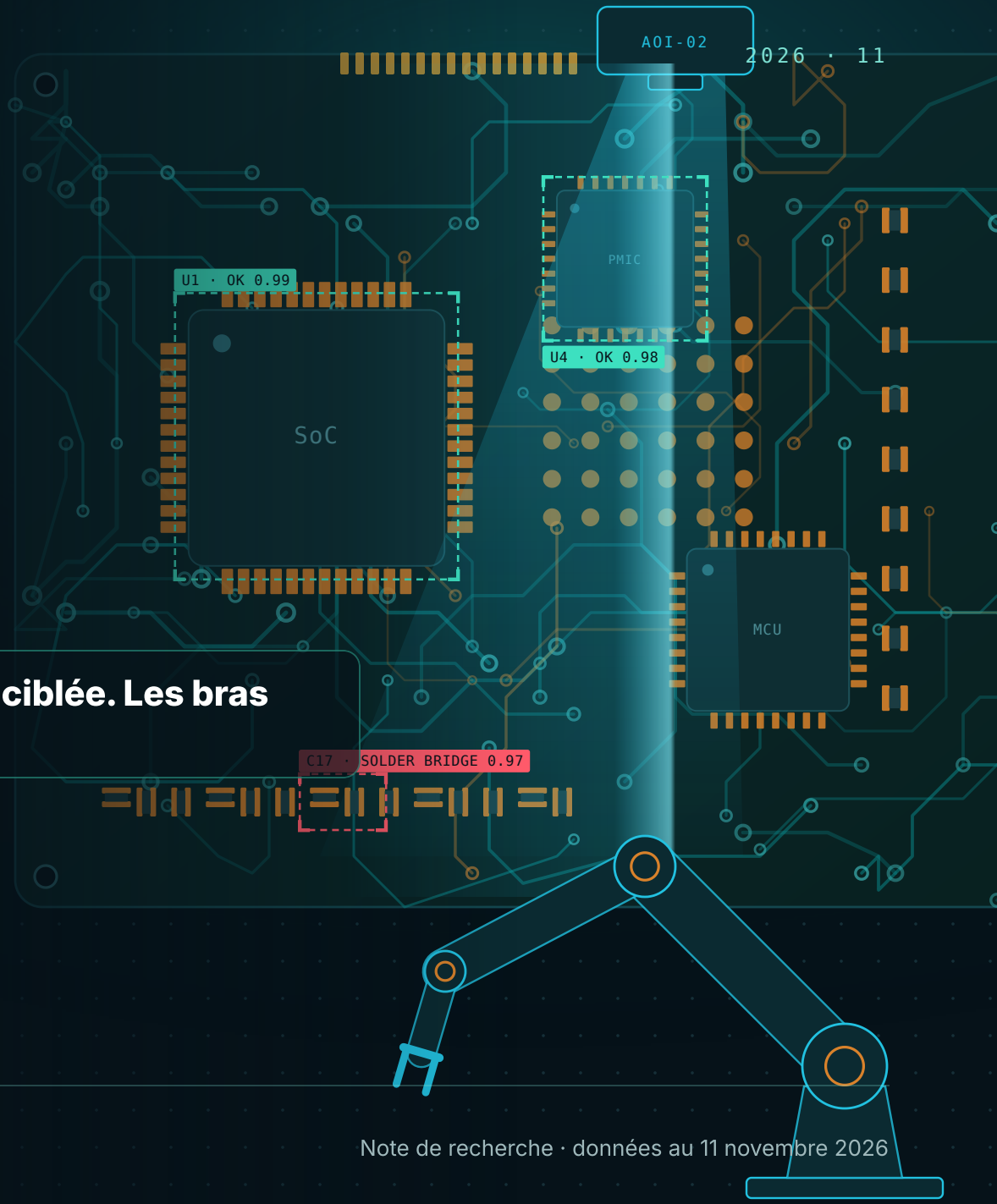
FABRICATION ÉLECTRONIQUE · INSPECTION PAR IA × ROBOTIQUE

L'IA sur la ligne SMT :

ce qui est rentable en 2026

Inspection visuelle par IA et bras robotisés dans les usines électroniques

Chine continentale · Taïwan · Asie du Sud-Est · États-Unis



CONSTAT CLÉ

La revue IA vaut la peine d'être menée, de manière ciblée. Les bras incarnés peuvent attendre.

Tan Shuai
tanshuai.comi@tanshuai.com

Note de recherche · données au 11 novembre 2026

Vaut la peine, mais de manière ciblée, maintenant et avec des partenaires

CONSTAT CLÉ

Ce qui est mature en 2026, et ce que les usines paieront, c'est la revue IA et l'inspection de niche sur images AOI, plus l'alimentation machine robotisée programmée, pas « un grand modèle qui regarde les cartes plus des robots incarnés qui les prélèvent ».

0.46–0.65 yr

Retour sur investissement de la revue IA aux États-Unis, à Taïwan et à Shenzhen, à 50 000 RMB par ligne

≥95% · <20 PPM

La revue IA interne d'un grand EMS : le niveau à atteindre pour vendre aux EMS de premier rang

100–700 RMB M

Marché de rétrofit tiers adressable en Chine continentale, ponctuel, fenêtre d'environ 5 ans

0 scores

Scores aux benchmarks de défauts industriels pour les modèles phares de novembre 2026

1.1–1.5 yr

Retour sur investissement d'une cellule cobot d'alimentation machine, deux équipes

0.003–0.14 FTE

Main-d'œuvre économisée par équipe par un bras robotisé de prélèvement : jamais rentabilisé

Source : résumé exécutif du rapport et chapitres 4–5 ; les chiffres de retour sur investissement et de marché sont des estimations de l'auteur.

Six questions, six réponses

① Ce que font tourner les usines

SPI 3D, AOI avant et après refusion, rayons X par échantillonnage, ICT/FCT. Une personne révise derrière chaque AOI ; takt ≈ 13–22 s par carte.

② Jusqu'où va la technologie

L'inspection est de niveau industriel sous forme détecteur + revue IA + quelques personnes ; pas zéro personne à aucun prix. Continent : modèles nationaux ; Taïwan : c'est le client qui décide ; États-Unis : droits de douane et certification.

③ Sur site et isolé (air gap)

Pratique ; les limites sont procédurales. Un boîtier sans ventilateur ≤ 100 W sur la ligne, des serveurs en salle des serveurs, 8 bits par défaut.

④ Enseignant et élève

Inspection : les modèles frontières pré-étiquettent, les personnes arbitrent, un jeu de référence valide. Bras : les enseignants restent les démonstrations humaines et la simulation.

