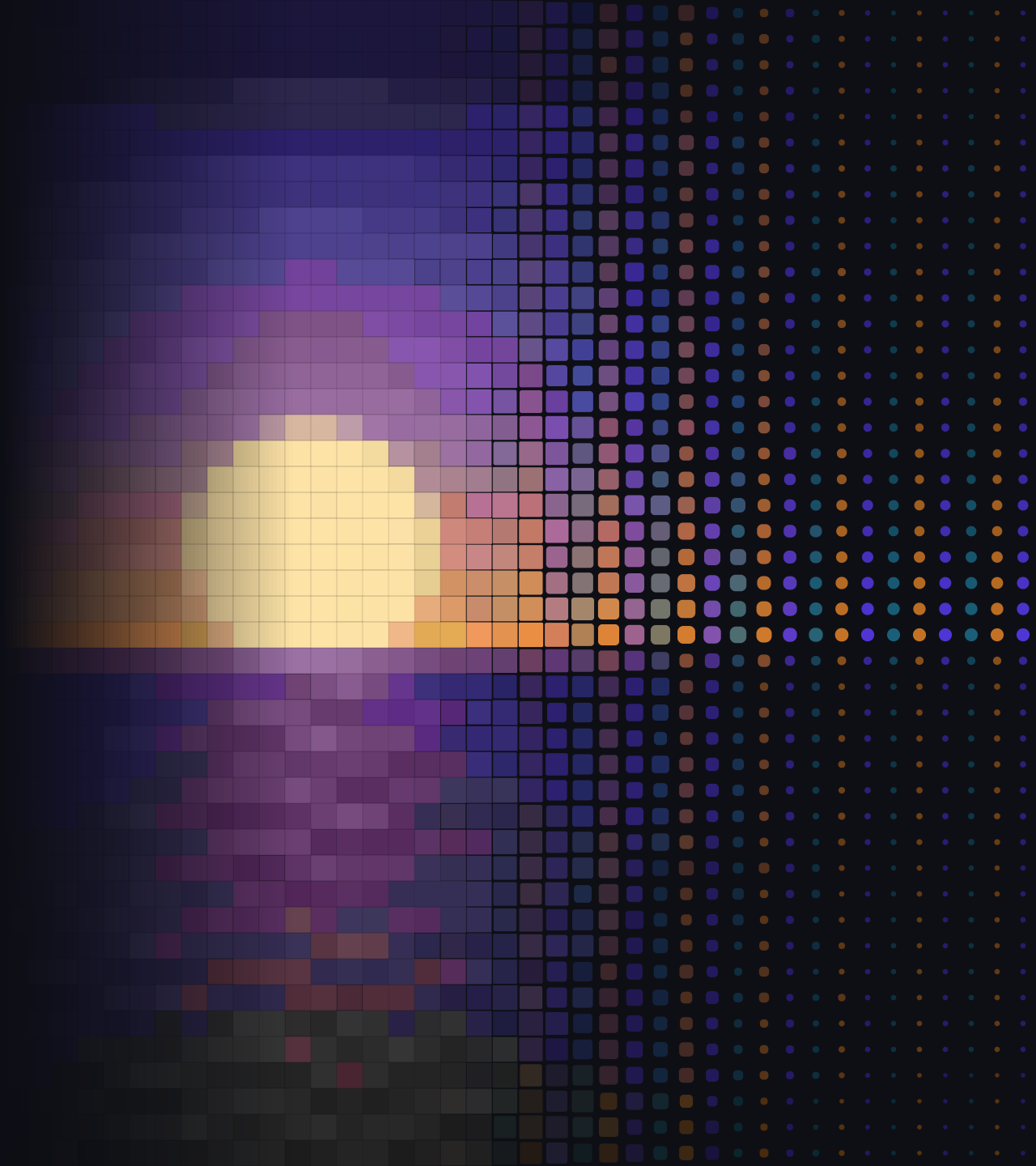
Note de recherche · novembre 2026

Le codage neuronal en 2026

Autour du codec, pas à sa place

Lecture vidéo · Photographie · Appels · IA générative · Jeux · Silicium

Tan Shuai
tanshuai.com · i@tanshuai.com



LE VERDICT

Déjà généralisé, mais autour du codec

01 Les réseaux de neurones sont courants aux deux extrémités des codecs normalisés

ISP des appareils photo, sous-échantillonnage côté encodage et super-résolution à la lecture. Entre les deux, le flux binaire reste en AV1, HEVC, AVC ou VVC, décodé par du matériel à fonction fixe.

02 Les codecs neuronaux de bout en bout restent en boucle fermée

Appels, tokeniseurs d'IA générative et un jeu AAA : une seule entreprise contrôle les deux extrémités et peut se replier sur un codec normalisé.

03 Les flux binaires neuronaux sur les écrans grand public attendent H.267

La version 1 est attendue fin 2029 ; premières puces vers 2031 et usage généralisé au milieu des années 2030 estimation.

≈ 75 %

du trafic de données mobiles est de la vidéo (fin 2025)

1,26 milliard

de smartphones livrés en 2025, qui décodent tous la vidéo avec du matériel à fonction fixe

≈ 12×

l'énergie par pixel d'un décodeur HEVC matériel, pour la seule conception publiée de puce de décodage neuronal

Trois façons pour un réseau de neurones d'entrer dans un codec



Autour du codec : aucune autorisation requise. Dans une norme : il faut une norme. À la place du codec : il faut contrôler les deux extrémités.

