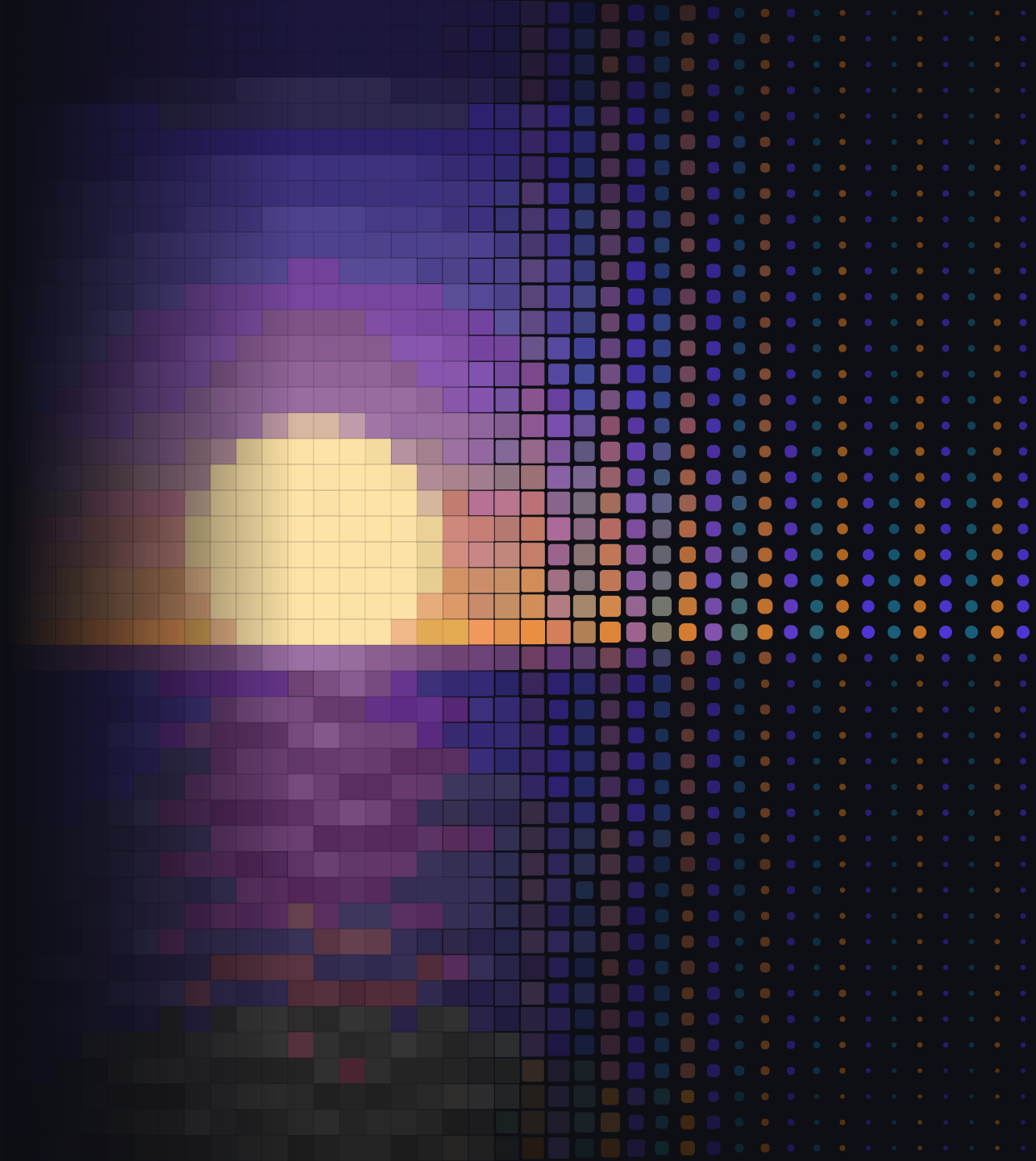
Forskningsnotat · november 2026

Neural kodning i 2026

Omkring codec'en, ikke i stedet for den

Afspilning · Fotografi · Opkald · Generativ AI · Spil · Silicium

Tan Shuai
tanshuai.com · i@tanshuai.com



KONKLUSIONEN

Allerede udbredt, men omkring codec'en

01 Neurale netværk er udbredte i begge ender af standardcodecs

Kamera-ISP'er, nedskalering ved kodningen og superopløsning ved afspilning. Bitstrømmen imellem er stadig AV1, HEVC, AVC eller VVC, dekodet af dedikeret hardware.

02 Ende-til-ende-neurale codecs bruges kun i lukkede kredsløb

Opkald, tokenizere i generativ AI og ét AAA-spil: én virksomhed styrer begge ender og kan falde tilbage på en standardcodec.

03 Neurale bitstrømme på massemarkedets skærme venter på H.267

Version 1 ventes ved udgangen af 2029; de første chips omkring 2031 og bred udbredelse i midten af 2030'erne estimat.

