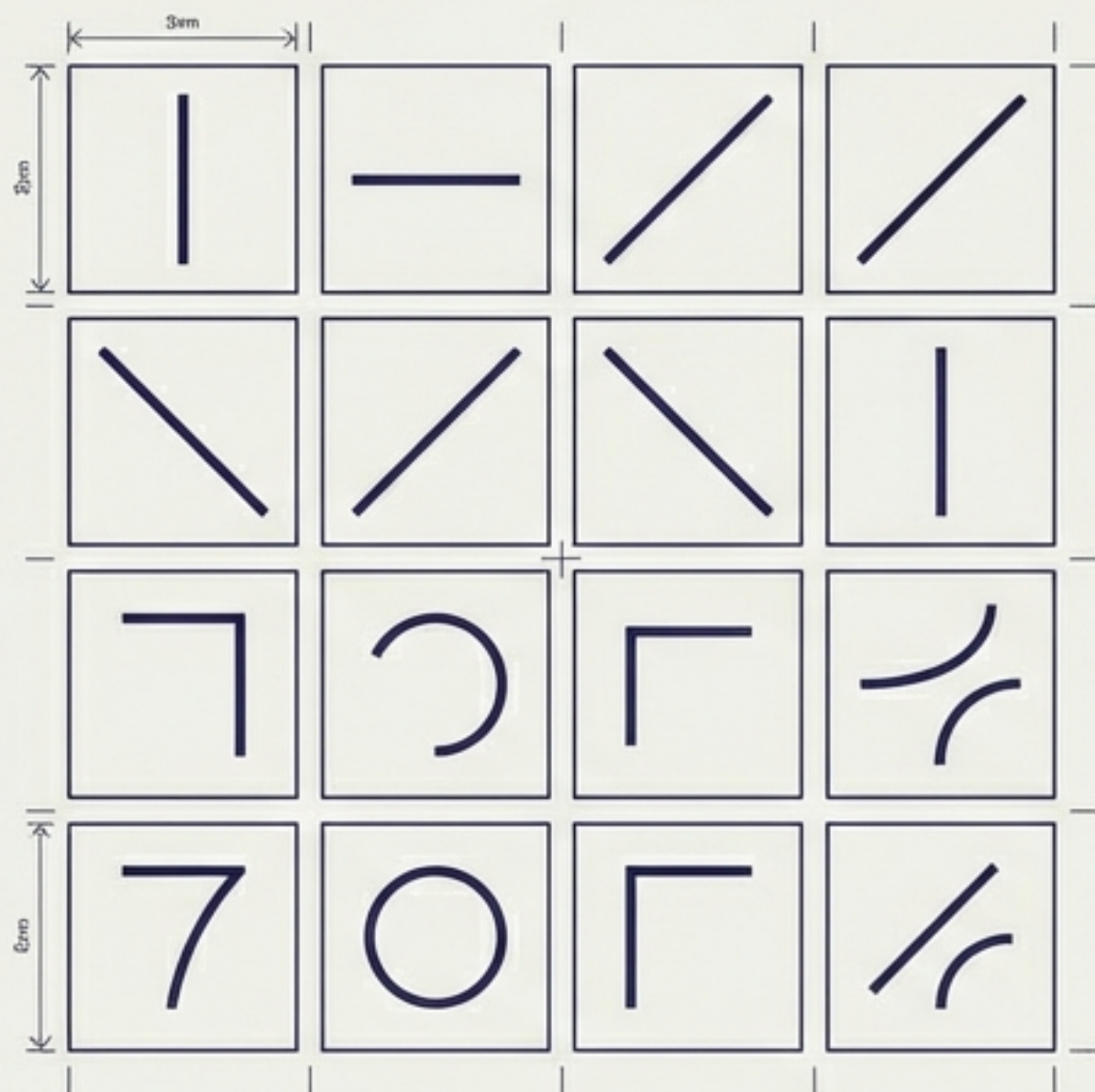
****L'espoir :** Que la 2e couche détecte des bords simples et que la 3e couche assemble ces bords en motifs (boucles, lignes).

Nous imaginions que chaque neurone de la 2e couche correspondait à une petite structure visuelle spécifique.