Où le codage neuronal est déployé en 2026

	Autour d'un codec normalisé prétraitement, amélioration, mise à l'échelle	Codec neuronal de bout en bout encodeur et décodeur appris	Dans une norme outils neuronaux dans un flux binaire ou une API normalisés
01 **Lecture et streaming**	**À GRANDE ÉCHELLE** Sous-échantillonneur de Netflix ; super-résolution chez YouTube et Meta ; mise à l'échelle par IA des téléviseurs	**NON DÉPLOYÉ** Aucun service de streaming n'envoie de format vidéo neuronal ; téléviseurs et téléphones décodent les codecs normalisés en matériel	**EN NORMALISATION** La version 1 de H.267 est attendue fin 2029 ; des outils neuronaux figurent parmi les candidats
02 **Photographie et imagerie**	**À GRANDE ÉCHELLE** ISP IA dans les téléphones haut de gamme ; un ISP entièrement neuronal présenté comme IP au CES 2026	**DÉMOS, PILOTES** JPEG AI est une norme ISO : démos sur téléphone, aucun produit	**NON DÉPLOYÉ** Le format qui progresse est classique : JPEG XL activé par défaut dans Chrome 155
03 **Appels et réunions**	**À GRANDE ÉCHELLE** Super-résolution côté récepteur et suppression neuronale du bruit	**À GRANDE ÉCHELLE** Audio neuronal dans les appels Teams à deux depuis 2021 ; vidéo MLVC en cours de déploiement dans Teams	**DÉPLOIEMENT LIMITÉ** Opus 1.5 ajoute des outils ML optionnels et rétrocompatibles
04 **IA générative**	—	**À GRANDE ÉCHELLE** Chaque modèle vidéo et vocal utilise un tokeniseur appris, par ex. un VAE vidéo 1:192	—
05 **Jeux et graphisme**	**À GRANDE ÉCHELLE** Mise à l'échelle neuronale et génération d'images : DLSS 4 seul dans plus de 250 jeux et applications	**DÉPLOIEMENT LIMITÉ** Textures neuronales livrées dans un jeu AAA ; le SDK de NVIDIA est en bêta	**EN NORMALISATION** DirectX ajoute des opérations matricielles pour les shaders (préversion)
06 **Liaisons spécialisées : satellites, robots**	**DÉMOS, PILOTES** Restauration côté serveur de flux à faible débit	**DÉMOS, PILOTES** Codecs neuronaux résistants aux pertes, en recherche et en pilote	—

> Les domaines sont classés selon le nombre de personnes qu'ils touchent.

Neuronal aux deux extrémités, normalisé au milieu

- **À l'encodeur**

  Le sous-échantillonneur neuronal de Netflix tourne depuis 2022 sans rien changer aux appareils ; environ 77 % des spectateurs l'ont préféré lors de ses tests selon le fournisseur.

- **À l'écran**

  YouTube Super Resolution sur les téléviseurs depuis octobre 2025 ; Meta applique la super-résolution sur ses serveurs et sur les téléphones.

- **Dans le téléviseur**

  Le processeur NQ8 AI Gen3 de Samsung revendique 768 réseaux de neurones pour la mise à l'échelle.

LE FLUX BINAIRE RESTE NORMALISÉ

30 %

des visionnages Netflix en AV1 (décembre 2025)

> 90 %

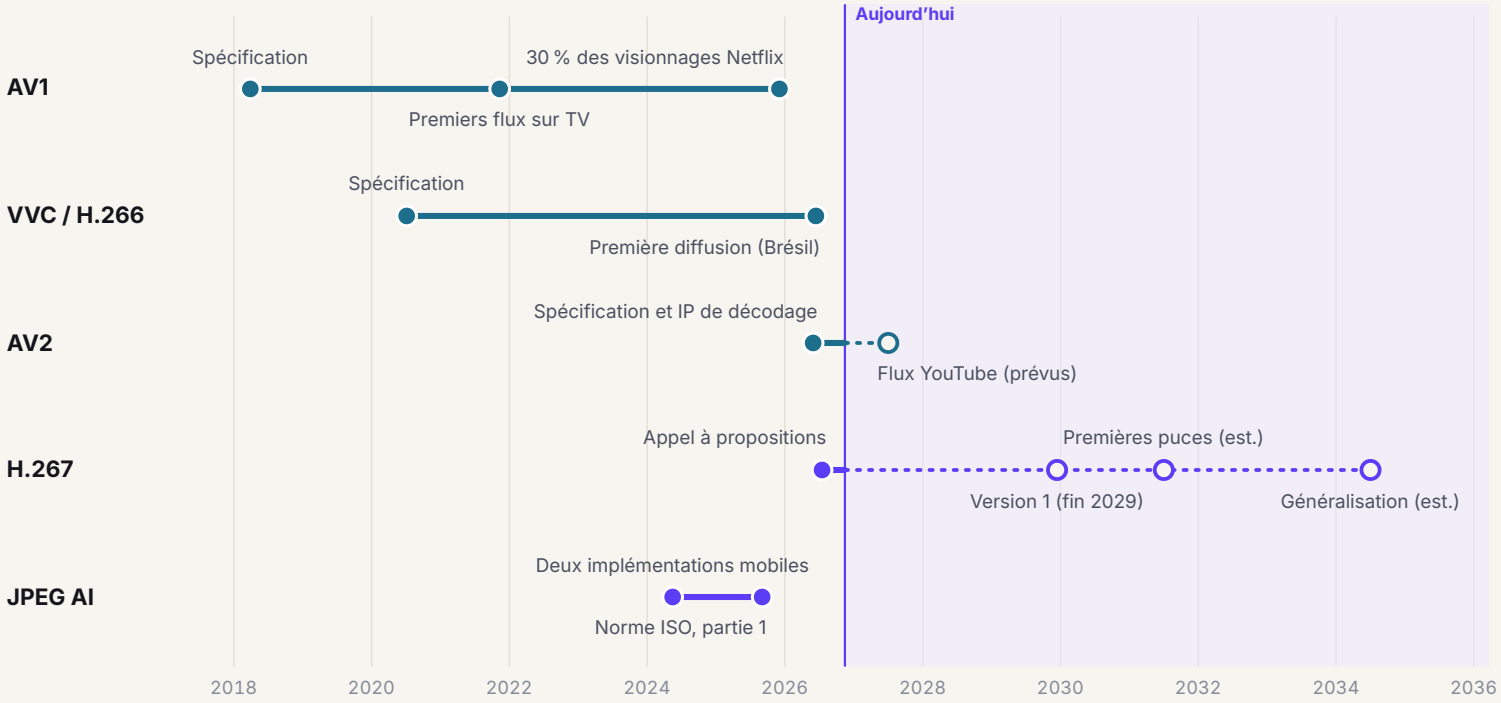
du catalogue YouTube, pondéré par le temps de visionnage, disponible en AV1 (mars 2026)