⑤ Coûts et retour sur investissement

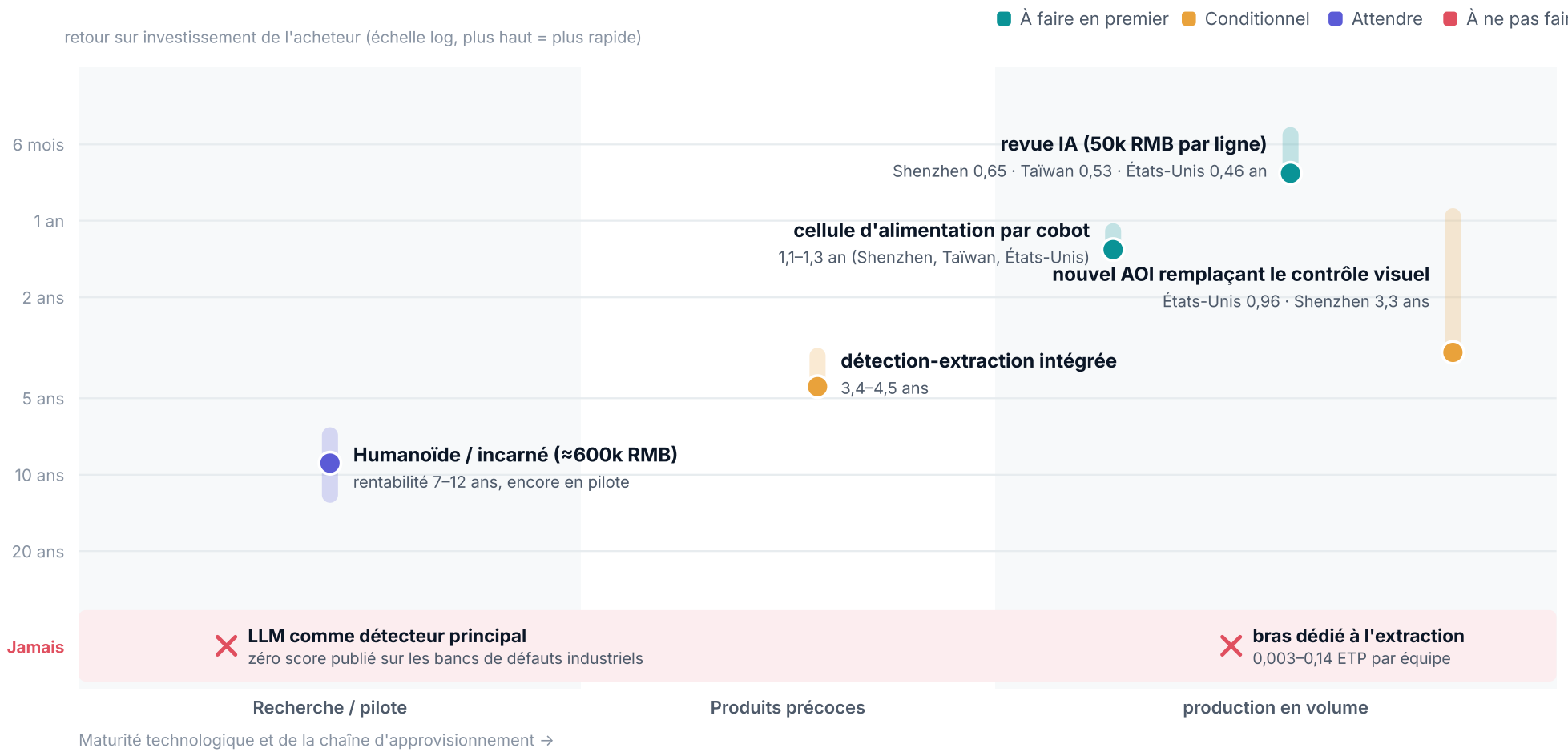
Revue IA en moins d'un an aux États-Unis, à Taïwan et à Shenzhen ; alimentation cobot 1,1–1,5 an ; Asie du Sud-Est 2–6 ans ; un bras de prélèvement, jamais.

⑥ Maintenant ou attendre

Demande réelle, offre encombrée ; l'IA des OEM laisse une fenêtre de rétrofit d'environ 5 ans. Inspection maintenant, de manière ciblée ; IA incarnée après les données de 2027–28.

Source : résumé exécutif du rapport.

Où l'IA s'insère aujourd'hui dans l'usine

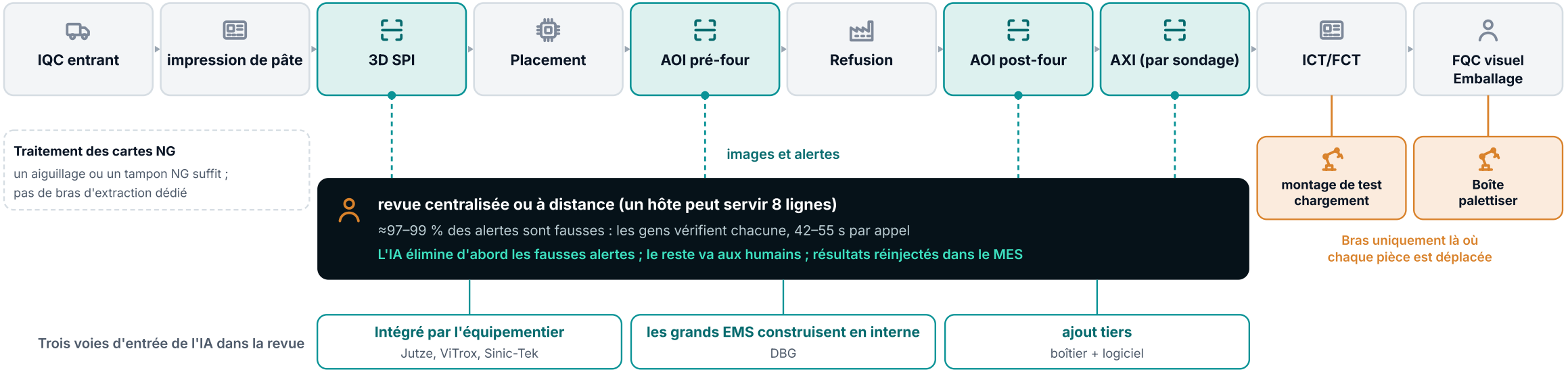


Source : Rentabilité du modèle de l'auteur (chapitre 4) ; l'axe de maturité est un jugement de l'auteur.

Les personnes sont derrière l'AOI. C'est là que l'IA entre en jeu

Une ligne SMT en volume

takt ≈13–22 s par carte · 3–5 AOI par ligne



Source : Documents publics et supports fournisseurs (Huaqin, DBG, Sinic-Tek et autres) ; compilé par l'auteur.

Environ 97–99 % des alertes AOI sont de fausses alertes

Pour 100 alertes AOI



Aujourd'hui : les gens vérifient chaque alerte

42–55 s chacun

≈448 fausses alertes par équipe sur une ligne chargée

→ 5,2–6,8 h par équipe ≈ 1 réviseur à plein temps

dotation du cas : 2 par équipe

revue IA : éliminer les fausses alertes, transmettre le reste aux humains

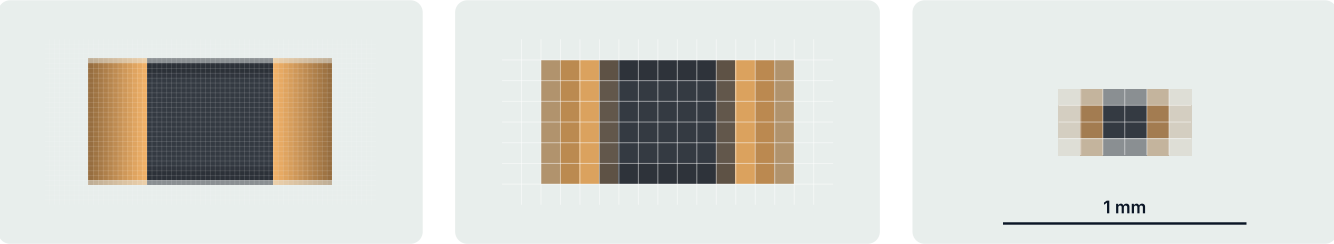
fausses alertes ou heures de revue réduites de 60–90 % (allégations fournisseur)

référence interne des grands EMS : couvre ≥95 % des composants signalés, échappements <20 PPM, revue humaine réduite ≥85 %

bascule ouverte vers les humains, la ligne continue

Source : Étude de cas Siemens, iFactory (allégation fournisseur), rapport annuel DBG 2025 ; heures de revue estimées par l'auteur.

Une caméra polyvalente n'est pas un AOI



AOI : 8–20 µm par pixel
0402 ≈ 100×50 px

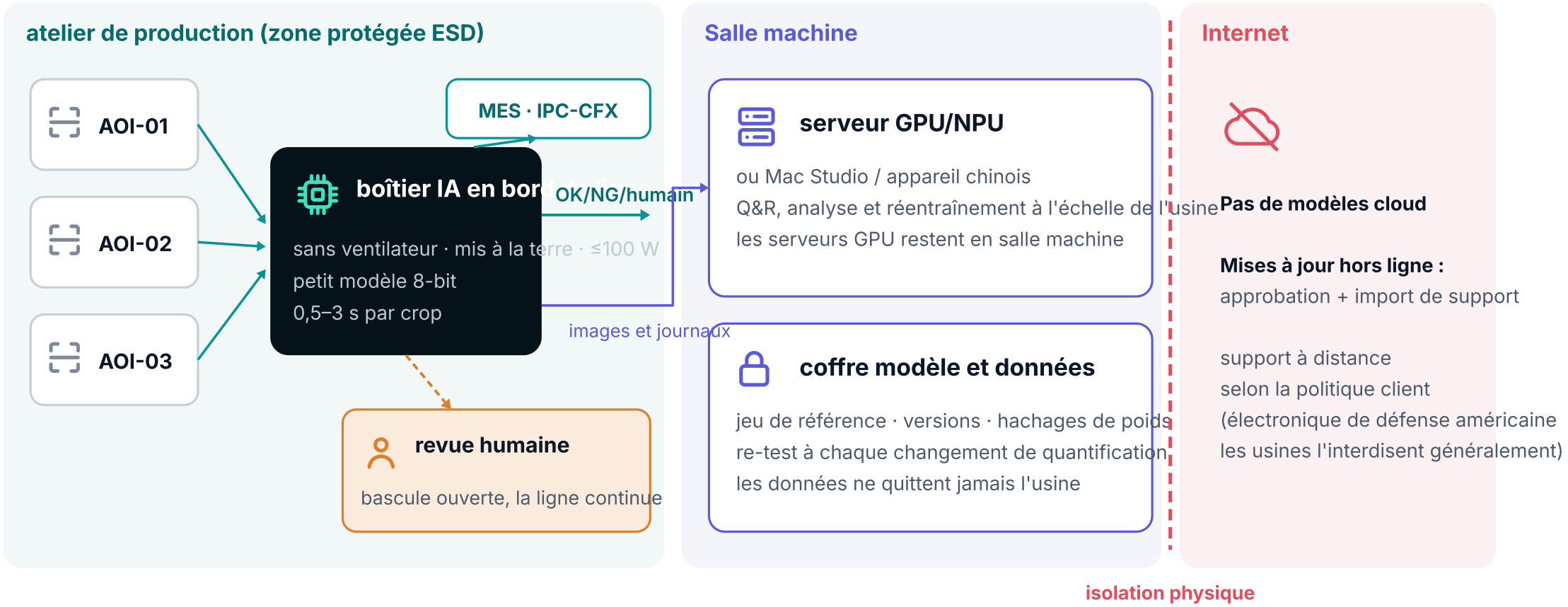
Caméra 4K sur un panneau de 300 mm : ≈78 µm/px
Le 0402 se réduit à ≈13×6 px : juste assez pour la présence et un décalage grossier

