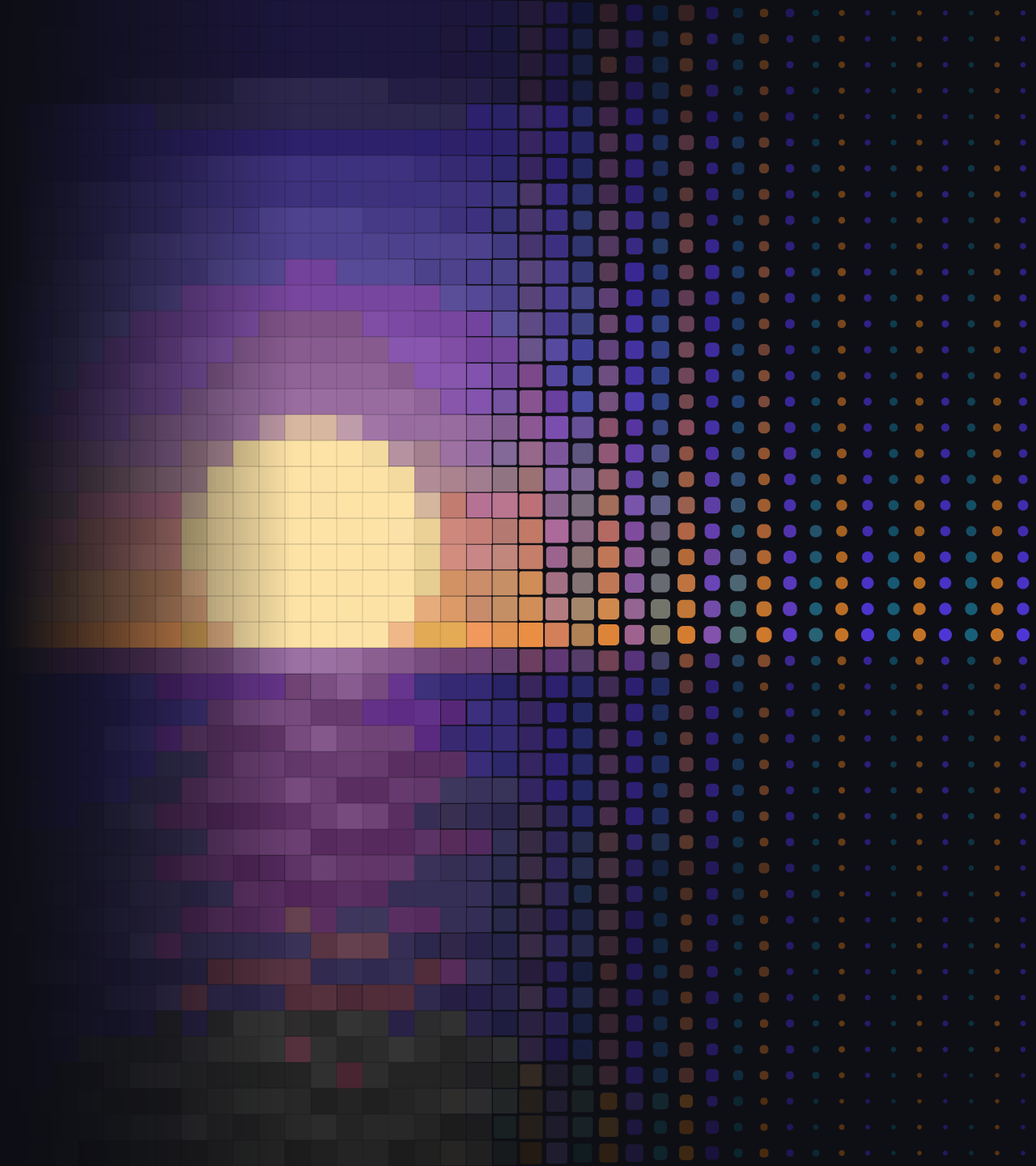
Аналитическая записка · ноябрь 2026

Нейросетевое кодирование в 2026 году

Вокруг кодека, а не вместо него

Воспроизведение · Фотография · Звонки · Генеративный ИИ · Игры · Микросхемы

Тан Шуай
tanshuai.com · i@tanshuai.com



ВЕРДИКТ

Уже мейнстрим, но вокруг кодека

01 Нейросети стали мейнстримом на обоих концах стандартных кодеков

ISP камер, даунскейлинг на стороне кодера и суперразрешение при воспроизведении. Битовый поток между ними — по-прежнему AV1, HEVC, AVC или VVC, и декодирует его жёсткая логика.

