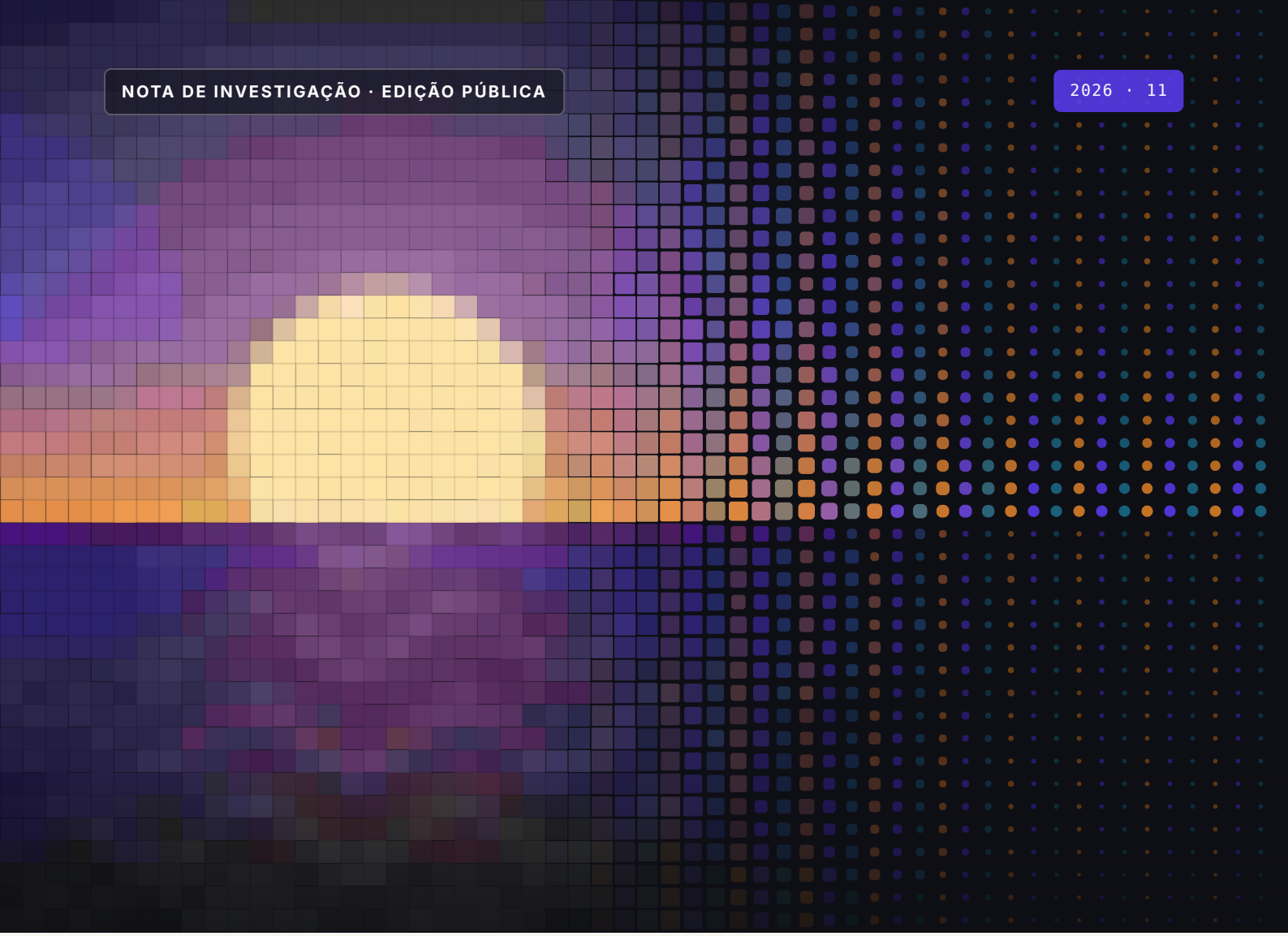


NOTA DE INVESTIGAÇÃO · EDIÇÃO PÚBLICA

2026 · 11

COMPRESSÃO APRENDIDA · REPRODUÇÃO × SILÍCIO

Codificação neural em 2026

Chega ao mercado em torno do codec, não em vez dele

Reprodução · Fotografia · Chamadas · IA generativa · Jogos · Silício

O VEREDICTO

As redes neuronais já trabalham nas duas pontas dos codecs que transportam o vídeo do mundo. Os codecs neurais de ponta a ponta só chegam ao mercado em circuitos fechados, e os ecrãs do grande público esperam pelo H.267.

Tan Shuai

Tecnólogo e fundador · Investigador e programador de IA

tanshuai.com · i@tanshuai.com

Dados a 12 de novembro de 2026

Valores em dólares dos EUA

ÍNDICE

Índice



◆ **Sumário executivo** — Já generalizada, mas em torno do codec — 04

01 **Âmbito e método** — O que significa aqui «codificação neural», e como a analisei — 07

02 **Reprodução e streaming** — Reprodução e streaming: neural nas duas pontas, normalizado no meio — 09

03 **Fotografia e imagem** — Fotografia e imagem: neural dentro da câmara, clássico na Web — 12

04 **Chamadas e reuniões** — Chamadas e reuniões: onde os codecs neurais de ponta a ponta chegam realmente ao mercado — 13

05 **IA generativa** — IA generativa: a maior implantação a que ninguém chama codec — 15

06 **Jogos e gráficos** — Jogos e gráficos: a reconstrução neural está generalizada, os recursos neurais estão a começar — 16

07 **Ligações especializadas** — Ligações especializadas: satélites, drones e robôs — 18

08 **Silício** — A perspetiva do silício: energia, computação, memória e exatidão — 19

09 **Normas e patentes** — Normas e patentes: onde as ferramentas neurais se tornam licenciáveis — 24

10 **Quem ganha dinheiro** — Quem ganha dinheiro, e onde uma empresa pode apostar — 25

11 **Alegações e evidência** — Alegações que vai ouvir, e o que diz a evidência — 27

12 **Marcos a acompanhar** — O que acompanhar, de 2027 a 2030 — 28

13 **Questões em aberto** — 29

◆ **Sobre o autor** — Tan Shuai — 30

◆ **Fontes** — 31

FIGURAS

Figuras

FIG. 01 Onde a codificação neural chega ao mercado em 2026 06

FIG. 02 Três formas de uma rede neuronal entrar num codec 08

FIG. 03 Entre uma norma concluída e os ecrãs do grande público passa a maior parte de uma década 11

FIG. 04 Em chamadas, um codec neural supera o HEVC em hardware em 65–76% 14

FIG. 05 Quando uma empresa controla toda a pilha, a reconstrução neural difunde-se depressa 17

FIG. 06 Energia por píxel descodificado: os descodificadores de função fixa estão uma ordem de grandeza à frente 20

FIG. 07 Uma ferramenta neural custa dezenas de milhares de MAC por píxel; um descodificador neural completo, centenas de milhares 21

FIG. 08 A divisão que chega ao mercado: hardware de função fixa para o fluxo de bits, um motor neural para os píxeis 22

FIG. 09 Quem ganha dinheiro com a codificação neural em 2026 25

Etiquetas no texto

estimativa — cálculo ou inferência do autor a partir de dados citados

alegação do fornecedor — publicado pela empresa que vende o produto, sem verificação independente

análise — interpretação da evidência pelo autor

Esta é uma tradução da edição em inglês, que prevalece em caso de divergência.

Dados a 12 de novembro de 2026 · Valores em dólares dos EUA

SUMÁRIO EXECUTIVO

Já generalizada, mas em torno do codec

O VEREDICTO

Em 2026, a codificação neural já está generalizada, mas funciona em torno dos codecs normalizados, não em vez deles.

≈75%
do tráfego de dados móveis é vídeo, final de 2025 (Ericsson)

1,26 mil milhões
de smartphones entregues em 2025, todos a descodificar vídeo em hardware de função fixa

≈12×
a energia por píxel de um chip descodificador HEVC medido com DRAM, no único projeto publicado de chip descodificador neural

≈22 TOPS
de computação sustentada para um descodificador neural da classe do DCVC-RT a 1080p30 estimativa

75,5%
menos débito binário do que o HEVC em hardware, para o MLVC a 360p em chamadas, no conjunto de testes da própria Microsoft alegação do fornecedor

Final de 2029
versão 1 do H.267, a primeira norma que deverá levar ferramentas neurais a todos os ecrãs

Onde chega ao mercado

As redes neuronais já funcionam nas duas pontas dos codecs que transportam quase todo o vídeo do mundo. À entrada, moldam a imagem: processadores de sinal de imagem com IA nas câmaras dos telemóveis e um redutor de resolução neural na cadeia de codificação da Netflix. À saída, reparam-na: o Super Resolution do YouTube nos televisores, a super-resolução na Meta e processadores de televisor que anunciam centenas de redes neuronais. Pelo meio, o fluxo de bits continua a ser AV1, HEVC, AVC ou VVC, descodificado por hardware de função fixa. Nenhum serviço de streaming envia um formato de vídeo neural para televisores ou telemóveis; onde hoje um descodificador de vídeo neural corre num dispositivo, corre dentro da aplicação própria de uma empresa.

Circuitos fechados

Os codecs neurais de ponta a ponta, em que tanto o codificador como o descodificador são aprendidos, só chegam ao mercado em circuitos fechados, onde uma empresa controla as duas pontas e pode recorrer a um codec normalizado. A Microsoft usa um codec de áudio com IA em todas as chamadas a dois do Teams e do Skype desde 2021 e, em julho de 2026, abriu o código do MLVC, um codec de vídeo neural que está agora a ser implementado nas chamadas ponto a ponto do Teams: a 360p, precisa de cerca de 75% menos débito binário do que o HEVC em hardware para a mesma qualidade avaliada. Todos os modelos de geração de voz e de vídeo assentam num codec aprendido, desde tokenizadores de voz a 12,5 Hz até VAE de vídeo com uma taxa de compressão de 1:192. A compressão neural de texturas já chegou a um jogo AAA. São implantações reais, e partilham uma condição: mais ninguém tem de descodificar o fluxo.

Calendário

A reprodução em massa de fluxos de bits neurais espera por uma norma e por um ciclo de silício. A primeira norma que deverá trazer ferramentas neurais, o sucessor do VVC a que chamarei H.267, teve o pedido de propostas lançado em julho de 2026, avalia as respostas em janeiro de 2027 e aponta para uma primeira versão no final de 2029. O AV1 precisou de quase oito anos desde a especificação concluída até 30% das visualizações da Netflix. Seguindo esse padrão, os primeiros chips H.267 chegam por volta de 2031 e o uso generalizado em meados da década de 2030 estimativa. Entretanto, o codec que vai chegar aos ecrãs a seguir é o AV2, concluído em maio de 2026 sem qualquer descodificador neural, e três fornecedores de IP anunciaram clientes para descodificadores AV2 em três semanas.

Silício

A razão está no silício. O único projeto publicado de chip descodificador de vídeo neural precisa de 14,6 nJ por píxel, cerca de 12 vezes o valor de um chip descodificador HEVC medido, contando com a sua DRAM, e cerca de 47 vezes o do seu núcleo isolado. O DCVC-RT, o codec de investigação em tempo real da Microsoft, precisa, pelas minhas contas a partir da sua própria tabela por módulo, de cerca de 175 mil multiplicações-acumulações por píxel descodificado, ou aproximadamente 22 TOPS de computação sustentada para 1080p30 estimativa: cerca de metade dos 40 TOPS que definem um PC com IA, gastos num único fluxo de vídeo. Um descodificador neural também tem de produzir resultados idênticos ao bit em todos os dispositivos, caso contrário a descodificação entrópica desmorona-se. É por isso que o MLVC abdica de compressão para descodificar de forma idêntica em diferentes NPU e funciona a 540p, não a 4K.

Negócio

Para uma empresa, o dinheiro segue as normas e as plataformas, não os codecs neurais independentes. As plataformas ficam com as poupanças do pré e pós-processamento neural. Os fornecedores de IP licenciam descodificadores de função fixa para cada nova norma, além de NPU para imagem. Os agrupamentos de patentes cobram por dispositivo: 0,50 dólares por telemóvel e 1,00–1,50 dólares por televisor 4K no VVC, nos EUA e nos outros mercados da Região 1. As três startups de codecs neurais mais conhecidas, WaveOne, iSIZE e Deep Render, foram compradas pela Apple, pela Sony e pela InterDigital. As oportunidades em aberto estão no pré e pós-processamento neural, nos produtos em circuito fechado, no silício que junta um codec de função fixa a uma NPU e no trabalho de avaliação e de normalização entre uns e outros.

Pergunta	Avaliação	Números-chave (fontes nos capítulos)
① **Onde chega hoje a codificação neural ao mercado em grande escala?**	Em torno do codec: ISP das câmaras, redução de resolução na codificação, super-resolução na reprodução. De ponta a ponta, só em circuitos fechados: chamadas, recursos de jogos, tokenizadores de IA generativa	Vídeo ≈75% dos dados móveis; 1,26 mil milhões de telemóveis e ≈200 milhões de televisores entregues em 2025, todos a descodificar em hardware de função fixa; DLSS 4 em 250+ jogos e aplicações
② **Os codecs neurais vão substituir o AV1, o HEVC e o VVC na reprodução? Quando?**	Não antes de uma norma e de um ciclo de dispositivos. O H.267 é o primeiro veículo realista, e deverá misturar ferramentas clássicas e neurais	Pedido de propostas do H.267 em julho de 2026, avaliação em janeiro de 2027, versão 1 no final de 2029; generalização em meados da década de 2030 estimativa
③ **Quanto custa em silício?**	Uma ordem de grandeza a mais em energia por píxel face à descodificação de função fixa atual, mais tráfego de memória e a exigência de exatidão ao bit	14,6 nJ/píxel (neural) face a 1,19 (HEVC, com DRAM); ≈175 kMAC por píxel descodificado ≈ 22 TOPS a 1080p30 estimativa

Pergunta	Avaliação	Números-chave (fontes nos capítulos)
④ **O que permitem as normas e as patentes?**	Os pós-filtros neurais já podem ser sinalizados em fluxos VVC; o JPEG AI é uma norma ISO; o H.267 vai decidir que ferramentas neurais se tornam licenciáveis	ECM (clássico) ≈28% face ao VVC; filtros em ciclo neurais 6–14% a 10–22+ kMAC/píxel; ByteDance com 645 famílias de patentes no trabalho pós-VVC
⑤ **Quem ganha dinheiro, e onde pode uma empresa apostar?**	Plataformas, fornecedores de IP, titulares de patentes. As startups dedicadas só a codecs neurais foram adquiridas. Apostar em pré e pós-processamento, circuitos fechados, silício que junta codec e NPU, avaliação	Agrupamento VVC: 0,50 dólares por telemóvel; IP de descodificador AV2 licenciado semanas após a especificação; três aquisições, em condições não divulgadas
⑥ **Que alegações devem ser desvalorizadas?**	Ganhos citados face a codecs antigos, em baixa resolução, sem custo de descodificação, ou de demonstrações em vez de produtos lançados	MLVC: 88% menos do que o H.264, mas 65–76% menos do que o HEVC em hardware; texturas: 85% de poupança de VRAM numa demonstração, cerca de 30% no jogo já lançado

FIG. 01

Onde a codificação neural chega ao mercado em 2026

As redes neuronais já funcionam nas duas pontas dos codecs dominantes. Os codecs neurais de ponta a ponta só chegam ao mercado onde uma empresa controla tanto o codificador como o descodificador.