Même caméra sur un 01005
≈5×3 px : effectivement invisible

- **AOI :** 8–20 µm par pixel, éclairage multi-angle multicolore et lumière structurée
- **4K sur un panneau de 300 mm :** ≈ 78 µm par pixel ; un 0402 fait ≈ 13×6 pixels, suffisant pour la présence, pas pour les joints
- **01005 :** ≈ 5×3 pixels, pratiquement invisible
- **Donc :** l'IA travaille sur les images propres de l'AOI ; elle ne remplace pas l'optique

Source : Circuits Assembly ; calcul de l'auteur pour un panneau de 300 mm et un capteur 4K. Les grilles de pixels sont illustratives.

Un déploiement isolé (air gap) : le boîtier décide, la salle des serveurs répond

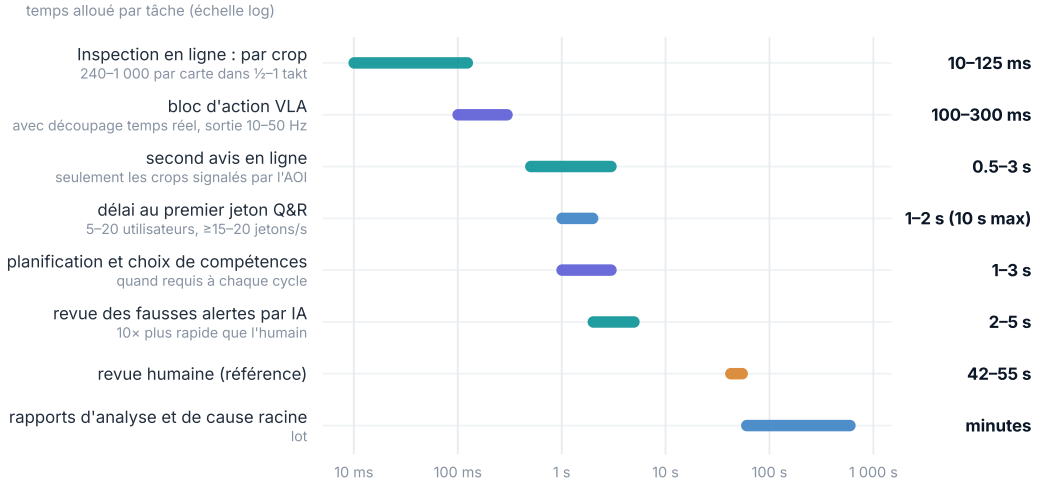


Source : Synthèse de l'auteur (chapitre 2).

Les zones propres et ESD déterminent où le calcul est placé ; les limites de vitesse dépendent de l'usage

Propreté →

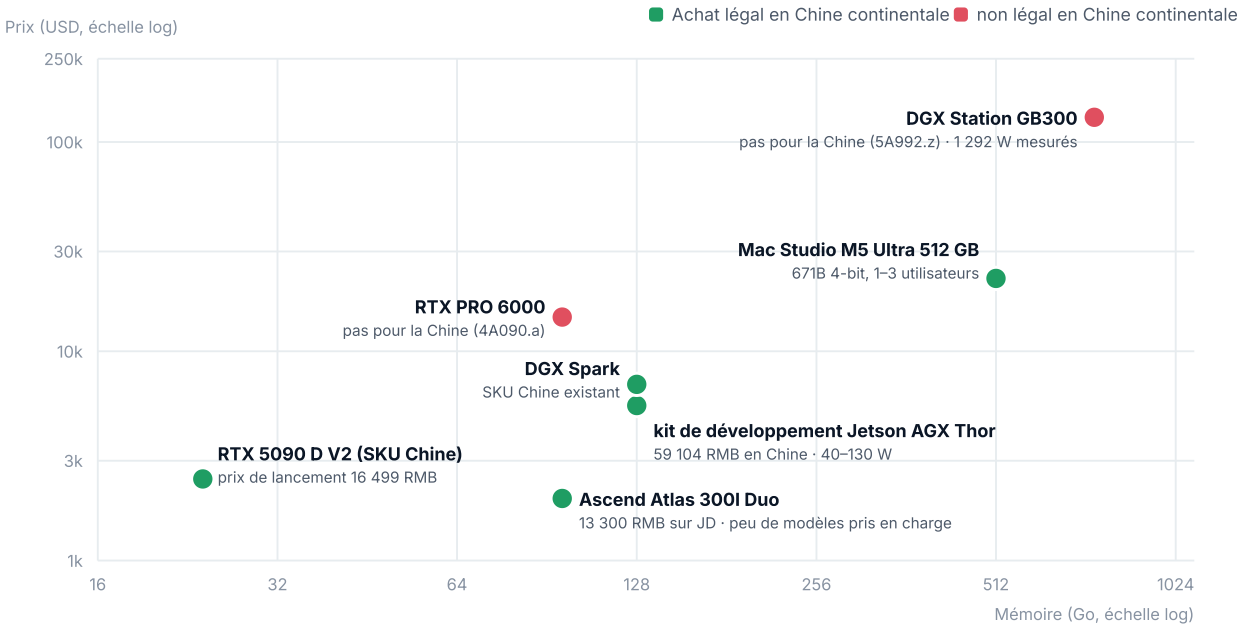
atelier SMT/PCBA ordinaire	salle blanche ISO 7–8	mini-environnement ISO 5–6	fab, ISO 5 ou plus propre
zone protégée ESD (≈ISO 8) HR 40–60 %	médical, modules caméra, VE salle blanche + ESD	puce nue, câblage, optique mini-environnement	SEMI E78 ESD · E187/E188 cyber plans spécifiques fab
calcul en bord de ligne boîtier sans ventilateur, blindage métallique mis à la terre, près de l'AOI ou du poste de revue	**calcul en bord de ligne** Sans ventilateur, étanche IP65/66, facile à nettoyer ; demander les données ISO 14644-14	**calcul en bord de ligne** Caméras et éclairages seuls ; le boîtier de calcul dehors	**calcul en bord de ligne** uniquement des contrôleurs intégrés aux outils et certifiés fab
serveur GPU/NPU salle machine ou armoire IT climatisée	**serveur GPU/NPU** hors de la salle blanche, réseau entrant	**serveur GPU/NPU** Salle machine	**serveur GPU/NPU** armoires refroidies par liquide en sous-fab ou centre de données



Point clé Un atelier SMT ordinaire est une zone protégée ESD, pas une salle blanche classée : un boîtier sans ventilateur dans un châssis métallique mis à la terre se place à l'AOI ou au poste de revue, et les serveurs GPU ne vont jamais sur la ligne.

Source : ANSI/ESD S20.20, IEC 61340-5-1, ISO 14644, normes SEMI et supports fournisseurs ; compilé par l'auteur. Synthèse de l'auteur (chapitres 1–2).

