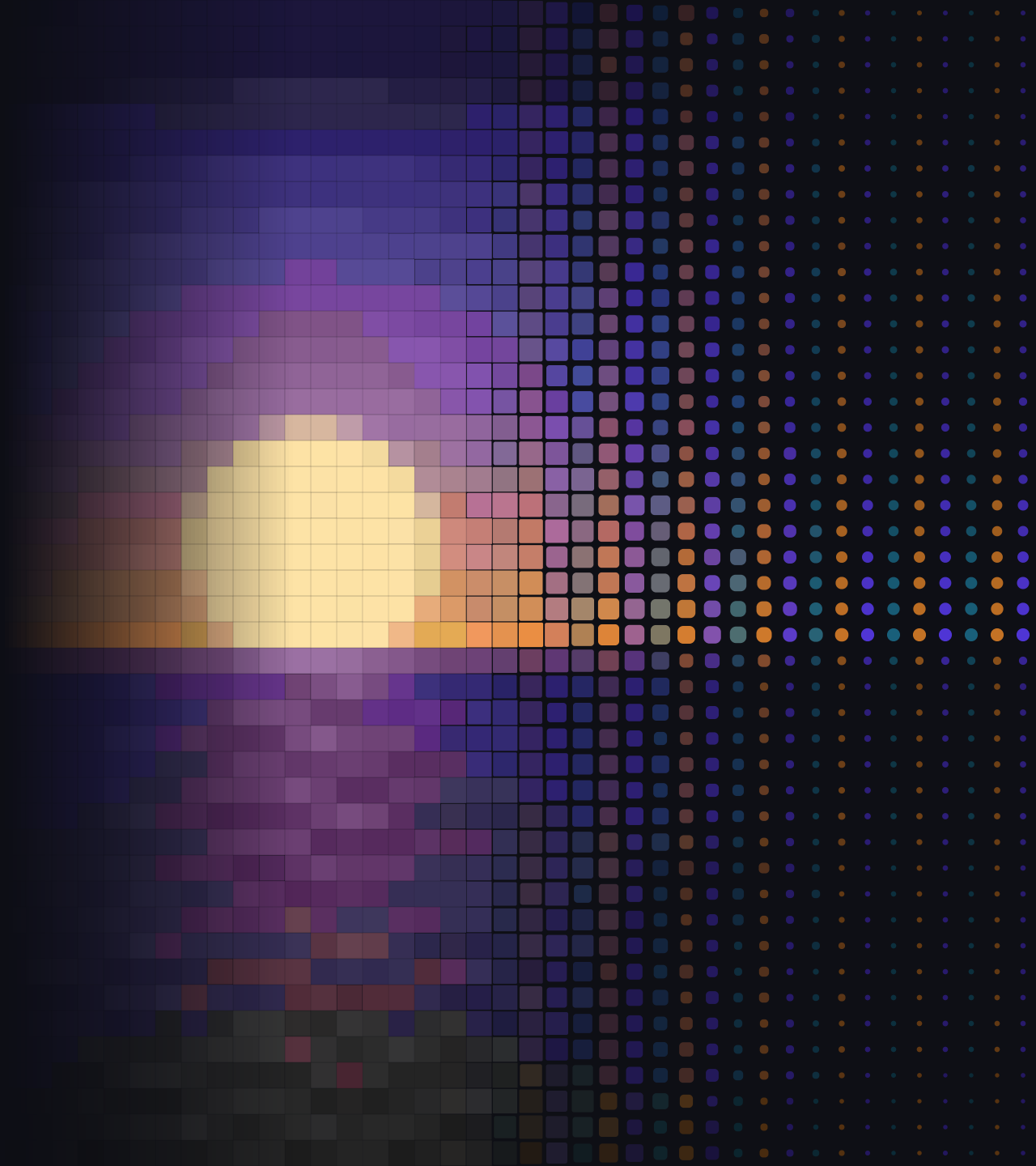
Studie · november 2026

Neural kodning 2026

Runt codecen, inte i stället för den

Uppspelning · Fotografi · Samtal · Generativ AI · Spel · Kisel

Tan Shuai
tanshuai.com · i@tanshuai.com



SLUTSATSEN

Redan etablerad, men runt codecen

01 Neurala nätverk är etablerade i båda ändar av standardcodecar

Kamera-ISP:er, nedskalning på kodarsidan och superupplösning vid uppspelning. Bitströmmen däremellan är fortfarande AV1, HEVC, AVC eller VVC, avkodad av hårdvara med fast funktion.

02 Helt neurala codecar används bara i slutna system

Samtal, tokeniserare för generativ AI och ett AAA-spel: ett företag styr båda ändarna och kan falla tillbaka på en standardcodec.

03 Neurala bitströmmar på massmarknadens skärmar väntar på H.267

Version 1 väntas i slutet av 2029; de första chipen omkring 2031 och massmarknaden i mitten av 2030-talet uppskattning.

≈75 %

av den mobila datatrafiken är video (slutet av 2025)

1,26 miljarder

smarttelefoner levererades 2025, och alla avkodar video i hårdvara med fast funktion

≈12×

energin per pixel hos en HEVC-avkodare i hårdvara, för den enda publicerade chipkonstruktionen av en neural avkodare

Tre vägar för neurala nätverk in i en codec



> ”Runt codecen” kräver ingens tillstånd. ”Inuti en standard” kräver en standard. ”I stället för codecen” kräver kontroll över båda ändarna.