3 semaines

entre la version 1.0.0 d'AV2 (28 mai 2026) et des clients pour l'IP de décodage chez trois fournisseurs

CALENDRIER

D’une norme finalisée aux écrans grand public, il faut près d’une décennie



Il a fallu près de huit ans à l’AV1 pour passer de sa spécification (2018) à 30 % des visionnages Netflix (décembre 2025). Les marqueurs creux sont des objectifs ou mes estimations.

H.267 décidera si les outils neuronaux atteignent tous les écrans

Octobre 2025
Appel à preuves : cinq réponses de 16 entreprises et laboratoires ; gains neuronaux les plus nets en 8K HDR

Juillet 2026
Appel à propositions lancé à Genève

Janvier 2027
Réponses évaluées à Brisbane

Second semestre 2027
Premier modèle de test

Fin 2029
Version 1 de la norme

≈ 28 %
économisés sur VVC par la piste classique ECM

6–14 %
économisés par les filtres en boucle neuronaux, pour 10 000 à plus de 22 000 MAC par pixel

SEI
les flux VVC peuvent déjà signaler un post-filtre neuronal à appliquer par les décodeurs

Ma lecture : un cœur classique avec quelques outils neuronaux là où les bits économisés justifient le coût au décodeur analyse.

Neuronal dans l’appareil photo, classique sur le Web

- **Appareil photo**

  Le Cognitive ISP de Qualcomm segmente l’image en temps réel dans les téléphones depuis 2023 ; un ISP IA complet tournant sur un NPU a été présenté au CES 2026.

- **JPEG AI**

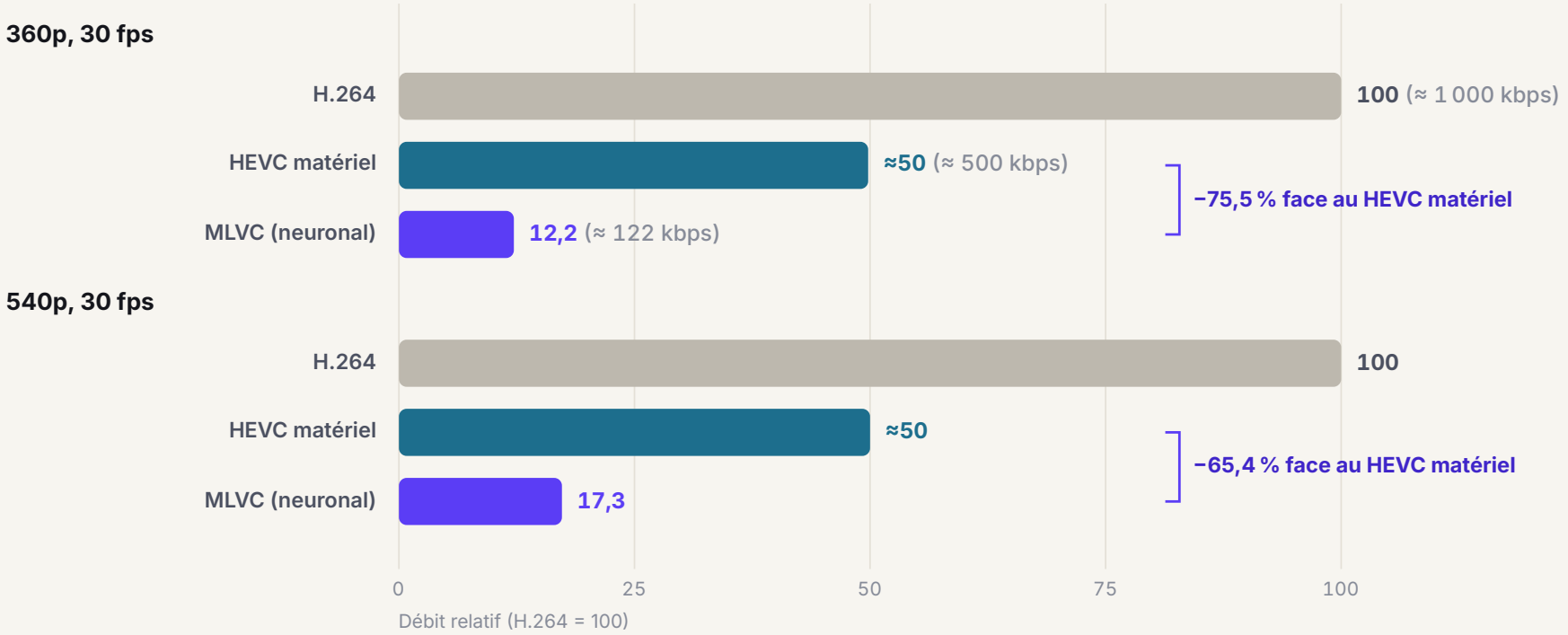
  Publiée en 2025 : 12,5–27,9 % de mieux que VVC intra pour 7 000 à 216 000 MAC par pixel. Je n’ai trouvé aucun produit qui l’utilise par défaut.