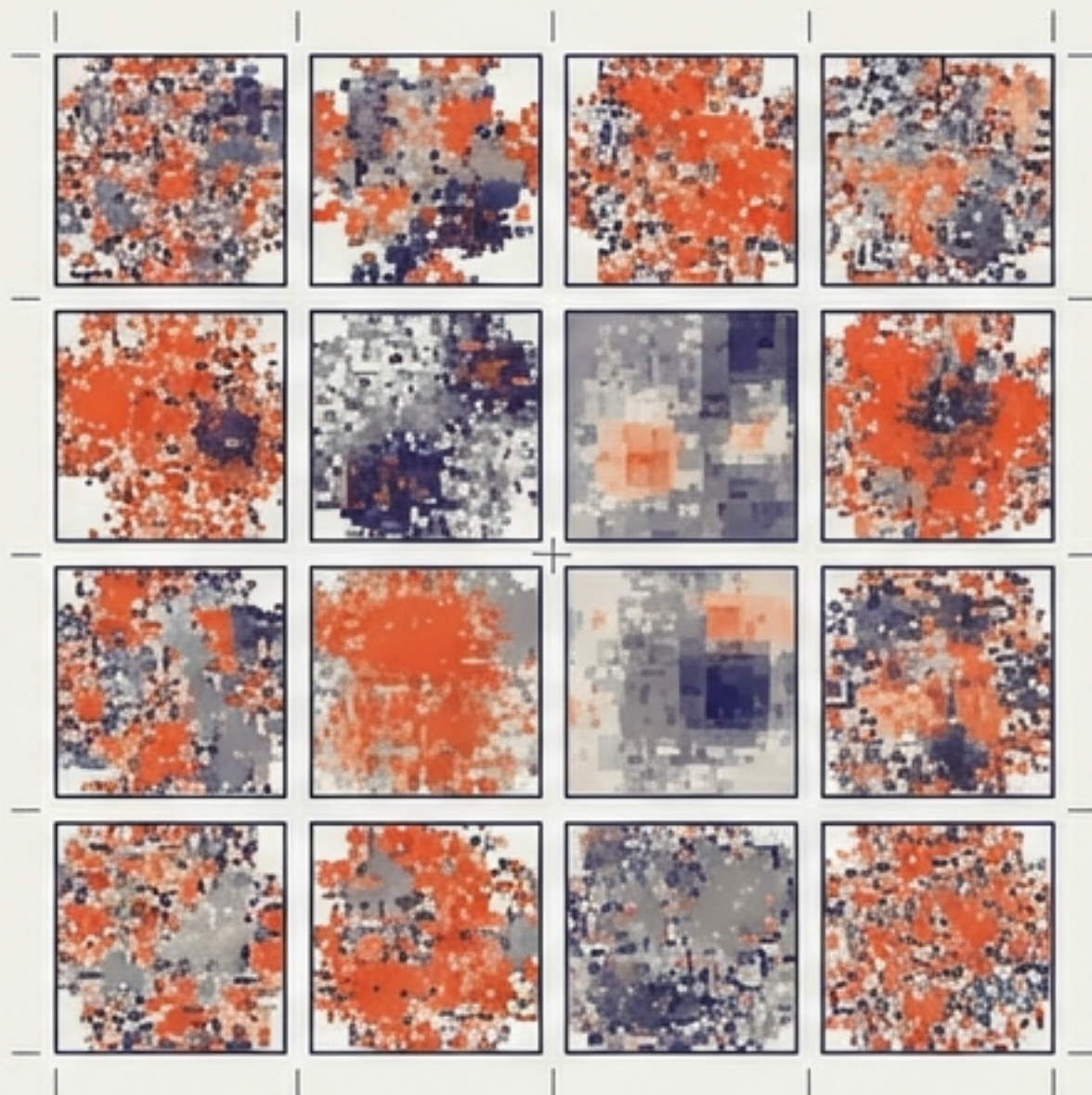
C'est ainsi que nous rationalisons souvent le fonctionnement de la vision par ordinateur.