Var neural kodning används 2026

	Runt en standardcodec förbehandling, förbättring, uppskalning	Helt neural codec inlärd kodare och avkodare	Inuti en standard neurala verktyg i en standardbitström eller ett API
01 **Uppspelning och streaming**	**I STOR SKALA** Netflix nedskalare; superupplösning hos YouTube och Meta; AI-uppskalare i tv-apparater	**INTE I BRUK** Ingen streamingtjänst sänder ett neuralt videoformat; tv-apparater och telefoner avkodar standardcodecar i hårdvara	**STANDARDISERAS** H.267 version 1 väntas i slutet av 2029; neurala verktyg finns bland kandidaterna
02 **Fotografi och bildbehandling**	**I STOR SKALA** AI-ISP:er i flaggskeppstelefoner; en helt neural ISP visad som IP på CES 2026	**DEMO, PILOT** JPEG AI är en ISO-standard: demonstrationer i telefoner, ingen produkt	**INTE I BRUK** Formatet som vinner mark är klassiskt: JPEG XL förvalt påslaget i Chrome 155
03 **Samtal och möten**	**I STOR SKALA** Superupplösning på mottagarsidan och neural brusdämpning	**I STOR SKALA** Neuralt ljud i tvåpartssamtal i Teams sedan 2021; MLVC-video rullas ut i Teams	**I BRUK, BEGRÄNSAT** Opus 1.5 lägger till valfria, bakåtkompatibla ML-verktyg
04 **Generativ AI**	—	**I STOR SKALA** Varje video- och röstmodell kör en inlärd tokeniserare, t.ex. en video-VAE med 1:192	—
05 **Spel och grafik**	**I STOR SKALA** Neural uppskalning och bildrutegenerering: enbart DLSS 4 i över 250 spel och appar	**I BRUK, BEGRÄNSAT** Neurala texturer i ett utgivet AAA-spel; Nvidias SDK är i betaversion	**STANDARDISERAS** DirectX får matrisoperationer för shaders (förhandsversion)
06 **Specialiserade länkar: satellit, robotar**	**DEMO, PILOT** Återställning av flöden med låg bithastighet på serversidan	**DEMO, PILOT** Förlusttåliga neurala codecar i forskning och piloter	—

Raderna är ordnade efter hur många människor varje område når.

Neuralt i båda ändar, standard i mitten

- **Vid kodaren**

  Netflix neurala nedskalare har körts sedan 2022 utan någon ändring i någon enhet; omkring 77 % av tittarna föredrog den i Netflix tester leverantörsuppgift.

- **Vid skärmen**

  YouTube Super Resolution på tv-apparater sedan oktober 2025; Meta kör superupplösning på sina servrar och i telefoner.

- **I tv:n**

  Samsungs processor NQ8 AI Gen3 marknadsförs med 768 neurala nätverk för uppskalning.

BITSTRÖMMEN FÖRBLIR STANDARD

30 %

av tittandet på Netflix i AV1 (december 2025)

>90 %

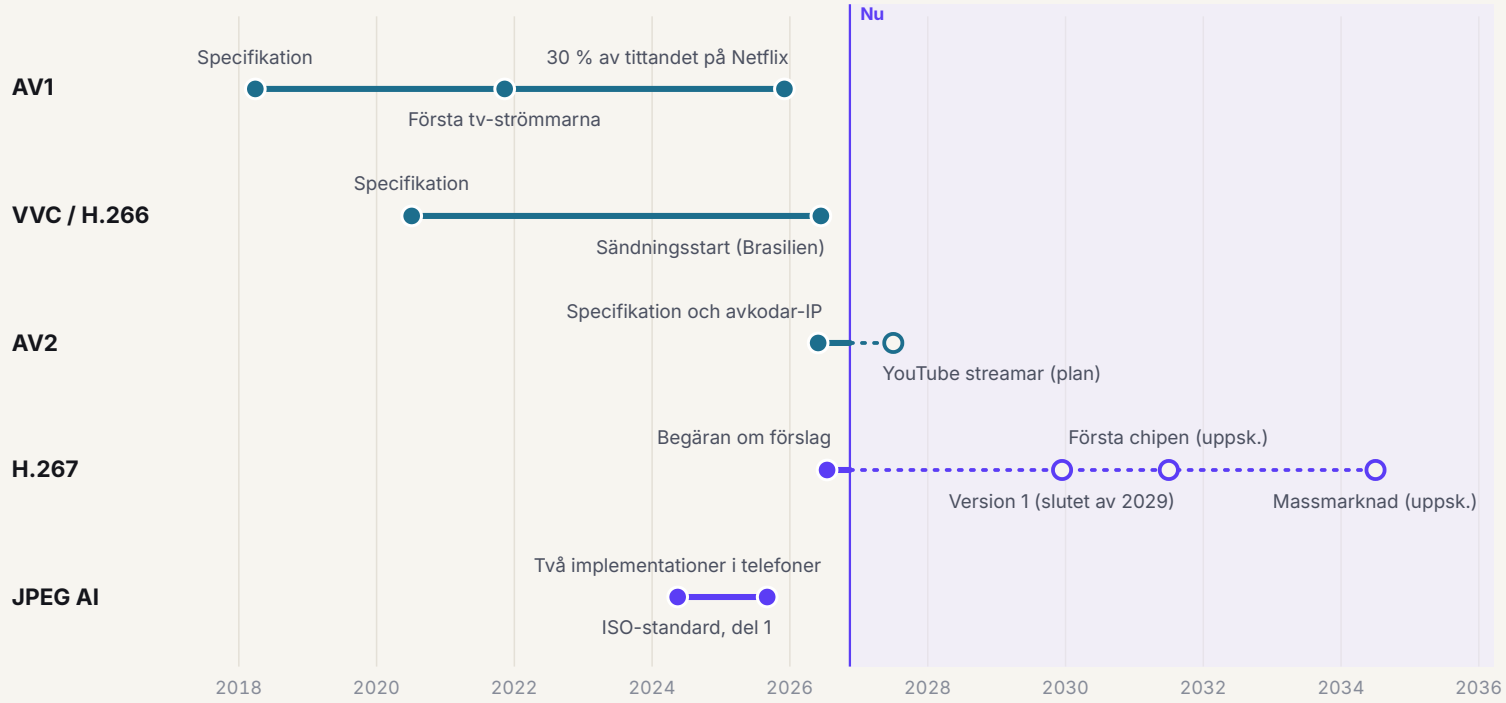
av YouTubes katalog, viktat efter visningstid, finns i AV1 (mars 2026)