	Em torno de um codec normalizado pré-processamento, melhoria, ampliação	**Codec neural de ponta a ponta** codificador e descodificador aprendidos	**Dentro de uma norma** ferramentas neurais num fluxo de bits ou API normalizados
01 **Reprodução e streaming**	● EM GRANDE ESCALA Redutor de resolução da Netflix; super-resolução do YouTube e da Meta; ampliação por IA nos televisores	● NÃO IMPLANTADO Nenhum serviço de streaming envia um formato de vídeo neural; televisores e telemóveis descodificam codecs normalizados em hardware	● EM NORMALIZAÇÃO A versão 1 do H.267 está prevista para o final de 2029; há ferramentas neurais entre as candidatas
02 **Fotografia e imagem**	● EM GRANDE ESCALA ISP com IA em telemóveis topo de gama; um ISP neural completo mostrado como IP na CES 2026	● DEMONSTRAÇÕES, PILOTOS O JPEG AI é uma norma ISO: demonstrações em telemóveis, nenhum produto	● NÃO IMPLANTADO O formato que ganha terreno é clássico: JPEG XL ativo por predefinição no Chrome 155
03 **Chamadas e reuniões**	● EM GRANDE ESCALA Super-resolução no recetor e supressão neural de ruído	● EM GRANDE ESCALA Áudio neural nas chamadas a dois do Teams desde 2021; vídeo MLVC em implementação no Teams	● USO LIMITADO O Opus 1.5 acrescenta ferramentas de ML opcionais e retrocompatíveis
04 **IA generativa**	—	● EM GRANDE ESCALA Cada modelo de vídeo e de voz usa um tokenizador aprendido, p. ex. um VAE de vídeo de 1:192	—
05 **Jogos e gráficos**	● EM GRANDE ESCALA Ampliação neural e geração de fotogramas: só o DLSS 4 está em 250+ jogos e aplicações	● USO LIMITADO Texturas neurais já lançadas num jogo AAA; o SDK da NVIDIA está em beta	● EM NORMALIZAÇÃO O DirectX acrescenta operações matriciais aos shaders (pré-visualização)
06 **Ligações especializadas: satélite, robôs**	● DEMONSTRAÇÕES, PILOTOS Restauro, no servidor, de fluxos de baixo débito	● DEMONSTRAÇÕES, PILOTOS Codecs neurais resistentes a perdas, em investigação e pilotos	—

Fonte: Documentos de empresas, de normalização e para programadores citados no relatório; situação a 12 de novembro de 2026. As linhas estão ordenadas pelo número de pessoas que cada domínio alcança.

ÂMBITO E MÉTODO

01 O que significa aqui «codificação neural», e como a analisei

Três formas de uma rede neuronal entrar num codec

«Codificação neural» abrange três coisas diferentes, e a maior parte das discussões sobre o tema nasce de as confundir.

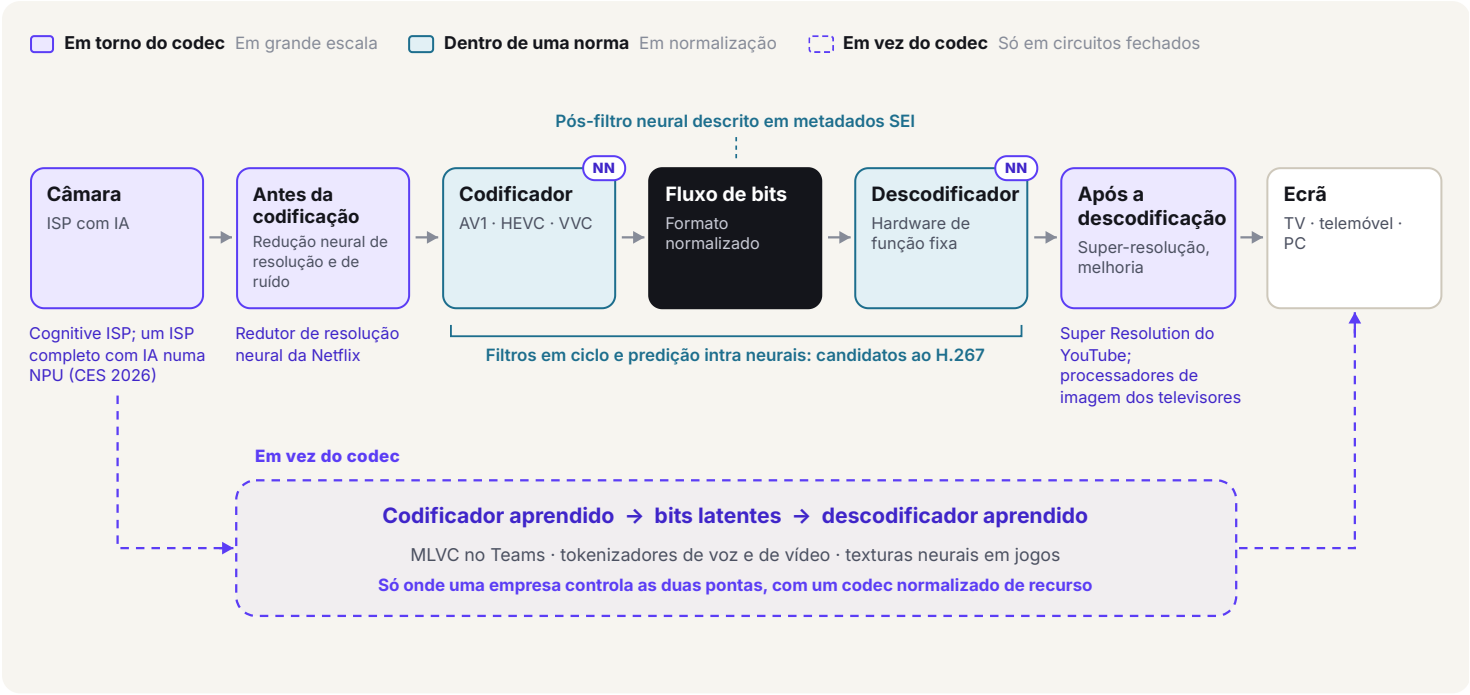
01 **Em torno do codec.** Uma rede processa píxeis antes da codificação ou depois da descodificação, e o fluxo de bits mantém-se normalizado. Exemplos: um processador de sinal de imagem (ISP) com IA numa câmara, o redutor de resolução neural da Netflix antes do seu codificador, a super-resolução num televisor. Qualquer dispositivo pode reproduzir o fluxo; a rede só muda o que entra ou o que sai.

02 **Dentro de uma norma.** Uma rede torna-se uma ferramenta de um codec normalizado, por exemplo um filtro em ciclo ou um preditor intra neural que todos os descodificadores têm de executar, pelo que tem de ser especificada com exatidão, normalmente em aritmética inteira. Uma forma mais leve é um pós-filtro neural que o fluxo de bits descreve em metadados e que os descodificadores podem optar por aplicar.

03 **Em vez do codec.** Tanto o codificador como o descodificador são aprendidos de ponta a ponta. O fluxo de bits é um conjunto de valores latentes quantizados que só a rede correspondente consegue descodificar. Exemplos: o MLVC da Microsoft no Teams, os tokenizadores dos modelos de voz, as texturas neurais num jogo.

FIG. 02 **Três formas de uma rede neuronal entrar num codec**

A maior parte do que chega ao mercado em 2026 funciona em torno de um codec normalizado. As ferramentas neurais dentro de uma norma estão a ser especificadas. Os codecs neurais de ponta a ponta só substituem o codec onde uma empresa controla as duas pontas.



Fonte: Síntese do autor a partir das fontes deste relatório; situação a 12 de novembro de 2026.

O primeiro tipo não precisa da autorização de ninguém. O segundo precisa de uma norma. O terceiro precisa do controlo das duas pontas. É por essa ordem que chegam às pessoas, e isso explica a maior parte do que se segue.

Ordenado pelo mundo real, não pela agenda de investigação

Os artigos de investigação estudam sobretudo o terceiro tipo, porque é aí que estão os recordes de compressão. Este relatório ordena os domínios pelo número de pessoas que alcançam, porque é isso que decide para onde vão o dinheiro e a engenharia: primeiro a reprodução e o streaming, depois a fotografia, as chamadas, a IA generativa, os jogos e, por último, as ligações especializadas (satélite, drones, robôs). O vídeo representa cerca de 75% de todo o tráfego de dados móveis, segundo as contas da Ericsson no final de 2025. Em 2025, a indústria entregou cerca de 1,26 mil milhões de smartphones (IDC) e aproximadamente 196–209 milhões de televisores (estimativas da TrendForce e da Omdia). Todos eles descodificam vídeo em hardware de função fixa.

Como trabalhei

Trata-se de investigação documental, feita como eu avaliaria um chip ou um plano de produto. Li documentos de normalização e relatórios de reuniões (JVET, MPEG, JPEG, AOMedia, IETF), blogues de engenharia e notas de versão de empresas, artigos sobre chips, tabelas de tarifas de agrupamentos de patentes, textos de controlo de exportações e a imprensa. Cada número indica a sua data. Quando eu próprio derivo um número, como a computação por píxel a partir das tabelas de um artigo, digo-o e assinalo-o como **estimativa**. As alegações que não pude confirmar numa fonte primária estão assinaladas como **alegação do fornecedor**. Os valores monetários estão em dólares dos EUA. A situação descrita é a de 12 de novembro de 2026.

Comparo o que é comparável sempre que posso, e digo-o quando não posso. Os ganhos de compressão pouco significam sem quatro coisas associadas: a âncora (ganho face a quê), a métrica (PSNR, VMAF ou um painel de

espetadores), o conteúdo e a resolução, e quanto custa descodificar. A maior parte das divergências sobre codecs neurais desaparece quando essas quatro coisas são indicadas.

CAPÍTULO 02

REPRODUÇÃO E STREAMING

02 Reprodução e streaming: neural nas duas pontas, normalizado no meio

O que os espetadores recebem continua a ser um fluxo de bits normalizado

Os maiores serviços de vídeo concluíram uma transição de codec e estão a começar a seguinte, e nenhuma das duas envolve um descodificador neural. Em dezembro de 2025, a Netflix informou que o AV1 transportava cerca de 30% das suas visualizações, sendo o seu segundo codec mais usado, depois do AVC. O AV1 só chegou aos televisores quando os descodificadores em hardware chegaram: a Netflix incluiu o AV1 na certificação de dispositivos em 2019, começou a transmiti-lo para televisores no final de 2021 e diz que 88% dos dispositivos de ecrã grande submetidos a certificação entre 2021 e 2025 o suportavam, quase todos desde 2023. O YouTube indicou que, a 19 de março de 2026, mais de 90% do seu catálogo, ponderado pelo tempo de visualização, estava disponível em AV1.

O AV2, o sucessor do AV1 criado pela Alliance for Open Media, chegou à versão 1.0.0 a 28 de maio de 2026, com uma compressão cerca de 30% melhor do que a do AV1, segundo as medições objetivas da AOMedia. O YouTube prevê emissões em AV2 na Web «já no início de 2027». O silício seguiu-se de imediato: a Chips&Media concluiu um descodificador AV2 em hardware em maio e licenciou-o em junho a uma grande tecnológica norte-americana não identificada; a Allegro DVT tinha entregado o seu IP de descodificador AV2 a três clientes até 16 de junho; a VeriSilicon tinha disponibilizado um IP de vídeo compatível com AV2 a vários clientes até 10 de junho. O AV2 é um codec clássico: os seus ganhos vêm de melhorias no particionamento, na predição, nas transformadas e nos filtros em ciclo (um deles, um filtro de detalhe guiado «orientado por dados»), não de uma rede de ponta a ponta.

O VVC, a atual norma da ITU/ISO, encontrou a sua primeira casa na radiodifusão no Brasil, cujo sistema DTV+ (TV 3.0) entrou em operação comercial em junho de 2026, a tempo do Mundial. Em julho de 2026, nenhum televisor à venda no Brasil tinha um recetor DTV+ integrado; os primeiros são esperados entre o final de 2026 e o início de 2027.

No codificador: o redutor de resolução neural da Netflix

O exemplo mais claro de codificação neural na reprodução funciona desde 2022 e não muda nada do que o dispositivo do espetador tem de fazer. A Netflix codifica cada título em várias resoluções, e o passo de redução de resolução que produz 1080p, 720p e 540p a partir de um master 4K usava tradicionalmente um filtro fixo, como o de Lanczos. A Netflix substituiu-o por uma pequena rede neuronal, um bloco de pré-processamento seguido de um bloco de redimensionamento, treinada para melhorar a qualidade observada depois de um filtro de ampliação convencional no dispositivo. A Netflix descreve-a como uma solução de substituição direta: nenhuma outra alteração do lado da codificação nem do lado do cliente, pelo que «milhões de dispositivos compatíveis com o streaming da Netflix beneficiam automaticamente». Nos testes de preferência da Netflix, cerca de 77% dos espetadores escolheram o resultado do redutor de resolução neural alegação do fornecedor. Para manter o custo aceitável à escala da Netflix, «bastaram apenas algumas camadas de rede neuronal».