La Réalité : Du bruit, pas des bords

Attente



Réalité

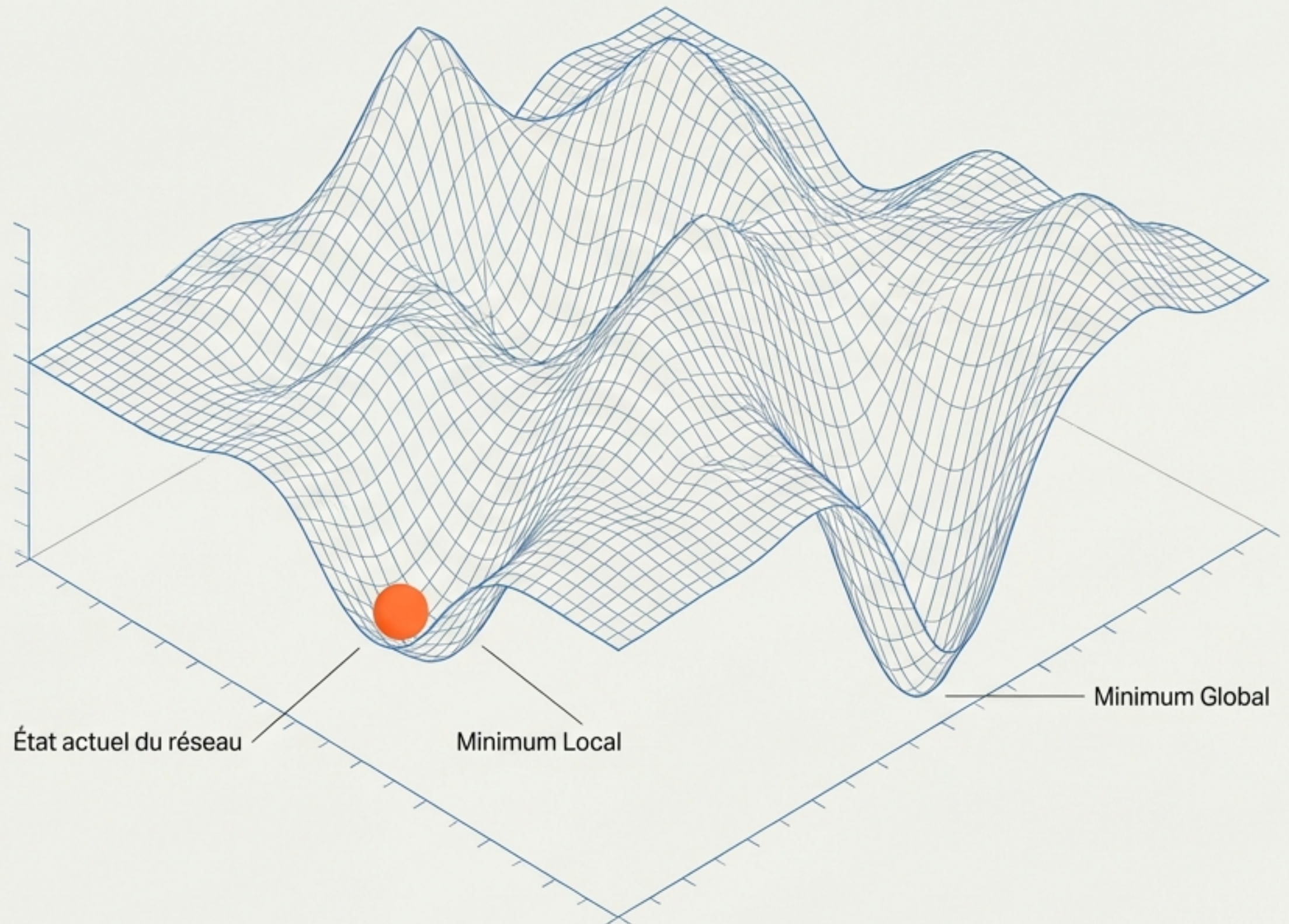


En visualisant les poids réels (les connexions entre la couche 1 et 2), nous ne voyons pas de détecteurs de bords propres. Les motifs semblent presque aléatoires, ressemblant à du bruit statique. Le réseau ne cherche pas de structure géométrique claire ; il cherche des corrélations 'floues' dans les pixels.

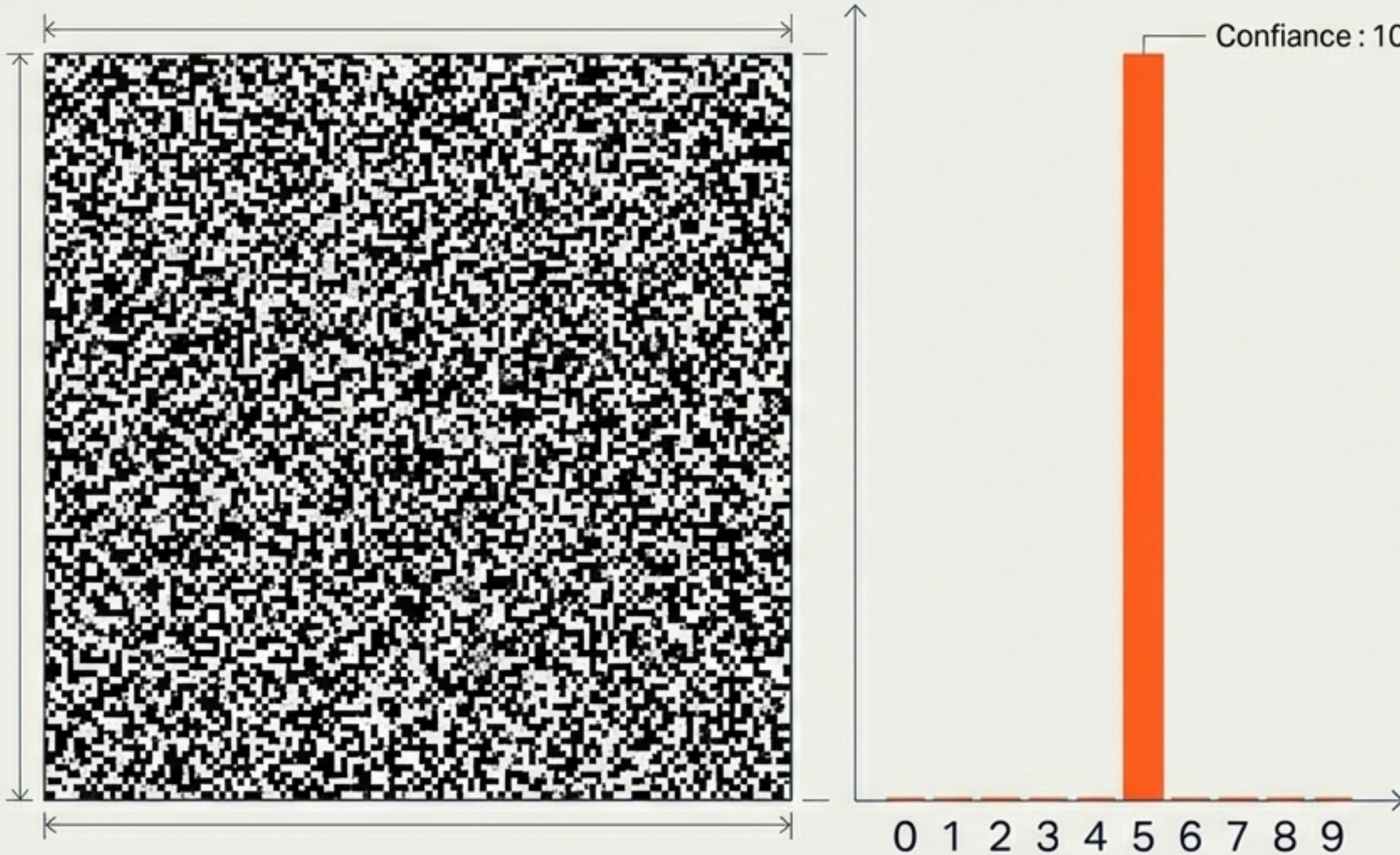
Le piège du minimum local

Dans un espace à **13 002 dimensions** (le nombre total de poids et de biais), le réseau a trouvé une solution mathématique qui minimise le coût.

Il s'agit d'un 'minimum local' : une configuration qui fonctionne pour classer les données, mais sans 'comprendre' la structure des chiffres. Le réseau a appris par cœur plutôt que par compréhension.