3 veckor

från AV2 version 1.0.0 (28 maj 2026) till kunder för avkodar-IP hos tre leverantörer

TIDSPLAN

Från färdig standard till massmarknadens skärmar går större delen av ett decennium



AV1 behövde nästan åtta år från specifikationen (2018) till 30 % av tittandet på Netflix (december 2025). Ofyllda markeringar är mål eller mina uppskattningar.

H.267 avgör om neurala verktyg når varje skärm

Oktober 2025
Begäran om belägg: fem svar från 16 företag och labb; neurala vinster tydligast på 8K HDR

Juli 2026
Begäran om förslag utfärdad i Genève

Januari 2027
Svaren utvärderas i Brisbane

Andra halvåret 2027
Första testmodellen

Slutet av 2029
Version 1 av standarden

≈28 %
sparat mot VVC av det klassiska ECM-spåret

6–14 %
sparat av neurala loopfilter vid 10 till över 22 tusen MAC per pixel

SEI
VVC-strömmar kan redan signalera ett neuralt efterfilter som avkodare kan tillämpa

Min tolkning: en klassisk kärna med några neurala verktyg där de sparade bitarna motiverar avkodarkostnaden analys **.**

Neuralt i kameran, klassiskt på webben

- **Kamera**

  Qualcomms Cognitive ISP har kört segmentering i realtid i telefoner sedan 2023; en fullständig AI-ISP som körs på en NPU visades på CES 2026.

- **JPEG AI**

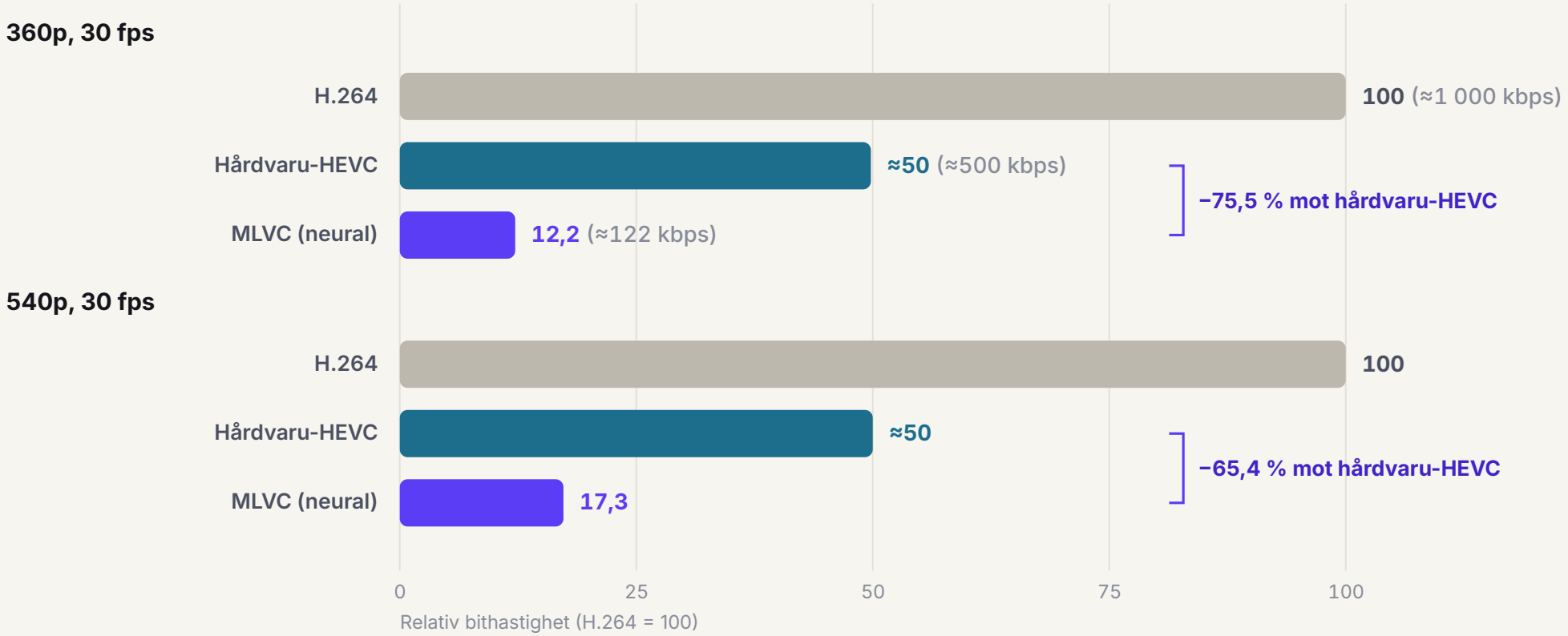
  Publicerad 2025: 12,5–27,9 % bättre än VVC intra vid 7 till 216 tusen MAC per pixel. Jag hittade ingen produkt som använder den som förval.