02 Сквозные нейросетевые кодеки применяются только в замкнутых системах

Звонки, токенизаторы генеративного ИИ и одна AAA-игра: одна компания контролирует оба конца и может переключиться на стандартный кодек.

03 Нейросетевые битовые потоки на массовых экранах ждут H.267

Версия 1 ожидается в конце 2029 года; первые чипы — около 2031 года, массовое применение — в середине 2030-х оценка.

≈75 %

мобильного трафика приходится на видео (конец 2025 года)

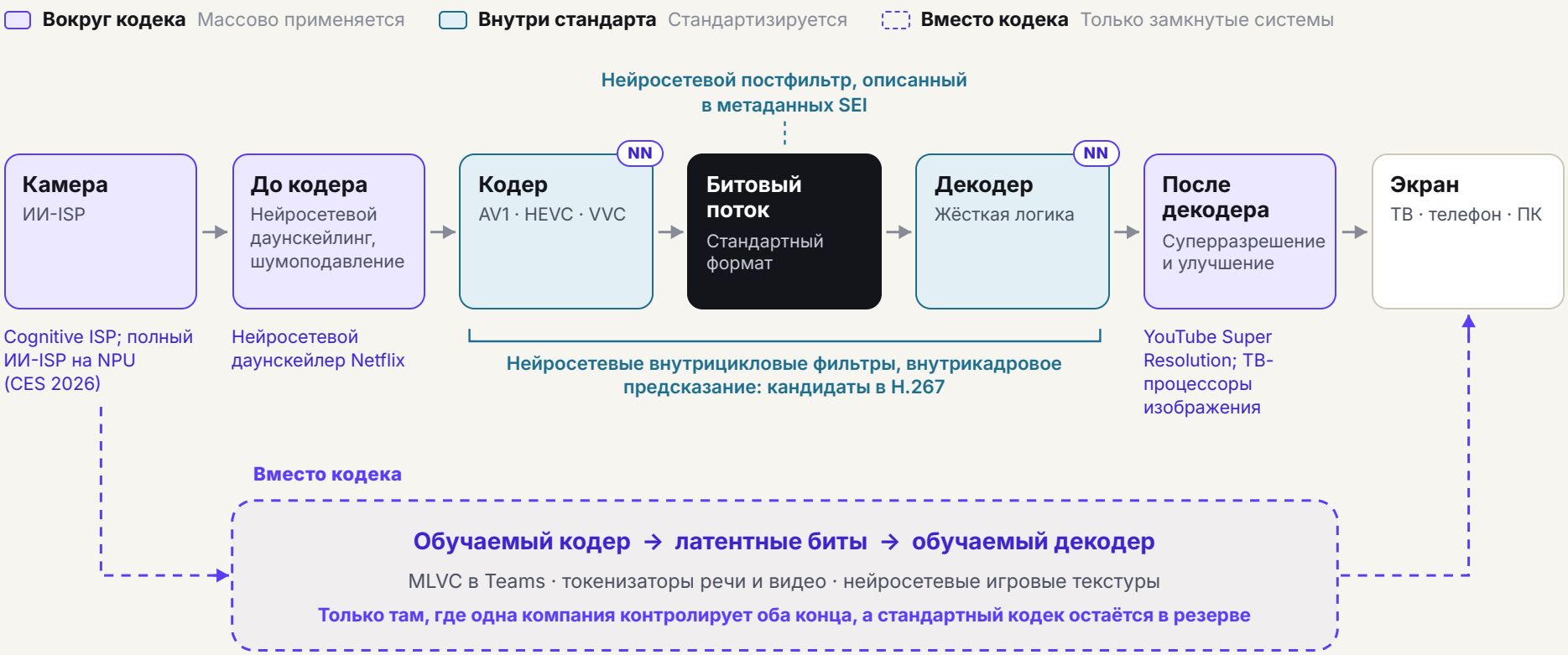
1,26 млрд

смартфонов отгружено в 2025 году, и все декодируют видео на жёсткой логике

≈12×

энергии на пиксель против аппаратного декодера HEVC у единственного опубликованного проекта чипа нейросетевого декодера

Три способа, которыми нейросеть попадает в кодек



> «Вокруг кодека» не требует ничьего разрешения. «Внутри стандарта» требует стандарта. «Вместо кодека» требует контроля над обоими концами.