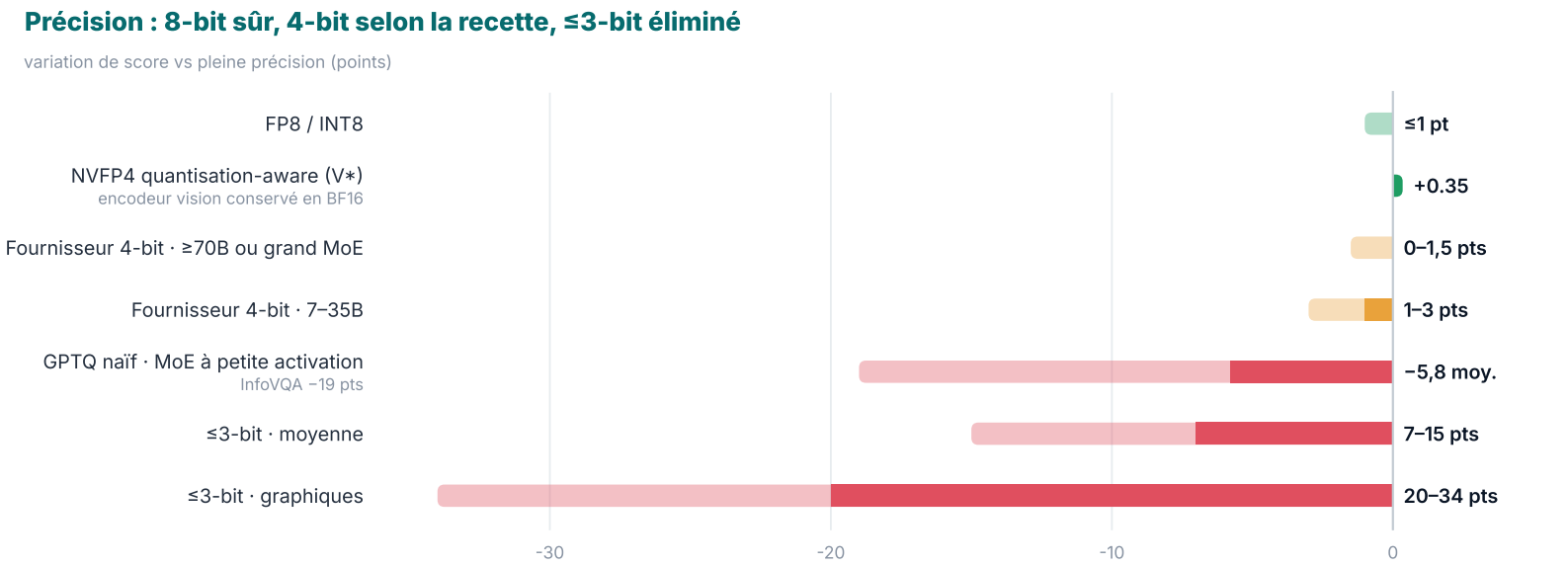
Les règles d'exportation déterminent la liste restreinte de matériel avant les benchmarks

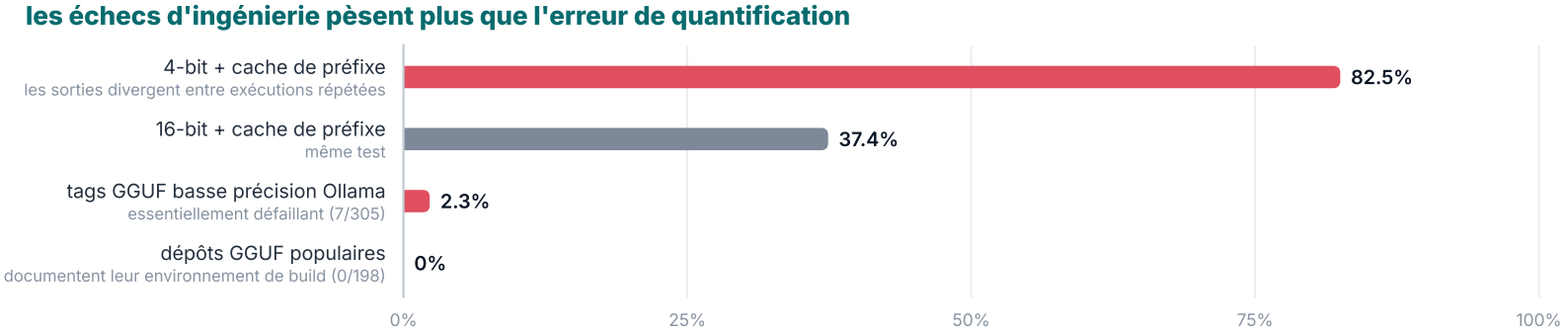


- **Jetson Thor :** 128 Go, 40–130 W ; modèles de vision 2–8B sur des recadrages AOI, VLA de classe 3B en temps réel ; 59 104 RMB en Chine
- **Mac Studio 512 Go :** DeepSeek 671B en 4 bits pour 1 à 3 utilisateurs
- **DGX Station :** 5A992.z, non destiné au continent ; un modèle 743B donne à 32 utilisateurs ≈ 4 jetons/s chacun
- **Accélérateurs nationaux :** deux serveurs Ascend à huit cartes pour des questions-réponses multi-utilisateurs pleine taille

Source : Classifications d'exportation NVIDIA, Apple, RS China, JD.com et prix de revue (novembre 2026). StorageReview (août et oct. 2026) ; ServeTheHome. Tests indépendants d'une seule source.

Quantification : 8-bit est sûr, 4-bit dépend de la recette, 3-bit et moins est éliminé

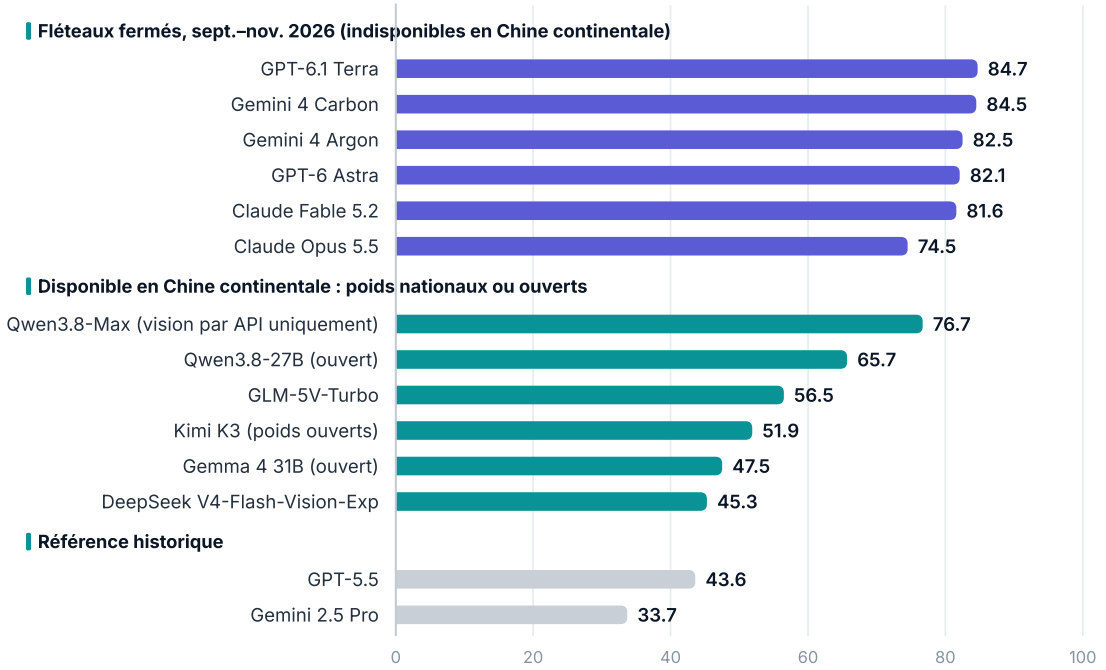




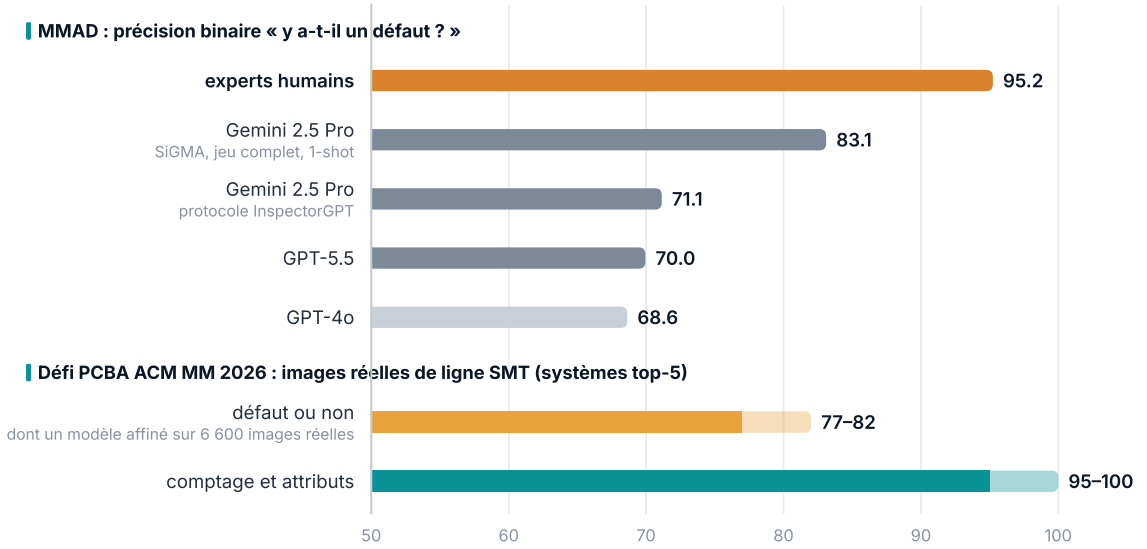
Source : Fiche modèle NVIDIA Nemotron, VLMQ, VEQ, arXiv 2609.05881, 2610.19498, 2609.04748 et autres.