≈75 %

af den mobile datatrafik er video (udgangen af 2025)

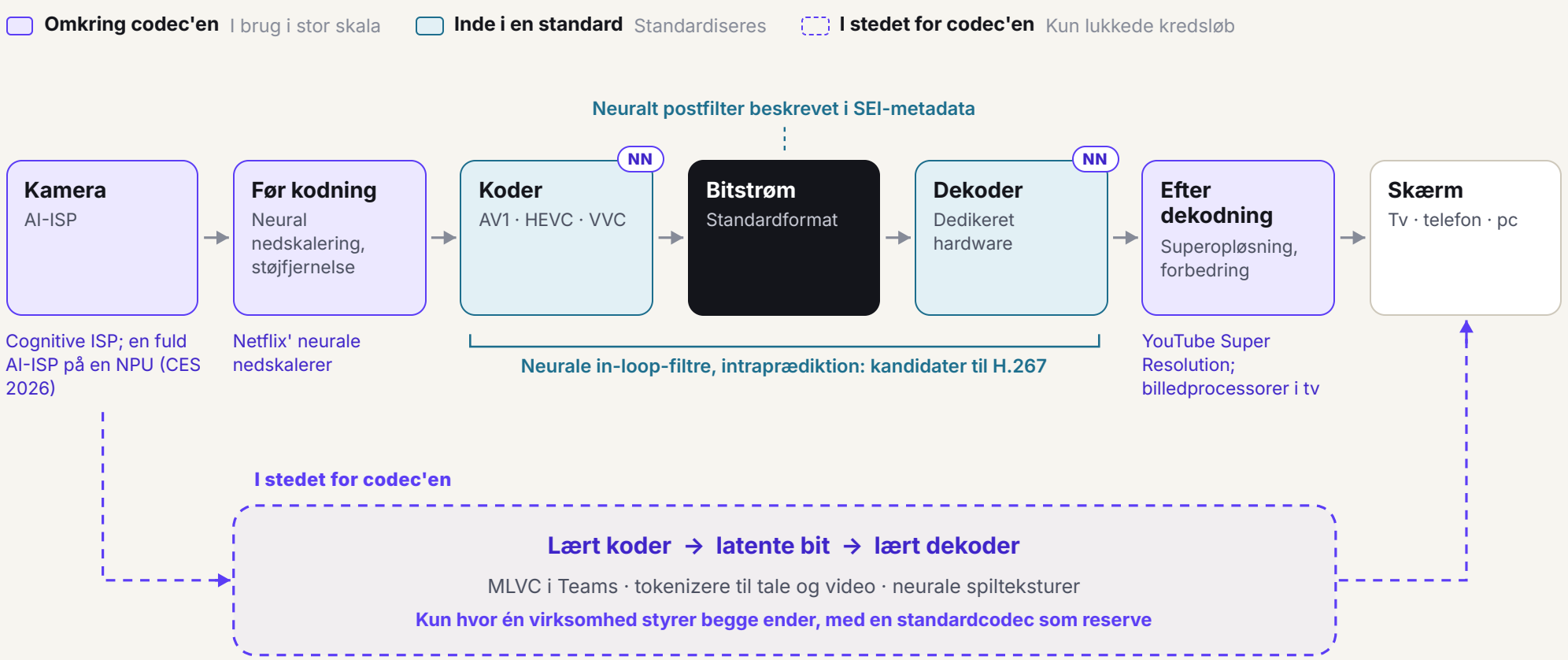
1,26 mia.

smartphones leveret i 2025, der alle dekoder video i dedikeret hardware

≈12×

energien pr. pixel i en HEVC-hardwaredekoder, for det eneste publicerede chipdesign til en neural dekoder

Tre veje ind i en codec for et neuralt netværk



Omkring codec'en: kræver ingens tilladelse. Inde i en standard: kræver en standard. I stedet for codec'en: kræver kontrol over begge ender.

Hvor neural kodning bruges i 2026

	Omkring en standardcodec forbehandling, forbedring, opskalering	Ende-til-ende-neural codec lært koder og dekoder	Inde i en standard neurale værktøjer i en standardbitstrøm eller et API
01 **Afspilning og streaming**	I STOR SKALA Netflix' nedskalerer; superopløsning hos YouTube og Meta; AI-opskalerere i tv	IKKE I BRUG Ingen streamingtjeneste sender et neuralt videoformat; tv og telefoner dekoder standardcodecs i hardware	STANDARDISERES H.267 version 1 ventes ved udgangen af 2029; neurale værktøjer er blandt kandidaterne
02 **Fotografi og billedbehandling**	I STOR SKALA AI-ISP'er i flagskibstelefoner; en fuld neural ISP vist som IP på CES 2026	DEMOER, PILOTFORSØG JPEG AI er en ISO-standard: demoer på telefoner, intet produkt	IKKE I BRUG Formatet, der vinder frem, er klassisk: JPEG XL slået til som standard i Chrome 155
03 **Opkald og møder**	I STOR SKALA Superopløsning på modtagersiden og neural støjdæmpning	I STOR SKALA Neural lyd i Teams-opkald mellem to parter siden 2021; MLVC-video rulles ud i Teams	BEGRÆNSET UDRULNING Opus 1.5 tilføjer valgfrie, bagudkompatible ML-værktøjer
04 **Generativ AI**	—	I STOR SKALA Alle video- og stemmemodeller kører en lært tokenizer, f.eks. en video-VAE på 1:192	—
05 **Spil og grafik**	I STOR SKALA Neural opskalering og frame generation: DLSS 4 alene i 250+ spil og apps	BEGRÆNSET UDRULNING Neurale teksturer udgivet i ét AAA-spil; NVIDIA's SDK er i beta	STANDARDISERES DirectX får matrixoperationer til shadere (preview)
06 **Specialforbindelser: satellitter, robotter**	DEMOER, PILOTFORSØG Gendannelse af feeds med lav bitrate på serversiden	DEMOER, PILOTFORSØG Tabsrobuste neurale codecs i forskning og pilotforsøg	—

Rækkerne er ordnet efter, hvor mange mennesker hvert område når.

Neural i begge ender, standard i midten

- **Ved koderen**

  Netflix' neurale nedskalerer har kørt siden 2022 uden ændringer på nogen enhed; omkring 77 % af seerne foretrak den i Netflix' test leverandørpåstand.

