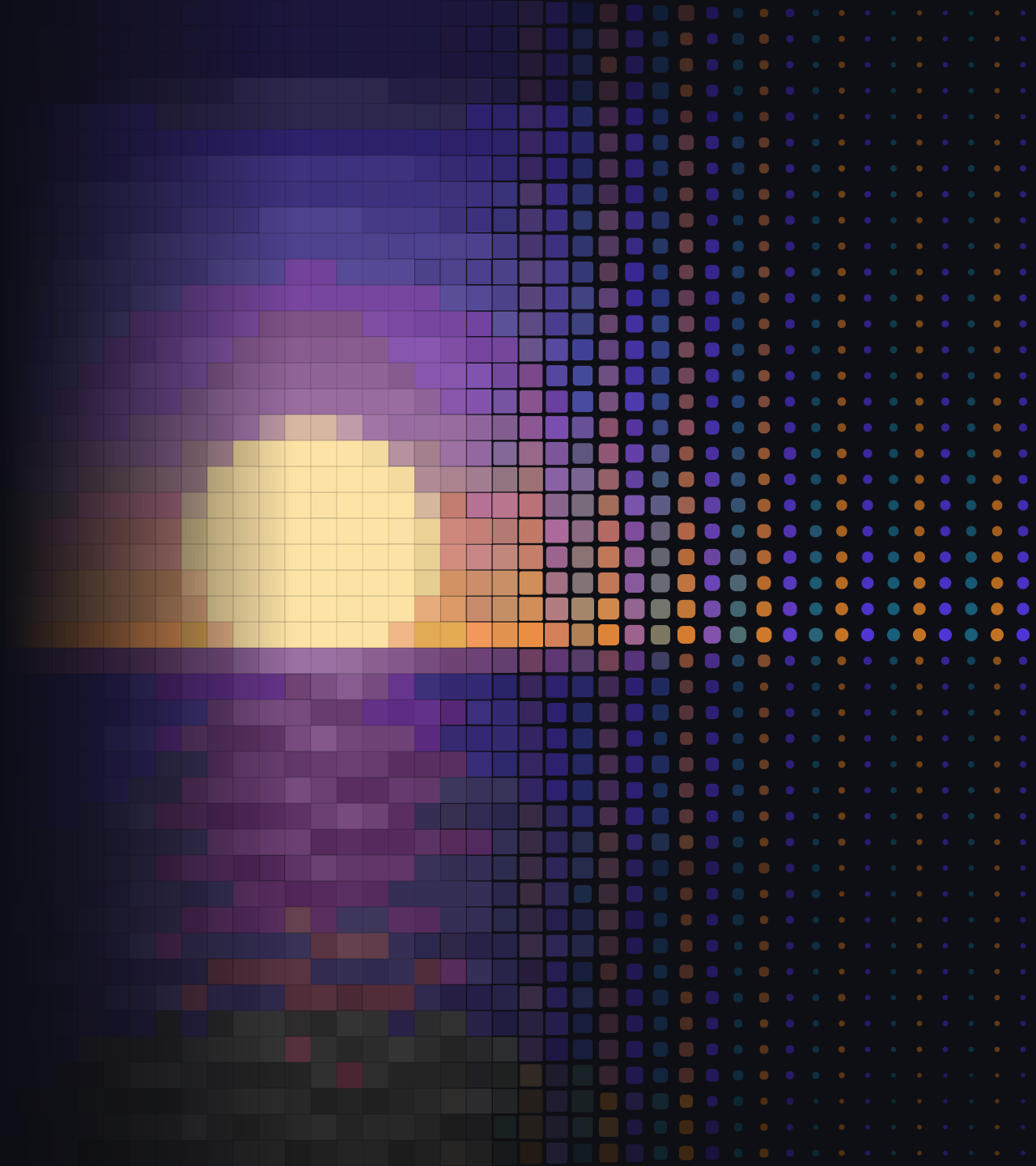
Nota de investigación · noviembre de 2026

Codificación neuronal en 2026

En torno al códec, no en su lugar

Reproducción · Fotografía · Llamadas · IA generativa · Videojuegos · Silicio

Tan Shuai
tanshuai.com · i@tanshuai.com



EL VEREDICTO

Ya generalizada, pero en torno al códec

01 Las redes neuronales están generalizadas en ambos extremos de los códecs estándar

ISP de cámara, reducción de resolución al codificar y superresolución al reproducir. El flujo de bits intermedio sigue siendo AV1, HEVC, AVC o VVC, decodificado por hardware de función fija.

02 Los códecs neuronales de extremo a extremo solo se despliegan en circuitos cerrados

Llamadas, tokenizadores de IA generativa y un juego AAA: una sola empresa controla ambos extremos y puede recurrir a un códec estándar.

03 Los flujos de bits neuronales en las pantallas del gran público esperan a H.267

Versión 1 a finales de 2029; primeros chips hacia 2031 y uso generalizado a mediados de la década de 2030 estimación.

≈75 %

del tráfico de datos móviles es vídeo (finales de 2025)

1260 millones

de smartphones comercializados en 2025, todos decodifican vídeo en hardware de función fija

≈12×

la energía por píxel de un decodificador HEVC por hardware, para el único diseño publicado de chip decodificador neuronal

Tres formas en que una red neuronal entra en un códec



> «En torno al códec» no necesita permiso de nadie. «Dentro de un estándar» necesita un estándar. «En lugar del códec» necesita controlar ambos extremos.