La vision générique a doublé en une génération ; les scores de défauts industriels manquent

détection générique Roboflow mAP@50



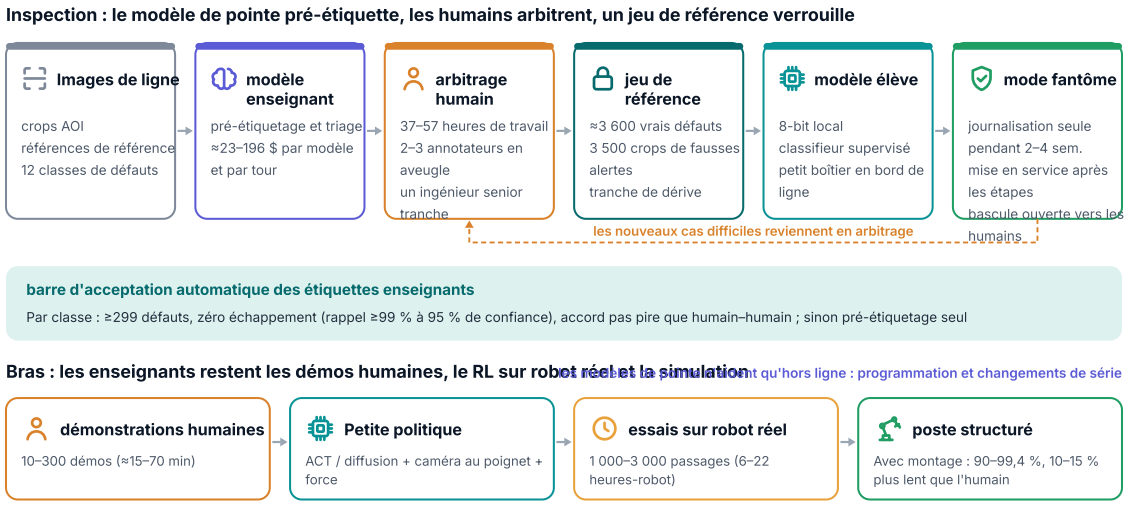
Précision ou score (%)

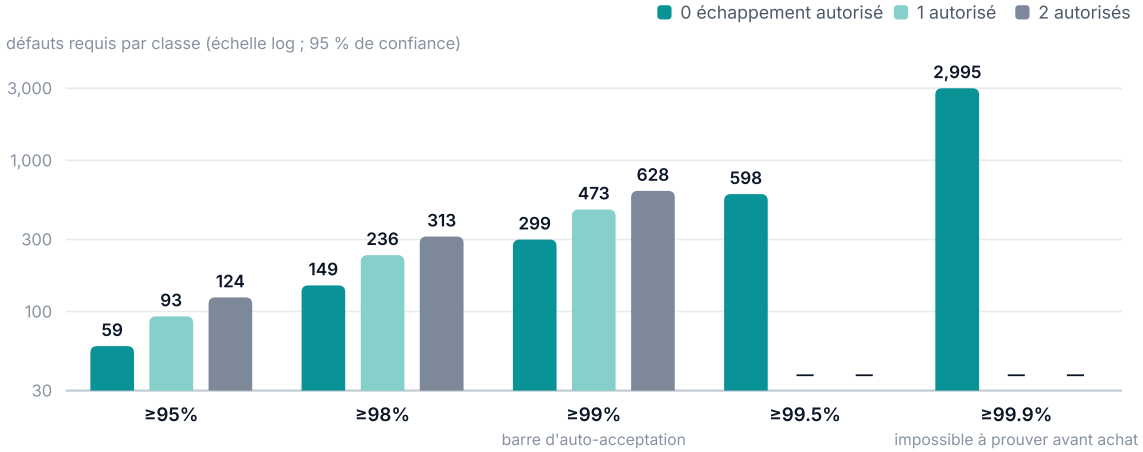


Bancs de défauts industriels (MMAD, MVTec AD 2, VisA, Real-IAD, UniPCB) : zéro score publié pour ces fléteaux 2026
les scores de détection génériques ont doublé en une génération, mais ce n'est pas un jugement de défaut ; testez sur votre propre jeu de référence

Source : Roboflow Vision Evals ; notes de version des modèles ; pages MMAD, MVTec AD 2 et autres bancs (au 11 nov. 2026). MMAD ; SiGMA ; InspectorGPT ; défi PCBA ACM MM 2026.

Les modèles frontières pré-étiquettent, les personnes arbitrent, un jeu de référence valide



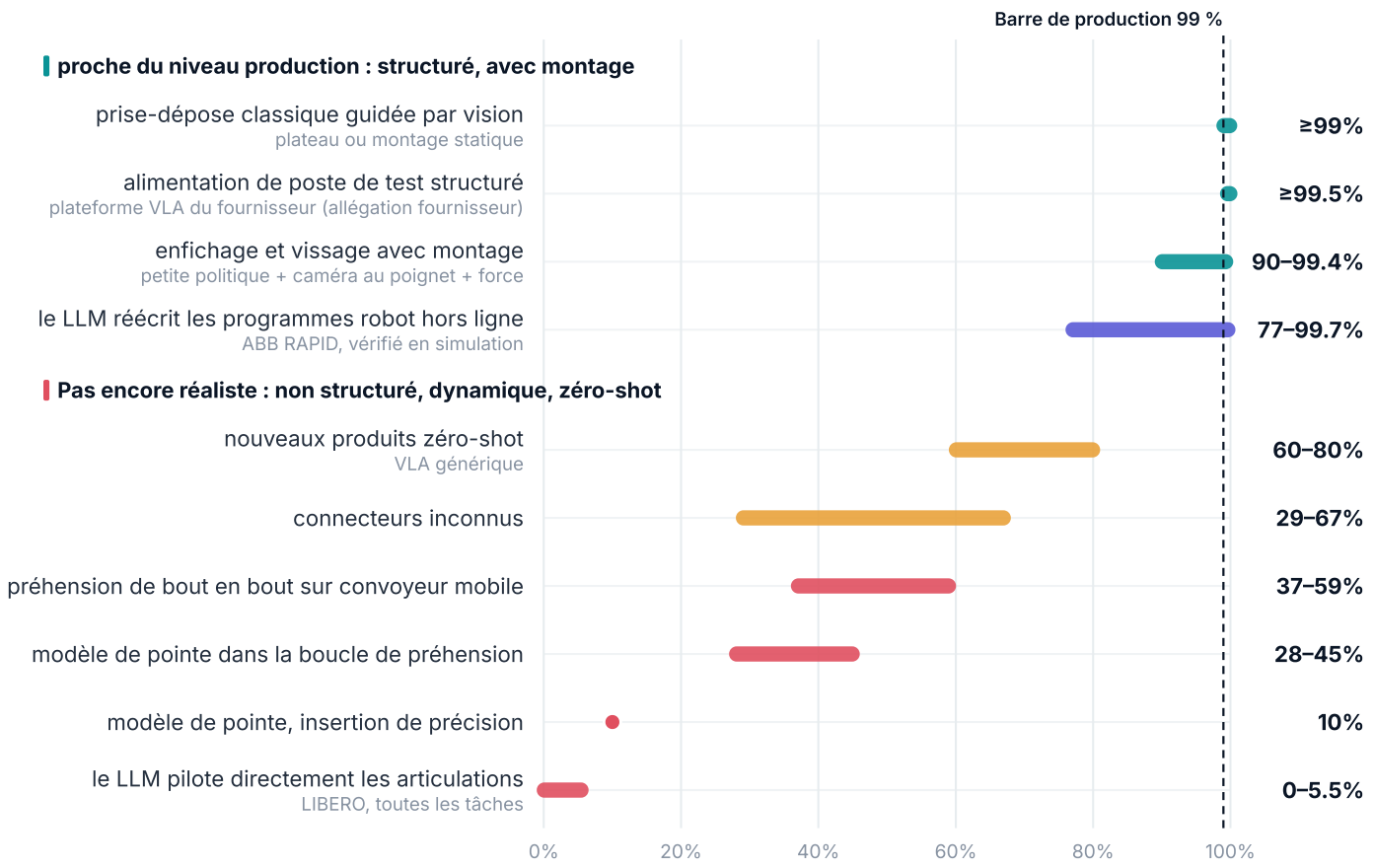


Point clé N'accepter automatiquement les étiquettes de l'enseignant que pour une classe sans aucun défaut manqué sur ≥ 299 défauts de référence (rappel de 99 % à 95 % de confiance) ; les classes critiques comme les courts-circuits, les pièces manquantes et la polarité exigent davantage.