É esse o padrão a reter. A rede corre uma vez por título no centro de dados, onde a computação é barata e pode ser planeada, e o resultado é um fluxo AV1 ou HEVC comum.

No ecrã: super-resolução na aplicação e no televisor

Na outra ponta, as redes reparam o que o codec e a fonte perderam. A 29 de outubro de 2025, o YouTube anunciou o Super Resolution para a sua aplicação de televisão: os vídeos carregados abaixo de 1080p, de 240p a 720p, são ampliados para HD com aprendizagem automática, sendo o 4K o objetivo declarado. As versões ampliadas são assinaladas, os espetadores podem voltar ao original e os criadores podem optar por não participar. A Meta descreveu em outubro de 2025 como usa super-resolução de vídeo tanto nos seus servidores, no carregamento, em CPU x86 comuns e em parceria com a Intel, como no cliente, na reprodução, e concluiu que as duas se complementam. Os fabricantes de televisores fazem o mesmo em silício: o processador 8K topo de gama da Samsung, o NQ8 AI Gen3, anuncia 768 redes neuronais para ampliação.

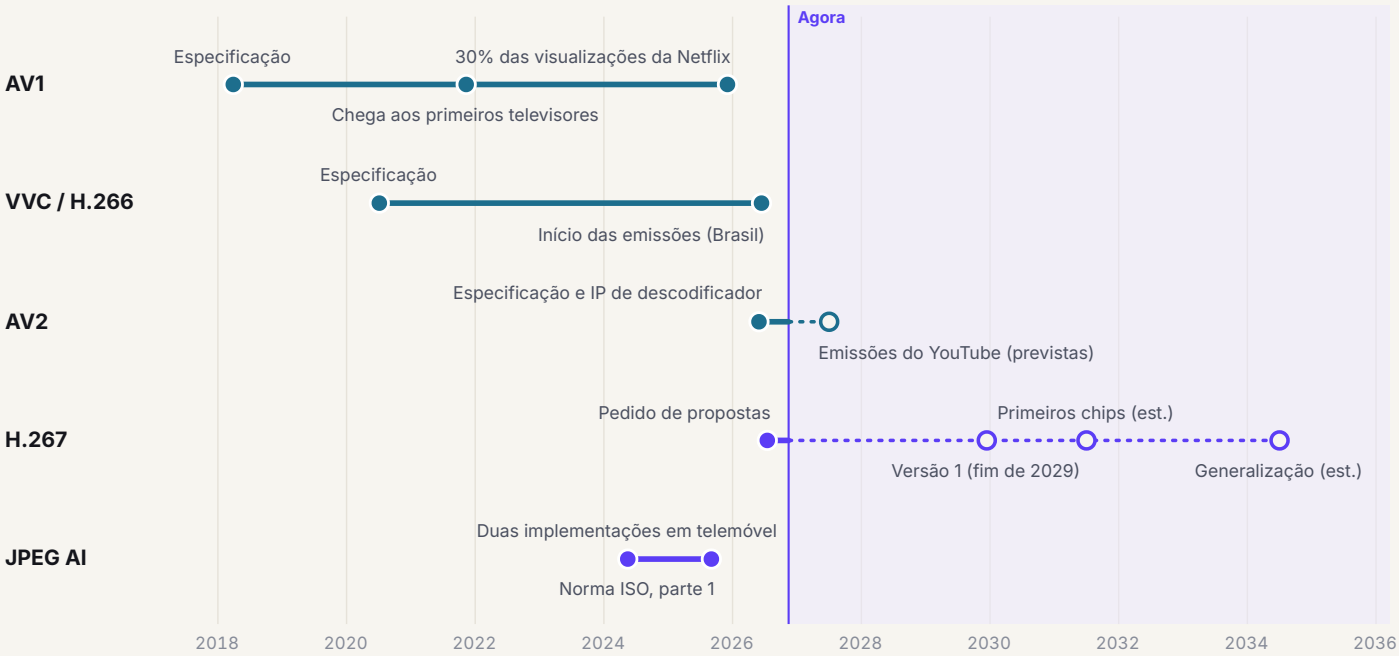
A economia muda consoante o sítio onde a rede corre. O processamento no servidor custa à plataforma computação por cada vídeo em que toca. O processamento no dispositivo não custa nada à plataforma; usa uma NPU que o comprador já pagou. É por isso que os maiores ganhos em 2026 vêm das duas coisas.

Porque é que nenhum televisor descodifica um fluxo de bits neural

Um formato de reprodução tem de ser descodificável em todos os dispositivos que alguma vez o vão reproduzir, e esses dispositivos duram anos. Isso torna lento o caminho de uma norma concluída até aos ecrãs do grande público, mesmo para os codecs clássicos. A especificação do AV1 ficou concluída em março de 2018; o descodificador de software dav1d chegou poucos meses depois; os televisores receberam-no em streaming a partir do final de 2021; chegou a cerca de 30% das visualizações da Netflix em dezembro de 2025, quase oito anos depois da especificação. O VVC ficou concluído em julho de 2020 e chegou ao seu primeiro serviço de radiodifusão em junho de 2026.

FIG. 03 **Entre uma norma concluída e os ecrãs do grande público passa a maior parte de uma década**

A reprodução precisa de um descodificador em hardware no dispositivo. O AV1 levou quase oito anos da especificação concluída aos 30% das visualizações da Netflix; a primeira norma que deverá trazer ferramentas neurais está prevista para o final de 2029.



Fonte: Netflix Tech Blog (dez. 2025); AOMedia; Allegro DVT; TV Technology; Fraunhofer IIS sobre o DTV+ do Brasil; MPEG-155; comité JPEG; ISO. Os marcadores vazios são metas ou estimativas do autor.

Um fluxo de bits neural acrescenta dois problemas que os codecs clássicos resolveram há muito: a energia por píxel e a descodificação exata ao bit. Ambos são tratados no capítulo 8. Juntos, significam que nenhum fabricante de televisores ou de telemóveis vai pôr um descodificador de vídeo neural em todos os dispositivos antes de uma norma lhe dizer exatamente o que construir.

O que o mudaria: o H.267

A norma que poderá levar ferramentas neurais a todos os ecrãs é o sucessor do VVC, desenvolvido em conjunto pelo ITU-T e pela ISO/IEC na Joint Video Experts Team (JVET). O nome ainda não é definitivo; vou chamar-lhe H.267, como faz grande parte da indústria.

Em outubro de 2025, a JVET avaliou as respostas a um pedido de evidências: cinco respostas de 16 empresas e laboratórios de investigação, com 20 configurações de codec testadas. O comité encontrou evidências de melhor compressão do que o VVC e considerou «interessantes» tanto a tecnologia convencional melhorada como a tecnologia de redes neuronais. As abordagens neurais mostraram ganhos percetuais claros, sobretudo em conteúdo 8K HDR, e ganhos menores em conteúdo de jogos. A InterDigital, uma das participantes, submeteu uma proposta baseada no Enhanced Compression Model (ECM) convencional e outra que sobrepunha ferramentas de IA ao VVC.

O pedido formal de propostas foi lançado na reunião de julho de 2026, em Genebra. As submissões principais tinham como prazo o final de outubro de 2026, um pacote complementar tem prazo em meados de dezembro, e as respostas serão avaliadas na reunião da JVET em Brisbane, em janeiro de 2027. A JVET espera um primeiro modelo de teste no segundo semestre de 2027 e uma primeira versão da norma até ao final de 2029, e descreve o trabalho como «provavelmente um esforço maior do que alguma vez se fez para arrancar com uma nova norma».

A via clássica põe a fasquia alta. O ECM, construído com processamento de sinal convencional, poupa aproximadamente 28% face ao VVC em configuração de acesso aleatório. A exploração de redes neuronais da JVET reporta filtros em ciclo que poupam cerca de 6% com 10 mil multiplicações-acumulações por píxel, 8% com 22 mil e 14% num ponto de funcionamento de alta complexidade. A minha leitura é que o H.267 será muito provavelmente um núcleo clássico com algumas ferramentas neurais onde os bits poupados justifiquem o custo no descodificador, e que a avaliação de janeiro de 2027 vai mostrar quantas análise.

Já existe uma ponte. Os fluxos VVC e HEVC podem transportar informação suplementar de melhoria (SEI) que descreve um pós-filtro neural a aplicar do lado do descodificador, e as novas edições do VVC e da sua norma de SEI (VSEI), concluídas em outubro de 2025, alargaram essas mensagens. Um fluxo pode indicar que rede executar depois da descodificação sem obrigar todos os descodificadores a executá-la.

O que isto significa para as empresas da reprodução

Venda já o que funciona em torno do codec: pré-processamento na codificação para os serviços, e blocos de pós-processamento e afinação para chips de televisores e de set-top boxes. Não construa um negócio de reprodução assente num fluxo de bits neural que tenha de ser descodificado em dispositivos de terceiros antes da década de 2030. Para quem faz silício, a receita a curto prazo está nos descodificadores AV2 e VVC; a parte neural é um caminho de NPU para melhoria ao lado deles (capítulo 8).

CAPÍTULO 03

FOTOGRAFIA E IMAGEM

03 Fotografia e imagem: neural dentro da câmara, clássico na Web

A cadeia da câmara já é neural

Todas as fotografias de um telemóvel topo de gama recente passaram por redes neuronais antes de serem comprimidas. O Snapdragon 8 Gen 2 da Qualcomm, nos telemóveis de 2023, introduziu um «Cognitive ISP» que faz segmentação semântica em tempo real durante a captura de fotografia e de vídeo, ajustando separadamente rostos, cabelo, céus e fundos; o Galaxy S23 Ultra da Samsung foi o primeiro telemóvel a usá-lo na câmara frontal. Em janeiro de 2026, a Chips&Media e a Visionary.ai mostraram na CES aquilo a que chamam o primeiro ISP completo baseado em IA: uma cadeia de imagem de ponta a ponta que corre inteiramente numa NPU, a WAVE-N da Chips&Media, em vírgula flutuante de 16 bits, sem um ISP de função fixa, e que pode ser atualizada remotamente. «As cadeias existentes acoplam blocos neurais a um ISP de função fixa. Nós substituímos totalmente o ISP convencional», afirmou o CEO da Visionary.ai. Foi proposto aos fabricantes para avaliação depois da feira.

Depois, o resultado é guardado em JPEG ou HEIF, formatos clássicos que qualquer dispositivo consegue abrir. O padrão é o mesmo da reprodução: neural dentro da cadeia de uma empresa, normalizado na fronteira.

JPEG AI: uma norma neural concluída, ainda sem produto

O JPEG AI é a primeira norma de codificação de imagem aprendida da ISO/IEC e do ITU-T (ISO/IEC 6048, ITU-T T.840). A parte 1, o sistema de codificação central, foi publicada em 2025. O comité JPEG reporta ganhos de compressão de 12,5% a 27,9% face à codificação intra do VVC, numa gama de complexidade do descodificador de 7 a 216 mil multiplicações-acumulações por píxel. A conceção prevê dois codificadores e três descodificado-

res, treinados em conjunto, para que um recetor possa implementar só um descodificador e ainda assim ler todos os fluxos. Em maio de 2024, o comité já reportava duas implementações em smartphones. Em paralelo, o IEEE publicou em dezembro de 2024 a IEEE 1857.11-2024, uma norma de codificação de imagem baseada em redes neuronais concebida para NPU.

Em novembro de 2026, não encontrei nenhuma câmara de telemóvel, aplicação de fotografia ou navegador que escreva ou leia JPEG AI por predefinição.

O que a Web adotou em vez dele: o JPEG XL

O formato de imagem que ganhou de facto terreno em 2026 é clássico. A 6 de outubro de 2026, a Google ativou por predefinição a descodificação de JPEG XL no Chrome 155, com um novo descodificador com segurança de memória escrito em Rust, quatro anos depois de ter retirado o suporte experimental. O Safari descodifica JPEG XL desde 2023, e a Mozilla tem seguido o mesmo caminho. Os argumentos da Google a favor são uma compressão melhor do que a do JPEG, a recompressão sem perdas de ficheiros JPEG existentes, HDR e descodificação progressiva.

A lição generaliza-se. Um formato que tem de abrir em todo o lado ganha com um descodificador rápido e determinístico, que corre em qualquer CPU e pode ser implementado com segurança de memória. Um descodificador de imagem neural precisa de uma NPU ou GPU e de aritmética exata entre dispositivos. A codificação neural ganha dentro da cadeia do próprio dispositivo e em sistemas onde a mesma empresa codifica e descodifica, como um serviço de fotografias na nuvem que nunca entrega o ficheiro comprimido a mais ninguém.

O que isto significa para as empresas de imagem

O dinheiro está na cadeia de captura: IP de ISP com IA, NPU dimensionadas para imagem e o trabalho de afinação que os fabricantes de câmaras compram. Para armazenamento e distribuição, um codec de imagem aprendido é hoje uma ferramenta para sistemas fechados, não um formato em que apostar.

CAPÍTULO 04

CHAMADAS E REUNIÕES

04 Chamadas e reuniões: onde os codecs neurais de ponta a ponta chegam realmente ao mercado

O áudio foi primeiro

A Microsoft apresentou o Satin, um codec de voz com IA, em fevereiro de 2021, e disse que já era usado em todas as chamadas a dois do Teams e do Skype. O Satin oferece voz em banda superlarga, até 16 kHz de largura de banda de áudio, a 6 kbps, quando o anterior codec Silk do Skype era de banda estreita a esse débito. O débito poupado é gasto em redundância, pelo que as chamadas resistem melhor à perda de pacotes.