L'illusion de la confiance

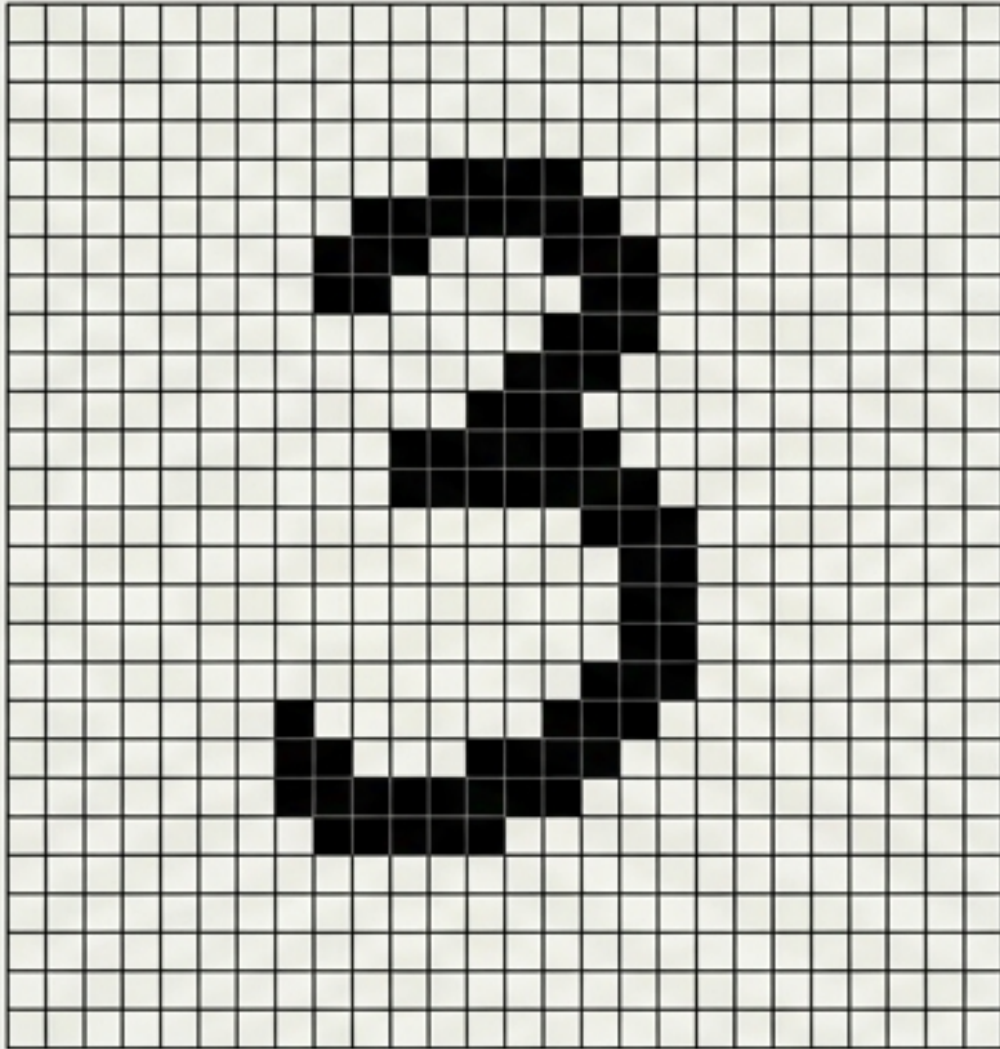


Si l'on montre du bruit aléatoire au réseau, il affirme avec confiance qu'il s'agit d'un '5'.

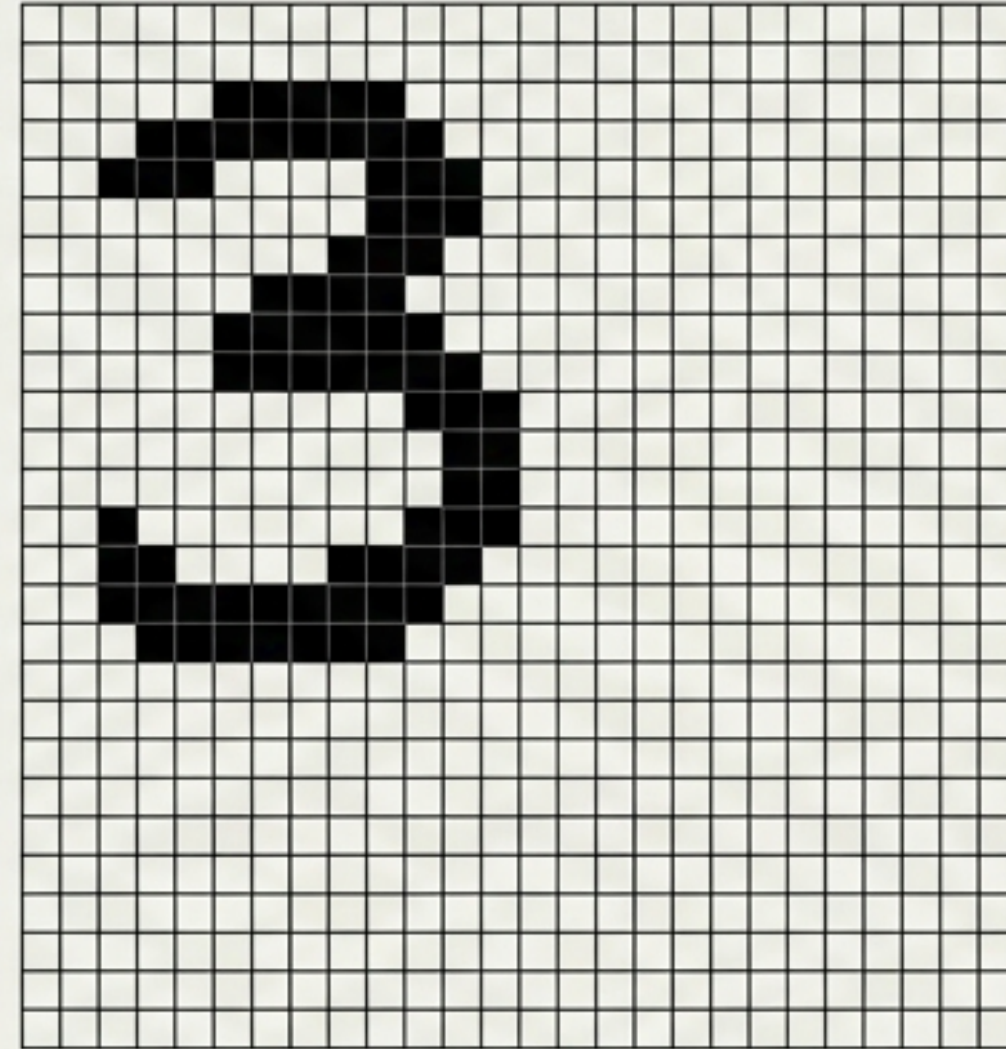
Une véritable intelligence serait incertaine. Notre réseau, lui, est conçu pour être confiant.