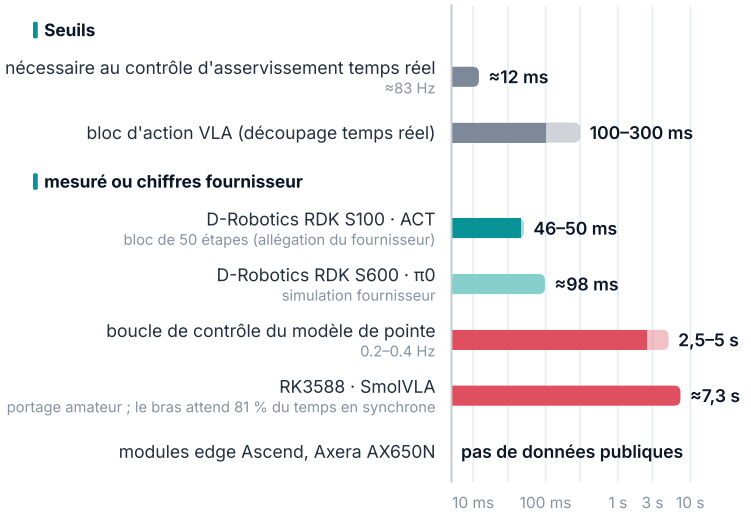
Source : Synthèse de l'auteur (chapitre 3) ; coûts et tailles d'échantillons sont des estimations. Intervalles de confiance binomiaux exacts (95 %).

Élèves robots : proches du niveau production seulement si le poste est structuré

taux de réussite de tâche

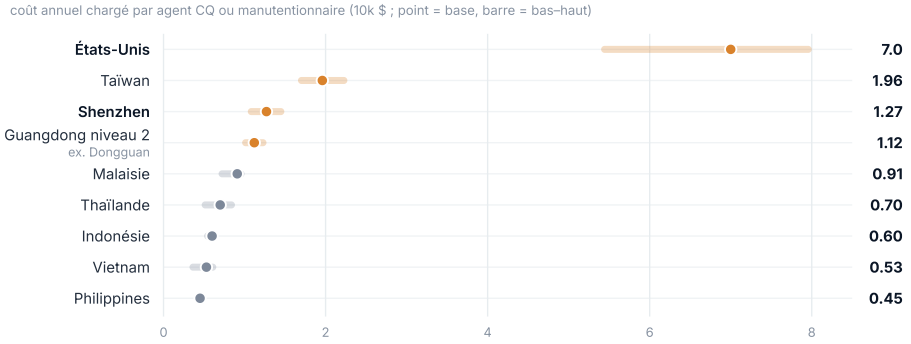


latence d'inférence par bloc d'action ou cycle de contrôle (échelle log)



Source : Articles et supports fournisseurs cités au chapitre 3 ; certains chiffres sont des allégations fournisseur. Documentation et forum D-Robotics (chiffres fournisseur) ; ports communautaires GitHub ; recherche Anthropic.

La revue IA est rentabilisée en quelques mois ; l'alimentation machine en environ un an



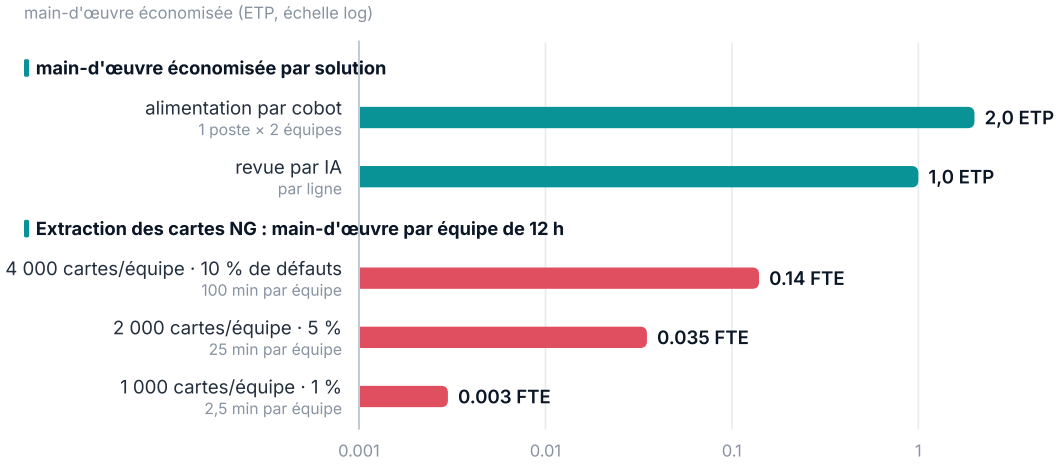


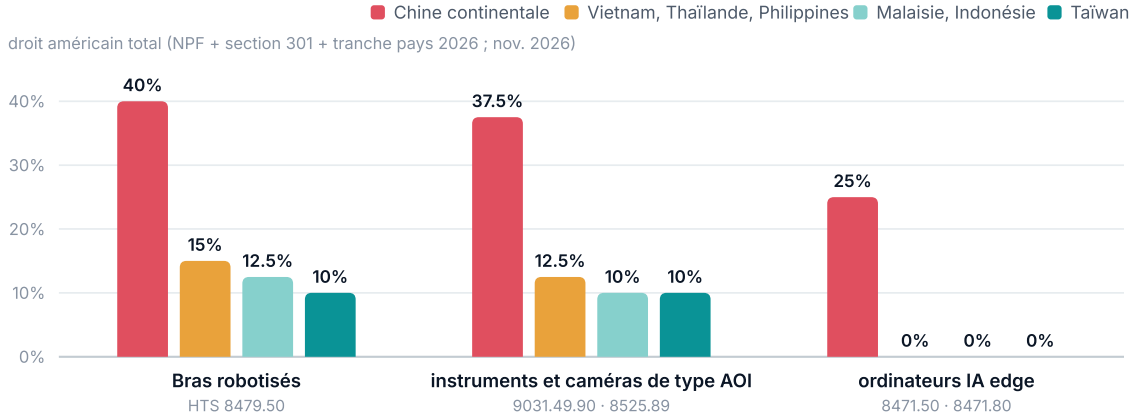
Un bras robot juste pour extraire les cartes NG ne rentabilise jamais, nulle part
Avec un rendement de premier passage >90 %, l'extraction ne prend que 0,003–0,14 ETP par équipe

Point clé À 50 000 RMB par ligne, la revue IA est rentabilisée en 0,46 an aux États-Unis, 0,53 à Taïwan, 0,65 à Shenzhen et 2,2 au Vietnam ; une cellule cobot d'alimentation en deux équipes en 1,1–1,5 an [estimations].

Source : Guidance salariale officielle de Shenzhen, ministère du Travail de Taïwan, JETRO, BLS ; estimations de l'auteur. Modèle de l'auteur : boîtier de revue IA 50k RMB par ligne ; cellule d'alimentation cobot à 200k RMB ou 140k $, deux équipes.

Un bras de prélèvement n'est jamais rentabilisé ; le lieu de fabrication d'un bras change ses droits de douane américains





Point clé Au-dessus de 90 % de rendement au premier passage, prélever les cartes NG ne mobilise que 0,003–0,14 personne par équipe ; un aiguillage NG ou un tampon suffit ; un bras fabriqué sur le continent paie aussi environ 40 % de droits de douane américains cumulés.

Source : Estimations de l'auteur. Chapitre 99 du tarif douanier harmonisé USITC ; Federal Register ; USTR (novembre 2026).