- **Web**

  Chrome 155 a activé JPEG XL par défaut le 6 octobre 2026, avec un décodeur à sûreté mémoire écrit en Rust.

> **Les formats qui doivent s’ouvrir partout restent classiques ; le codage neuronal gagne dans la chaîne d’une seule entreprise.**

Les appels, là où les codecs neuronaux de bout en bout sont déployés



Satin tourne dans tous les appels Teams et Skype à deux depuis 2021. MLVC, open source sous licence MIT depuis le 24 juillet 2026, se déploie dans les appels Teams pair à pair jusqu'en 540p30, avec au plus la moitié du NPU **selon le fournisseur**.

Toutes les plateformes ne sont pas passées au neuronal, et leurs raisons sont instructives

DANS UNE NORME, EN OPTION

Opus 1.5 et 1.6

Amélioration neuronale, masquage des pertes de paquets, redondance profonde et extension de bande passante, toutes désactivées par défaut et entièrement compatibles avec la norme de 2012.

CLASSIQUE, PAR CHOIX

Meta MLow (2024)

Plus de 20 % des appels de Meta passent par des téléphones ARMv7, et des dizaines de millions d'appels WhatsApp par jour par des téléphones de plus de dix ans. MLow offre environ deux fois la qualité d'Opus à 6 kbps avec 10 % de calcul en moins.

Le budget de complexité du parc installé décide. Les codecs neuronaux arrivent d'abord là où les appareils sont récents et où un seul fournisseur contrôle les deux extrémités.

Le plus grand déploiement, que personne n’appelle un codec

- **Des codecs appris partout**

  Chaque modèle vocal et vidéo lit et écrit les jetons ou les latents d’un codec appris.

- **Pas des formats de distribution**

  Chaque modèle a son propre espace latent : ni régulation de débit, ni test de conformité, ni second décodeur.

- **Une ligne de coût**

  Pour quiconque fournit de la vidéo générative, un latent plus compact réduit le coût de génération plus que proportionnellement analyse.

DEUX EXEMPLES

12,5 Hz · 1,1 kbps

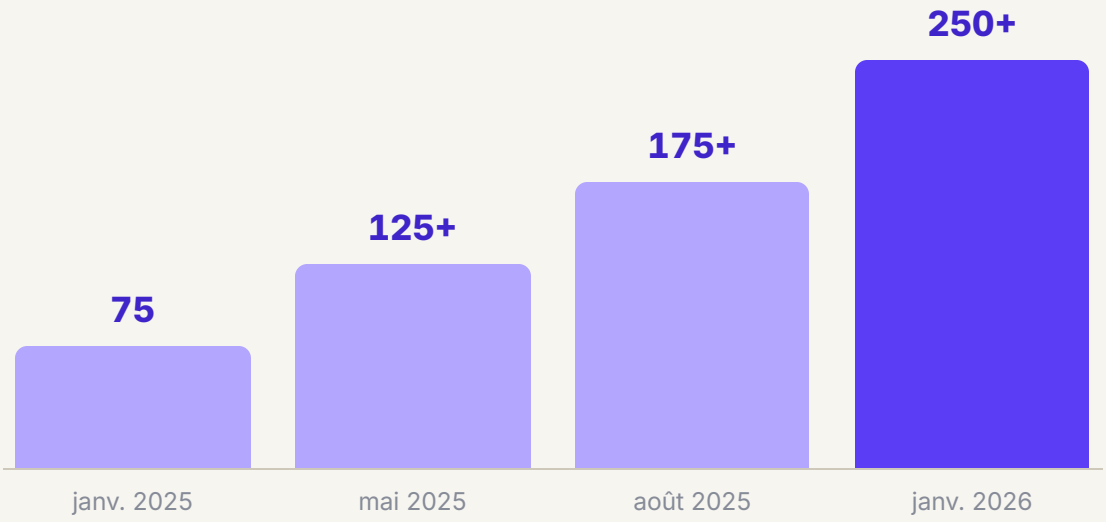
jetons de parole Mimi de Kyutai, sous tout codec de téléphonie