Где нейросетевое кодирование применяется в 2026 году

	Вокруг стандартного кодека предобработка, улучшение, апскейлинг	Сквозной нейросетевой кодек обучаемые кодер и декодер	Внутри стандарта нейросетевые инструменты в стандартном битовом потоке или API
01 **Воспроизведение и стриминг**	**МАССОВО** Даунскейлер Netflix; суперразрешение YouTube и Meta; ИИ-апскейлеры в телевизорах	**НЕ ВНЕДРЕНО** Ни один стриминговый сервис не передаёт нейросетевой видеоформат; телевизоры и телефоны декодируют стандартные кодеки аппаратно	**СТАНДАРТИЗАЦИЯ** Версия 1 H.267 ожидается в конце 2029 года; нейросетевые инструменты — среди кандидатов
02 **Фотография и изображения**	**МАССОВО** ИИ-ISP во флагманских смартфонах; полностью нейросетевой ISP показан как IP-блок на CES 2026	**ДЕМО, ПИЛОТЫ** JPEG AI — стандарт ISO: демо на телефонах, продуктов нет	**НЕ ВНЕДРЕНО** Распространяется классический формат: JPEG XL включён по умолчанию в Chrome 155
03 **Звонки и конференции**	**МАССОВО** Суперразрешение на стороне получателя и нейросетевое шумоподавление	**МАССОВО** Нейросетевой звук в звонках Teams с двумя участниками с 2021 года; видео MLVC внедряется в Teams	**ОГРАНИЧЕННО** Opus 1.5 добавляет необязательные обратно совместимые ML-инструменты
04 **Генеративный ИИ**	—	**МАССОВО** Каждая видео- и голосовая модель работает на обучаемом токенизаторе, например на видео-VAE со сжатием 1:192	—
05 **Игры и графика**	**МАССОВО** Нейросетевой апскейлинг и генерация кадров: один только DLSS 4 — в 250+ играх и приложениях	**ОГРАНИЧЕННО** Нейросетевые текстуры вышли в одной AAA-игре; SDK от NVIDIA в бета-версии	**СТАНДАРТИЗАЦИЯ** DirectX добавляет матричные операции для шейдеров (предварительная версия)
06 **Специализированные каналы: спутники, роботы**	**ДЕМО, ПИЛОТЫ** Серверное восстановление низкобитрейтных потоков	**ДЕМО, ПИЛОТЫ** Устойчивые к потерям нейросетевые кодеки в исследованиях и пилотах	—

Строки упорядочены по числу людей, которых охватывает каждая область.

Нейросети на обоих концах, стандарт посередине

- **На стороне кодера**

  Нейросетевой даунскейлер Netflix работает с 2022 года без изменений на каких-либо устройствах; в тестах Netflix его предпочли около 77 % зрителей данные компании.

- **На экране**

  YouTube Super Resolution на телевизорах с октября 2025 года; Meta применяет суперразрешение на своих серверах и на телефонах.

- **В телевизоре**

  Процессор Samsung NQ8 AI Gen3 рекламирует 768 нейросетей для апскейлинга.

БИТОВЫЙ ПОТОК ОСТАЁТСЯ СТАНДАРТНЫМ

30 %

просмотров Netflix — в AV1 (декабрь 2025 года)