MLVC: um codec de vídeo neural em chamadas reais

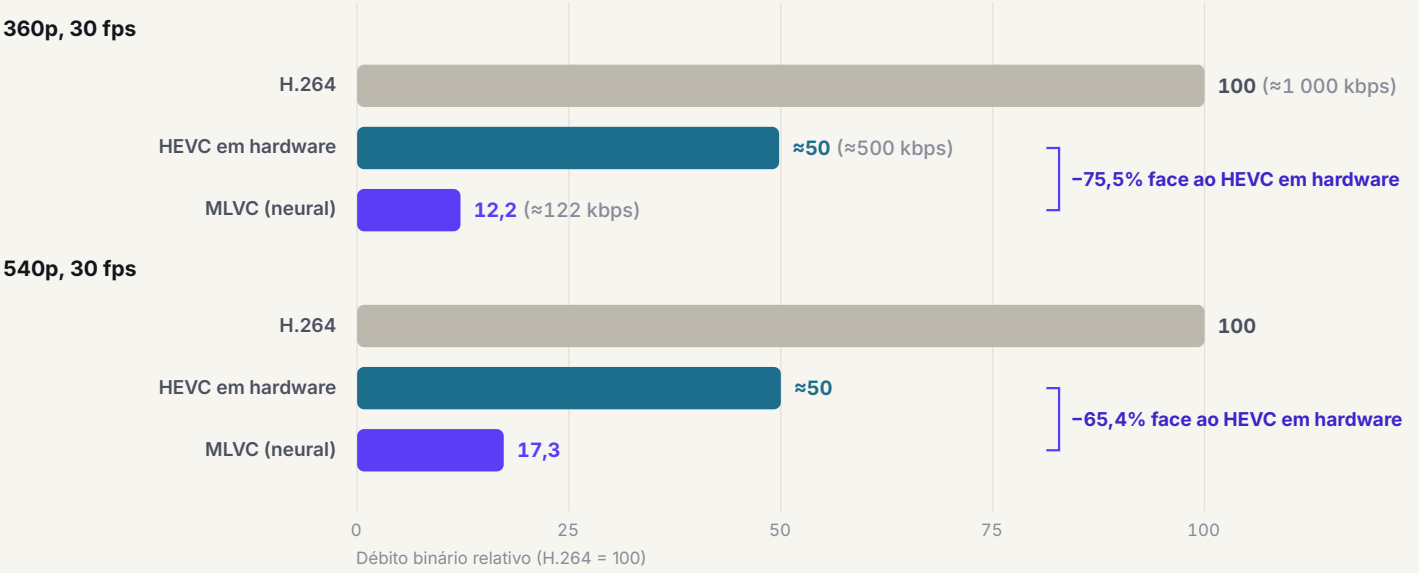
A 24 de julho de 2026, a Microsoft disponibilizou o MLVC, o seu codec de vídeo de aprendizagem automática, em código aberto sob a licença MIT, com o código do modelo, os pesos treinados, os scripts de treino e as ferramentas de conversão para diferentes NPU. O MLVC é a versão de produto da família DCVC que a Microsoft Rese-

arch publica desde 2021. A Microsoft diz que está a implementar o MLVC no Teams para chamadas ponto a ponto, com telemetria em tempo real, testes A/B e os codecs convencionais como recurso.

O MLVC funciona a 540p e a 360p, a 30 fotogramas por segundo, e foi concebido para não usar mais de metade da NPU no Neural Engine da Apple e nas NPU da Qualcomm e da Intel, repartindo o trabalho entre CPU e NPU. Em testes subjetivos segundo a ITU-T P.910, no Video Conferencing Dataset da própria Microsoft, precisou de menos 87,8% de débito binário do que o H.264 e de menos 75,5% do que o HEVC em hardware a 360p, e de menos 82,7% e 65,4% a 540p alegação do fornecedor. A 360p, isso corresponde a aproximadamente 122 kbps, quando o H.264 precisa de cerca de 1 Mbps.

FIG. 04 **Em chamadas, um codec neural supera o HEVC em hardware em 65–76%**

Débito binário necessário para a mesma qualidade avaliada no conjunto de testes de videoconferência da Microsoft, H.264 = 100. O MLVC funciona nas NPU de portáteis e telemóveis até 540p30, usando no máximo metade da NPU.



Fonte: Microsoft, lançamento do MLVC (24 de julho de 2026): testes subjetivos ITU-T P.910 no seu Video Conferencing Dataset. Barras do HEVC derivadas das poupanças comunicadas. Alegações do fornecedor, não medições independentes.

Duas coisas moderam o número de destaque. Os 88% são face ao H.264; face ao HEVC em hardware que os mesmos portáteis já contêm, são 65–76%, o que continua a ser muito. E o MLVC abdicou de compressão para poder ser implantado: uma análise independente nota que um DCVC-RT treinado percetualmente obtém 81,9% na mesma medida em que o MLVC obtém 75,5%. O preço de descodificar de forma idêntica em hardware que não é da Microsoft foi de cerca de seis pontos.

Porque é que as chamadas são o habitat natural

Uma chamada é um circuito fechado. Os dois terminais correm a mesma aplicação, negociam o que conseguem fazer no início da chamada e podem recorrer a um codec normalizado a meio dela. As resoluções são modestas, as redes são muitas vezes fracas, e os portáteis das duas pontas já vêm com NPU, como os telemóveis há anos. A Microsoft fixa 40 TOPS de NPU como fasquia para os seus PC com IA, e a Apple atribui 38 TOPS ao Neural Engine do M4. Todas as condições que travam um codec neural na reprodução estão aqui ausentes.

Os contraexemplos: Opus e MLow

Nem todas as plataformas de chamadas fizeram a mesma escolha, e as razões são instrutivas.

O Opus, o codec de áudio aberto do WebRTC e da maioria das aplicações de chamadas, acrescentou aprendizagem profunda na versão 1.5 (março de 2024), a primeira vez que usou redes para processar ou gerar o próprio sinal. Mas acrescentou-as como opções: melhoria do lado do descodificador (LACE e NoLACE), que não exige alterações ao fluxo de bits, uma ocultação neural de perda de pacotes e um esquema de redundância profunda (DRED); tudo vem desativado por predefinição, para não sobrecarregar os dispositivos antigos. O Opus 1.6 acrescentou uma extensão neural de largura de banda e um modelo DRED três vezes mais pequeno, e mantém-se totalmente compatível com a norma de 2012. A abordagem é «dentro de uma norma, opcional e retrocompatível».

A Meta recuou ainda mais. Em junho de 2024, lançou o MLow, um codec de voz de baixo débito construído sobre processamento de sinal clássico em vez de uma rede, hoje presente em todas as chamadas do Instagram e do Messenger e em implementação no WhatsApp. As razões da Meta foram os seus dispositivos: mais de 20% das suas chamadas são feitas em telemóveis ARMv7, e dezenas de milhões de chamadas de WhatsApp por dia correm em telemóveis com mais de dez anos. Os codecs neurais, incluindo o próprio EnCodec da Meta, só funcionam de forma fiável em aparelhos caros, escreveu a Meta. O MLow oferece cerca do dobro da qualidade do Opus a 6 kbps, usando menos 10% de computação.

Para um serviço cujos utilizadores vão de portáteis topo de gama a telemóveis com uma década, quem decide é o orçamento de complexidade. Os codecs neurais chegam primeiro ao mercado onde a base instalada é recente e o fornecedor controla as duas pontas.

O que isto significa para as empresas de comunicações

A videoconferência empresarial, a telepresença, a assistência remota, as ligações de contribuição e as redes privadas são os sítios onde um codec neural de ponta a ponta é hoje um produto, desde que as duas pontas sejam suas. Planeie três coisas desde o início: um codec de recurso, testes de descodificação em todas as NPU que suportar e uma avaliação de qualidade sobre o seu próprio conteúdo, e não sobre o conjunto de dados de um fornecedor.

CAPÍTULO 05

IA GENERATIVA

05 IA generativa: a maior implantação a que ninguém chama codec

Todos os modelos de voz e de vídeo assentam num codec aprendido

A maior implantação de compressão neural de ponta a ponta não está em nenhuma norma de codecs. Está dentro da IA generativa. Um modelo de voz não ouve nem produz formas de onda diretamente: lê e escreve os tokens discretos de um codec de áudio aprendido. O Mimi, da Kyutai, lançado em 2024, transforma a voz em tokens a 12,5 tramas por segundo e 1,1 kbps, abaixo de qualquer codec de telefonia, combinando informação semântica e acústica num único fluxo. O EnCodec da Meta, publicado em 2022, é outro; os próprios engenheiros da Meta descreveram mais tarde estes codecs como excelentes a baixos débitos, mas demasiado pesados para telemóveis mais antigos (capítulo 4).

A geração de vídeo funciona da mesma forma. Um modelo de difusão ou um transformer gera no espaço latente de um autocodificador de vídeo aprendido (VAE), e o descodificador do VAE transforma o resultado em píxeis. O VAE do LTX-Video comprime o vídeo 1:192: cada token latente representa um bloco de 32×32 píxeis ao longo de

8 fotogramas, com 128 canais. O seu descodificador também executa o último passo de redução de ruído, pelo que em parte gera detalhe em vez de apenas o reconstruir.

Porque é que não são formatos de distribuição

Estes codecs correm em cada pedido a um modelo de voz ou de vídeo, mas nenhum deles é um formato que mais alguém descodifique. Cada modelo tem o seu próprio espaço latente; não há controlo de débito, teste de conformidade nem garantia de que outro descodificador produziria a mesma imagem. São concebidos para a qualidade de geração, não para a fidelidade a um dado débito binário, e descodificam em GPU de centros de dados. O seu sucesso mostra que a compressão aprendida funciona em grande escala quando uma empresa é dona do codificador, do descodificador e do hardware.

O que isto significa para as empresas

Dois pontos práticos. Primeiro, a eficiência do tokenizador é uma linha de custo direta para quem serve vídeo ou voz generativos: num transformer cuja atenção abrange todos os tokens, a computação cresce mais depressa do que o número de tokens, pelo que uma representação latente mais compacta reduz o custo de geração mais do que proporcionalmente análise. Os engenheiros de codecs que percebem de conceção de espaços latentes, de modelos de entropia e de eficiência de descodificadores são valiosos para as equipas de IA generativa. Segundo, a fronteira está a mexer-se: os organismos de normalização começaram a trabalhar em codificação para máquinas em vez de pessoas. O Video Coding for Machines do MPEG (ISO/IEC 23888-2) chegou a projeto de norma internacional no início de 2026 e deverá ficar concluído no final de 2026, e o JPEG AI foi concebido desde o início para servir tarefas de visão por computador, além da visualização humana.

CAPÍTULO 06

JOGOS E GRÁFICOS

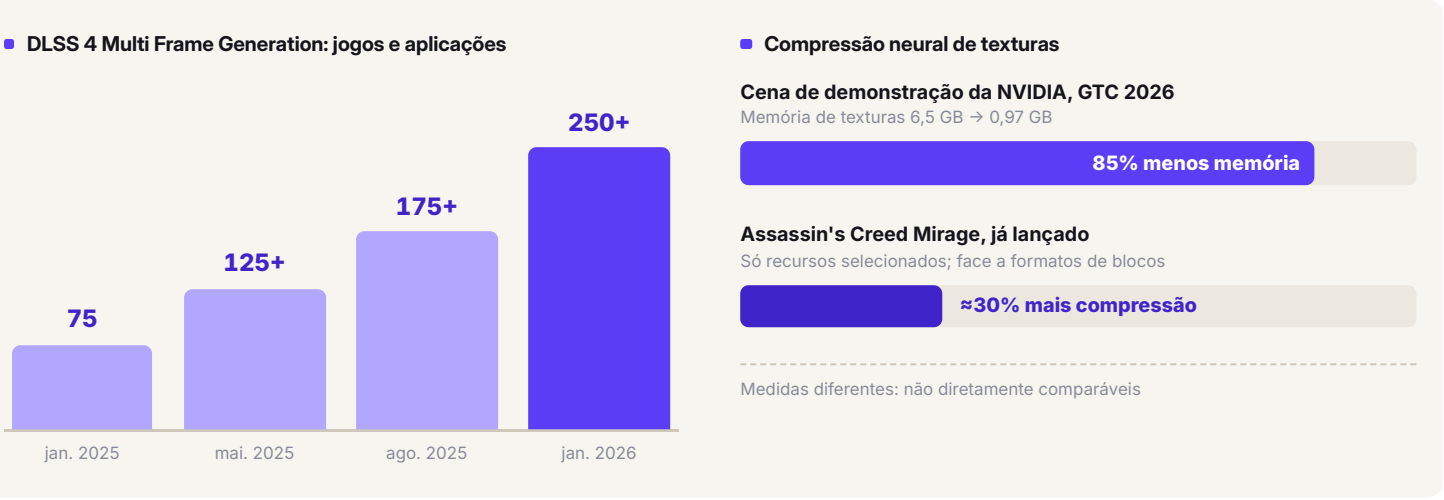
06 Jogos e gráficos: a reconstrução neural está generalizada, os recursos neurais estão a começar

Renderizar menos píxeis e deixar uma rede preencher o resto

Nos gráficos em tempo real, uma rede já reconstrói a maior parte do que os jogadores veem. O DLSS da NVIDIA renderiza menos píxeis e menos fotogramas e usa redes para produzir o resto. O DLSS 4 com Multi Frame Generation foi lançado em janeiro de 2025 com suporte em 75 jogos e aplicações; a NVIDIA contava mais de 125 em maio de 2025, mais de 175 em agosto e mais de 250 em janeiro de 2026, quando anunciou o DLSS 4.5, com um modelo transformer de segunda geração para super-resolução. É um descodificador em tudo menos no nome: o «fluxo de bits» é um fotograma de baixa resolução mais dados de movimento, e a rede reconstrói a imagem completa.

FIG. 05 **Quando uma empresa controla toda a pilha, a reconstrução neural difunde-se depressa**

À esquerda: jogos e aplicações compatíveis com o DLSS 4 Multi Frame Generation da NVIDIA. À direita: compressão neural de texturas numa cena de demonstração e num jogo já lançado, que só a usou em recursos selecionados devido ao custo em tempo de execução.



Fonte: NVIDIA (janeiro de 2025 a janeiro de 2026; demonstração na GTC 2026); Ubisoft La Forge (fevereiro de 2026). Alegações dos fornecedores; os dois valores de texturas medem coisas diferentes.