1:192

le VAE vidéo de LTX-Video : un jeton latent pour 32×32 pixels sur 8 images

La reconstruction neuronale est généralisée ; les assets neuronaux démarrent

DLSS 4 Multi Frame Generation : jeux et applications



Compression neuronale de textures

Scène de démo NVIDIA, GTC 2026

Mémoire de textures 6,5 GB → 0,97 GB

85 % de mémoire en moins

Assassin's Creed Mirage, commercialisé

Certains assets seulement ; face aux formats par blocs

≈ 30 % de compression en plus

Mesures différentes : pas directement comparables

DirectX Linear Algebra est entré en préversion publique en avril 2026 : un seul chemin de code pour le décodage neuronal dans les shaders, sur le GPU de tout fournisseur.

Satellites, drones et robots : le meilleur terrain, un petit marché

- **Pourquoi c’est adapté**

  Un seul propriétaire contrôle les deux extrémités, la bande passante est rare, et le contenu assez restreint pour entraîner un modèle.

- **Chine**

  L’annonce n° 49 de 2026, en vigueur le 18 novembre, révise le contrôle des exportations de drones ; une rubrique retient la compression sémantique d’image comme seuil.

- **États-Unis**

  La FCC a ajouté les drones de fabrication étrangère et leurs composants critiques à sa Covered List le 22 décembre 2025.

> **Un créneau solide pour un spécialiste capable de gérer la conformité à l’export, pas un marché de masse.**

Un ordre de grandeur de plus en énergie par pixel

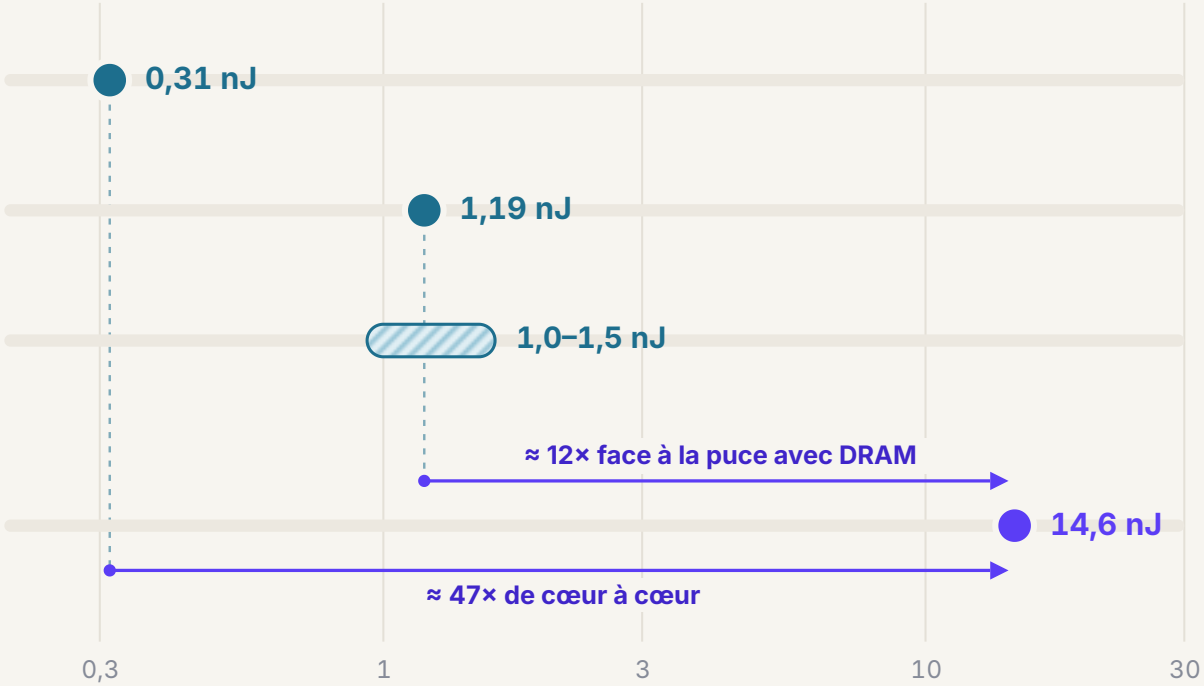
Puce de décodage HEVC, cœur
40 nm, 4K30, mesuré

Puce de décodage HEVC, avec DRAM
40 nm, 4K30, cœur mesuré et DRAM simulée

Décodeur VVC
28 nm, 4K60, projection

Décodeur vidéo neuronal (NVCA)
28 nm, 1080p25, synthèse seule ; poids et DRAM non comptés

En 1080p30, cela représente environ 74 mW contre 908 mW.



En 1080p30 : environ 74 mW pour la puce HEVC avec DRAM, contre environ 900 mW pour NVCA. Nœud plus ancien pour la puce HEVC, poids et DRAM non comptés pour NVCA : la comparaison avantage la conception neuronale.