- **Webben**

  Chrome 155 slog på JPEG XL som förval den 6 oktober 2026, med en minnessäker avkodare skriven i Rust.

> **Format som måste gå att öppna överallt förblir klassiska; neural kodning vinner inom ett enskilt företags kedja.**

Det är i samtal som helt neurala codecar används



Satin har körts i alla tvåpartssamtal i Teams och Skype sedan 2021. MLVC, öppen källkod under MIT sedan 24 juli 2026, rullas ut i peer-to-peer-samtal i Teams med upp till 540p30 och högst halva NPU:n **leverantörsuppgift**.

Alla plattformar valde inte neuralt, och skälen är lärorika

INUTI EN STANDARD, VALFRITT

Opus 1.5 och 1.6

Neural förbättring, maskering av paketförluster, djup redundans och bandbreddsutvidgning, allt avstängt som förval och helt kompatibelt med standarden från 2012.

KLASSISKT, MED AVSIKT

Meta MLow (2024)

Mer än 20 % av Metas samtal sker på ARMv7-telefoner, och tiotals miljoner WhatsApp-samtal om dagen på telefoner som är över tio år gamla. MLow ger ungefär dubbelt så hög kvalitet som Opus vid 6 kbps med 10 % mindre beräkning.

Den installerade basens komplexitetsbudget avgör. Neurala codecar används först där enheterna är nya och en leverantör styr båda ändarna.

GENERATIV AI

Den största utrullningen som ingen kallar en codec

- **Inlärda codecar överallt**

  Varje röst- och videomodell läser och skriver en inlärd codecs tokens eller latenta värden.

- **Inga distributionsformat**

  Varje modell har sitt eget latenta rum: ingen bithastighetsstyrning, inget konformitetstest, ingen andra avkodare.

- **En kostnadspost**

  För den som erbjuder generativ video sänker en mer komprimerande latent representation genereringskostnaden mer än proportionellt analys.

TVÅ EXEMPEL

12,5 Hz · 1,1 kbps

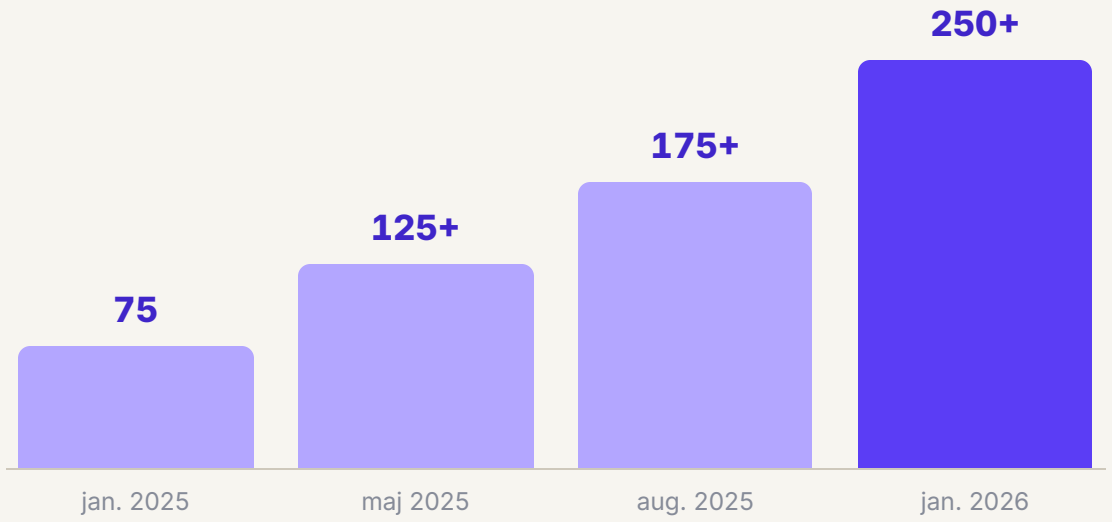
Kyutais taltokens i Mimi, lägre än någon telefonicodec