As texturas neurais chegaram ao mercado, de forma seletiva

As texturas são comprimidas em formatos de blocos fixos (BCn) que as GPU descodificam em hardware. A compressão neural de texturas substitui um conjunto de texturas de materiais por uma representação latente compacta e uma pequena rede que reconstrói os texels quando são precisos. Em fevereiro de 2026, o laboratório La Forge da Ubisoft descreveu como a lançou em Assassin's Creed Mirage, um jogo de 2023, com uma compressão efetiva cerca de 30% superior à dos formatos convencionais para a mesma qualidade visual alegação do fornecedor. Não foi usada em todo o lado: «a utilização foi limitada pelo desempenho em tempo de execução», pelo que a Ubisoft a aplicou a um subconjunto de recursos com muitas instâncias e forte pressão sobre a memória.

O SDK RTX Neural Texture Compression da NVIDIA, em beta no GitHub, mostra o teto. Numa demonstração na GTC 2026, uma cena de uma vivenda que usava 6,5 GB de memória de texturas com BCn precisou de 970 MB com texturas neurais, menos 85% alegação do fornecedor. O SDK oferece duas vias: descodificar para BCn no carregamento, o que poupa espaço em disco e tamanho de transferência, mas não memória de vídeo, ou descodificar na amostragem, o que poupa a memória, mas custa tempo de shader em cada fotograma.

A API está a chegar

As API gráficas estão a acrescentar as operações matriciais de que a descodificação neural precisa dentro dos shaders. A Microsoft lançou o Shader Model 6.9 em fevereiro de 2026; a funcionalidade de vetores cooperativos que apresentou em pré-visualização em 2025 foi integrada numa conceção unificada de álgebra linear para um Shader Model posterior, com o DirectX Linear Algebra a entrar em pré-visualização pública em abril de 2026. Quando isso estiver em todos os controladores, um descodificador de texturas neural poderá correr na GPU de qualquer fabricante a partir de um único caminho de código.

O que isto significa para as empresas

Os jogos já trazem descodificação neural porque se verificam as condições de um circuito fechado: um só motor, uma só cadeia de produção de recursos, hardware tensorial na GPU e um formato de reserva. O trabalho em aberto está nas ferramentas, nas cadeias de produção que codificam materiais neurais, nos ambientes de execu-

ção compatíveis com vários fabricantes e, sobretudo, nos dispositivos móveis, onde quem decide é o custo energético por texel descodificado.

CAPÍTULO 07

07

LIGAÇÕES ESPECIALIZADAS

Ligações especializadas: satélites, drones e robôs

São os casos em que, no papel, a codificação neural de ponta a ponta melhor se adequa. Um único dono controla as duas pontas da ligação, a largura de banda é escassa e sujeita a perdas, o conteúdo é suficientemente restrito para o treino, e uma imagem fiel do ponto de vista percetual, ou até semântico, muitas vezes basta. Os codecs neurais resistentes a perdas para estas ligações são uma área de investigação ativa, e o restauro, no servidor, de fluxos de baixo débito já é usado.

São também um mercado pequeno e cada vez mais regulado, comparado com o vídeo de consumo análise. Na China, o Anúncio n.º 49 de 2026 do Ministério do Comércio e de outros dois organismos, publicado a 19 de outubro e com efeitos a partir de 18 de novembro de 2026, revê os controlos de exportação de 13 categorias de artigos relacionados com drones. Entre as entradas de tecnologia, uma abrange plataformas terrestres não tripuladas definidas em parte por limiares como 75% de compressão semântica de imagem e reconstrução de imagem em 150 milissegundos, segundo a imprensa especializada. A própria compressão semântica tornou-se um parâmetro de controlo de exportações. Nos Estados Unidos, a Comissão Federal de Comunicações acrescentou os drones de fabrico estrangeiro e os seus componentes críticos à sua Covered List a 22 de dezembro de 2025, pelo que os novos modelos estrangeiros já não podem receber autorização de equipamento.

O veredicto: um nicho sólido para uma empresa especializada, com clientes industriais ou governamentais e capacidade para gerir o cumprimento das regras de exportação dos dois lados, mas não um mercado de massas, nem o sítio para perceber se a codificação neural funciona em grande escala.

CAPÍTULO 08

SILÍCIO

08 A perspetiva do silício: energia, computação, memória e exatidão

Este é o capítulo em que a minha experiência mais conta. Tenho muitos anos de I&D e de gestão de chips fabless, incluindo conceção full-custom de caminhos de dados de ASIC e projetos de chips multimédia para câmaras, como aceleradores de codecs de vídeo, e as perguntas que uma equipa de chips faz sobre um codec novo não são sobre BD-rate. São estas: quantos joules por píxel, quantos bytes de tráfego de DRAM por píxel, quantas portas lógicas, e todos os descodificadores do mundo conseguem calcular exatamente os mesmos bits? Nas quatro, um descodificador neural está em 2026 onde o HEVC em hardware estava antes dos primeiros chips: promissor em software, caro em silício.

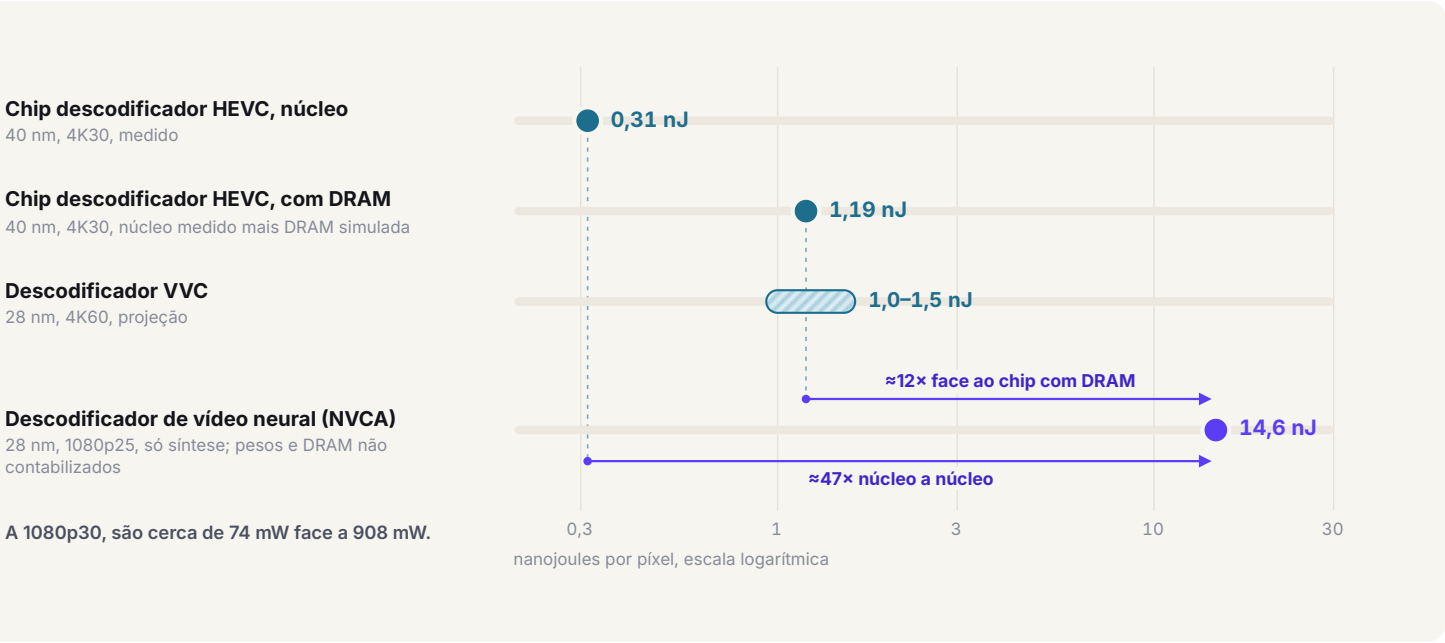
Energia por píxel: uma ordem de grandeza de distância

Poucos descodificadores neurais foram concebidos como chips, pelo que a comparação assenta num punhado de projetos publicados. A referência do lado clássico é o chip descodificador HEVC do MIT (Tikekar et al., ISSCC 2013 e IEEE JSSC 2014): 40 nm, 1,77 mm², a descodificar 4K a 30 fotogramas por segundo a 200 MHz. O seu núcleo consome 78 mW, ou 0,31 nJ por píxel; contando com a potência simulada da DRAM, o sistema precisa de 1,19 nJ por píxel. Para o VVC, os números de produção são proprietários; uma análise publicada na Electronic Design em 2026, que escala o chip do MIT para 28 nm e lhe aplica a maior complexidade do VVC, projeta 1–1,5 nJ por píxel para um descodificador 4K60 com cerca de um milhão de portas lógicas.

Do lado neural, o único projeto publicado de ASIC descodificador de vídeo neural de ponta a ponta, segundo as contas da Electronic Design, é o NVCA (Zhang et al., DATE 2024), sintetizado no processo de 28 nm da TSMC com 5,01 milhões de portas lógicas, sem contar os pesos, que ficam fora do chip. Descodifica 1080p a 25 fotogramas por segundo com 760 mW, ou 14,6 nJ por píxel.

FIG. 06 **Energia por píxel descodificado: os descodificadores de função fixa estão uma ordem de grandeza à frente**

O único projeto publicado de chip descodificador de vídeo neural precisa de cerca de 12 vezes a energia por píxel de um chip descodificador HEVC medido, contando com a sua DRAM, e de cerca de 47 vezes a do núcleo isolado. O tráfego de pesos e de DRAM do próprio projeto neural não é contabilizado.



Fonte: Tikekar et al., IEEE JSSC 2014 (MIT); NVCA, Zhang et al., DATE 2024 (arXiv 2312.10716); o intervalo do VVC é uma projeção da Electronic Design (2026). O chip HEVC usa um nó mais antigo (40 nm face a 28 nm), o que favorece o projeto neural.

Isso é cerca de 12 vezes o valor do chip HEVC contando com a sua DRAM, e cerca de 47 vezes o do seu núcleo isolado. A comparação favorece o NVCA de três formas: o chip HEVC está num nó mais antigo, o valor do NVCA deixa de fora o tráfego de pesos e de DRAM, e o NVCA é um resultado de síntese e não silício medido. A 1080p30, a diferença é de cerca de 74 mW face a aproximadamente 900 mW: um descodificador que ninguém nota, face a outro que se faz sentir na bateria e no orçamento térmico de um telemóvel.

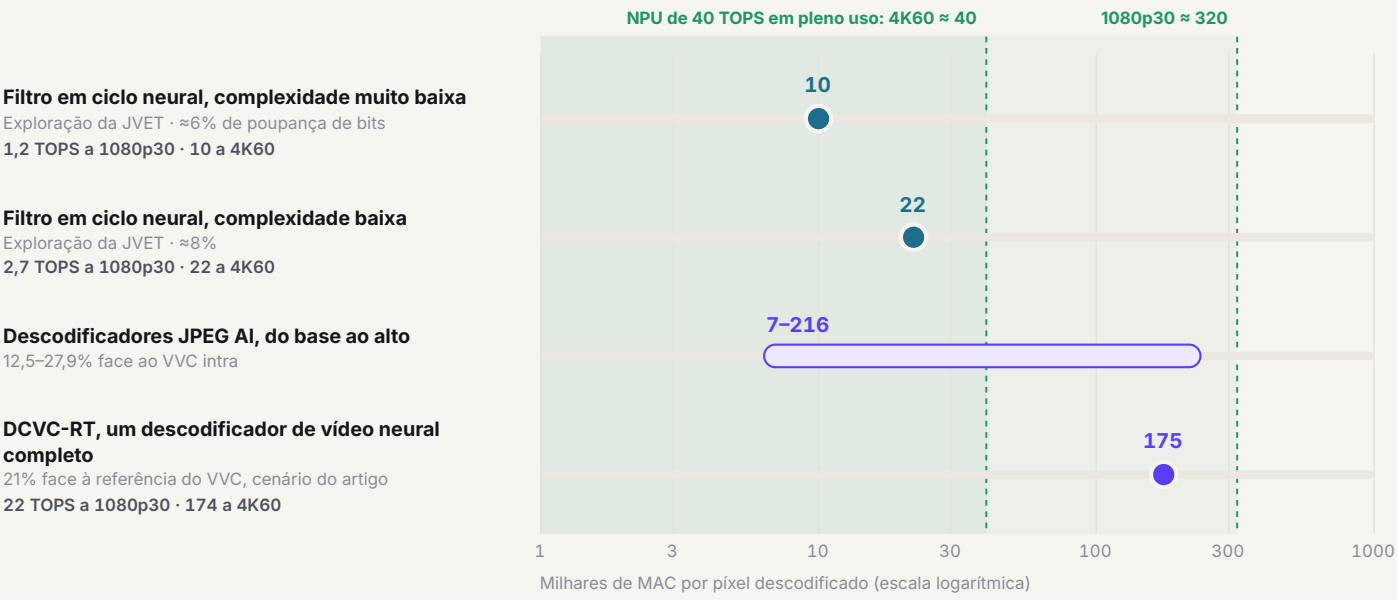
Computação por píxel: o que um descodificador neural pede a uma NPU

O codec de vídeo neural publicado mais rápido que encontrei é o DCVC-RT da Microsoft Research (Jia et al., CVPR 2025). Codifica e descodifica 1080p a 125 e 113 fotogramas por segundo numa GPU NVIDIA A100, poupando 21% de débito binário face ao codificador de referência do VVC no cenário de teste do artigo, e usa aritmética inteira para que dispositivos diferentes descodifiquem de forma consistente. O seu material suplementar indica a complexidade de cada módulo. Somando os quatro módulos que um descodificador tem de executar (descodificador, extrator de características, modelo de entropia e reconstrução), obtenho cerca de 175 mil multiplicações-acumulações por píxel descodificado estimativa.