- **Ved skærmen**

  YouTube Super Resolution på tv siden oktober 2025; Meta kører superopløsning på sine servere og på telefoner.

- **I tv'et**

  Samsungs NQ8 AI Gen3-processor reklamerer med 768 neurale netværk til opskalering.

BITSTRØMMEN FORBLIVER STANDARD

30 %

af Netflix' seertid i AV1 (december 2025)

Over 90 %

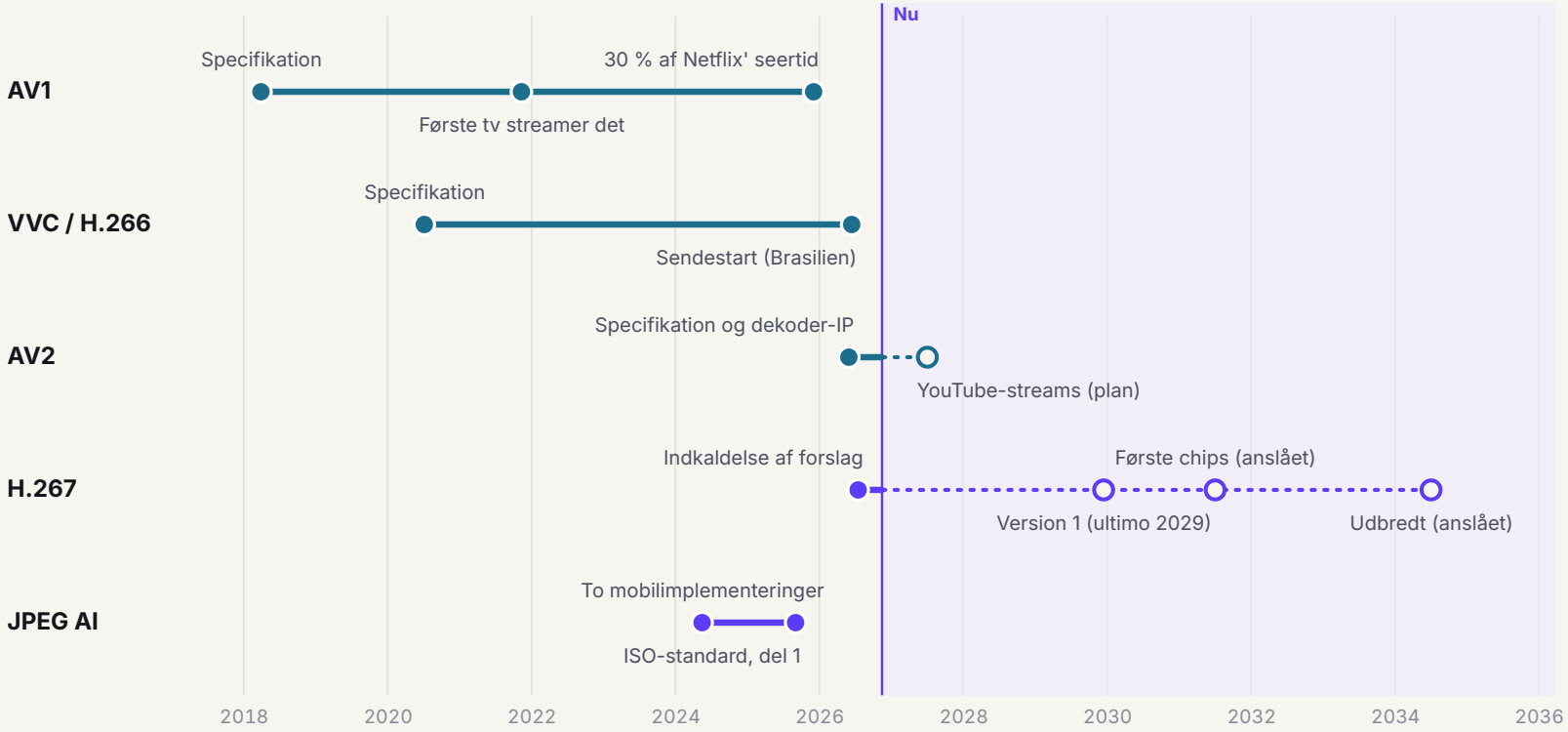
af YouTubes katalog, målt på seertid, findes i AV1 (marts 2026)

3 uger

fra AV2 version 1.0.0 (28. maj 2026) til kunder på dekoder-IP hos tre leverandører

TIDSHORISONT

Fra færdig standard til massemarkedets skærme går det meste af et årti



AV1 tog næsten otte år fra specifikationen (2018) til 30 % af Netflix' seertid (december 2025). Åbne markeringer er mål eller mine estimater.

H.267 afgør, om neurale værktøjer når ud på alle skærme

Oktober 2025
Indkaldelse af dokumentation: fem svar fra 16 virksomheder og laboratorier; neurale gevinster tydeligst på 8K HDR

Juli 2026
Indkaldelse af forslag udsendt i Genève

Januar 2027
Svarene evalueres i Brisbane

Andet halvår 2027
Første testmodel

Ultimo 2029
Version 1 af standarden

≈28 %
sparet mod VVC af det klassiske ECM-spor

6–14 %
sparet af neurale in-loop-filtre ved 10 til 22+ tusind MAC'er pr. pixel

SEI
VVC-strømme kan allerede signalere et neuralt postfilter, som dekoderne kan anvende

Min læsning: en klassisk kerne med nogle få neurale værktøjer, hvor de sparede bit retfærdiggør dekoderomkostningen analyse.

Neural inde i kameraet, klassisk på nettet

- **Kamera**

  Qualcomms Cognitive ISP har kørt semantisk segmentering i realtid i telefoner siden 2023; en fuld AI-ISP, der kører på en NPU, blev vist på CES 2026.