>90 %

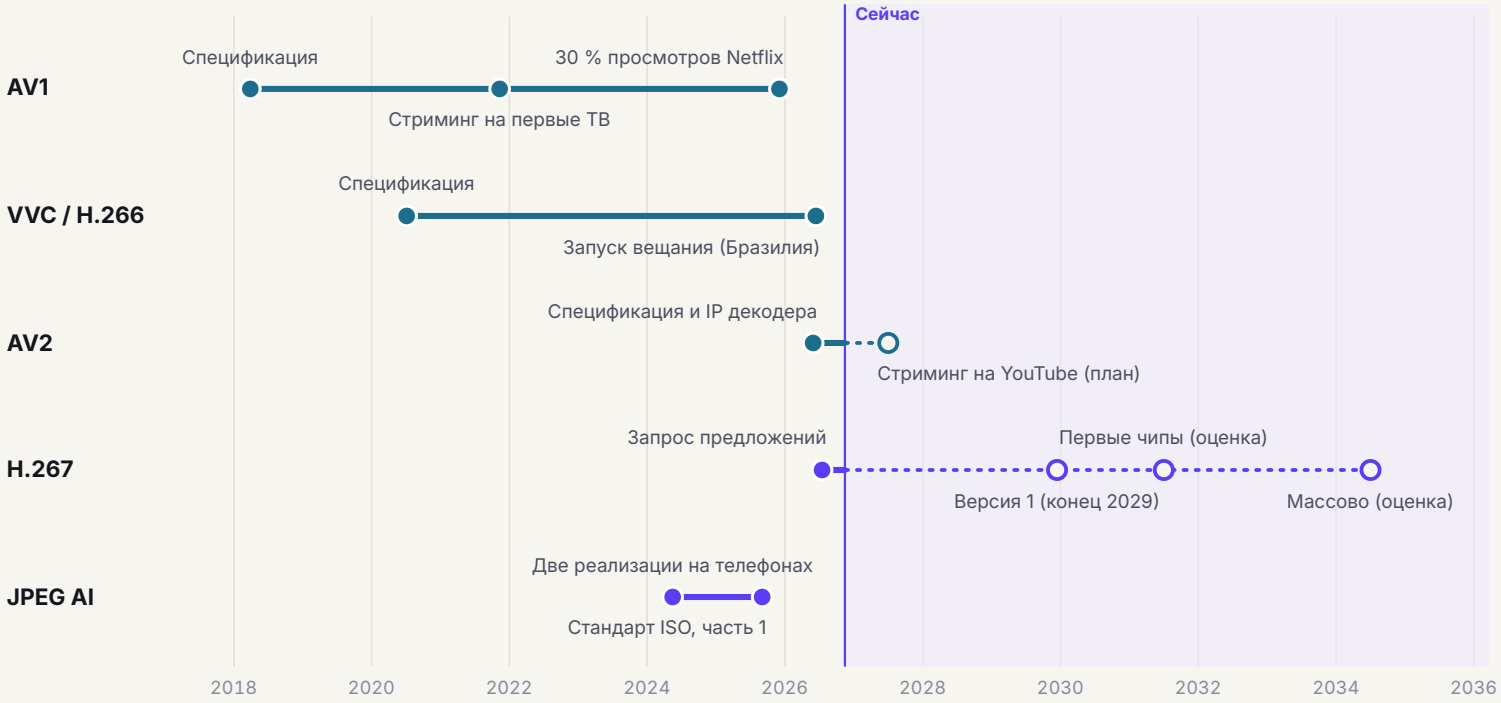
каталога YouTube по времени просмотра доступно в AV1 (март 2026 года)

3 недели

от AV2 версии 1.0.0 (28 мая 2026 года) до заказчиков IP декодера у трёх поставщиков

СРОКИ

От готового стандарта до массовых экранов — большая часть десятилетия



AV1 понадобилось почти восемь лет от спецификации (2018) до 30 % просмотров Netflix (декабрь 2025 года). Полые метки — целевые сроки или мои оценки.

H.267 решит, дойдут ли нейросети до каждого экрана

Октябрь 2025
Запрос свидетельств: пять ответов от 16 компаний и лабораторий; выигрыш нейросетей виднее всего на 8K HDR

Июль 2026
Запрос предложений объявлен в Женеве

Январь 2027
Оценка ответов в Брисбене

Второе полугодие 2027
Первая тестовая модель

Конец 2029
Версия 1 стандарта

≈28 %
экономит относительно VVC классический трек ECM

6–14 %
экономят нейросетевые внутрицикловые фильтры при 10–22+ тыс. MAC на пиксель

SEI
в потоках VVC уже можно указать нейросетевой постфильтр, который применит декодер

На мой взгляд: классическое ядро и несколько нейросетевых инструментов там, где сэкономленные биты оправдывают затраты декодера анализ**.**

Нейросети — внутри камеры, классика — в вебе

- **Камера**

  Cognitive ISP от Qualcomm с 2023 года выполняет в телефонах сегментацию в реальном времени; полный ИИ-ISP, работающий на NPU, показали на CES 2026.