Dónde se despliega la codificación neuronal en 2026

	En torno a un códec estándar preprocesado, mejora, reescalado	Códec neuronal de extremo a extremo codificador y decodificador aprendidos	Dentro de un estándar herramientas neuronales en un flujo de bits o API estándar
01 **Reproducción y streaming**	**A GRAN ESCALA** Reductor de resolución de Netflix; superresolución de YouTube y Meta; reescalado por IA en televisores	**SIN DESPLIEGUE** Ningún servicio de streaming envía un formato de vídeo neuronal; televisores y móviles decodifican códecs estándar por hardware	**EN ESTANDARIZACIÓN** La versión 1 de H.267 está prevista para finales de 2029; entre las candidatas hay herramientas neuronales
02 **Fotografía e imagen**	**A GRAN ESCALA** ISP con IA en móviles de gama alta; un ISP neuronal completo, presentado como IP en el CES 2026	**DEMOS Y PILOTOS** JPEG AI es un estándar ISO: demos en móviles, ningún producto	**SIN DESPLIEGUE** El formato que gana terreno es clásico: JPEG XL, activado por defecto en Chrome 155
03 **Llamadas y reuniones**	**A GRAN ESCALA** Superresolución en el receptor y supresión neuronal de ruido	**A GRAN ESCALA** Audio neuronal en las llamadas de Teams entre dos personas desde 2021; despliegue en curso del vídeo MLVC en Teams	**DESPLIEGUE LIMITADO** Opus 1.5 añade herramientas de ML opcionales y retrocompatibles
04 **IA generativa**	—	**A GRAN ESCALA** Todo modelo de vídeo y de voz usa un tokenizador aprendido, p. ej., un VAE de vídeo 1:192	—
05 **Videojuegos y gráficos**	**A GRAN ESCALA** Reescalado neuronal y generación de fotogramas: DLSS 4, por sí solo, está en más de 250 juegos y aplicaciones	**DESPLIEGUE LIMITADO** Texturas neuronales en un juego AAA ya publicado; el SDK de NVIDIA, en beta	**EN ESTANDARIZACIÓN** DirectX incorpora operaciones matriciales para shaders (versión preliminar)
06 **Enlaces especializados: satélite, robots**	**DEMOS Y PILOTOS** Restauración en servidor de señales con baja tasa de bits	**DEMOS Y PILOTOS** Códecs neuronales resistentes a pérdidas, en investigación y pilotos	—

Las filas se ordenan por el número de personas al que llega cada ámbito.

Neuronal en ambos extremos, estándar en el medio

- **En el codificador**

  El reductor de resolución neuronal de Netflix funciona desde 2022 sin cambios en ningún dispositivo; cerca del 77 % de los espectadores lo prefirió en las pruebas de Netflix dato del proveedor.

- **En la pantalla**

  Super Resolution de YouTube en televisores desde octubre de 2025; Meta aplica superresolución en sus servidores y en los móviles.

- **En el televisor**

  El procesador NQ8 AI Gen3 de Samsung anuncia 768 redes neuronales para el reescalado.

EL FLUJO DE BITS SIGUE SIENDO ESTÁNDAR

30 %

del visionado de Netflix, en AV1 (diciembre de 2025)

>90 %

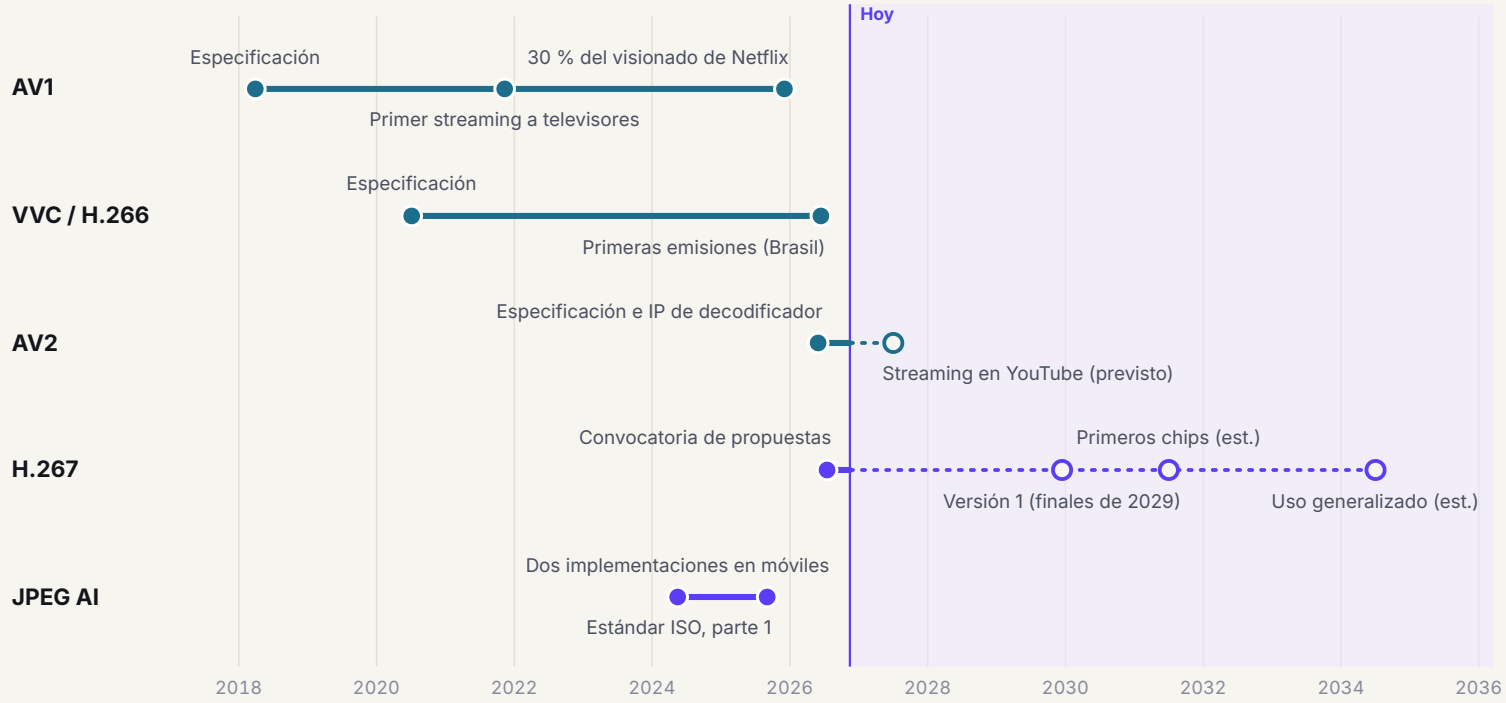
del catálogo de YouTube, por tiempo de visualización, disponible en AV1 (marzo de 2026)