- **JPEG AI**

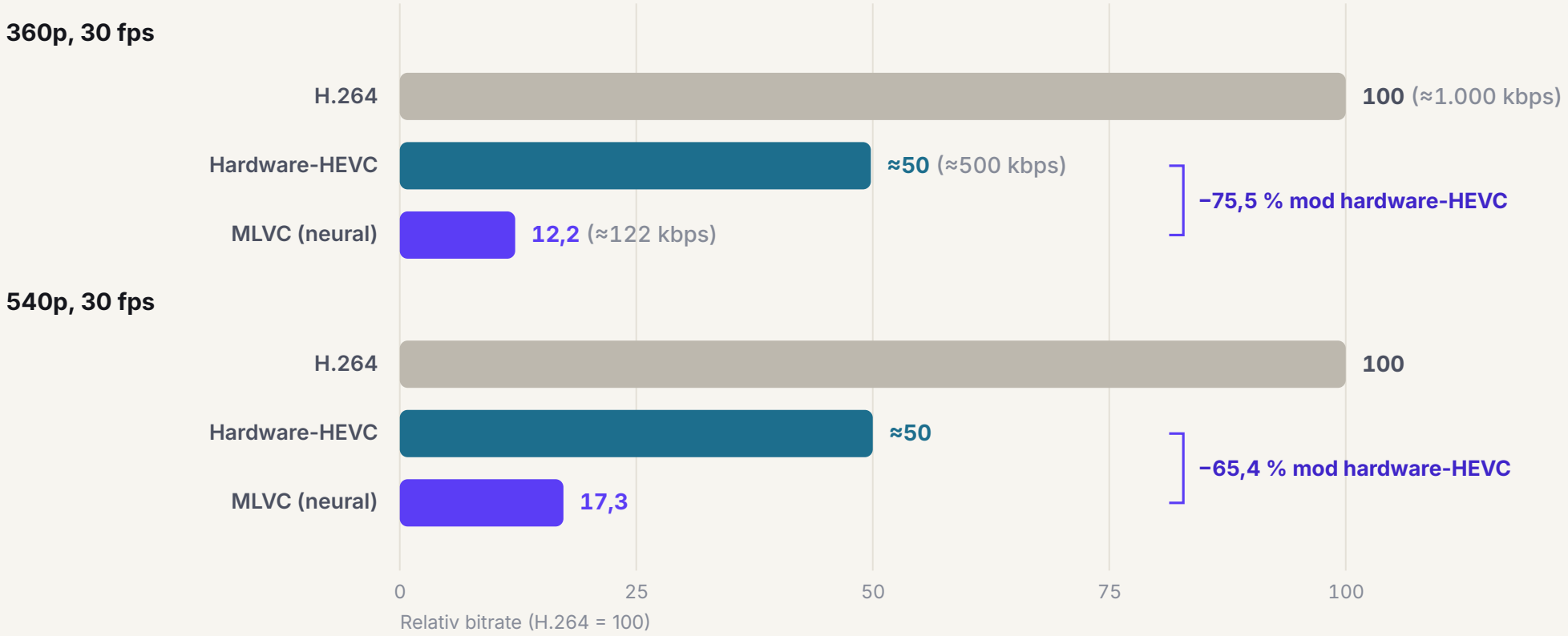
  Udgivet i 2025: 12,5–27,9 % bedre end VVC intra ved 7 til 216 tusind MAC'er pr. pixel. Jeg fandt intet produkt, der bruger det som standardformat.

- **Nettet**

  Chrome 155 slog JPEG XL til som standard den 6. oktober 2026 med en hukommelsessikker dekoder skrevet i Rust.

> **Formater, der skal kunne åbnes overalt, forbliver klassiske; neural kodning vinder inden for én virksomheds pipeline.**

I opkald bruges ende-til-ende-neurale codecs



Satin har kørt i alle Teams- og Skype-opkald mellem to parter siden 2021. MLVC, open source under MIT siden 24. juli 2026, rulles ud i peer-to-peer-opkald i Teams ved op til 540p30 og bruger højst halvdelen af NPU'en leverandørpåstand.

Ikke alle platforme valgte neuralt, og grundene er lærerige

INDE I EN STANDARD, VALGFRIT

Opus 1.5 og 1.6

Neural forbedring, maskering af pakketab, dyb redundans og båndbreddeudvidelse, alt sammen som udgangspunkt slået fra og fuldt kompatibelt med standarden fra 2012.

KLASSISK, MED VILJE

Meta MLow (2024)

Over 20 % af Metas opkald kører på ARMv7-telefoner, og titals millioner WhatsApp-opkald om dagen på telefoner, der er over ti år gamle. MLow giver omkring dobbelt så god kvalitet som Opus ved 6 kbps med 10 % mindre regnekraft.

Kompleksitetsbudgettet i den installerede base afgør det. Neurale codecs kommer først i brug, hvor enhederne er nye, og én leverandør styrer begge ender.

Den største udrulning, som ingen kalder en codec

- **Lærte codecs overalt**

  Alle stemme- og videomodeller læser og skriver tokens eller latente værdier fra en lært codec.

- **Ikke distributionsformater**

  Hver model har sit eget latente rum: ingen bitratestyring, ingen konformitetstest, ingen anden dekoder.

- **En omkostningspost**

  For udbydere af generativ video sænker en mere komprimerende latent repræsentation genereringsomkostningen mere end proportionalt analyse.