- **JPEG AI**

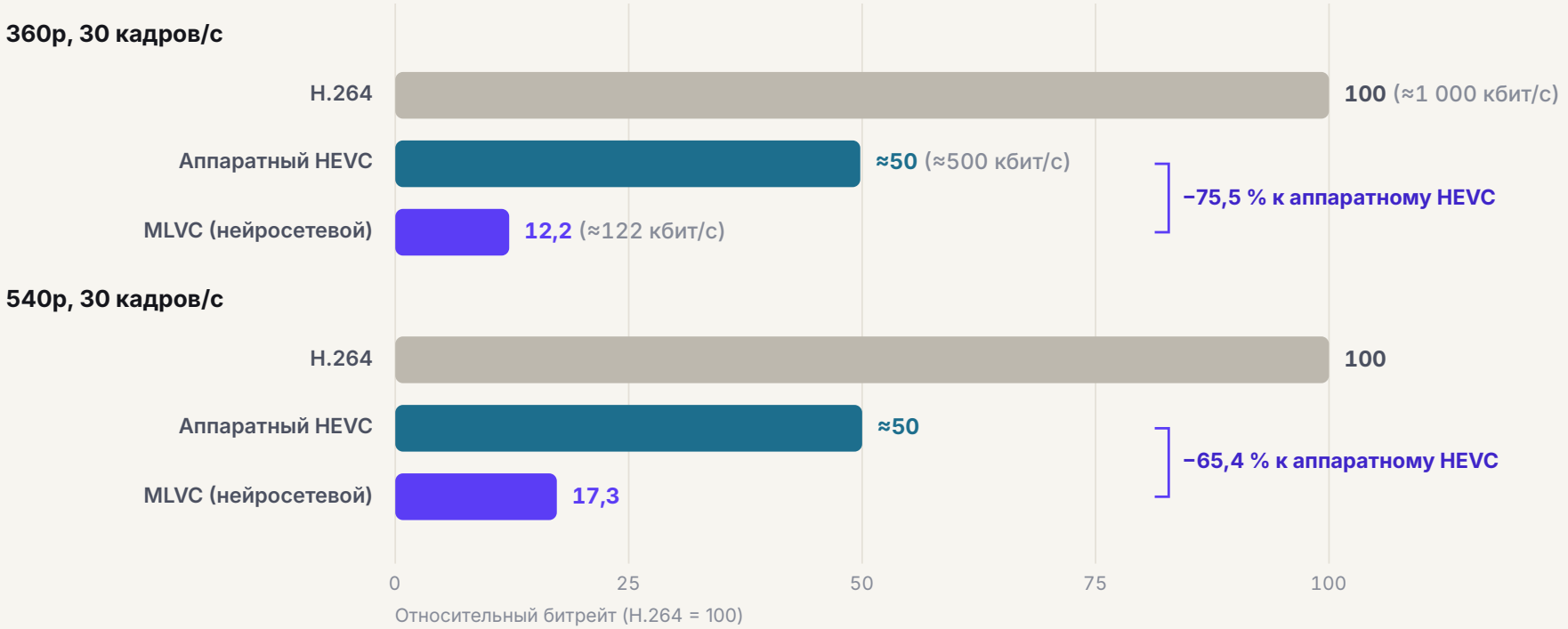
  Опубликован в 2025 году: 12,5–27,9 % относительно VVC intra при 7–216 тыс. MAC на пиксель. Я не нашёл продуктов, где он используется по умолчанию.

- **Веб**

  Chrome 155 включил JPEG XL по умолчанию 6 октября 2026 года, с безопасным по памяти декодером на Rust.

Форматы, которые должны открываться везде, остаются классическими; нейросетевое кодирование выигрывает внутри конвейера одной компании.

Именно в звонках применяются сквозные нейросетевые кодеки



Satin работает во всех звонках Teams и Skype с двумя участниками с 2021 года. MLVC, открытый под лицензией MIT с 24 июля 2026 года, внедряется в одноранговых звонках Teams с разрешением до 540p30 и занимает не более половины NPU данные компании.

Не все платформы перешли на нейросети, и причины поучительны

ВНУТРИ СТАНДАРТА, ОПЦИОНАЛЬНО

Opus 1.5 и 1.6

Нейросетевое улучшение, маскирование потерь пакетов, глубокая избыточность и расширение полосы — всё выключено по умолчанию и полностью совместимо со стандартом 2012 года.

КЛАССИКА — СОЗНАТЕЛЬНО

Meta MLow (2024)

Более 20 % звонков Meta идут с телефонов на ARMv7, а десятки миллионов звонков WhatsApp в день — с телефонов старше десяти лет. MLow даёт примерно вдвое лучшее качество, чем Opus, при 6 кбит/с и требует на 10 % меньше вычислений.