1:192

LTX-Videos video-VAE: en latent token för 32×32 pixlar över 8 bildrutor

Neural rekonstruktion är etablerad; neurala tillgångar är på väg in

DLSS 4 Multi Frame Generation: spel och appar



Neural texturkomprimering

Nvidias demoscen, GTC 2026
Texturminne 6,5 GB → 0,97 GB

85 % mindre minne

Assassin's Creed Mirage, utgivet
Bara utvalda tillgångar; mot blockformat

≈30 % bättre komprimering

Olika mått: inte direkt jämförbara

DirectX Linear Algebra blev en offentlig förhandsversion i april 2026: en enda kodväg för neural avkodning i shaders på alla tillverkares GPU:er.

Satelliter, drönare och robotar: bäst lämpade, liten marknad

- **Varför det passar**

  En ägare styr båda ändarna, bandbredden är knapp och innehållet är tillräckligt avgränsat för att träna på.

- **Kina**

  Tillkännagivande nr 49 från 2026, i kraft från den 18 november, reviderar exportkontrollen för drönare; en post har semantisk bildkomprimering som tröskelvärde.

- **USA**

  FCC förde upp utlandstillverkade drönare och deras kritiska komponenter på sin Covered List den 22 december 2025.

> **En sund nisch för en specialist som klarar exportkontrollen, inte en massmarknad.**

En storleksordning mer energi per pixel

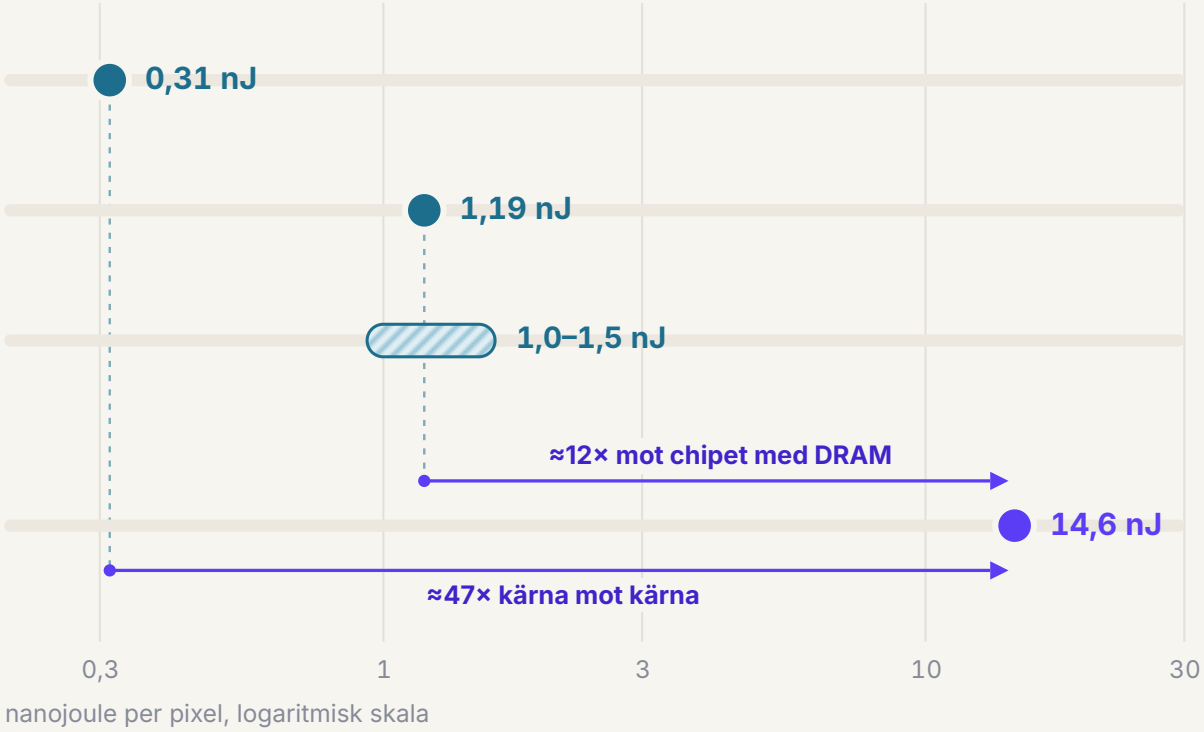
HEVC-avkodarchip, kärna
40 nm, 4K30, uppmätt

HEVC-avkodarchip, med DRAM
40 nm, 4K30, uppmätt kärna plus simulerat DRAM

VVC-avkodare
28 nm, 4K60, prognos

Neural videoavkodare (NVCA)
28 nm, 1080p25, endast syntes; vikter och DRAM ej medräknade

Vid 1080p30: omkring 74 mW mot 908 mW.



Vid 1080p30: omkring 74 mW för HEVC-chipet med DRAM mot ungefär 900 mW för NVCA. En äldre nod för HEVC-chipet, vikter och DRAM ej medräknade för NVCA: jämförelsen gynnar den neurala konstruktionen.