Réel mais petit : 16 000 lignes adressables, ≈ 320 M RMB ponctuels, une fenêtre de cinq ans

lignes SMT en Chine (2026)

40 000 lignes

35k–46k

× 80% × 50%

pas encore d'IA : 75–85 %
accessibles : 40–60 %

Lignes servies

16 000 lignes

10,5k–23,5k

× 20k RMB

10–30k RMB ponctuel

Marché ponctuel

320 M RMB

105–705 M RMB

÷ ≈5 ans

rétrécit de 8–12 %/an

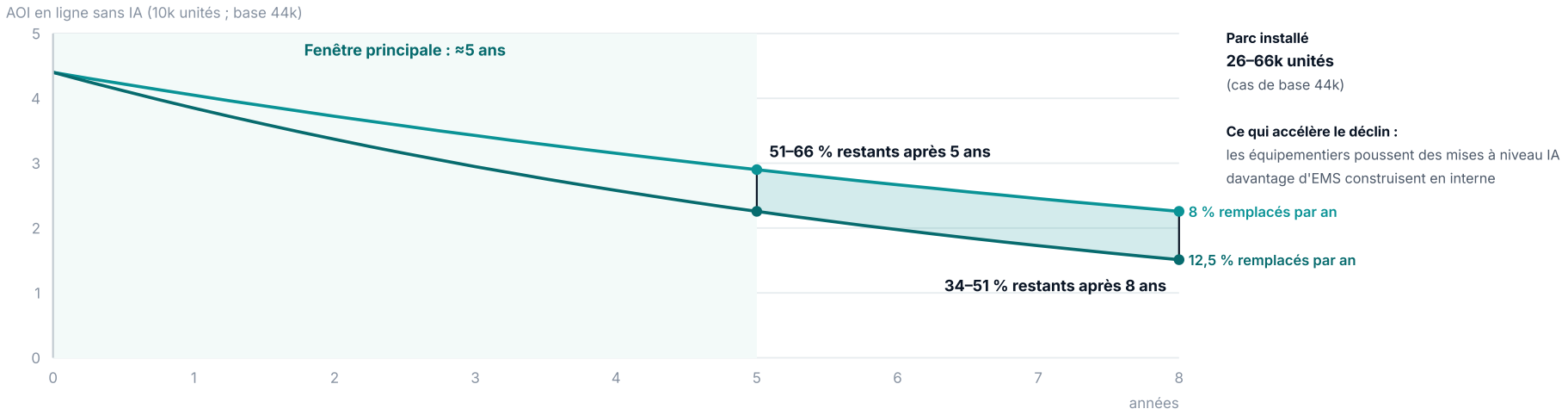
fenêtre de 5 ans

RMB 64M/yr

21–141 M RMB

Un nouvel entrant gagnant 2–5 % des lignes servies en 3 ans : ≈320–800 lignes, ≈2–5 M RMB par an

En abonnement (5k RMB par ligne et par an) : 1,6–4 M RMB récurrents, seulement 0,3–2,3 % de la prévision IDC du CQ industriel IA en Chine



Source : NVR (base 2020) ; estimations de l'auteur. Estimations de l'auteur ; parc et parts non vérifiés indépendamment.

Le créneau de la revue est encerclé ; la marge se situe dans le matériel industrialisé

les grands EMS construisent en interne

DBG (2025) : couvre ≥95 % des composants signalés
échappements <20 PPM, revue humaine −85 %
c'est la barre pour vendre aux EMS de rang 1

IA intégrée à l'AOI par les équipementiers

Jutze AI 3D AOI en production, « supprime la revue humaine à l'inspection »
ViTrox V510Ai ; Sinic-Tek codéveloppe avec Tencent
les machines neuves intègrent l'IA ; l'existant nécessite des mises à niveau

revue IA tierce

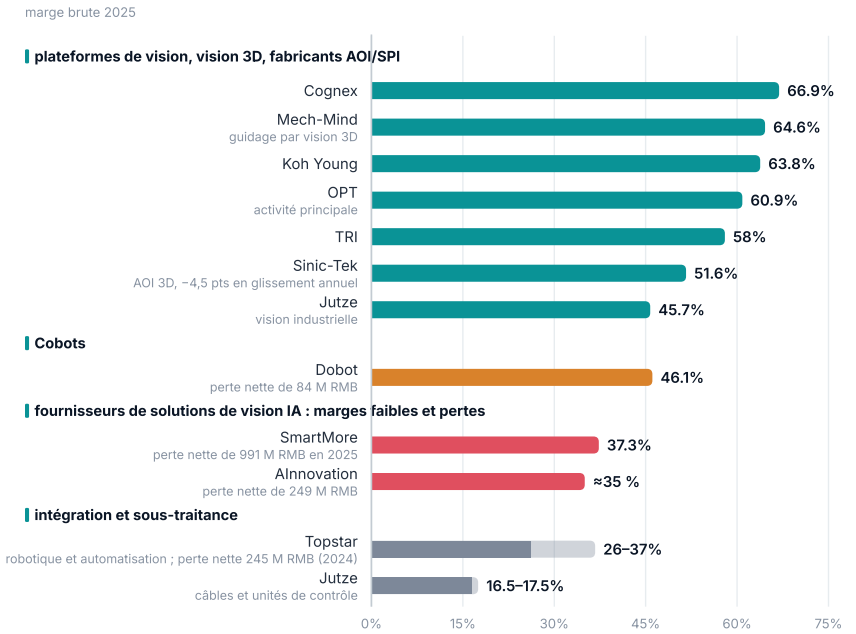
boîtier + logiciel boulonné sur l'AOI existant

fournisseurs de plateforme

Hikvision Guanlan et SmartMore IndustryGPT se vendent
fournisseur IA de tête : revenu 2025 de 1,086 Md RMB
perte nette de 991 M RMB : guerre des prix

Ce qui reste aux tiers

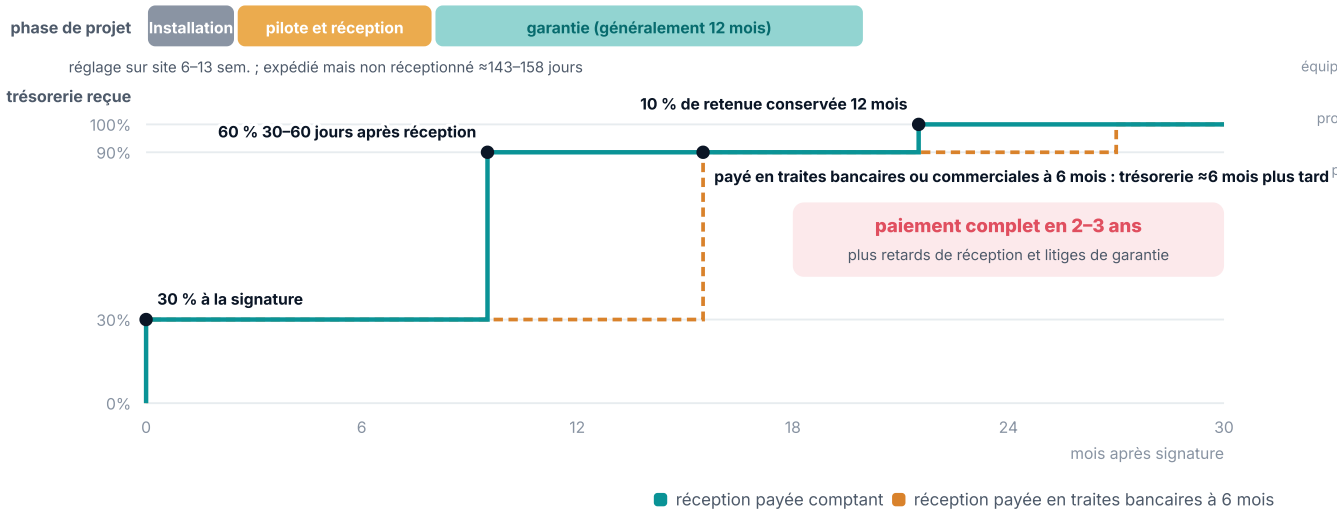
EMS de taille moyenne (≈10–40 lignes) avec AOI ancien ou multi-marques sans mise à niveau IA équipementier
Niches mal couvertes par l'AOI : THT et soudure manuelle, vernis de tropicalisation, câbles/connecteurs/FPC, aspect final et vis