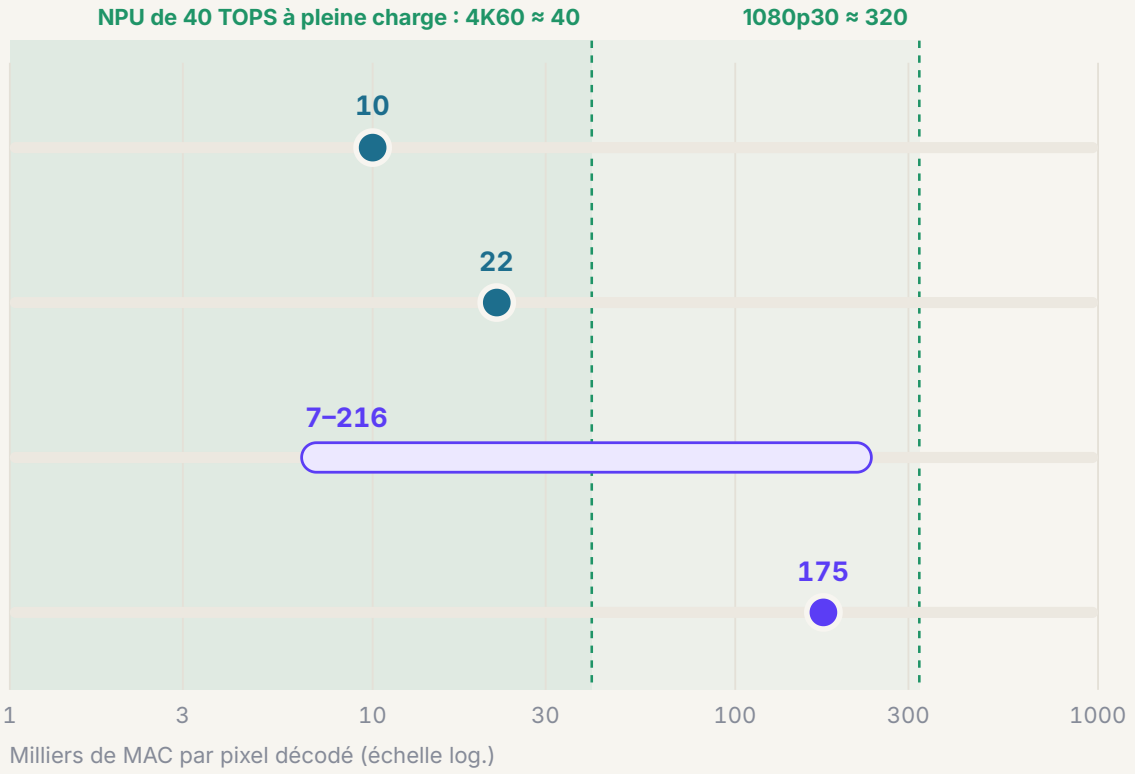
Ce qu'un décodeur neuronal demande à un NPU

Filtre en boucle neuronal, très faible complexité
Exploration JVET · ≈ 6 % de bits économisés
1,2 TOPS en 1080p30 · 10 en 4K60

Filtre en boucle neuronal, faible complexité
Exploration JVET · ≈ 8 %
2,7 TOPS en 1080p30 · 22 en 4K60

Décodeurs JPEG AI, de base à élevé
12,5–27,9 % face à VVC intra

DCVC-RT, un décodeur vidéo neuronal complet
21 % face à la référence VVC, réglages de l'article
22 TOPS en 1080p30 · 174 en 4K60



DCVC-RT nécessite environ 175 000 MAC par pixel : environ 22 TOPS en 1080p30 et 174 en 4K60, à pleine utilisation **estimation**. Un NPU de 40 TOPS pleinement utilisé permet environ 40 000 en 4K60.

Le trafic mémoire et l'exactitude au bit près sont les vraies limites

LA DRAM DOMINE

La mémoire avant les MAC

Dans la puce HEVC du MIT, la DRAM pèse environ les trois quarts de l'énergie du système : 0,88 sur 1,19 nJ par pixel. Les décodeurs neuronaux transportent de grandes cartes de caractéristiques d'une image à l'autre ; selon les auteurs de DCVC-RT, ce sont les E/S mémoire, pas l'arithmétique, qui limitent la vitesse.