TO EKSEMPLER

12,5 Hz · 1,1 kbps

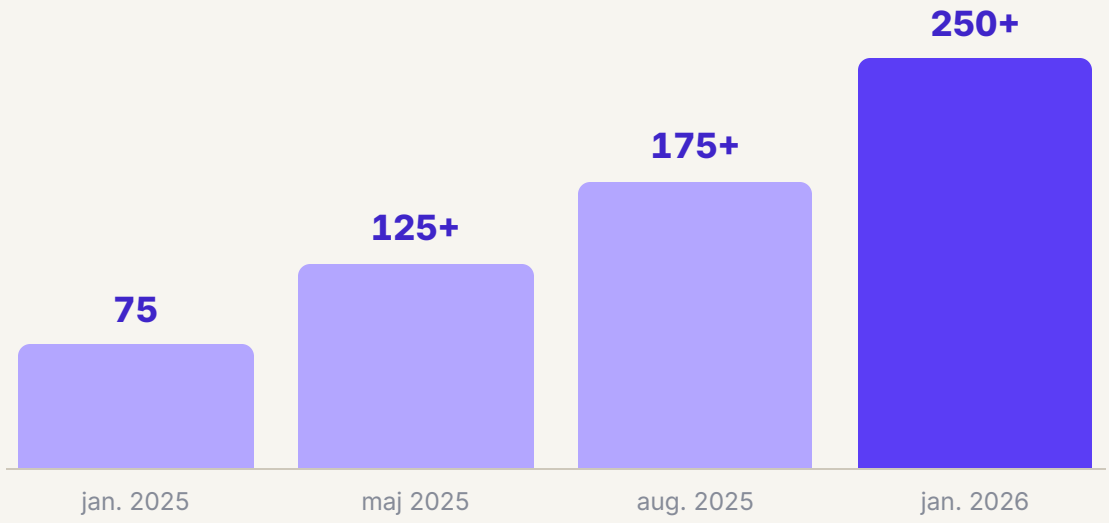
Kyutais Mimi-taletokens, under enhver telefonicodec

1:192

LTX-Videos video-VAE: ét latent token for 32×32 pixels over 8 billeder

Neural rekonstruktion er udbredt; neurale assets er på vej

DLSS 4 Multi Frame Generation: spil og apps



Neural teksturkomprimering

Scene i NVIDIA-demo, GTC 2026

Teksturhukommelse 6,5 GB → 0,97 GB

85 % mindre hukommelse

Assassin's Creed Mirage, udgivet

Kun udvalgte assets; mod blokformater

≈30 % bedre komprimering

Forskellige mål: ikke direkte sammenlignelige

DirectX Linear Algebra kom i offentlig preview i april 2026: én kodesti til neural dekodning i shadere på alle leverandørers GPU'er.

SPECIALFORBINDELSER

Satellitter, droner og robotter: det bedste match, et lille marked

- **Hvorfor det passer**

  Én ejer styrer begge ender, båndbredden er knap, og indholdet er snævert nok til at træne på.

- **Kina**

  Meddelelse nr. 49 af 2026, gældende fra 18. november, reviderer eksportkontrollen for droner; ét punkt bruger semantisk billedkomprimering som tærskel.

- **USA**

  FCC føjede udenlandskproducerede droner og deres kritiske komponenter til sin Covered List den 22. december 2025.

> **En fornuftig niche for en specialist med kapacitet til at håndtere eksportregler, ikke et massemarked.**

En størrelsesorden mere energi pr. pixel

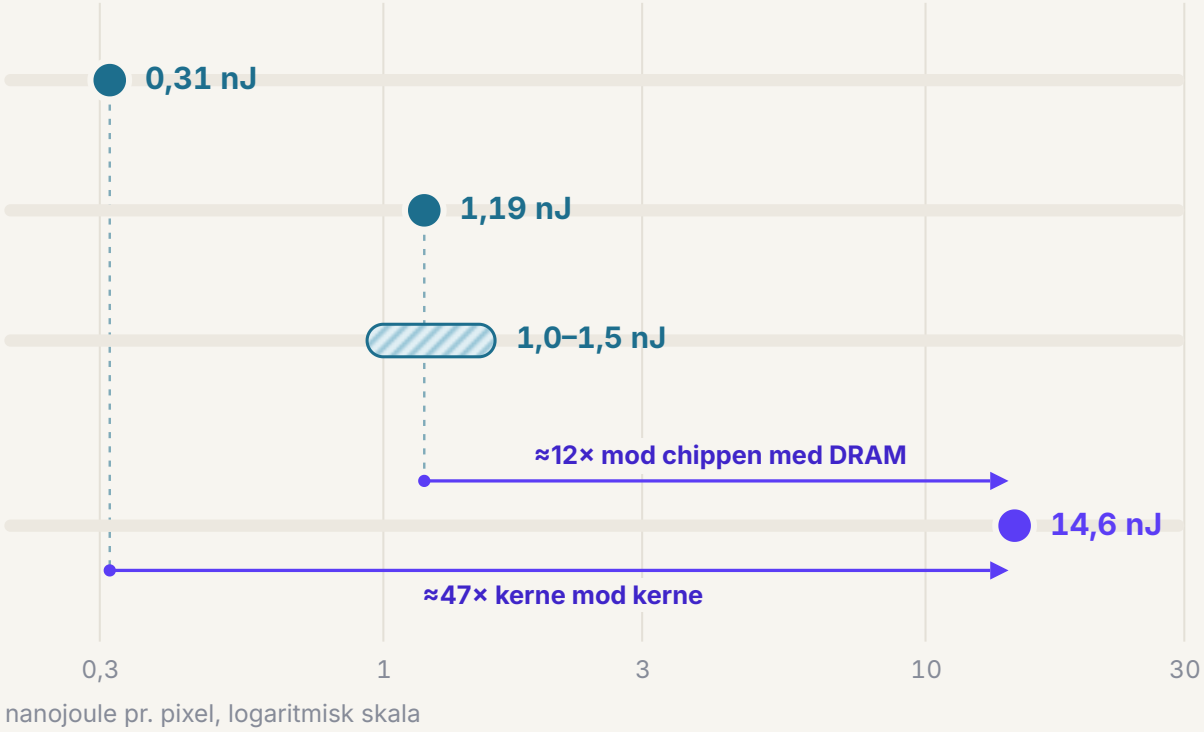
HEVC-dekoderchip, kerne
40 nm, 4K30, målt