3 semanas

desde la versión 1.0.0 de AV2 (28 de mayo de 2026) hasta clientes de IP de decodificador en tres proveedores

PLAZOS

Entre un estándar terminado y las pantallas del gran público pasa buena parte de una década



AV1 tardó casi ocho años desde su especificación (2018) hasta el 30 % del visionado de Netflix (diciembre de 2025). Las marcas huecas son objetivos o estimaciones mías.

H.267 decide si las herramientas neuronales llegan a todas las pantallas

Octubre de 2025
Convocatoria de evidencias: cinco respuestas de 16 empresas y laboratorios; ganancias neuronales más claras en 8K HDR

Julio de 2026
Convocatoria de propuestas, emitida en Ginebra

Enero de 2027
Evaluación de las respuestas en Brisbane

Segundo semestre de 2027
Primer modelo de prueba

Finales de 2029
Versión 1 del estándar

≈28 %
de ahorro frente a VVC en la vía clásica (ECM)

6–14 %
de ahorro con filtros en bucle neuronales, con entre 10.000 y más de 22.000 MAC por píxel

SEI
los flujos VVC ya pueden señalizar un posfiltro neuronal para que lo apliquen los decodificadores

Mi lectura: un núcleo clásico con algunas herramientas neuronales allí donde los bits ahorrados justifiquen el coste del decodificador análisis.

Neuronal dentro de la cámara, clásica en la web

- **Cámara**

  El Cognitive ISP de Qualcomm aplica segmentación en tiempo real en móviles desde 2023; en el CES 2026 se mostró un ISP completo con IA que funciona en una NPU.

- **JPEG AI**

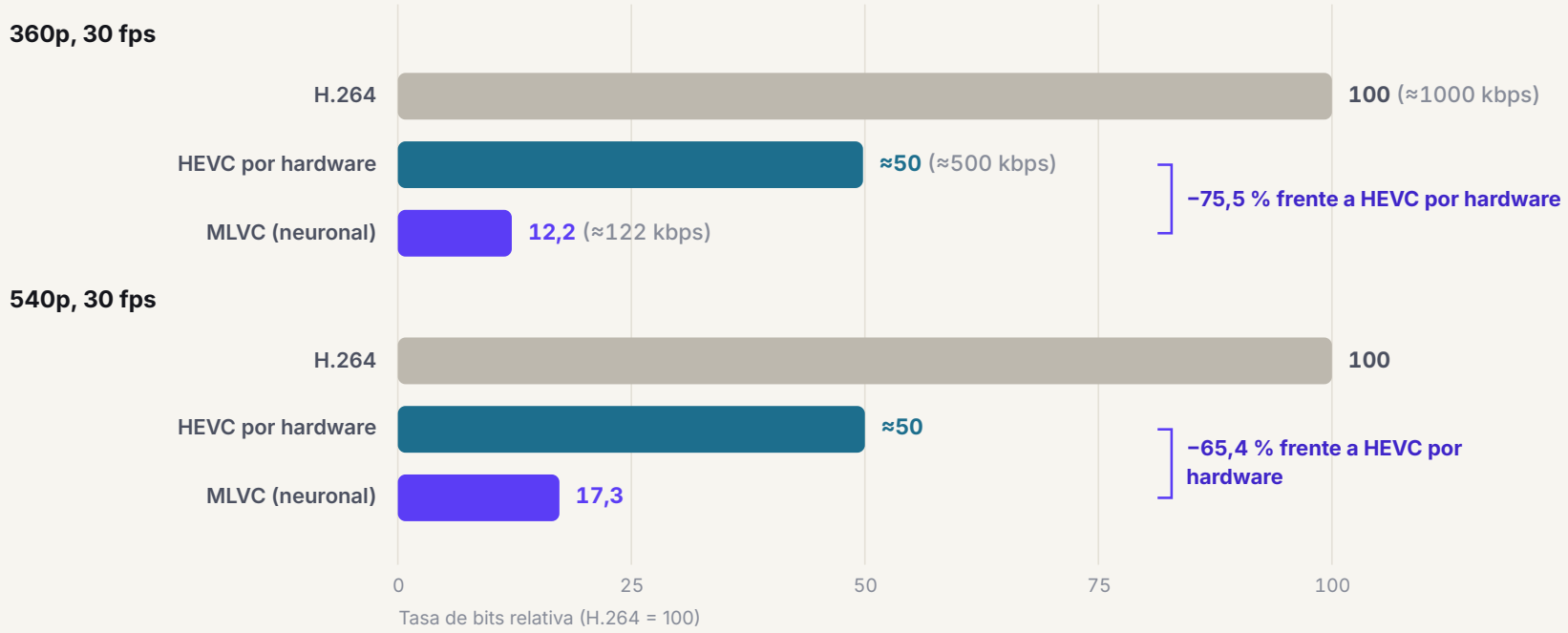
  Publicado en 2025: un 12,5–27,9 % mejor que VVC intra, con entre 7000 y 216.000 MAC por píxel. No encontré ningún producto que lo use por defecto.