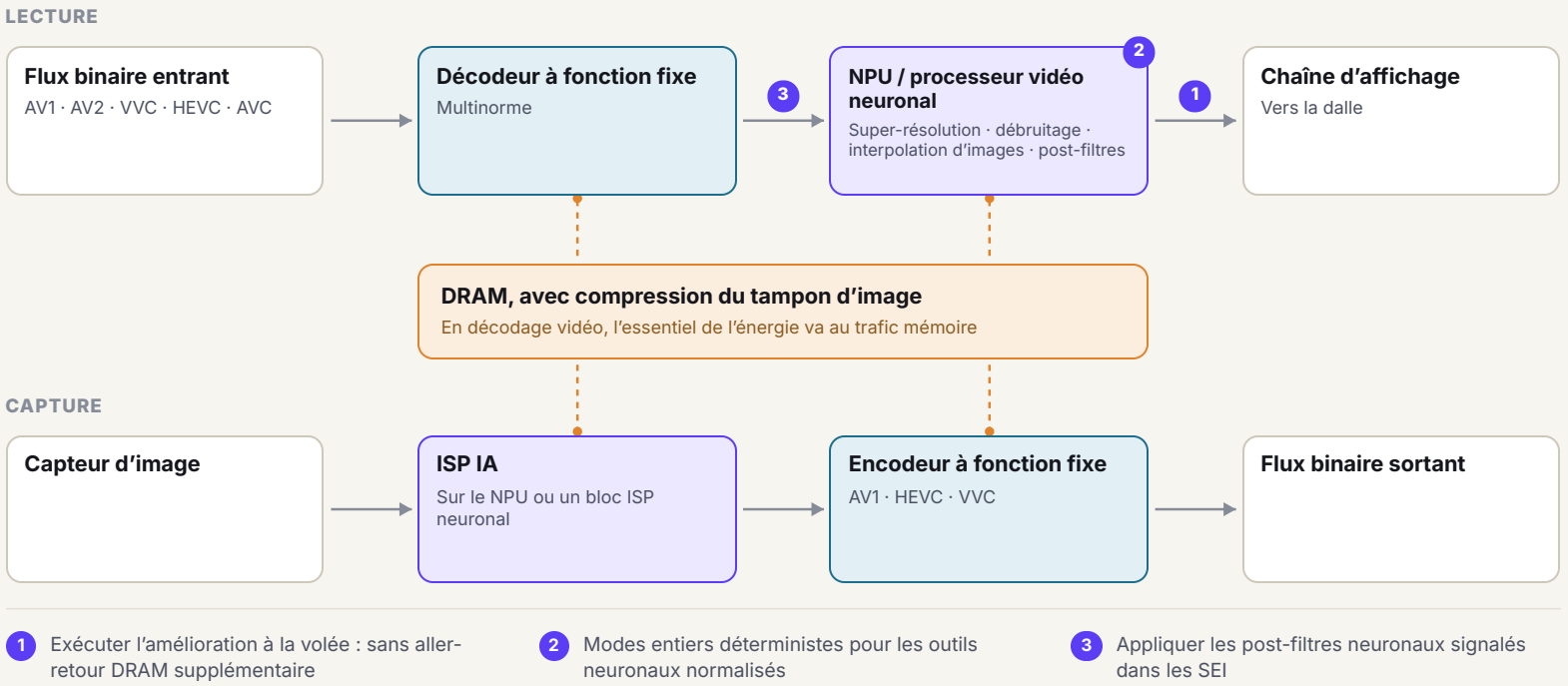
EXACTITUDE

Chaque décodeur doit calculer les mêmes bits

Le décodage entropique échoue si deux appareils calculent une probabilité différemment. Les parades : modèles entiers (DCVC-RT), hyperprior à échelle partagée (MLVC), inférence en entiers 16 bits (SADL du JVET), points de conformité (JPEG AI).

Pour une équipe puce, c'est de nouveau la discipline des transformées entières de HEVC : la norme doit définir l'arithmétique exactement.

Le partage des tâches dans les produits : un codec à fonction fixe et un NPU



> Canghai de Tencent : plus de 100 000 puces de première génération déployées ; la deuxième génération ajoute H.266 et la super-résolution sur puce. Chips&Media et Allegro DVT vendent côte à côte des IP de codec et des IP vidéo neuronales.

Ce que je construirais, selon le rôle

Équipes SoC et puces TV

- Décodeurs à fonction fixe pour chaque norme à lire, AV2 dès maintenant
- Un NPU dimensionné pour l'amélioration à la résolution d'affichage, sans aller-retour DRAM supplémentaire
- SEI de post-filtre neuronal et modes entiers déterministes
- H.267 : début de conception en 2029 ou après

Fournisseurs d'IP

- Décodeurs AV2 et VVC dès maintenant
- Accélérateurs de post-filtres et compression du tampon d'image
- Outils de conformité pour le traitement neuronal

Plateformes vidéo

- Le prétraitement neuronal à l'encodeur est rentabilisé en premier
- Super-résolution côté client là où les appareils le permettent
- Codecs neuronaux de bout en bout : seulement dans vos boucles fermées

Où les outils neuronaux passent sous licence

Norme	Contenu neuronal	Statut, novembre 2026
H.267 (nom de travail)	Ouvert : outils classiques (ECM) et neuronaux sont tous deux candidats	Réponses évaluées en janvier 2027 ; version 1 fin 2029
VVC et HEVC, avec VSEI	Post-filtre neuronal décrit dans les SEI ; optionnel pour les décodeurs	En vigueur ; étendu en octobre 2025
AV2	Classique, sans redevances	Version 1.0.0 le 28 mai 2026 ; flux Web YouTube prévus début 2027
JPEG AI	Codec d'image appris de bout en bout	Partie 1 publiée en 2025 ; aucune prise en charge par défaut constatée
IEEE 1857.11	Codage d'image neuronal conçu pour les NPU	Publiée en décembre 2024
Video Coding for Machines	Codage pour l'analyse par machine plutôt que pour le visionnage	Projet de norme internationale début 2026
Extensions ML d'Opus	Optionnelles et rétrocompatibles	Dans libopus 1.5 et 1.6

645
familles de brevets post-VVC détenues par ByteDance (GreyB, 2025)