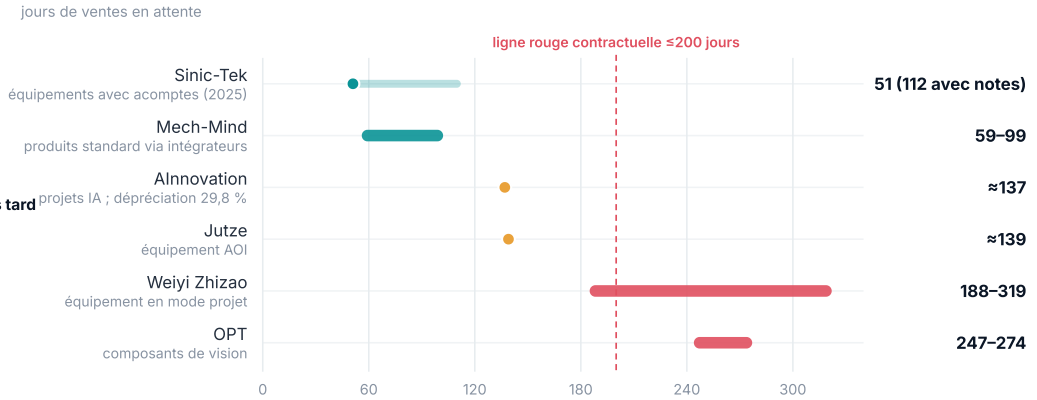


Preuves Le principal fournisseur de vision IA a réalisé un chiffre d'affaires 2025 de 1,086 milliard de RMB et une perte nette de 991 millions de RMB ; le chiffre d'affaires 3C du comparable coté le plus proche a chuté de 79,9 % sur un an en janvier–mai 2026.

Source : Rapports annuels Jutze, Sinic-Tek et DBG ; supports publics Hikvision et SmartMore ; compilé par l'auteur. Rapports annuels ou prospectus 2025.

Les acheteurs paient 30/60/10 ; le paiement intégral prend deux à trois ans





Point clé 30 % à la commande, 60 % environ quatre à neuf mois plus tard et les 10 % restants environ 16 à 21 mois plus tard ; entre un premier pilote chez un grand EMS et le paiement intégral, il peut s'écouler deux à trois ans [estimate].

Source : Réponse d'enquête Sinic-Tek, Mech-Mind, GKG, JT et autres dépôts ; schéma sur conditions usuelles, calendrier estimé. Dépôts d'entreprise ; créances (y c. actifs contractuels) ÷ revenu × 365.

Planifier hors ligne, agir avec de petites politiques, garder le mouvement et la sécurité déterministes

Couche	Ce qu'il fait, et jusqu'où va la preuve	Échelle de temps · qui la maîtrise
Planification L4	**modèles de pointe (hors ligne)** écrire ou modifier les programmes robot hors ligne, changements de série, expliquer les pannes réécrit les programmes ABB RAPID à 77–99,7 %, déployé après simulation boucle de contrôle à seulement 0,2–0,4 Hz : reste hors temps réel	secondes à minutes géants des modèles
Compétences / poli L3	**ACT, politiques de diffusion, petits VLA** postes structurés avec montage 90–99,4 % puces edge : Jetson Thor, D-Robotics RDK, Ascend, Rockchip	blocs d'action 100–300 ms fabricants de puces · fabricants de robots
contrôle de mouver L2	**Contrôleur robot** trajectoires, asservissement, contrôle de force, détection de collision code déterministe ; les politiques apprises ne fixent que les objectifs	millisecondes (le temps réel exige >83 Hz) fabricants de robots
Sécurité L1	**PLC de sécurité, scanners, limites de force et de vitesse** certifié selon ISO 10218-1/-2:2025 et les règles locales Pas de ML dans les fonctions de sécurité : évaluation par organisme notifié UE à partir du 20 jan 2027	temps réel strict robot certifié + intégrateur

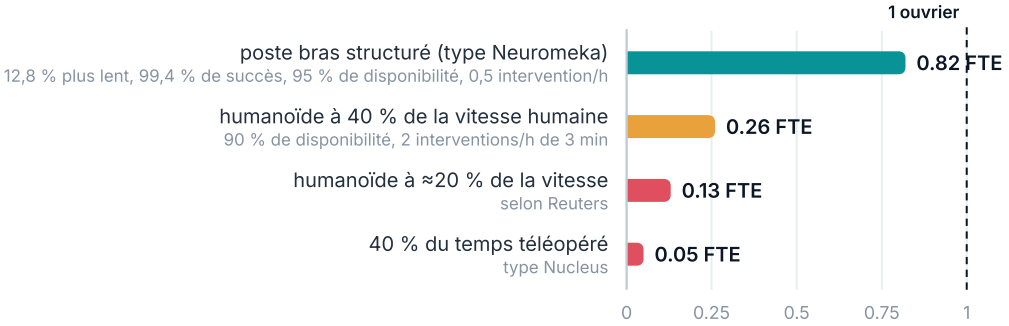
Les petites entreprises gagnent dans l'intégration et le matériel de détection : intégrer la perception apprise et de petites politiques dans des cellules robot certifiées, un poste à la fois

	taux de succès	Takt	Interventions / téléopération	Disponibilité / pannes	Volume
Figure 02 @ BMW	objectif seul >99 %/équip	objectif seul 84 s	objectif seul 0/équipe	n.d.	1 250 h
Agility Digit @ GXO	n.d.	n.d.	n.d.	n.d.	100 000 bacs
Apptronik Apollo @ Mercedes	n.d.	n.d.	n.d.	n.d.	n.d.
Agile ONE @ Audi	n.d.	n.d.	n.d.	n.d.	n.d.
Nucleus (usine non nommée)	n.d.	n.d.	≈40 % téléopéré*	n.d.	≈2 h de vidéo
AgiBot G2 @ Longcheer	99.99%	18–20 s par étape	n.d.	une panne, 2 min	64+ h
cellule hybride Neuromeka	322/324	12,8 % plus lent	n.d.	n.d.	324 cycles
π0.6 @ site client	—	—	96,4 % autonome	n.d.	1 équipe

* méthode non divulguée

Divulgué (surtout allégations fournisseur) · objectifs seuls · n.d.

main-d'œuvre nette remplacée par robot et par équipe



Source : synthèse de l'auteur (chapitres 3, 6) et estimations ; ISO 10218:2025 ; règlement européen sur les machines ; documents publics de BMW, Figure, Agility, Mercedes-Benz et autres.

Construire de manière ciblée, le prouver sur les images propres de l'usine, puis passer à l'échelle

Construire	Revue IA et classification des défauts sur des images AOI pour les EMS de taille moyenne disposant d'AOI anciens ou de marques mixtes ; inspection que l'AOI couvre mal ; cellules cobots où chaque pièce est manipulée
Ne pas construire	Une plateforme générale IA-plus-bras ; un LLM comme inspecteur principal ; un bras de prélèvement sur une ligne SMT ; des humanoïdes et des robots-en-tant-que-service ; des modèles frontières dans des boucles de contrôle ou des fonctions de sécurité

Pilote

1. **Semaines 0–2 :** accord sur les données, classes de défauts, référence
2. **Semaines 1–4 :** jeu de référence (golden set), étiqueté en aveugle et arbitré
3. **Semaines 2–4+ :** mode fantôme (shadow), comparé par classe
4. **Semaines 4–8 :** exécution assistée sur les classes validées uniquement
5. **Ensuite :** acceptation écrite, montée en charge ligne par ligne, revalidation à chaque changement

Sécurité Pas d'apprentissage automatique dans les fonctions de sécurité : l'UE exige une évaluation par un organisme notifié à partir du 20 janvier 2027, et les voies de certification de l'IA ne se stabiliseront pas avant 2027–2029.

Source : chapitres 6–7 du rapport et résumé exécutif ; ISO, IEC, UL, EUR-Lex, SAC et bases de données réglementaires nationales.

À PROPOS DE L'AUTEUR

Tan Shuai

Technologue et fondateur, avec plus de vingt ans d'expérience dans les puces, le matériel et les logiciels système, et 20 brevets délivrés. Son parcours l'a mené à travers la Silicon Valley, New York, Chicago, Taïwan, la Grande Baie (Shenzhen et Hong Kong), le Vietnam et la Malaisie, avec une attention de longue date portée à la fabrication avancée et aux chaînes d'approvisionnement. Aujourd'hui, il travaille comme chercheur et développeur en IA, en se concentrant sur la manière dont l'IA est utilisée en atelier.

Le rapport complet (toutes les figures, tableaux et sources) et l'article sont sur tanshuai.com. Il accueille volontiers les échanges avec les personnes travaillant sur l'IA dans la fabrication : i@tanshuai.com