A 1080p30, isso dá cerca de 11 biliões de multiplicações-acumulações por segundo, ou cerca de 22 TOPS a duas operações por MAC, antes de qualquer perda de utilização; 4K60 exige oito vezes mais estimativa. Como termo de comparação, a Microsoft define os seus PC com IA por uma NPU de pelo menos 40 TOPS, e a Apple atribui 38 TOPS ao Neural Engine do M4. São valores de pico, partilhados com tudo o resto que o dispositivo faz.

FIG. 07 **Uma ferramenta neural custa dezenas de milhares de MAC por píxel; um descodificador neural completo, centenas de milhares**

Multiplicações-acumulações do descodificador por píxel, em escala logarítmica, face ao que uma NPU de 40 TOPS em pleno uso conseguiria comportar a 4K60 e a 1080p30. Os ganhos referem-se a âncoras diferentes e não são diretamente comparáveis.



Fonte: JVET, através do presidente do SC 29 (abril de 2026); comité JPEG (2024); DCVC-RT, material suplementar da CVPR 2025, soma dos módulos do descodificador feita pelo autor; definição de PC Copilot+ da Microsoft. Cálculos do autor: duas operações por MAC, utilização plena.

A escada da figura 7 é a forma útil de ler esta área. As ferramentas neurais que custam dezenas de milhares de MAC por píxel, como os filtros em ciclo da JVET (cerca de 6% de poupança com 10 mil, 8% com 22 mil), são plausíveis dentro de um futuro descodificador normalizado. O JPEG AI cobre a gama, de 7 a 216 mil MAC por píxel, para 12,5–27,9% face ao VVC intra. Um descodificador de vídeo neural completo, com umas duas centenas de milhares de MAC por píxel, tem o seu lugar numa NPU, a uma resolução modesta, num circuito fechado. É exatamente aí que está o MLVC: 540p30, usando no máximo metade da NPU.

O tráfego de memória, e não os MAC, é muitas vezes o verdadeiro limite

Os autores do DCVC-RT concluíram que o que limitava a velocidade não era a aritmética, mas o «custo operacional»: a entrada e saída de memória e o número de chamadas de funções. Removeram os módulos explícitos de movimento e o processamento multiescala em grande parte por essa razão. Os projetistas do NVCA construíram um fluxo de dados especial para reduzir o tráfego de memória fora do chip associado às características de movimento e de resíduo. Um projetista de chips reconhecerá o padrão. No chip HEVC do MIT, a DRAM representa cerca de três quartos da energia do sistema (0,88 de 1,19 nJ por píxel), mesmo depois de uma cache que reduziu em 67% a largura de banda da DRAM análise. Um descodificador neural que guarda grandes mapas de características multicanal entre fotogramas agrava o problema da memória, em vez de o atenuar.

Exatidão: todos os descodificadores têm de calcular os mesmos bits

A codificação entrópica só funciona se o modelo probabilístico do descodificador coincidir exatamente com o do codificador. Num codec clássico, esse modelo é uma tabela ou um pequeno conjunto de regras inteiras. Num codec neural, é uma rede, e uma diferença de vírgula flutuante entre dois dispositivos, devida a um kernel diferente,

a uma ordem de acumulação diferente ou a uma NPU diferente, pode levar o descodificador a ler os símbolos errados e a perder o fluxo a partir desse ponto.

Todos os que lançaram ou normalizaram um codec neural tiveram de resolver isto por engenharia. O DCVC-RT converte o modelo para inteiros, para uma codificação consistente entre dispositivos. O MLVC usa uma conceção de hyperprior de escala partilhada para resistir às discrepâncias entre plataformas, o que provavelmente explica em parte porque abdica de compressão face ao DCVC-RT análise. A exploração de redes neuronais da JVET faz passar todas as redes pelo SADL, uma pequena biblioteca com inferência inteira de 16 bits. O JPEG AI define pontos de conformidade nos domínios latente, residual e de píxel. Para uma equipa de chips, é a mesma disciplina das transformadas inteiras do HEVC: a norma tem de definir a aritmética com exatidão, o que restringe a conceção da rede e exclui atalhos específicos de cada fornecedor.

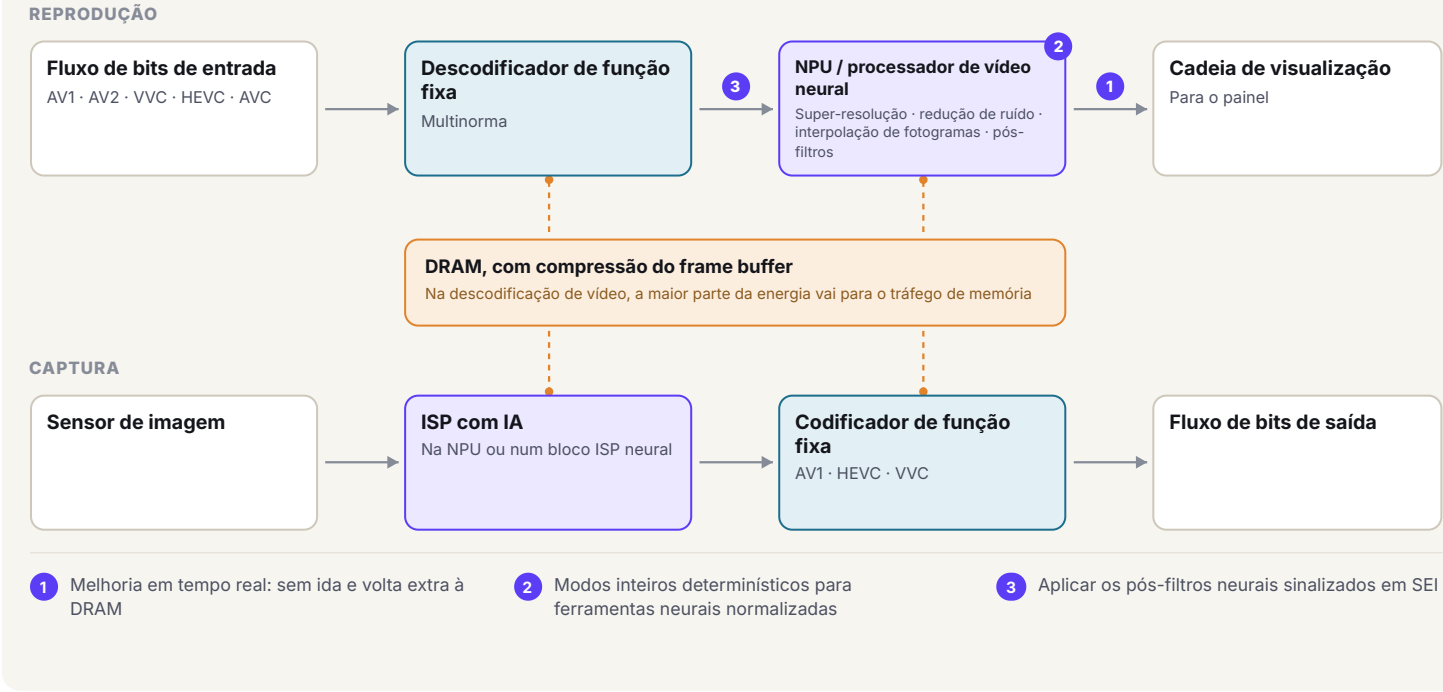
A área não é o obstáculo; a verificação e a vida útil, sim

Os 5 milhões de portas lógicas do NVCA, face a cerca de um milhão na projeção para um descodificador VVC, parecem muito, mas num SoC moderno de telemóvel ou de televisor que já inclui uma NPU de vários TOPS, uns quantos milhões de portas não são o que decide o projeto. O que custa dinheiro a sério é a verificação face a uma bateria de testes de conformidade, manter um descodificador programável correto ao longo das atualizações de software durante a vida útil do dispositivo, e a potência a 4K60 HDR. Esses custos só fazem sentido quando uma norma fixa o que tem de ser construído.

A arquitetura que ganha hoje: um codec de função fixa mais uma NPU

FIG. 08 **A divisão que chega ao mercado: hardware de função fixa para o fluxo de bits, um motor neural para os píxeis**

Um SoC multimédia de 2026. O bloco de codec descodifica todas as normas com uma fração de nanojoule por píxel; a NPU faz aquilo em que as redes são boas e que as normas não restringem. O tráfego de DRAM é o custo que conta.



Fonte: Visão de arquitetura do autor, com base em informação de produto divulgada pela Tencent (Canghai V2), Chips&Media, Allegro DVT e Samsung (2026).

O silício que vende em 2026 junta um codec multinorma de função fixa a um motor neural ao lado. O bloco de codec trata do AV1, do AV2, do VVC, do HEVC e do AVC com uma fração de nanojoule por píxel. O motor neural trata daquilo em que as redes são boas e que as normas não restringem: a redução de ruído e o ISP na captura, a super-resolução, a interpolação de fotogramas e a melhoria na visualização. Os dois partilham a compressão do frame buffer e o escalonamento da DRAM, porque o tráfego de memória é o custo que conta.

O mercado mostra a divisão. O ASIC de codec de vídeo Canghai (沧海) da Tencent, desenvolvido internamente desde 2019, tem mais de 100 000 chips de primeira geração implantados na Tencent Cloud e nos seus próprios serviços de emissão em direto, a pedido e de jogos na nuvem; a segunda geração, anunciada em maio de 2026 com os primeiros chips já em funcionamento e a entrar em produção em série para serviço no segundo semestre, suporta H.265 e H.266 numa só arquitetura, com uma compressão mais de 10% e 30% melhor do que a da antecessora, cerca do dobro da capacidade de processamento e melhoria de imagem no chip, como a super-resolução. A Chips&Media vende lado a lado IP de codec e uma NPU orientada para vídeo, a WAVE-N. A Allegro DVT vende descodificadores AV2 e acrescentou o seu primeiro IP de processamento de vídeo neural. Os processadores de televisor da Samsung juntam descodificadores normalizados a centenas de redes de ampliação.

O que eu construiria, por perfil

Equipas de SoC e de chips de TV. Manter descodificadores de função fixa para todas as normas que for preciso reproduzir, incluindo desde já o AV2, e o VVC onde a radiodifusão o exigir. Dimensionar a NPU para pós-processamento à resolução e à taxa de fotogramas do ecrã, e conceber o caminho do descodificador para a NPU e daí para o ecrã de forma que a melhoria não acrescente uma ida e volta à DRAM. Suportar o SEI de pós-filtro neural. Oferecer modos inteiros determinísticos na NPU. Tratar o H.267 como um início de projeto para 2029 ou depois, e manter alguma programabilidade para as suas ferramentas neurais.

Fornecedores de IP. Vender já descodificadores AV2 e VVC, mais as peças de que o processamento neural precisa à sua volta: aceleradores de pós-filtros, compressão do frame buffer e ferramentas de conformidade.

Plataformas de vídeo. O pré-processamento neural no codificador é o primeiro a compensar, porque corre uma vez por título e não precisa de nada dos dispositivos. Usar super-resolução no cliente onde os dispositivos o permitirem. Reservar os codecs neurais de ponta a ponta para circuitos fechados sob controlo próprio.

CAPÍTULO 09

NORMAS E PATENTES

09 Normas e patentes: onde as ferramentas neurais se tornam licenciáveis

A cadeia de normalização

Norma	Organismo	Conteúdo neural	Situação, novembro de 2026
H.267 (nome provisório; pós-VVC)	JVET (ITU-T e ISO/IEC)	Em aberto. As ferramentas clássicas (ECM) e as neurais (filtros em ciclo, predição intra, super-resolução) são ambas candidatas	Pedido de propostas em julho de 2026; respostas avaliadas em janeiro de 2027; primeiro modelo de teste no segundo semestre de 2027; versão 1 no final de 2029
VVC e HEVC, com VSEI	JVET	Um pós-filtro neural pode ser descrito em metadados SEI; aplicá-lo é opcional para o descodificador	Em vigor; mensagens alargadas nas edições de outubro de 2025
AV2	AOMedia	Ferramentas clássicas, isentas de royalties	Versão 1.0.0 a 28 de maio de 2026; IP de descodificador licenciado desde junho de 2026; emissões do YouTube na Web previstas para o início de 2027
JPEG AI	JPEG (ISO/IEC e ITU-T)	Codec de imagem aprendido de ponta a ponta, para visão humana e de máquina	Parte 1 publicada em 2025; uma segunda versão em discussão; sem suporte por predefinição encontrado em produtos
IEEE 1857.11	IEEE	Codificação de imagem por redes neuronais concebida para NPU	Publicada em dezembro de 2024
Video Coding for Machines	MPEG	Codificação para análise por máquinas e não para visualização	Projeto de norma internacional no início de 2026; versão final prevista para o final de 2026
Extensões de ML do Opus	IETF	Redundância profunda, melhoria do lado do descodificador e extensão de largura de banda, opcionais	No libopus 1.5 e 1.6; redundância profunda em normalização

Destacam-se duas coisas. Primeiro, o conteúdo neural entra nas normas pelas margens antes de chegar ao núcleo: como pós-filtros opcionais, melhoria opcional e codificação de imagem, tudo coisas que deixam inalterado o caminho principal de descodificação. Segundo, a única norma que poderia pôr ferramentas neurais no caminho de descodificação de todos os ecrãs, o H.267, está a três anos de uma primeira versão.

Quem está posicionado nas patentes

As ferramentas normalizadas tornam-se patentes licenciáveis, pelo que o processo do H.267 é também uma corrida às patentes, e as ferramentas neurais fazem parte dela. O panorama de 2025 da GreyB, com 3 916 patentes pós-VVC, coloca a ByteDance à frente, com 645 famílias de patentes e 251 contribuições para a JVET, e destaca

parcerias entre universidades e indústria, como a da Xiaomi com a Universidade de Xidian, com mais de 20 famílias. A MLex noticiou em março de 2026 que a Tencent, a Alibaba e a ByteDance estão a afirmar-se como participantes ativos em agrupamentos de patentes de codecs de vídeo. Em maio de 2026, a Sauvegarder Investment Management comprou à Alibaba 126 ativos de patentes de codecs de vídeo que abrangem o AV1, o AV2 e o Video Coding for Machines. A compra da Deep Render pela InterDigital, em outubro de 2025, acrescentou ao portefólio desta as patentes de codificação de vídeo com IA da Deep Render.