- **Web**

  Chrome 155 activó JPEG XL por defecto el 6 de octubre de 2026, con un decodificador con seguridad de memoria escrito en Rust.

> **Los formatos que deben abrirse en todas partes siguen siendo clásicos; la codificación neuronal gana dentro de la cadena de una sola empresa.**

Es en las llamadas donde se despliegan los códecs neuronales de extremo a extremo



Satin funciona en todas las llamadas de Teams y Skype entre dos personas desde 2021. MLVC, de código abierto con licencia MIT desde el 24 de julio de 2026, se está desplegando en las llamadas entre pares de Teams hasta 540p30 y usa como máximo la mitad de la NPU **dato del proveedor**.

No todas las plataformas eligieron lo neuronal, y los motivos son instructivos

DENTRO DE UN ESTÁNDAR, OPCIONAL

Opus 1.5 y 1.6

Mejora neuronal, ocultación de pérdida de paquetes, redundancia profunda y extensión de ancho de banda, todo desactivado por defecto y plenamente compatible con el estándar de 2012.

CLÁSICO, POR DECISIÓN PROPIA

Meta MLow (2024)

Más del 20 % de las llamadas de Meta se hacen en móviles ARMv7, y decenas de millones de llamadas diarias de WhatsApp, en móviles de más de diez años. MLow ofrece cerca del doble de calidad que Opus a 6 kbps con un 10 % menos de cómputo.

Decide el presupuesto de complejidad de la base instalada. Los códecs neuronales se despliegan primero donde los dispositivos son nuevos y un solo proveedor controla ambos extremos.

El mayor despliegue que nadie llama códec

- **Códecs aprendidos en todas partes**

  Todo modelo de voz y de vídeo lee y escribe los tokens o latentes de un códec aprendido.

- **No son formatos de distribución**

  Cada modelo tiene su propio espacio latente: sin control de tasa, sin prueba de conformidad, sin un segundo decodificador.

- **Una partida de costes**

  Para quien sirve vídeo generativo, un latente que comprime más reduce el coste de generación más que proporcionalmente análisis.

DOS EJEMPLOS

12,5 Hz · 1,1 kbps

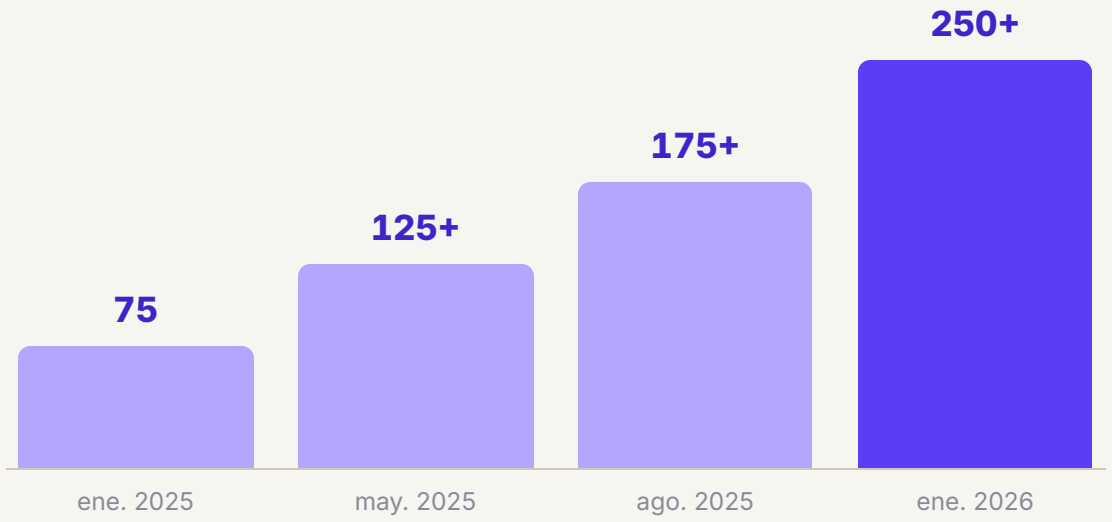
tokens de voz de Mimi, de Kyutai, por debajo de cualquier códec telefónico