Conclusion : Il sait reconnaître des motifs, mais ne sait pas comment les *valider* ou les dessiner.

La fragilité du positionnement



Sortie : 3 (Correct)

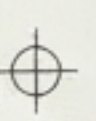
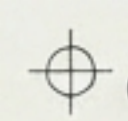


Sortie : 8 (Incorrect)

Le réseau traite chaque pixel individuellement. Il ne sait pas qu'un '3' en haut à gauche est le même qu'un '3' au centre.

Il manque d'**invariance par translation**.

Pour le réseau, un chiffre déplacé est une image totalement nouvelle et inconnue.



Un univers restreint

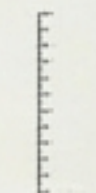
Crimson Pro

Du point de vue du réseau,
l'univers entier se résume à des
chiffres centrés sur une grille
minuscule.

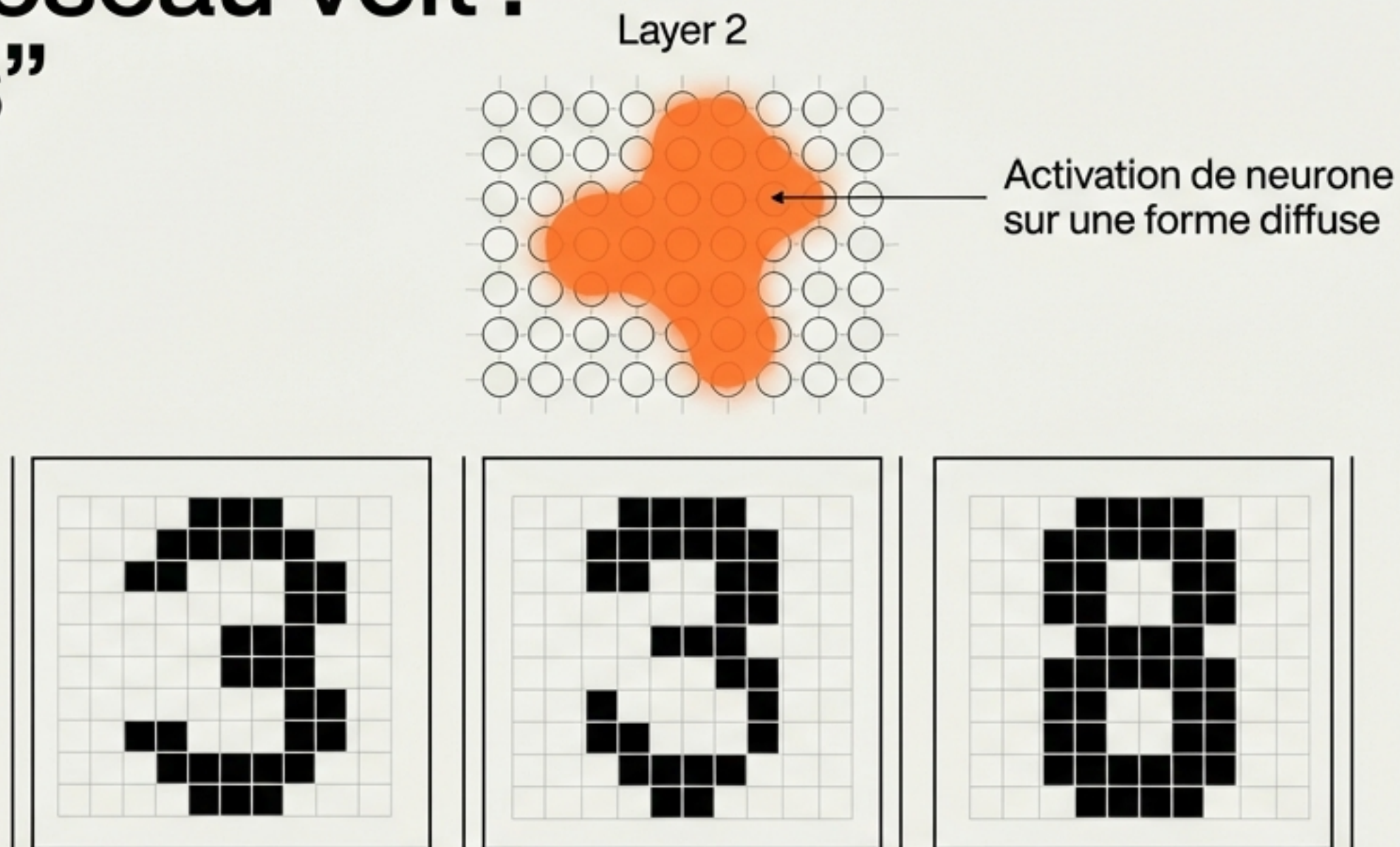
Sa fonction de coût ne l'a jamais
incité à être flexible.