O dinheiro por dispositivo vê-se facilmente nas tabelas de tarifas dos agrupamentos. As tarifas da Access Advance para o VVC (a 1 de julho de 2026, para licenciados em situação regular e que usem a sua marca) são de 0,50 dólares por telemóvel, tablet ou portátil nos mercados da Região 1, como os EUA, a UE e o Japão, e de 0,25 dólares na Região 2; para televisores 4K, de 1,00 a 1,50 dólares, consoante o preço, na Região 1; com um teto anual por empresa de 75 milhões de dólares. O AV1 e o AV2 são isentos de royalties por parte dos membros da AOMedia. Seja qual for o conteúdo do H.267, as suas ferramentas terão preços definidos da mesma forma, e é por isso que as empresas com patentes de codificação neural estão já a contribuir para a JVET análise.

CAPÍTULO 10

QUEM GANHA DINHEIRO

10 Quem ganha dinheiro, e onde uma empresa pode apostar

O dinheiro em 2026

FIG. 09 **Quem ganha dinheiro com a codificação neural em 2026**

O dinheiro segue as normas e as plataformas. Não encontrei nenhuma empresa independente que viva de licenciar um codec de vídeo neural de ponta a ponta.

JÁ A FATURAR

Plataformas

Ficam com as poupanças

Redutor de resolução neural da Netflix; super-resolução do YouTube e da Meta; a Microsoft abriu o código do MLVC sob licença MIT

JÁ A FATURAR

Fabricantes de dispositivos e chips

Vendem diferenciação

ISP de câmaras; processadores de ampliação para televisores; chips de codec Canghai, próprios da Tencent

JÁ A FATURAR

Fornecedores de IP

Licenciam cada nova norma

Clientes para descodificadores AV2 na Chips&Media, Allegro DVT e VeriSilicon nas três semanas seguintes à especificação

JÁ A FATURAR

Fabricantes de GPU

Criam dependência da plataforma

DLSS 4 em 250+ jogos e aplicações; SDK de texturas neurais em beta

POR DISPOSITIVO

Titulares e agrupamentos de patentes

Cobram por dispositivo

VVC: 0,50 USD por telemóvel, 1,00–1,50 USD por TV 4K; ferramentas do H.267 após 2029

SAÍDA POR AQUISIÇÃO

Startups de codecs neurais

Adquiridas

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

Fonte: Informação de empresas e de agrupamentos de patentes citada no relatório; situação a 12 de novembro de 2026. Tarifas do agrupamento: Access Advance, VVC, Região 1, licenciado em situação regular e com uso da marca, a 1 de julho de 2026.

As plataformas ficam com as poupanças. A Netflix, o YouTube, a Meta, a Microsoft e a Tencent desenvolvem o processamento neural internamente e ficam com o benefício sob a forma de menos largura de banda, menos ar-

mazenamento e melhor qualidade por bit. Quando isso ajuda a adoção, oferecem o código: a Microsoft disponibilizou o MLVC sob a licença MIT.

Os fabricantes de dispositivos e de chips vendem diferenciação. Os ISP das câmaras e os processadores de imagem dos televisores são o sítio onde as redes neuronais chegam a todos os utilizadores, e são vendidos como funcionalidades: o Cognitive ISP da Qualcomm, os processadores de ampliação da Samsung.

Os fornecedores de IP vendem descodificadores para cada nova norma, além de NPU. Nas três semanas seguintes à especificação do AV2, a Chips&Media, a Allegro DVT e a VeriSilicon todas anunciaram clientes para descodificadores AV2. Os mesmos fornecedores vendem hoje blocos de processamento neural ao lado dos codecs.

Os fabricantes de GPU constroem plataformas. O DLSS e a compressão neural de texturas tornam o hardware tensorial da NVIDIA parte da forma como os jogos são feitos.

Os titulares de patentes cobram por dispositivo, através de agrupamentos para o VVC e o HEVC e, depois de 2029, para o H.267.

As startups de codecs neurais foram adquiridas. A WaveOne foi para a Apple em 2023, a iSIZE para a Sony Interactive Entertainment em 2023 e a Deep Render para a InterDigital em 2025, todas em condições não divulgadas. Em cada caso, uma plataforma ou um licenciante comprou a equipa e as patentes. Não encontrei em 2026 nenhuma empresa independente que viva de licenciar um codec de vídeo neural de ponta a ponta.

Onde uma empresa pode apostar

Aposta	Quem compra	Porquê agora	Principal risco
Pré-processamento neural no codificador (redução de resolução, redução de ruído, afinação por título)	Plataformas de streaming e de vídeo de utilizadores, radiodifusores	Não exige alterações nos dispositivos; provado na Netflix	As grandes plataformas fazem-no em casa
IP e afinação para melhoria no ecrã (super-resolução, interpolação de fotogramas, qualidade de imagem)	Fabricantes de chips para televisores, set-top boxes e telemóveis	As NPU já estão nos dispositivos; diferenciação visível	Banalização; fabricantes de chips com soluções internas
Codecs neurais em circuito fechado (videoconferência, operações remotas, ligações privadas)	Fornecedores de comunicações empresariais, industriais e de telepresença	O MLVC prova a viabilidade, e o seu código é aberto	É preciso controlar as duas pontas; fragmentação das NPU
Silício ou IP de codec mais NPU (AV2, VVC e pós-processamento)	Fabricantes de SoC	O ciclo do AV2 está a começar	Incumbentes fortes
Ferramentas de avaliação, conformidade e inferência determinística	Quem lança ferramentas neurais	A exatidão ao bit é um requisito estrito	Mercado pequeno, salvo em pacote
Contribuições normativas para o H.267	Licenciantes, grandes plataformas	As respostas são avaliadas em janeiro de 2027	Longo, caro, incerto

Para a maioria das empresas, é nas duas primeiras linhas que há receita disponível em 2027. A terceira é a mais interessante para uma empresa de produto focada. A última exige paciência e uma estratégia de licenciamento.

CAPÍTULO 11

11 ALEGAÇÕES E EVIDÊNCIA
Alegações que vai ouvir, e o que diz a evidência

Alegação	O que diz a evidência
«Os codecs neurais já batem o VVC.»	Em laboratório, sim. O DCVC-RT poupa 21% face ao codificador de referência do VVC a 1080p, no cenário do seu artigo; o JPEG AI poupa 12,5–27,9% face ao VVC intra. O descodificador custa cerca de 175 mil MAC por píxel, ou a energia de uma NPU, face a uma fração de nanojoule em hardware de função fixa.
«Os codecs de IA vão substituir o H.26x e o AV1 dentro de poucos anos.»	Não na reprodução. A primeira norma capaz de trazer ferramentas neurais aponta para o final de 2029, os chips seguem-se por volta de 2031 e o uso generalizado em meados da década de 2030 estimativa. O codec que está a chegar agora, o AV2, é clássico.
«O MLVC reduz a largura de banda em 88%.»	Face ao H.264 a 360p, no conjunto de testes da própria Microsoft. Face ao HEVC em hardware, são 75,5% a 360p e 65,4% a 540p, em chamadas em que a Microsoft controla as duas pontas.
«As texturas neurais reduzem a memória de vídeo em 85–96%.»	Em cenas de demonstração. O jogo já lançado ganhou cerca de 30% de compressão efetiva e só usou texturas neurais em recursos selecionados, devido ao custo em tempo de execução.
«O JPEG AI é uma norma ISO, por isso câmaras e navegadores vão adotá-lo.»	Não encontrei suporte por predefinição em produtos. O formato que a Web adotou em 2026, o JPEG XL no Chrome 155, é clássico.
«A descodificação neural é demasiado cara para telemóveis.»	Em parte, é verdade. O MLVC descodifica 540p30 com metade de uma NPU; 4K60 com um modelo da classe do DCVC-RT exigiria cerca de 175 TOPS de computação sustentada estimativa.
«As ligações de satélites e drones são o grande mercado do vídeo neural.»	Tecnicamente, é onde os codecs neurais melhor se adequam, mas o mercado é pequeno e cada vez mais sujeito a controlos de exportação dos dois lados do Pacífico.
«As grandes plataformas vão licenciar um codec neural de uma startup.»	Desenvolveram o seu próprio (Netflix, Microsoft, Tencent) ou compraram a startup (Apple, Sony, InterDigital).

Um teste simples desmonta a maior parte dos argumentos de venda: pergunte pela âncora, pela métrica, pela resolução e pelo custo de descodificação. Uma alegação que não consegue indicar as quatro é um resultado de investigação, não um produto.

CAPÍTULO 12

12 MARCOS A ACOMPANHAR
O que acompanhar, de 2027 a 2030

Quando	Marco	Porque é importante
Janeiro de 2027	A JVET avalia as respostas ao H.267 em Brisbane	Quantas das melhores respostas dependem de ferramentas neurais, e a que custo por píxel
Início de 2027	Primeiras emissões AV2 do YouTube na Web; descodificadores AV2 em SoC anunciados	O próximo ciclo clássico define os orçamentos dos dispositivos até cerca de 2030
Segundo semestre de 2027	Primeiro modelo de teste do H.267	Se as ferramentas neurais estão no núcleo da conceção ou só em exploração
2027	O MLVC além das chamadas ponto a ponto do Teams; outros fornecedores a adotar o código aberto	O primeiro teste de um codec neural aberto fora da empresa que o detém
2027–2028	Um primeiro produto que escreva ou leia JPEG AI por predefinição	O primeiro formato neural usado para lá das fronteiras de uma empresa
2027–2028	DirectX Linear Algebra nas versões finais dos controladores; texturas neurais em mais jogos e nos dispositivos móveis	Descodificação neural multifabricante dentro dos gráficos
Em curso	Orçamentos de NPU nos telemóveis e nos PC	Com cerca de 100 TOPS sustentados, um descodificador 1080p30 da classe do DCVC-RT ocuparia cerca de um quinto da NPU, em vez de metade análise
Final de 2029	Versão 1 do H.267	O início do ciclo de silício para as ferramentas neurais na reprodução para o grande público

CAPÍTULO 13

13 Questões em aberto

QUESTÕES EM ABERTO

Várias questões deste relatório ainda não têm resposta pública, e são as que acho mais interessantes do ponto de vista técnico.

Exatidão em escala. Até que ponto um descodificador neural se consegue aproximar de uma aritmética só com inteiros e exata ao bit, em NPU de fornecedores diferentes, antes de abdicar da maior parte da sua vantagem de compressão? A troca de seis pontos do MLVC é um dado; uma norma precisa de uma resposta geral.

Memória antes dos MAC. Quanto custa o pós-processamento neural em tráfego de DRAM por píxel quando corre em tempo real entre um descodificador em hardware e o ecrã, e quão pequeno pode ser o seu conjunto de trabalho?

O preço de uma ferramenta neural no H.267. A partir de quantos MAC e nanojoules por píxel é que um filtro em ciclo neural deixa de se pagar com os bits que poupa num descodificador 4K60?

Circuitos fechados para além das chamadas. Que outros sistemas fechados, do armazenamento de fotografias na nuvem e do streaming de jogos aos robôs e às ligações privadas de contribuição, podem adotar codecs neurais de ponta a ponta antes de o H.267 chegar?

Descodificadores que geram. Quando um descodificador gera detalhe em vez de o reconstruir, como os VAE de vídeo já fazem, como se deve medir e especificar a fidelidade?

São áreas que podemos explorar juntos nos próximos anos, através de intercâmbio técnico e de colaboração técnica: i@tanshuai.com.

Limitações deste relatório

Trata-se de investigação documental baseada em documentos públicos até 12 de novembro de 2026. Para esta edição, não fiz entrevistas primárias nem medições independentes. Vários números-chave são alegações de fornecedores ou de autores sobre os seus próprios conjuntos de testes, como as poupanças de débito binário do MLVC e as taxas de compressão neural de texturas; estão assinalados como tal. As minhas estimativas de computação para o DCVC-RT somam os valores por módulo que os seus autores publicam e ignoram as perdas de utilização, pelo que a carga real na NPU será maior. A comparação energética assenta em poucos projetos publicados, em nós de processo diferentes; os números de produção dos descodificadores VVC e AV2 são proprietários. Os calendários das normas são as metas dos próprios comités. As datas do H.267 posteriores a 2029 são estimativas minhas a partir dos historiais do AV1 e do VVC.

Sobre o autor

Tan Shuai é tecnólogo e fundador, com mais de vinte anos de trabalho em chips, hardware e software de sistema, e 20 patentes concedidas. Tem muitos anos de I&D e de gestão de chips fabless, incluindo conceção full-custom de caminhos de dados de ASIC e projetos de chips multimédia para câmaras, como aceleradores de codecs de vídeo, e é essa a lente deste relatório: energia por píxel, tráfego de memória e o que uma norma permite a um descodificador fixar em silício. O seu trabalho levou-o a passar por Silicon Valley, Nova Iorque, Chicago, Taiwan, a Área da Grande Baía (Shenzhen e Hong Kong), o Vietname e a Malásia. Hoje trabalha como investigador e programador de IA.

Discussão técnica: i@tanshuai.com · tanshuai.com

Notas e aviso legal

Este relatório é investigação documental independente baseada em documentos públicos disponíveis até 12 de novembro de 2026. É fornecido apenas a título informativo e não constitui aconselhamento de investimento, jurídico ou de engenharia. Os produtos, os calendários das normas, as tarifas dos agrupamentos de patentes e as regras de exportação mudam; confirme na fonte primária tudo aquilo em que se baseie. Os valores assinalados como alegação do fornecedor, estimativa ou análise devem ser lidos como tal. Os nomes de empresas e de produtos pertencem aos respetivos titulares e são usados apenas para identificação; nenhuma empresa reviu ou patrocinou este relatório.

© 2026 Tan Shuai. Pode partilhar este relatório sem alterações, com indicação da autoria.

Fontes

Todas consultadas entre 3 e 12 de novembro de 2026. Agrupadas por capítulo.