1:192

VAE de vídeo de LTX-Video: un token latente por cada 32×32 píxeles a lo largo de 8 fotogramas

La reconstrucción neuronal está generalizada; los recursos neuronales empiezan

DLSS 4 Multi Frame Generation: juegos y aplicaciones



Compresión neuronal de texturas

Escena de demostración de NVIDIA, GTC 2026

Memoria de texturas: 6,5 GB → 0,97 GB

85 % menos memoria

Assassin's Creed Mirage, ya publicado

Solo en algunos recursos; frente a formatos de bloques

≈30 % más de compresión

Medidas distintas: no son directamente comparables

> DirectX Linear Algebra entró en versión preliminar pública en abril de 2026: una sola ruta de código para la decodificación neuronal en shaders, en la GPU de cualquier fabricante.

Satélites, drones y robots: el mejor encaje, un mercado pequeño

- **Por qué encaja**

  Un solo propietario controla ambos extremos, el ancho de banda escasea y el contenido es lo bastante acotado para entrenar con él.

- **China**

  El Anuncio n.º 49 de 2026, en vigor desde el 18 de noviembre, revisa los controles de exportación de drones; una entrada usa como umbral la compresión semántica de imagen.

- **Estados Unidos**

  La FCC añadió los drones de fabricación extranjera y sus componentes críticos a su Covered List el 22 de diciembre de 2025.

> **Un nicho sólido para un especialista con capacidad de cumplimiento en exportaciones, no un mercado masivo.**

Un orden de magnitud más de energía por píxel

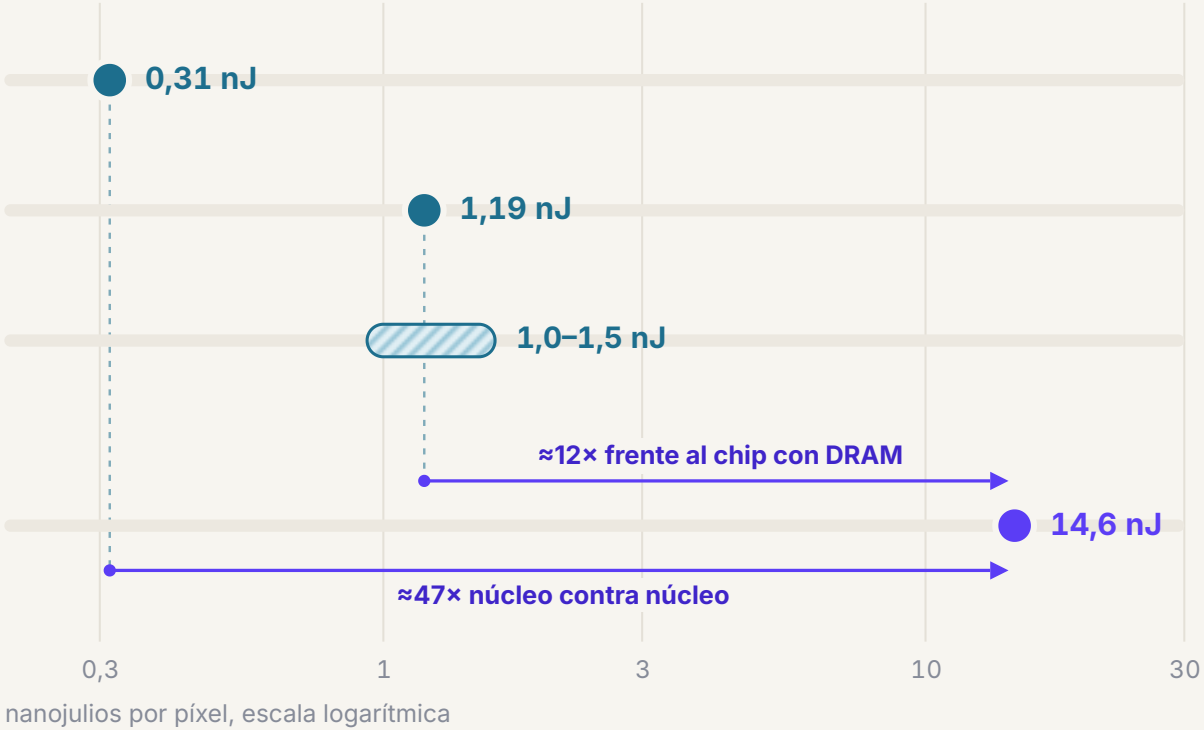
Chip decodificador HEVC, núcleo
40 nm, 4K30, medido

Chip decodificador HEVC, con DRAM
40 nm, 4K30, núcleo medido más DRAM simulada

Decodificador VVC
28 nm, 4K60, proyección

Decodificador de vídeo neuronal (NVCA)
28 nm, 1080p25, solo síntesis; sin contar pesos ni DRAM

A 1080p30, son unos 74 mW frente a 908 mW.



A 1080p30: unos 74 mW para el chip HEVC con DRAM frente a cerca de 900 mW para NVCA. Un nodo más antiguo para el chip HEVC, y pesos y DRAM sin contar en NVCA: la comparación favorece al diseño neuronal.