Il n'a jamais appris à gérer le
changement d'échelle ou de
position car il n'en a jamais eu
besoin pendant l'entraînement.

9	0	1	2	3	6	7	8	9	0
1	0	1	2	3	6	9	8	9	1
7	1	2	3	5	6	7	8	0	2
3	1	3	4	5	6	4	9	2	9
4	2	4	5	3	9	1	7	3	6
5	3	3	8	4	9	1	5	9	5
6	1	6	7	0	1	0	6	7	6
7	9	4	0	2	7	1	7	8	4
8	7	3	1	5	2	1	8	8	9
0	2	1	3	5	6	7	8	9	0



Ce que le réseau voit : Des “Blobs”



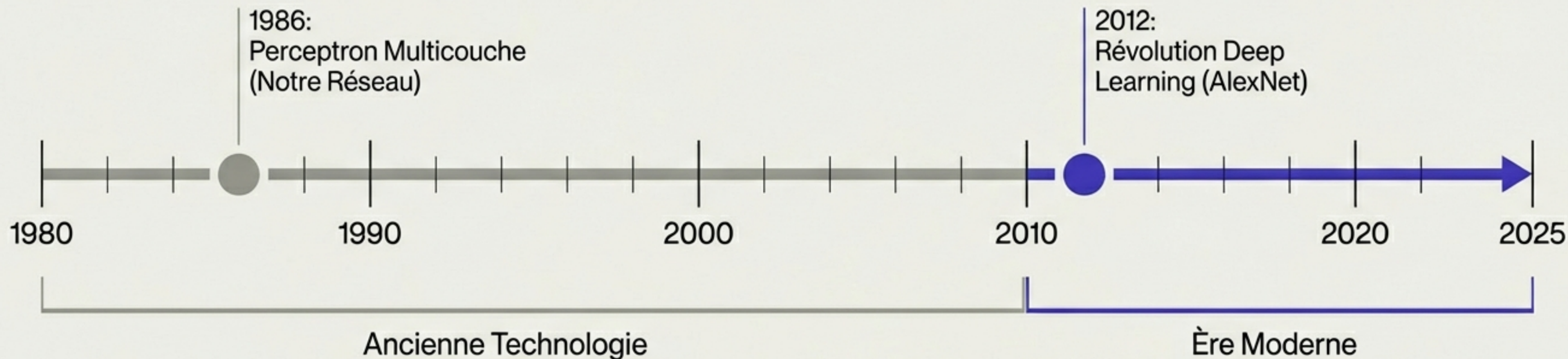
En observant l'activation des neurones en temps réel, on constate qu'ils réagissent à des formes amorphes (des blobs) plutôt qu'à des traits nets. Lorsque l'on transforme un 3 en 8, les neurones changent d'état de manière non intuitive. C'est la signature d'une reconnaissance basée sur la densité de pixels plutôt que sur la géométrie.