HEVC-dekoderchip med DRAM
40 nm, 4K30, målt kerne plus simuleret DRAM

VVC-dekoder
28 nm, 4K60, fremskrivning

Neural videodekoder (NVCA)
28 nm, 1080p25, kun syntese; vægte og DRAM ikke medregnet

Ved 1080p30 svarer det til omkring 74 mW mod 908 mW.



Ved 1080p30: omkring 74 mW for HEVC-chippen med DRAM mod cirka 900 mW for NVCA. En ældre procesnode for HEVC-chippen, og vægte og DRAM er ikke medregnet for NVCA: sammenligningen favoriserer det neurale design.

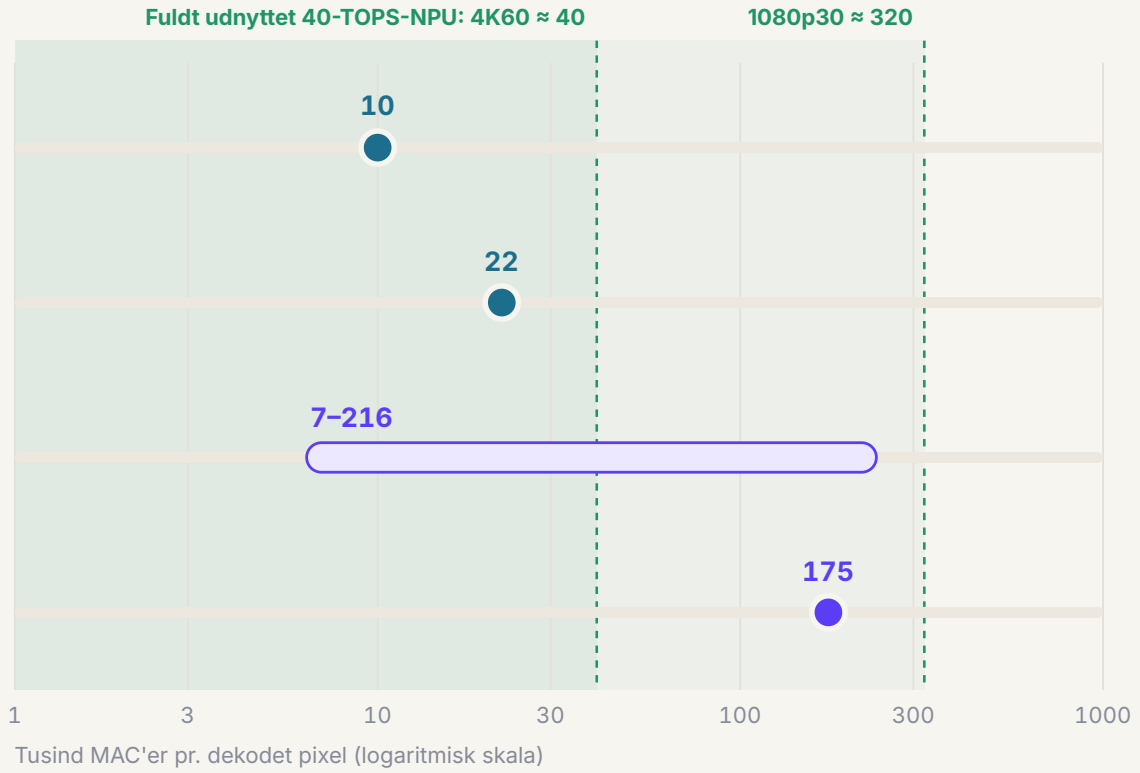
Hvad en neural dekoder kræver af en NPU

Neuralt in-loop-filter, meget lav kompleksitet
JVET-udforskning · ≈6 % bitbesparelse
1,2 TOPS ved 1080p30 · 10 ved 4K60

Neuralt in-loop-filter, lav kompleksitet
JVET-udforskning · ≈8 %
2,7 TOPS ved 1080p30 · 22 ved 4K60

JPEG AI-dekodere, basis til høj
12,5–27,9 % mod VVC intra

DCVC-RT, en hel neural videodekoder
21 % mod VVC's reference, artiklens opsætning
22 TOPS ved 1080p30 · 174 ved 4K60



> DCVC-RT kræver omkring 175 tusind MAC'er pr. pixel: omkring 22 TOPS ved 1080p30 og 174 ved 4K60 ved fuld udnyttelse **estimat**. En fuldt udnyttet NPU på 40 TOPS har råd til omkring 40 tusind ved 4K60.

Hukommelsestrafik og bitnøjagtighed er de hårde grænser

DRAM DOMINERER

Hukommelse før MAC'er

I MIT's HEVC-chip står DRAM for omkring tre fjerdedele af systemets energi: 0,88 af 1,19 nJ pr. pixel. Neurale dekodere bærer store featurekort mellem billederne, og DCVC-RT's forfattere fandt, at hukommelses-I/O, ikke aritmetik, begrænsede hastigheden.