0,50 $ · 1,00–1,50 $
tarif du pool VVC par téléphone et par TV 4K, Région 1

0 $
de redevances pour AV1 et AV2 de la part des membres d'AOMedia

L'argent suit les normes et les plateformes

REVENUS ACTUELS

Plateformes

Gardent les économies

Sous-échantillonneur neuronal de Netflix ; super-résolution chez YouTube et Meta ; MLVC publié en open source par Microsoft sous licence MIT

REVENUS ACTUELS

Fabricants d'appareils et de puces

Vendent de la différenciation

ISP des appareils photo ; processeurs de mise à l'échelle pour téléviseurs ; puces codec Canghai maison de Tencent

REVENUS ACTUELS

Fournisseurs d'IP

Vendent des licences à chaque nouvelle norme

Clients pour des décodeurs AV2 chez Chips&Media, Allegro DVT et VeriSilicon dans les trois semaines suivant la spécification

REVENUS ACTUELS

Fournisseurs de GPU

Créent une dépendance à leur plateforme

DLSS 4 dans plus de 250 jeux et applications ; SDK de textures neuronales en bêta

PAR APPAREIL

Détenteurs de brevets et pools

Perçoivent par appareil

VVC : 0,50 $ par téléphone, 1,00–1,50 $ par TV 4K ; outils H.267 après 2029

SORTIE PAR RACHAT

Start-up de codecs neuronaux

Rachetées

WaveOne → Apple (2023) ; iSIZE → Sony (2023) ; Deep Render → InterDigital (2025)

> Les trois start-up de codecs neuronaux les plus connues, WaveOne, iSIZE et Deep Render, ont été rachetées par Apple, Sony et InterDigital.

Six créneaux, et le risque de chacun

Créneau	Qui achète	Pourquoi maintenant	Risque principal
Prétraitement neuronal à l'encodeur	Plateformes de streaming et de vidéos d'utilisateurs, diffuseurs	Aucun changement d'appareil ; éprouvé chez Netflix	Les grandes plateformes le font elles-mêmes
IP et réglage d'amélioration côté affichage	Fabricants de puces pour TV, boîtiers et téléphones	Les NPU sont déjà dans les appareils	Banalisation ; équipes internes
Codecs neuronaux en boucle fermée	Fournisseurs de communications d'entreprise, industriels et de téléprésence	MLVC prouve que cela fonctionne, et son code est ouvert	Il faut contrôler les deux extrémités ; fragmentation des NPU
Silicium ou IP codec plus NPU	Fabricants de SoC	Le cycle AV2 commence	Acteurs établis puissants
Évaluation, conformité et inférence déterministe	Quiconque livre des outils neuronaux	L'exactitude au bit près est impérative	Petit marché, sauf en offre groupée
Contributions normatives à H.267	Concédants de licences, grandes plateformes	Les réponses seront évaluées en janvier 2027	Long, coûteux, incertain

CE QUE VOUS ENTENDREZ

Relativisez toute affirmation qui ne précise pas quatre choses

Affirmation	Ce que disent les faits
« Les codecs neuronaux battent déjà VVC. »	En laboratoire. DCVC-RT économise 21 % par rapport à l'encodeur de référence de VVC ; son décodeur coûte environ 175 000 MAC par pixel.
« Les codecs IA remplaceront H.26x et AV1 d'ici quelques années. »	Pas dans la lecture vidéo : version 1 de H.267 fin 2029, puces vers 2031 estimation. Le codec qui arrive maintenant, AV2, est classique.
« MLVC réduit la bande passante de 88 %. »	Par rapport à H.264 en 360p. Par rapport au HEVC matériel : 75,5 % en 360p et 65,4 % en 540p, sur le propre jeu de test de Microsoft.
« Les textures neuronales réduisent la mémoire vidéo de 85–96 %. »	Dans des scènes de démo. Le jeu commercialisé a gagné environ 30 %, sur certains assets seulement.
« JPEG AI est une norme ISO, donc les appareils photo et les navigateurs vont l'adopter. »	Aucune prise en charge par défaut trouvée dans les produits. Le Web a adopté JPEG XL, un format classique.

Demandez la référence, la métrique, la résolution et le coût de décodage. Une affirmation qui ne peut pas préciser ces quatre éléments est un résultat de recherche, pas un produit.

À surveiller, et questions ouvertes

JALONS

Janvier 2027	Évaluation des réponses H.267 : combien reposent sur des outils neuronaux, et à quel coût par pixel
Début 2027	Premiers flux Web AV2 de YouTube ; décodeurs AV2 dans les SoC annoncés
Second semestre 2027	Premier modèle de test de H.267
2027	MLVC au-delà des appels Teams pair à pair
2027–2028	Un premier produit qui utilise JPEG AI par défaut
Fin 2029	La version 1 de H.267 lance le cycle silicium

QUESTIONS OUVERTES

01 L'exactitude à grande échelle : un décodage entier exact au bit près sur les NPU de tous les fournisseurs, à quel coût en compression ?

02 La mémoire avant les MAC : le trafic DRAM par pixel du post-traitement neuronal

03 Le coût par pixel à partir duquel un outil neuronal cesse d'être rentable dans H.267

04 Quelles boucles fermées, hors appels, adopteront d'abord les codecs neuronaux de bout en bout

05 Comment mesurer la fidélité quand un décodeur génère du détail

Ce sont des domaines que nous pouvons explorer ensemble, par des échanges techniques et une collaboration technique.

i@tanshuai.com · tanshuai.com