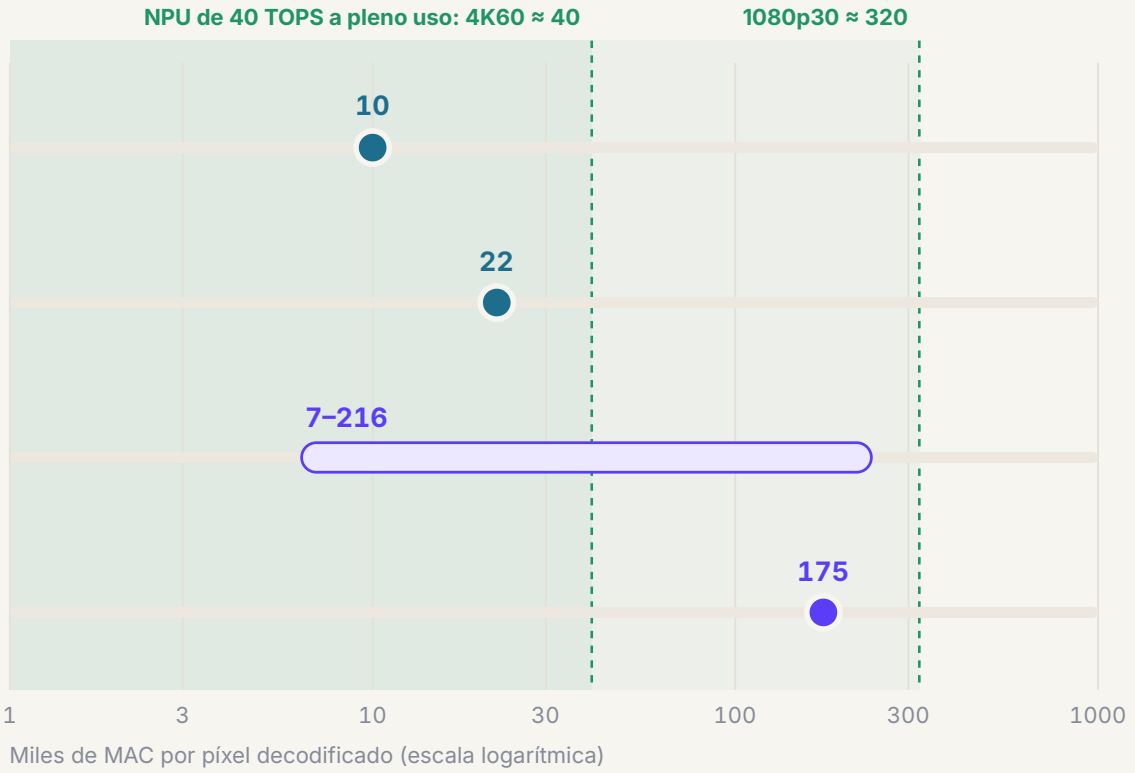
Lo que un decodificador neuronal le pide a una NPU

Filtro en bucle neuronal, complejidad muy baja
Exploración de JVET · ≈6 % de ahorro de bits
1,2 TOPS a 1080p30 · 10 a 4K60

Filtro en bucle neuronal, complejidad baja
Exploración de JVET · ≈8 %
2,7 TOPS a 1080p30 · 22 a 4K60

Decodificadores JPEG AI, del nivel base al alto
12,5–27,9 % frente a VVC intra

DCVC-RT, un decodificador de vídeo neuronal completo
21 % frente a la referencia de VVC, condiciones del artículo
22 TOPS a 1080p30 · 174 a 4K60



> DCVC-RT necesita unos 175.000 MAC por píxel: unos 22 TOPS a 1080p30 y 174 a 4K60, con utilización completa **estimación**. Una NPU de 40 TOPS a pleno uso da para unos 40.000 a 4K60.

El tráfico de memoria y la exactitud bit a bit son los límites duros

DOMINA LA DRAM

La memoria antes que los MAC

En el chip HEVC del MIT, la DRAM supone unas tres cuartas partes de la energía del sistema: 0,88 de 1,19 nJ por píxel. Los decodificadores neuronales arrastran grandes mapas de características entre fotogramas, y los autores de DCVC-RT vieron que la velocidad la limitaba la E/S de memoria, no la aritmética.