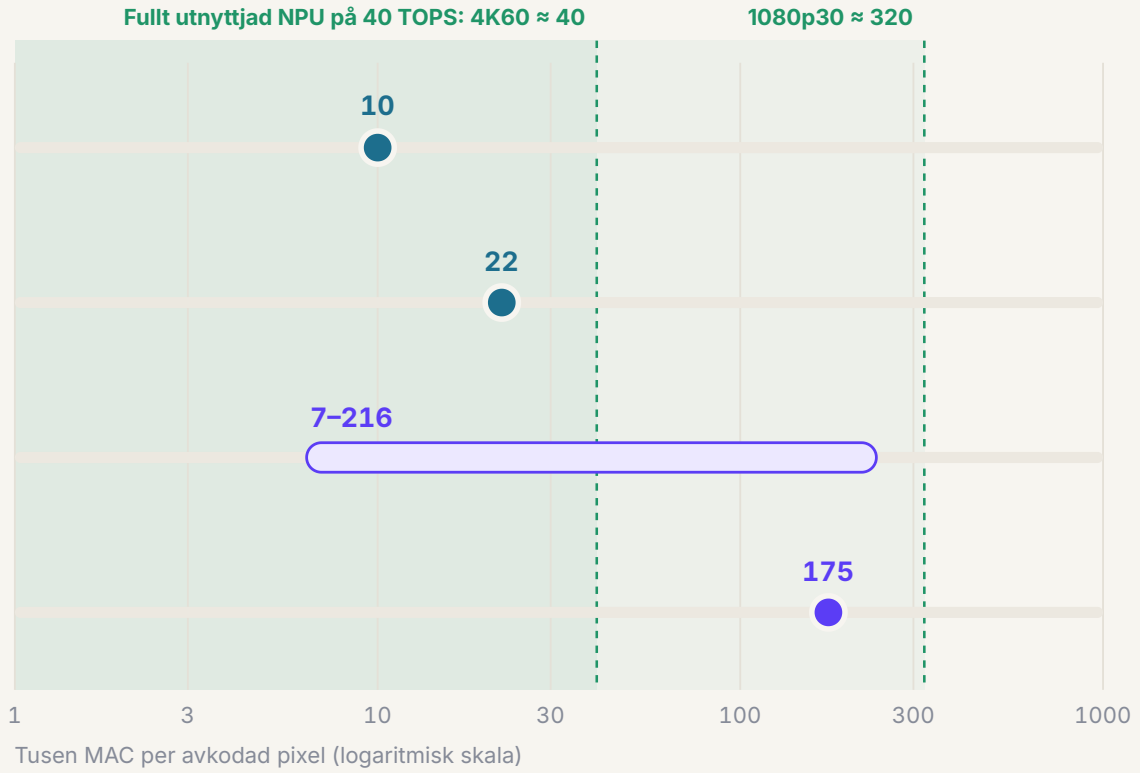
Vad en neural avkodare kräver av en NPU

Neuralt loopfilter, mycket låg komplexitet
JVET-utforskning · ≈6 % bitbesparing
1,2 TOPS vid 1080p30 · 10 vid 4K60

Neuralt loopfilter, låg komplexitet
JVET-utforskning · ≈8 %
2,7 TOPS vid 1080p30 · 22 vid 4K60

JPEG AI-avkodare, från bas till hög
12,5–27,9 % mot VVC intra

DCVC-RT, en hel neural videoavkodare
21 % mot VVC:s referens, artikelns testvillkor
22 TOPS vid 1080p30 · 174 vid 4K60



DCVC-RT behöver omkring 175 tusen MAC per pixel: omkring 22 TOPS vid 1080p30 och 174 vid 4K60, vid fullt utnyttjande **uppskattning**. En fullt utnyttjad NPU på 40 TOPS räcker till omkring 40 tusen vid 4K60.

Minnestrafik och bitexakthet är de hårda gränserna

DRAM DOMINERAR

Minne före MAC

I MIT:s HEVC-chip står DRAM för omkring tre fjärdedelar av systemets energi: 0,88 av 1,19 nJ per pixel. Neurala avkodare bär stora särdragskartor mellan bildrutor, och författarna bakom DCVC-RT fann att minnes-I/O, inte aritmetik, begränsade hastigheten.