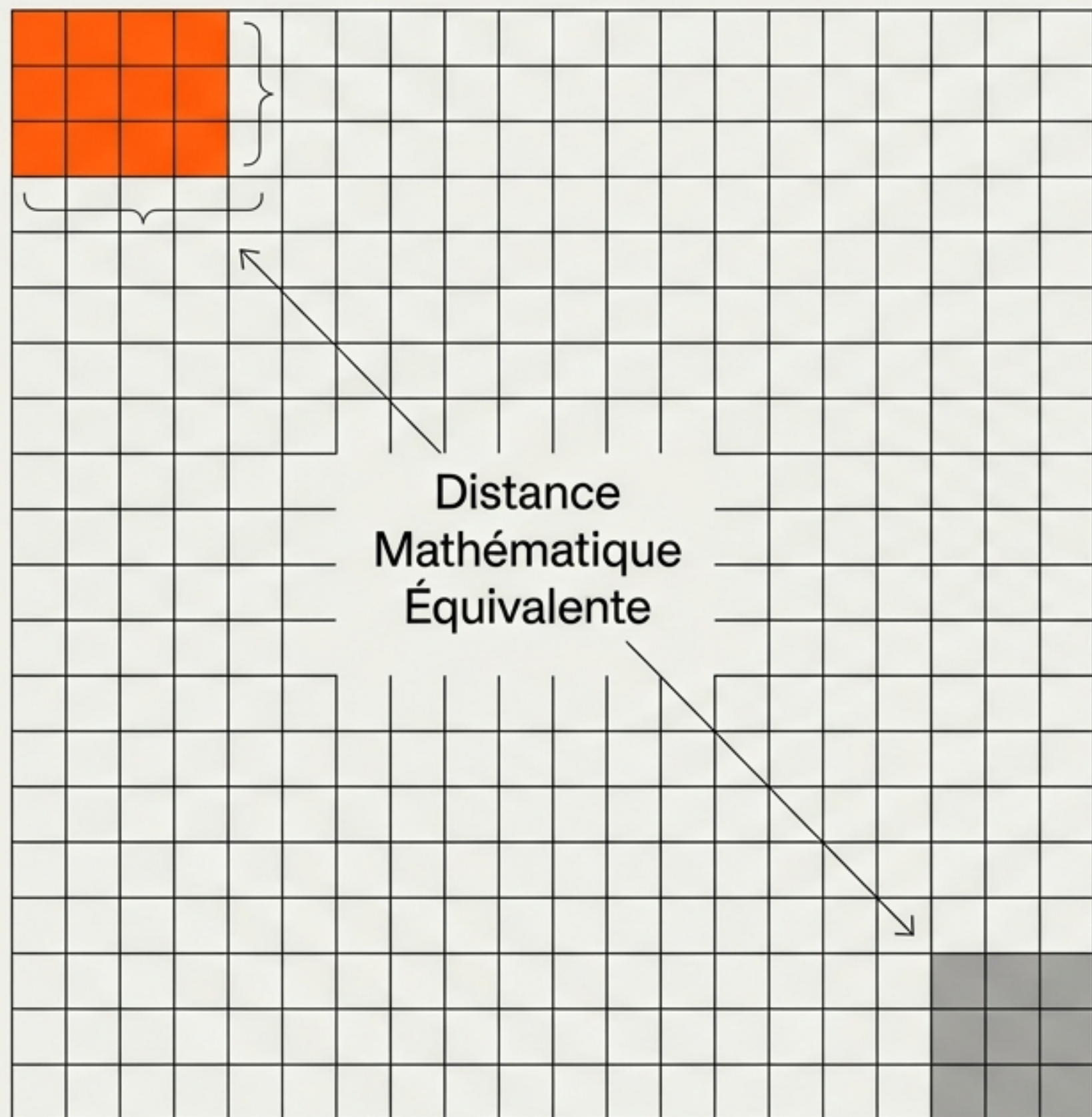
Une technologie des années 80

Ce perceptron multicouche est une technologie ancienne.

Bien qu'il soit capable de résoudre des problèmes intéressants, il est fragile.

Pour comprendre les IA modernes, il faut d'abord comprendre les limites de ces modèles classiques.