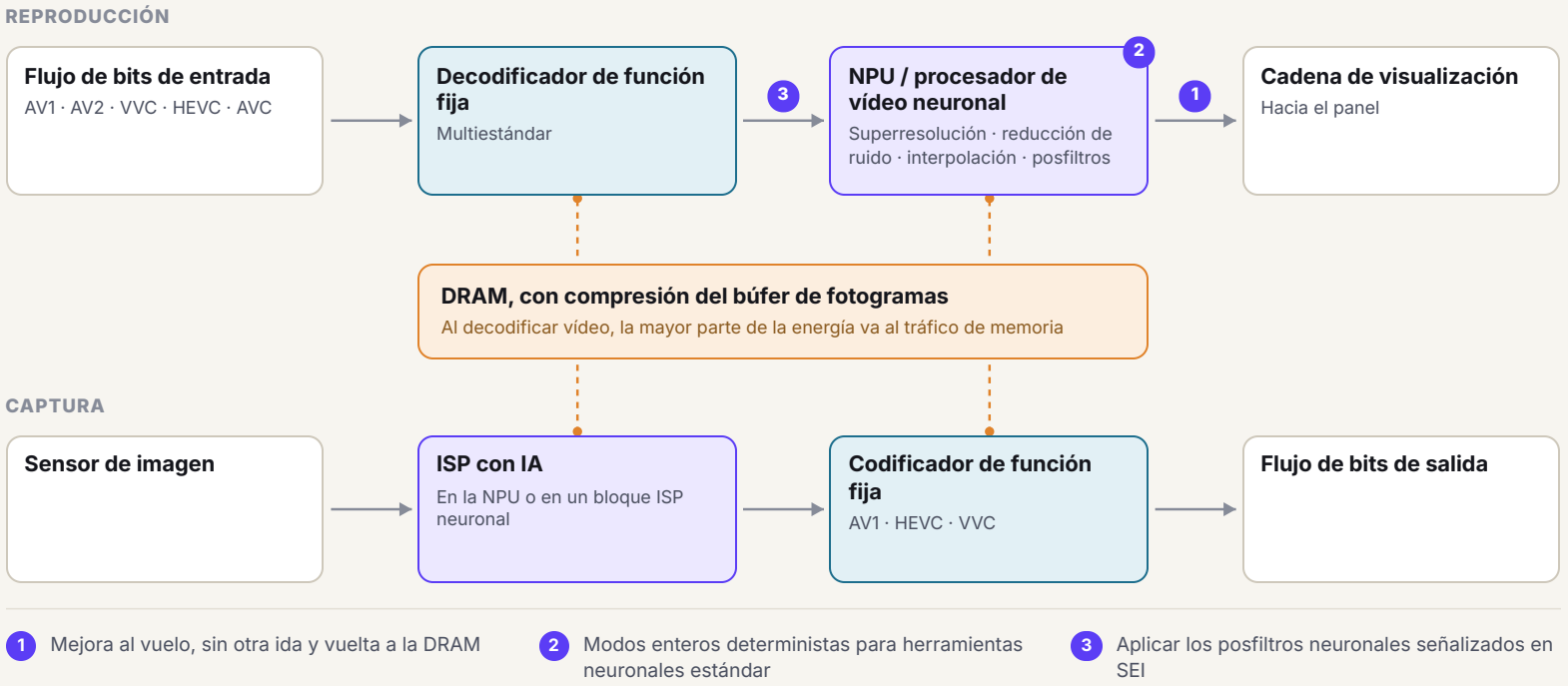
EXACTITUD

Todos los decodificadores deben calcular los mismos bits

La decodificación entrópica falla si dos dispositivos calculan una probabilidad de forma distinta. Las soluciones: modelos enteros (DCVC-RT), un hiperprior de escala compartida (MLVC), inferencia con enteros de 16 bits (SADL de JVET), puntos de conformidad (JPEG AI).

Para un equipo de chips es de nuevo la disciplina de la transformada entera de HEVC: el estándar debe definir la aritmética con exactitud.

La división que se comercializa: un códec de función fija más una NPU



> Canghai de Tencent: más de 100.000 chips de primera generación desplegados; la segunda generación añade H.266 y superresolución en el chip. Chips&Media y Allegro DVT venden IP de códec e IP de vídeo neuronal en paralelo.

Lo que construiría, por perfil

Equipos de SoC y de chips para televisores

- Decodificadores de función fija para cada estándar que haya que reproducir, AV2 ya
- Una NPU dimensionada para la mejora a la resolución de pantalla, sin otra ida y vuelta a la DRAM
- SEI de posfiltro neuronal y modos enteros deterministas
- H.267 como inicio de diseño en 2029 o después

Proveedores de IP

- Decodificadores AV2 y VVC, ya
- Aceleradores de posfiltro y compresión del búfer de fotogramas
- Herramientas de conformidad para el procesado neuronal

Plataformas de vídeo

- El preprocesado neuronal en el codificador se amortiza antes
- Superresolución en el cliente donde los dispositivos lo permitan
- Códecs neuronales de extremo a extremo solo en circuitos cerrados propios

Dónde las herramientas neuronales pasan a ser licenciables

Estándar	Contenido neuronal	Situación, noviembre de 2026
H.267 (nombre provisional)	Abierto: las herramientas clásicas (ECM) y las neuronales son candidatas	Respuestas evaluadas en enero de 2027; versión 1 a finales de 2029
VVC y HEVC, con VSEI	Posfiltro neuronal descrito en SEI; opcional para los decodificadores	En vigor; ampliado en octubre de 2025
AV2	Clásico, libre de regalías	Versión 1.0.0 el 28 de mayo de 2026; streaming web de YouTube previsto para principios de 2027
JPEG AI	Códec de imagen aprendido de extremo a extremo	Parte 1 publicada en 2025; no se ha encontrado soporte por defecto en productos
IEEE 1857.11	Codificación neuronal de imagen diseñada para NPU	Publicado en diciembre de 2024
Video Coding for Machines	Codificación para el análisis automático en lugar de la visualización	Borrador de estándar internacional a principios de 2026
Extensiones de ML de Opus	Opcionales y retrocompatibles	En libopus 1.5 y 1.6

645

familias de patentes «más allá de VVC» en manos de ByteDance (GreyB, 2025)