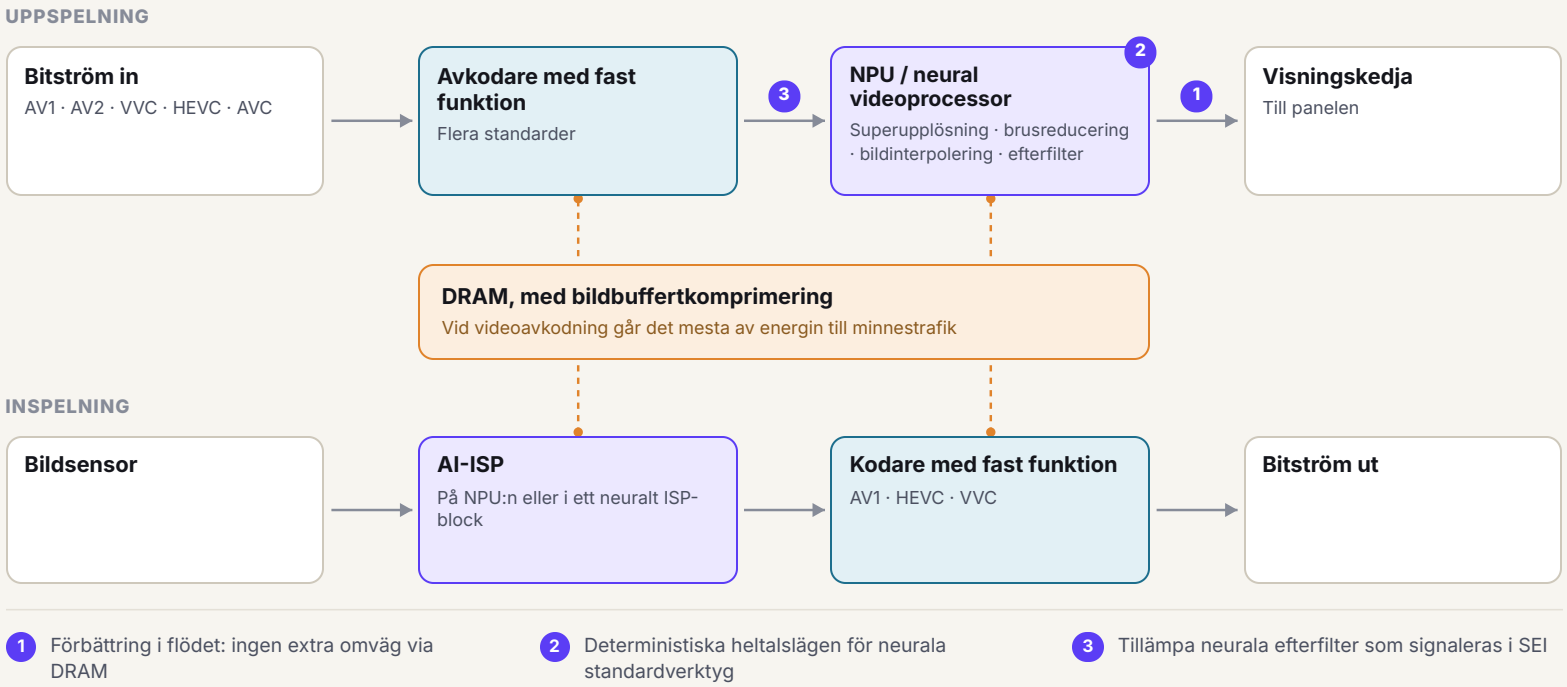
EXAKTHET

Varje avkodare måste räkna fram samma bitar

Entropiavkodningen misslyckas om två enheter beräknar en sannolikhet olika. Lösningarna: heltalsmodeller (DCVC-RT), en hyperprior med delad skala (MLVC), inferens med 16-bitars heltal (JVET:s SADL), konformitetspunkter (JPEG AI).

För ett chipteam är det HEVC:s disciplin med heltalstransformer igen: standarden måste definiera aritmetiken exakt.

Uppdelningen i dagens produkter: en codec med fast funktion plus en NPU



> Tencents Canghai: över 100 000 chip av första generationen i drift; andra generationen lägger till H.266 och superupplösning på chipet. Chips&Media och Allegro DVT säljer codec-IP och IP för neural video sida vid sida.

Vad jag skulle bygga, efter roll

SoC- och tv-chipteam

- Avkodare med fast funktion för varje standard du måste spela upp, AV2 nu
- En NPU dimensionerad för förbättring i skärmupplösning, utan extra omväg via DRAM
- SEI för neurala efterfilter och deterministiska heltalslägen
- H.267 som konstruktionsstart 2029 eller senare

IP-leverantörer

- AV2- och VVC-avkodare nu
- Acceleratorer för efterfilter och bildbuffertkomprimering
- Konformitetsverktyg för neural bearbetning

Videoplattformar

- Neural förbehandling vid kodaren lönar sig först
- Superupplösning på klientsidan där enheterna tillåter
- Helt neurala codecar bara i slutna system som du styr

Där neurala verktyg blir licensierbara

Standard	Neuralt innehåll	Status, november 2026
H.267 (arbetsnamn)	Öppet: klassiska (ECM) och neurala verktyg är båda kandidater	Svaren utvärderas i januari 2027; version 1 i slutet av 2029
VVC och HEVC, med VSEI	Neuralt efterfilter beskrivet i SEI; valfritt för avkodare	Gäller; utökat i oktober 2025
AV2	Klassiskt, royaltyfritt	Version 1.0.0 den 28 maj 2026; YouTube-strömmar på webben planerade till början av 2027
JPEG AI	Helt inlärd bildcodec	Del 1 publicerad 2025; inget förvalt stöd i produkter hittat
IEEE 1857.11	Neural bildkodning utformad för NPU:er	Publicerad i december 2024
Video Coding for Machines	Kodning för maskinanalys i stället för tittande	Utkast till internationell standard i början av 2026
Opus ML-tillägg	Valfria och bakåtkompatibla	I libopus 1.5 och 1.6

645

patentfamiljer bortom VVC som ByteDance innehar (GreyB, 2025)