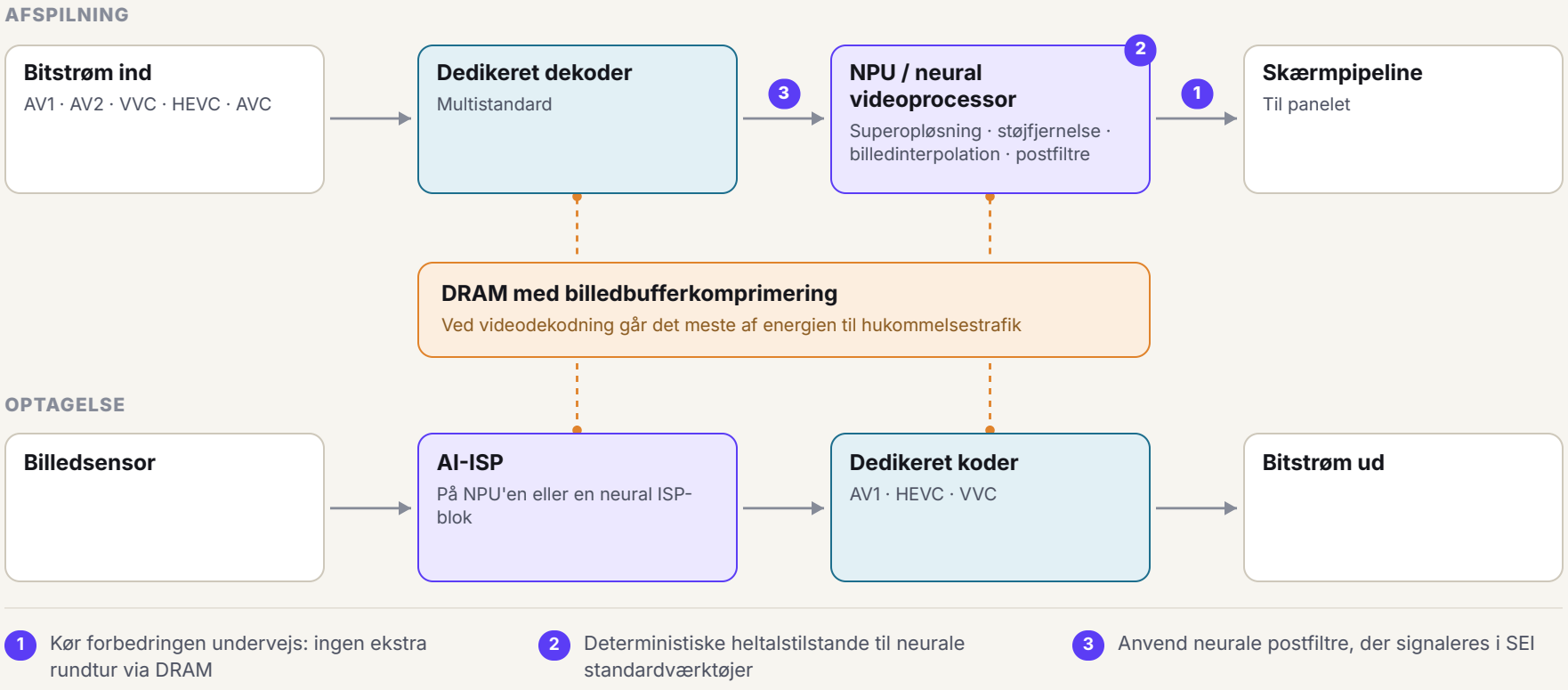
NØJAGTIGHED

Alle dekodere skal beregne de samme bit

Entropidekodning fejler, hvis to enheder beregner en sandsynlighed forskelligt. Løsningerne: heltalsmodeller (DCVC-RT), en hyperprior med delt skala (MLVC), 16-bit heltalsinferens (JVET's SADL), konformitetspunkter (JPEG AI).

For et chipteam er det HEVC's disciplin med heltalstransformationer igen: standarden skal definere aritmetikken nøjagtigt.

Opdelingen, der bruges: en dedikeret codec plus en NPU



> Tencents Canghai: over 100.000 chips af første generation i drift; anden generation tilføjer H.266 og superopløsning på chippen. Chips&Media og Allegro DVT sælger codec-IP og neural video-IP side om side.

Hvad jeg ville bygge, efter rolle

SoC- og tv-chipteams

- Dedikerede dekodere til alle de standarder, du skal afspille, AV2 nu
- En NPU dimensioneret til forbedring ved skærmopløsning, uden ekstra rundtur via DRAM
- SEI for neurale postfiltre og deterministiske heltalstilstande
- H.267 som designstart i 2029 eller senere

IP-leverandører

- AV2- og VVC-dekodere nu
- Postfilteracceleratorer og billedbufferkomprimering
- Konformitetsværktøjer til neural behandling

Videoplatforme

- Neural forbehandling ved koderen tjener sig først hjem
- Superopløsning på klienten, hvor enhederne tillader det
- Ende-til-ende-neurale codecs kun i lukkede kredsløb, du selv styrer

Hvor neurale værktøjer bliver licenserbare

Standard	Neuralt indhold	Status, november 2026
H.267 (arbejdsnavn)	Åben: klassiske (ECM) og neurale værktøjer er begge kandidater	Svarene evalueres i januar 2027; version 1 ved udgangen af 2029
VVC og HEVC med VSEI	Neuralt postfilter beskrevet i SEI; valgfrit for dekodere	I kraft; udvidet i oktober 2025
AV2	Klassisk, royaltyfri	Version 1.0.0 den 28. maj 2026; YouTube-streams på nettet planlagt i starten af 2027
JPEG AI	Ende-til-ende-lært billedcodec	Del 1 udgivet 2025; ikke fundet som standard i noget produkt
IEEE 1857.11	Neural billedkodning designet til NPU'er	Udgivet december 2024
Video Coding for Machines	Kodning til maskinanalyse frem for visning	Udkast til international standard i starten af 2026
Opus ML-udvidelser	Valgfrie og bagudkompatible	I libopus 1.5 og 1.6