0,50 $ · 1,00–1,50 $

tarifa del consorcio VVC por móvil y por televisor 4K, Región 1

0 $

en regalías por AV1 y AV2 de los miembros de AOMedia

QUIÉN COBRA

El dinero sigue a los estándares y a las plataformas

INGRESOS HOY

Plataformas

Se quedan el ahorro

Reductor de resolución neuronal de Netflix; superresolución de YouTube y Meta; Microsoft liberó MLVC con licencia MIT

INGRESOS HOY

Fabricantes de dispositivos y chips

Venden diferenciación

ISP de cámara; procesadores de reescalado para TV; chips de códec Canghai, de diseño propio de Tencent

INGRESOS HOY

Proveedores de IP

Licencian cada nuevo estándar

Clientes de decodificadores AV2 en Chips&Media, Allegro DVT y VeriSilicon a menos de tres semanas de la especificación

INGRESOS HOY

Fabricantes de GPU

Crean dependencia de su plataforma

DLSS 4 en más de 250 juegos y aplicaciones; SDK de texturas neuronales en beta

POR DISPOSITIVO

Titulares de patentes y consorcios

Cobran por dispositivo

VVC: 0,50 $ por móvil, 1,00–1,50 $ por televisor 4K; herramientas de H.267 después de 2029

SALIDA POR ADQUISICIÓN

Startups de códecs neuronales

Adquiridas

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

> Las tres startups de códecs neuronales más conocidas, WaveOne, iSIZE y Deep Render, las compraron Apple, Sony e InterDigital.

Seis apuestas y el riesgo de cada una

Apuesta	Quién compra	Por qué ahora	Riesgo principal
Preprocesado neuronal en el codificador	Plataformas de streaming y de vídeo de usuarios, radiodifusores	Sin cambios en los dispositivos; probado en Netflix	Las grandes plataformas lo desarrollan ellas mismas
IP y ajuste de mejora de imagen en la pantalla	Fabricantes de chips para televisores, set-top boxes y móviles	Las NPU ya están en los dispositivos	Comoditización; equipos internos
Códecs neuronales en circuito cerrado	Proveedores de comunicaciones empresariales, industriales y de telepresencia	MLVC demuestra que funciona, y su código es abierto	Hay que controlar ambos extremos; fragmentación de NPU
Silicio o IP de códec más NPU	Fabricantes de SoC	Empieza el ciclo de AV2	Competidores consolidados
Evaluación, conformidad e inferencia determinista	Cualquiera que despliegue herramientas neuronales	La exactitud bit a bit es un requisito estricto	Mercado pequeño, salvo en paquete
Contribuciones al estándar H.267	Licenciantes, grandes plataformas	Las respuestas se evalúan en enero de 2027	Largo, costoso e incierto

AFIRMACIONES HABITUALES

Cautela ante toda afirmación que no precise cuatro cosas

Afirmación	Lo que dice la evidencia
«Los códecs neuronales ya superan a VVC».	En el laboratorio. DCVC-RT ahorra un 21 % frente al codificador de referencia de VVC; su decodificador cuesta unos 175.000 MAC por píxel.
«Los códecs de IA sustituirán a H.26x y AV1 en pocos años».	No en reproducción: versión 1 de H.267 a finales de 2029 y chips hacia 2031 estimación. El códec que llega ahora, AV2, es clásico.
«MLVC reduce el ancho de banda un 88 %».	Frente a H.264 a 360p. Frente a HEVC por hardware: un 75,5 % a 360p y un 65,4 % a 540p, con el propio conjunto de pruebas de Microsoft.
«Las texturas neuronales reducen la memoria de vídeo un 85–96 %».	En escenas de demostración. El juego ya publicado ganó cerca de un 30 %, y solo en algunos recursos.
«JPEG AI es un estándar ISO, así que cámaras y navegadores lo adoptarán».	No se ha encontrado soporte por defecto en productos. La web adoptó JPEG XL, un formato clásico.

Hay que pedir la referencia, la métrica, la resolución y el coste de decodificación. Una afirmación que no pueda precisar las cuatro es un resultado de investigación, no un producto.

QUÉ VIGILAR

Qué vigilar y preguntas abiertas

HITOS

Enero de 2027	Evaluación de las respuestas a H.267: cuántas dependen de herramientas neuronales y a qué coste por píxel
Principios de 2027	Primeras emisiones web de YouTube en AV2; decodificadores AV2 en los SoC anunciados
Segundo semestre de 2027	Primer modelo de prueba de H.267
2027	MLVC más allá de las llamadas entre pares de Teams
2027–2028	Un primer producto que use JPEG AI por defecto
Finales de 2029	La versión 1 de H.267 inicia el ciclo del silicio

PREGUNTAS ABIERTAS

01 Exactitud a escala: decodificación entera exacta bit a bit en NPU de distintos fabricantes, ¿con qué pérdida de compresión?

02 La memoria antes que los MAC: el tráfico de DRAM por píxel del posprocesado neuronal

03 El coste por píxel a partir del cual una herramienta neuronal deja de compensar en H.267

04 Qué circuitos cerrados, más allá de las llamadas, adoptarán primero los códecs neuronales de extremo a extremo

05 Cómo medir la fidelidad cuando un decodificador genera detalle

Son áreas que podemos explorar juntos, mediante el intercambio técnico y la colaboración técnica.

i@tanshuai.com · tanshuai.com