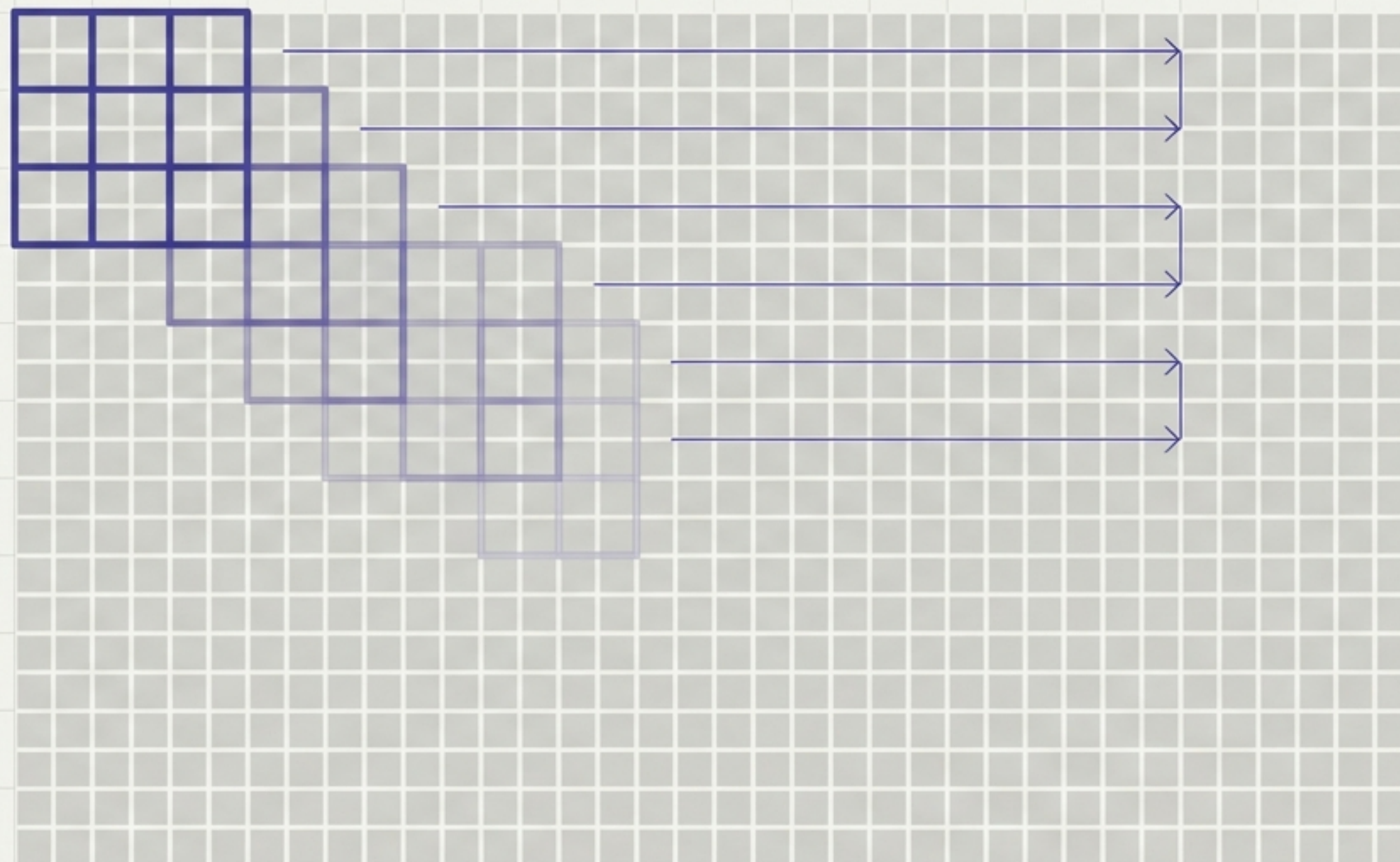


L'absence de contexte spatial

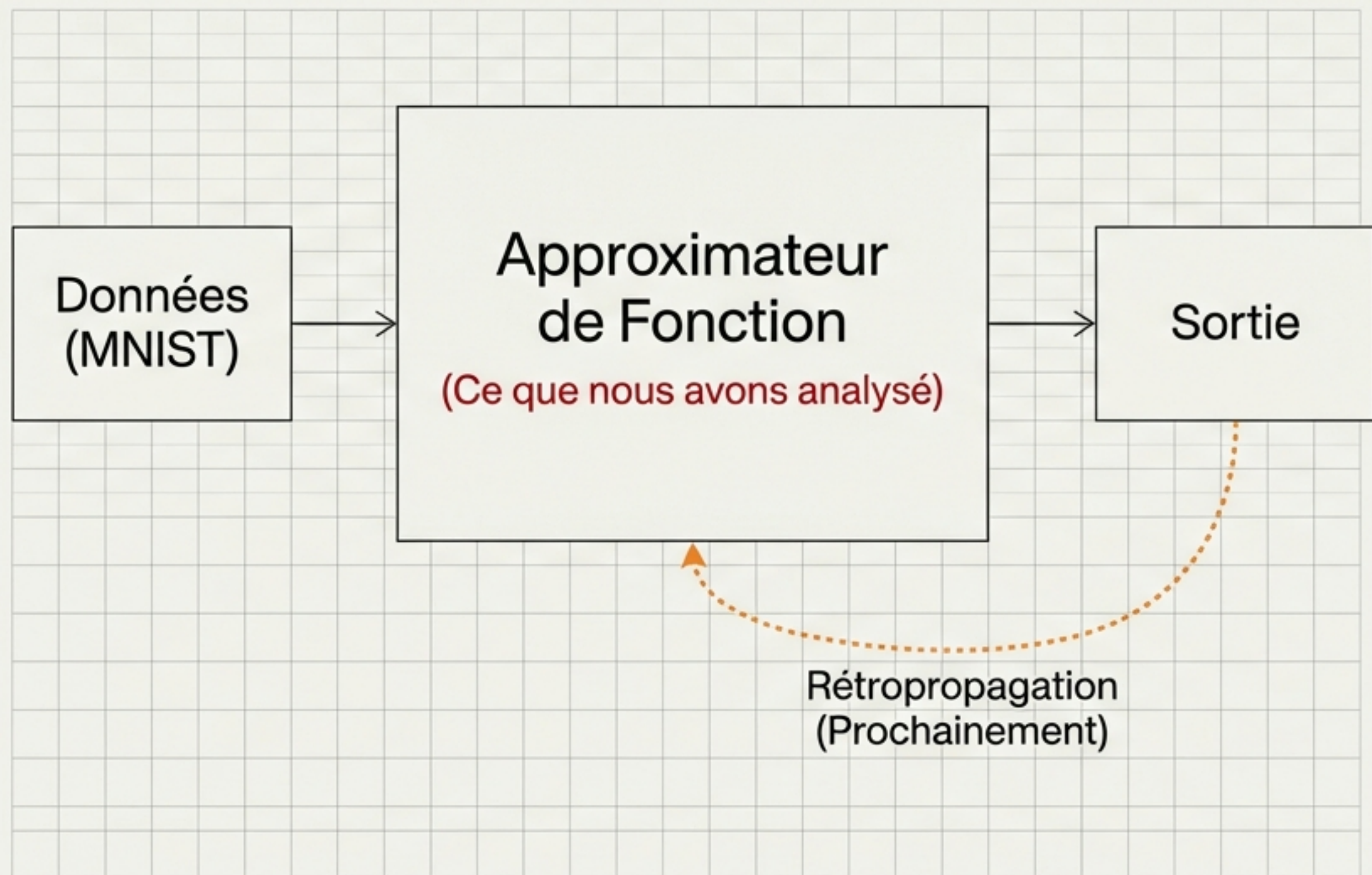
Notre algorithme n'utilise pas le fait que certains pixels sont voisins. Si l'on mélangeait l'ordre des pixels de toutes les images (entraînement et test), le réseau apprendrait tout aussi bien (ou mal). Il lui manque la capacité de transférer la connaissance d'une zone de l'image à une autre.

La solution : Réseaux de Neurones Convolutifs

Pour résoudre le problème de translation, nous avons besoin de réseaux qui analysent des *motifs* locaux, peu importe où ils se trouvent. Les **CNN** (Convolutional Neural Networks) permettent de transférer la reconnaissance d'un motif d'une partie de l'image à une autre.



En résumé



Nous avons construit un **approximateur de fonction** performant mais fragile, pas un cortex visuel.

Il a trouvé un minimum local efficace dans un espace à haute dimension.

Prochaine étape : Comment l'algorithme trouve-t-il ce minimum ? Introduction à la **Rétropropagation** (Backpropagation).