Tráfego, dispositivos e âmbito

- Ericsson, *Mobility Report*, junho de 2026 (vídeo, cerca de 75% do tráfego de dados móveis no final de 2025): https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2026 ; resumo: https://telecomlead.com/5g/ericsson-mobility-report-2026-global-mobile-data-traffic-to-reach-515-eb-by-2031-as-india-leads-smartphone-data-consumption-126290
- IDC, entregas mundiais de smartphones, 4.º trimestre e ano completo de 2025 (1,26 mil milhões), 13 de janeiro de 2026: https://www.idc.com/resource-center/press-releases/q425mobilephonetop5/
- TrendForce, previsão de entregas de televisores para 2025 (195,71 milhões): https://iconnect007.com/article/146406/early-shipments-by-tv-brands-push-global-1h25-volume-up-2-but-raise-concerns-for-secondhalf-peak-season/146403/ein ; Omdia, previsão de televisores para 2025 (208,7 milhões): https://iconnect007.com/article/145954/omdia-forecasts-flat-tv-shipments-in-2025-but-mini-led-poised-for-explosive-growth-at-829/145951/ein

Reprodução e streaming

- Netflix Technology Blog, «AV1 — Now Powering 30% of Netflix Streaming», dezembro de 2025: https://netflixtechblog.com/av1-now-powering-30-of-netflix-streaming-02f592242d80
- Netflix Technology Blog, «For your eyes only: improving Netflix video quality with neural networks», novembro de 2022: https://netflixtechblog.com/for-your-eyes-only-improving-netflix-video-quality-with-neural-networks-5b8d032da09c
- AOMedia, AV2 Bitstream and Decoding Process Specification v1.0.0, 28 de maio de 2026: https://av2.aomedia.org/v1.0.0/index.html ; «AV2: What the Final Specification Means for Implementers»: https://aomedia.org/blog%20posts/AV2-What-the-Final-Specification-Means-for-Implementers/
- TV Technology, «YouTube to Commence AV2 Streaming in 2027» (quota do AV1 no catálogo do YouTube): https://www.tvtechnology.com/platform/streaming/youtube-to-commence-av2-streaming-in-2027
- Google Research, «The in-loop filtering pipeline in AV2», 2026: https://research.google/pubs/the-in-loop-filtering-pipeline-in-av2/
- Fraunhofer IIS, «The Future of Broadcast in Brazil» (operação comercial do DTV+, junho de 2026): https://mpegh.com/brazil/ ; CartaCapital sobre os recetores nos televisores: https://www.cartacapital.com.br/?p=628337 ; Olhar Digital sobre o VVC na TV 3.0: https://olhardigital.com.br/2026/05/01/pro/em-breve-a-tv-aberta-vai-mudar-no-brasil/
- 9to5Google, «YouTube adds AI-upscaled 'Super Resolution' for low-quality videos», 29 de outubro de 2025: https://9to5google.com/2025/10/29/youtube-adds-ai-upscaled-super-resolution-for-low-quality-videos/
- Meta, «How Meta Deployed Super Resolution at Scale to Transform Video Quality», @Scale, 16 de outubro de 2025: https://atscaleconference.com/how-meta-deployed-super-resolution-at-scale-to-transform-video-quality/
- Samsung, destaques dos televisores 8K (NQ8 AI Gen3, 768 redes neuronais): https://www.samsung.com/us/tvs/8k-tv/highlights/

Fotografia e imagem

- Qualcomm, «Meet the Snapdragon 8 Gen 2 Super Cameras» (Cognitive ISP): https://www.qualcomm.com/news/onq/2023/07/meet-the-snapdragon-8-gen-2-super-cameras
- Chips&Media e Visionary.ai, ISP completo baseado em IA, 5 de janeiro de 2026: https://www.edge-ai-vision.com/2026/01/chipsmedia-and-visionary-ai-unveil-the-worlds-first-ai-based-full-image-signal-processor-redefining-the-future-of-image-quality/ ; entrevista da PetaPixel: https://petapixel.com/2026/01/18/a-look-at-the-worlds-first-full-ai-based-image-signal-processor
- Comité JPEG, comunicado de imprensa da 103.ª reunião (ganhos, complexidade e implementações em smartphones do JPEG AI): https://jpeg.org/items/20240507_press.html
- G. J. Sullivan, introdução ao ISO/IEC JTC 1/SC 29, abril de 2026 (situação do JPEG AI; números do ECM e do NNVC): https://www.mpeg.org/introduction-to-iso-iec-jtc-1-sc-29-coding-of-audio-picture-multimedia-and-hypermedia-information-the-world-of-jpeg-and-mpeg/
- Grupo de trabalho IEEE 1857.11 (codificação de imagem baseada em redes neuronais): https://sagroups.ieee.org/fvc
- Chrome 155 e JPEG XL, 6 de outubro de 2026: https://xenospectrum.com/en/chrome-155-jpeg-xl-rust/

Chamadas e reuniões

- Microsoft, «Satin: Microsoft's latest AI-powered audio codec for real-time communications», 17 de fevereiro de 2021: https://techcommunity.microsoft.com/t5/microsoft-teams-blog/satin-microsoft-s-latest-ai-powered-audio-codec-for-real-time/ba-p/2141382
- Microsoft, «Announcing the Open-Source Release of ML Video Codec (MLVC)», 24 de julho de 2026: https://techcommunity.microsoft.com/blog/linuxandopensourceblog/announcing-the-open-source-release-of-ml-video-codec-mlvc/4539875 ; código: https://github.com/microsoft/mlvc
- Streaming Learning Center, «Microsoft MLVC: Interoperability, Power Realities, and the Neural Compression Patent Landscape»: https://streaminglearningcenter.com/articles/microsoft_mlvc.html ; análise da iReplay: https://ireplay.tv/blog/mlvc-microsoft-neural-video-codec-open-source-streaming/

- Notas de versão do Opus 1.5: https://opus-codec.org/demo/opus-1.5 ; Opus 1.6: https://opus-codec.org/demo/opus-1.6
- Meta Engineering, «MLow: Meta's low bitrate audio codec», 13 de junho de 2024: https://engineering.fb.com/2024/06/13/web/mlow-metas-low-bitrate-audio-codec/
- Microsoft, PC Copilot+ (NPU de 40+ TOPS): https://www.microsoft.com/en-us/windows/copilot-plus-pcs ; Apple, anúncio do M4 (Neural Engine de 38 TOPS): https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip/

IA generativa

- Kyutai, ficha do modelo Mimi (12,5 Hz, 1,1 kbps): https://huggingface.co/kyutai/mimi
- Lightricks, «LTX-Video: Realtime Video Latent Diffusion» (VAE de vídeo de 1:192): https://arxiv.org/abs/2501.00103
- Comunicado de imprensa da 153.ª reunião do MPEG (Video Coding for Machines em DIS): https://www.mpeg.org/wp-content/uploads/2026/03/press_release_153.pdf

Jogos e gráficos

- NVIDIA, adoção do DLSS 4: lançamento e Computex 2025 (75 e 125+): https://www.nvidia.com/en-us/geforce/news/125-dlss-4-multi-frame-gen-games-more-announced-computex-2025/ ; Gamescom 2025 (175+): https://www.nvidia.com/en-us/geforce/news/gamescom-2025-dlss-4-full-ray-tracing-game-announcements/ ; CES 2026 (250+, DLSS 4.5): https://www.nvidia.com/en-us/geforce/news/dlss-4-rtx-path-tracing-game-announcements-ces-2026/
- Ubisoft La Forge, «Shipping Neural Texture Compression in Assassin's Creed Mirage», fevereiro de 2026: https://www.ubisoft.com/en-us/studio/laforge/news/415slmB3ZGzv8d2LhO9QT/shipping-neural-texture-compression-in-assassins-creed-mirage
- TweakTown sobre a demonstração de texturas neurais da NVIDIA na GTC 2026 (de 6,5 GB para 970 MB): https://www.tweaktown.com/news/110886/nvidias-neural-texture-compression-cuts-vram-usage-from-6-5gb-down-to-970mb/index.html
- Blogue do Microsoft DirectX, versão final do Shader Model 6.9, 26 de fevereiro de 2026: https://devblogs.microsoft.com/directx/shader-model-6-9-retail-and-more/ ; sobre os vetores cooperativos: https://devblogs.microsoft.com/directx/shader-model-6-9-and-the-future-of-cooperative-vector/ ; planos para a GDC 2026: https://windowsreport.com/microsoft-showcases-neural-rendering-and-dxr-2-0-plans-for-directx-at-gdc-2026/

Ligações especializadas

- Global Trade Alert sobre o Anúncio n.º 49 de 2026 da China: https://globaltradealert.org/intervention/161396-china-government-revises-export-controls-on-drone-related-items ; DroneXL: https://dronexl.co/2026/10/19/china-drone-export-controls-motors-props-november-18/ ; Army Recognition (limiares para plataformas terrestres não tripuladas): https://www.armyrecognition.com/news/army-news/2026/china-tightens-drone-export-controls-raising-supply-chain-risks-for-the-us-europe-and-ukraine
- DroneLife, a FCC acrescenta drones e componentes de fabrico estrangeiro à Covered List, 22 de dezembro de 2025: https://dronelife.com/2025/12/22/fcc-adds-foreign-made-drones-and-components-to-covered-list-citing-national-security-risks/

Silício

- M. Tikekar et al., «A 249-Mpixel/s HEVC Video-Decoder Chip for 4K Ultra-HD Applications», IEEE JSSC 2014: https://dspace.mit.edu/handle/1721.1/93876 ; página do chip no MIT (78 mW a 4K30): https://eecsg.mit.edu/chips/hevc-video-decoder-for-4k-ultra-hd-applications/
- S. Zhang et al., NVCA, DATE 2024: https://arxiv.org/abs/2312.10716
- Electronic Design, «Why Traditional VVC Will Outlive the Neural-Codec Hype on Silicon», 2026: https://www.electronicdesign.com/home/article/55414955/why-traditional-vvc-will-outlive-the-neural-codec-hype-on-silicon
- Z. Jia et al., «Towards Practical Real-Time Neural Video Compression» (DCVC-RT), CVPR 2025: https://openaccess.thecvf.com/content/CVPR2025/html/Jia_Towards_Practical_Real-Time_Neural_Video_Compression_CVPR_2025_paper.html ; material suplementar (complexidade por módulo): https://openaccess.thecvf.com/content/CVPR2025/supplemental/Jia_Towards_Practical_Real-Time_CVPR_2025_supplemental.pdf
- NNVC e a biblioteca de inferência inteira SADL: https://arxiv.org/abs/2309.05846
- Tencent Canghai V2, 27 de maio de 2026 (Shanghai Securities News, via 10jqka): https://stock.10jqka.com.cn/20260527/c677016020.shtml ; resumo das funcionalidades: https://www.besthub.dev/articles/tencent-s-canghai-v2-chip-enters-mass-production-what-the-new-capabilities-mean-for-video-encoding-bcfc18dbfa29
- Chips&Media, IP de descodificador AV2, 12 de maio de 2026: https://www.design-reuse.com/news/202530501-chips-media-completes-development-of-next-gen-av2-hw-decoder-ip/ ; licença AV2, 17 de junho de 2026: https://www.edge-ai-vision.com/?p=58270
- Allegro DVT, IP de descodificador AV2, 16 de junho de 2026: https://www.edge-ai-vision.com/2026/06/allegro-dvts-pulsar-decoder-ip-adds-support-for-av2-video-codec/
- VeriSilicon, VC9800D com AV2, 10 de junho de 2026: https://www.businesswirechina.com/en/news/63176.html

Normas e patentes

- Comunicado de imprensa da 155.ª reunião do MPEG (pedido conjunto de propostas para além do VVC), 5 de agosto de 2026: https://www.mpeg.org/wp-content/uploads/2026/08/ag3w254SN26474Press-Release-MPEG155.pdf
- Comunicado de imprensa da 152.ª reunião do MPEG (resultados do pedido de evidências; edições do VVC e da VSEI), 30 de outubro de 2025: https://www.mpeg.org/wp-content/uploads/2025/11/Press-Release-MPEG152.pdf ; coluna MPEG da ACM SIGMM: https://records.sigmm.org/2025/12/09/mpeg-column-152nd-mpeg-meeting/

- Streaming Media, «AI Helps InterDigital Reach Beyond VVC in Race to Develop Next-Gen Codec»: https://www.streamingmedia.com/Articles/News/Online-Video-News/AI-Helps-InterDigital-Reach-Beyond-VVC-in-Race-to-Develop-Next-Gen-Codec-171971.aspx
- GreyB, panorama de patentes pós-VVC (H.267): https://greyb.com/blog/exclusive-video-codec-landscape/ ; sobre o licenciamento do AV2 e a venda do portefólio da Alibaba: https://greyb.com/blog/av2-licensing/
- MLex, «Fragmented codec licensing exposes need for clearer rules for Chinese tech majors», 27 de março de 2026: https://www.mlex.com/mlex/articles/2458402/fragmented-codec-licensing-exposes-need-for-clearer-rules-for-chinese-tech-majors
- Access Advance, tarifas de royalties do MCBA e do VVC Advance a 1 de julho de 2026: https://accessadvance.com/wp-content/uploads/2025/07/MCBA-Program-Overview-September-2026.pdf

Dinheiro e empresas

- A InterDigital adquire a Deep Render, 30 de outubro de 2025: https://www.advanced-television.com/2025/10/30/interdigital-acquires-deep-render
- TechCrunch, «Apple acquired a startup using AI to compress videos», 27 de março de 2023: https://techcrunch.com/2023/03/27/apple-acquired-a-startup-using-ai-to-compress-videos
- Neowin, «Sony acquires AI video startup iSIZE», novembro de 2023: https://www.neowin.net/news/sony-acquires-ai-video-startup-isize-to-enhance-playstation-streaming/

COMPRESSÃO APRENDIDA · REPRODUÇÃO × SILÍCIO

Codificação neural em 2026

Chega ao mercado em torno do codec, não em vez dele

Cada número-chave deste relatório indica a fonte e a data. Quando o autor deriva um número, este é assinalado como estimativa; quando só quem vende o produto o comunica, é assinalado como alegação do fornecedor.

Tan Shuai

Tecnólogo e fundador · Investigador e programador de IA

tanshuai.com · i@tanshuai.com

EDIÇÃO PÚBLICA · Dados a 12 de novembro de 2026