645

ByteDances patentfamilier inden for kodning efter VVC (GreyB, 2025)

0,50 · 1,00–1,50 dollar

VVC-puljesats pr. telefon og pr. 4K-tv, Region 1

0 dollar

i royalties for AV1 og AV2 fra AOMedias medlemmer

HVEM TJENER PENGENE

Pengene følger standarderne og platformene

TJENER NU

Platforme

Beholder besparelserne

Netflix' neurale nedskalerer; superopløsning hos YouTube og Meta; Microsoft har frigivet MLVC som open source under MIT

TJENER NU

Enheds- og chipproducenter

Sælger differentiering

Kamera-ISP'er; opskaleringsprocessorer til tv; Tencents egenudviklede Canghai-codecchips

TJENER NU

IP-leverandører

Licenserer hver ny standard

Kunder til AV2-dekodere hos Chips&Media, Allegro DVT og VeriSilicon inden for tre uger efter specifikationen

TJENER NU

GPU-leverandører

Binder kunderne til platformen

DLSS 4 i 250+ spil og apps; SDK til neurale teksturer i beta

PR. ENHED

Patentindehavere og puljer

Opkræver pr. enhed

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

EXIT VIA OPKØB

Startups med neurale codecs

Opkøbt

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

> De tre bedst kendte startups med neurale codecs, WaveOne, iSIZE og Deep Render, blev købt af Apple, Sony og InterDigital.

HVOR EN VIRKSOMHED KAN BYDE IND

Seks muligheder og risikoen ved hver

Mulighed	Hvem køber	Hvorfor nu	Hovedrisiko
Neural forbehandling ved koderen	Streaming- og brugervideoplatforme, tv-selskaber	Ingen ændring i enhederne; bevist hos Netflix	Store platforme bygger det selv
IP og tuning til forbedring på skærmsiden	Chipproducenter til tv, set-top-bokse og telefoner	NPU'erne sidder allerede i enhederne	Kommodificering; interne teams
Neurale codecs i lukkede kredsløb	Leverandører af virksomhedskommunikation, industri og telepresence	MLVC viser, at det virker, og koden er åben	Du skal styre begge ender; fragmenterede NPU'er
Silicium eller IP med codec plus NPU	SoC-producenter	AV2-cyklussen er ved at begynde	Stærke etablerede aktører
Evaluering, konformitet og deterministisk inferens	Alle, der leverer neurale værktøjer	Bitnøjagtighed er et hårdt krav	Lille marked, medmindre det bundtes
Standardbidrag til H.267	Licensgivere, store platforme	Svarene evalueres i januar 2027	Langt, dyrt, usikkert

Tag påstande med forbehold, indtil de oplyser fire ting

Påstand	Hvad dokumentationen viser
»Neurale codecs slår allerede VVC.«	I laboratoriet. DCVC-RT sparer 21 % i forhold til VVC's referencekoder; dekoderen koster omkring 175 tusind MAC'er pr. pixel.
»AI-codecs vil erstatte H.26x og AV1 inden for få år.«	Ikke i afspilning: H.267 version 1 ved udgangen af 2029, chips omkring 2031 estimat. Den codec, der kommer nu, AV2, er klassisk.
»MLVC skærer båndbredden med 88 %.«	I forhold til H.264 ved 360p. I forhold til hardware-HEVC: 75,5 % ved 360p og 65,4 % ved 540p, på Microsofts eget testsæt.
»Neurale teksturer skærer videohukommelsen med 85–96 %.«	I demonstrationsscener. Det udgivne spil opnåede omkring 30 %, kun på udvalgte assets.
»JPEG AI er en ISO-standard, så kameraer og browsere vil tage den i brug.«	Ingen produkter med den som standardformat fundet. Nettet tog JPEG XL i brug, et klassisk format.

Spørg efter referencen, kvalitetsmålet, opløsningen og dekodningsomkostningen. En påstand, der ikke kan oplyse alle fire, er et forskningsresultat, ikke et produkt.

VÆRD AT FØLGE

Værd at følge, og åbne spørgsmål

PEJLEMÆRKER

Januar 2027	H.267-svarene evalueres: hvor mange bygger på neurale værktøjer, og til hvilken pris pr. pixel
Primo 2027	YouTubes første AV2-streams på nettet; AV2-dekodere i annoncerede SoC'er
Andet halvår 2027	H.267's første testmodel
2027	MLVC ud over peer-to-peer-opkald i Teams
2027–2028	Et første produkt med JPEG AI som standardformat
Ultimo 2029	H.267 version 1 starter chipcyklussen

ÅBNE SPØRGSMÅL

01 Nøjagtighed i stor skala: bitnøjagtig heltalsdekodning på alle leverandørers NPU'er, med hvilket komprimeringstab?

02 Hukommelse før MAC'er: DRAM-trafikken pr. pixel ved neural efterbehandling

03 Den pris pr. pixel, hvor et neuralt værktøj holder op med at betale sig i H.267

04 Hvilke lukkede kredsløb ud over opkald der først indfører ende-til-ende-neurale codecs

05 Hvordan gengivelsestroskab måles, når en dekoder genererer detaljer

Det er områder, vi kan udforske sammen gennem teknisk udveksling og teknisk samarbejde.

i@tanshuai.com · tanshuai.com