0,50 · 1,00–1,50 dollar

VVC-poolens avgift per telefon och per 4K-tv, Region 1

0 dollar

i royalty för AV1 och AV2 från AOMedias medlemmar

VEM SOM FÅR BETALT

Pengarna följer standarderna och plattformarna

TJÄNAR NU

Plattformar

Behåller besparingarna

Netflix neurala nedskalare; superupplösning hos YouTube och Meta; Microsoft släppte MLVC som öppen källkod under MIT

TJÄNAR NU

Enhets- och chiptillverkare

Säljer differentiering

Kamera-ISP:er; uppskalningsprocessorer för tv; Tencents egenutvecklade codec-chip Canghai

TJÄNAR NU

IP-leverantörer

Licensierar varje ny standard

Kunder för AV2-avkodare hos Chips&Media, Allegro DVT och VeriSilicon inom tre veckor efter specifikationen

TJÄNAR NU

GPU-leverantörer

Bygger plattformsinlåsning

DLSS 4 i över 250 spel och appar; SDK för neurala texturer i betaversion

PER ENHET

Patentinnehavare och pooler

Tar betalt per enhet

VVC: 0,50 dollar per telefon, 1,00–1,50 dollar per 4K-tv; H.267-verktyg efter 2029

EXIT GENOM FÖRVÄRV

Startupbolag inom neurala codecar

Uppköpta

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

> De tre mest kända startupbolagen inom neurala codecar, WaveOne, iSIZE och Deep Render, köptes av Apple, Sony och InterDigital.

VAR ETT FÖRETAG KAN SATSA

Sex satsningar, och risken i var och en

Satsning	Vem köper	Varför nu	Huvudrisk
Neural förbehandling vid kodaren	Streaming- och användarvideoplattformar, tv-bolag	Ingen ändring i enheterna; beprövat hos Netflix	Stora plattformar bygger det själva
IP och finjustering för förbättring på skärmsidan	Chiptillverkare för tv, digitalboxar och telefoner	NPU:erna finns redan i enheterna	Kommoditisering; egna team
Neurala codecar i slutna system	Leverantörer inom företagskommunikation, industri och telenärvaro	MLVC visar att det fungerar, och koden är öppen	Du måste styra båda ändarna; splittrat NPU-landskap
Kisel eller IP för codec plus NPU	SoC-tillverkare	AV2-cykeln börjar	Starka etablerade aktörer
Utvärdering, konformitet och deterministisk inferens	Alla som levererar neurala verktyg	Bitexakthet är ett hårt krav	Liten marknad utan paketering
Standardbidrag till H.267	Licensgivare, stora plattformar	Svaren utvärderas i januari 2027	Långt, dyrt, osäkert

Var skeptisk tills påståendet anger fyra saker

Påstående	Vad beläggen säger
"Neurala codecar slår redan VVC."	I labbet. DCVC-RT sparar 21 % mot VVC:s referenskodare; dess avkodare kostar omkring 175 tusen MAC per pixel.
"AI-codecar kommer att ersätta H.26x och AV1 inom några år."	Inte i uppspelning: H.267 version 1 i slutet av 2029, chip omkring 2031 uppskattning. Codecen som kommer nu, AV2, är klassisk.
"MLVC minskar bandbredden med 88 %."	Mot H.264 vid 360p. Mot hårdvaru-HEVC: 75,5 % vid 360p och 65,4 % vid 540p, på Microsofts egna testdata.
"Neurala texturer minskar videominnet med 85–96 %."	I demoscener. Det utgivna spelet vann omkring 30 %, och bara för utvalda tillgångar.
"JPEG AI är en ISO-standard, så kameror och webbläsare kommer att införa den."	Inget förvalt stöd hittat i produkter. Webben valde JPEG XL, ett klassiskt format.

Fråga efter jämförelsebasen, måttet, upplösningen och avkodningskostnaden. Ett påstående som inte kan ange alla fyra är ett forskningsresultat, inte en produkt.

ATT BEVAKA

Att bevaka, och öppna frågor

HÅLLPUNKTER

Januari 2027	H.267-svaren utvärderas: hur många som bygger på neurala verktyg, till vilken kostnad per pixel
Början av 2027	YouTubes första AV2-strömmar på webben; AV2-avkodare i annonserade SoC:er
Andra halvåret 2027	H.267:s första testmodell
2027	MLVC bortom peer-to-peer-samtal i Teams
2027–2028	En första produkt som använder JPEG AI som förval
Slutet av 2029	H.267 version 1 startar kiselcykeln

ÖPPNA FRÅGOR

01 Exakthet i stor skala: bitexakt heltalsavkodning på olika leverantörers NPU:er, till vilken förlust av komprimering?

02 Minne före MAC: DRAM-trafiken per pixel för neural efterbearbetning

03 Kostnaden per pixel där ett neuralt verktyg slutar löna sig i H.267

04 Vilka slutna system bortom samtal som först inför helt neurala codecar

05 Hur man mäter återgivningstrohet när en avkodare genererar detaljer

Det här är områden vi kan utforska tillsammans, genom tekniskt utbyte och tekniskt samarbete.

i@tanshuai.com · tanshuai.com