Решает бюджет сложности парка устройств. Нейросетевые кодеки появляются сначала там, где устройства новые, а оба конца контролирует один поставщик.

Крупнейшее внедрение, которое никто не называет кодеком

- **Обучаемые кодеки повсюду**

  Каждая голосовая и видеомодель читает и пишет токены или латентные представления обучаемого кодека.

- **Не форматы распространения**

  У каждой модели своё латентное пространство: ни управления битрейтом, ни теста на соответствие, ни второго декодера.

- **Статья затрат**

  Для тех, кто обслуживает генеративное видео, более компактное латентное представление снижает стоимость генерации более чем пропорционально анализ.

ДВА ПРИМЕРА

12,5 Гц · 1,1 кбит/с

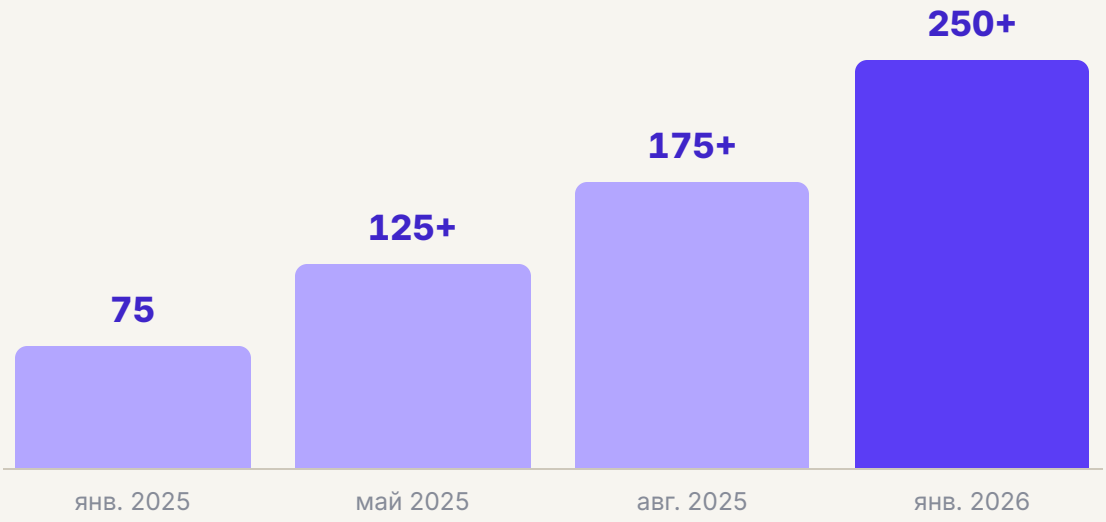
речевые токены Mimi от Kyutai — ниже любого телефонного кодека

1:192

видео-VAE LTX-Video: один латентный токен на 32×32 пикселя за 8 кадров

Нейросетевая реконструкция — мейнстрим, нейросетевые ассеты — только начало

DLSS 4 Multi Frame Generation: игры и приложения



Нейросетевое сжатие текстур

Демосцена NVIDIA, GTC 2026

Текстурная память 6,5 ГБ → 0,97 ГБ

На 85 % меньше памяти

Assassin's Creed Mirage, вышедшая игра

Только отдельные ассеты; относительно блочных форматов

Сжатие лучше на ≈30 %

Разные показатели: напрямую не сопоставимы

DirectX Linear Algebra вышла в публичную предварительную версию в апреле 2026 года: единый путь кода для нейросетевого декодирования в шейдерах на GPU любого производителя.

Спутники, дроны и роботы: лучшее соответствие, малый рынок

- **Почему подходит**

  Один владелец контролирует оба конца, полосы мало, а контент достаточно узок, чтобы под него обучить модель.

- **Китай**

  Объявление № 49 за 2026 год, вступающее в силу 18 ноября, пересматривает экспортный контроль над дронами; в одной из позиций порогом служит семантическое сжатие изображений.

- **США**

  FCC 22 декабря 2025 года внесла дроны иностранного производства и их критические компоненты в свой перечень Covered List.

> **Разумная ниша для специализированной компании, способной обеспечить экспортный комплаенс, но не массовый рынок.**

На порядок больше энергии на пиксель

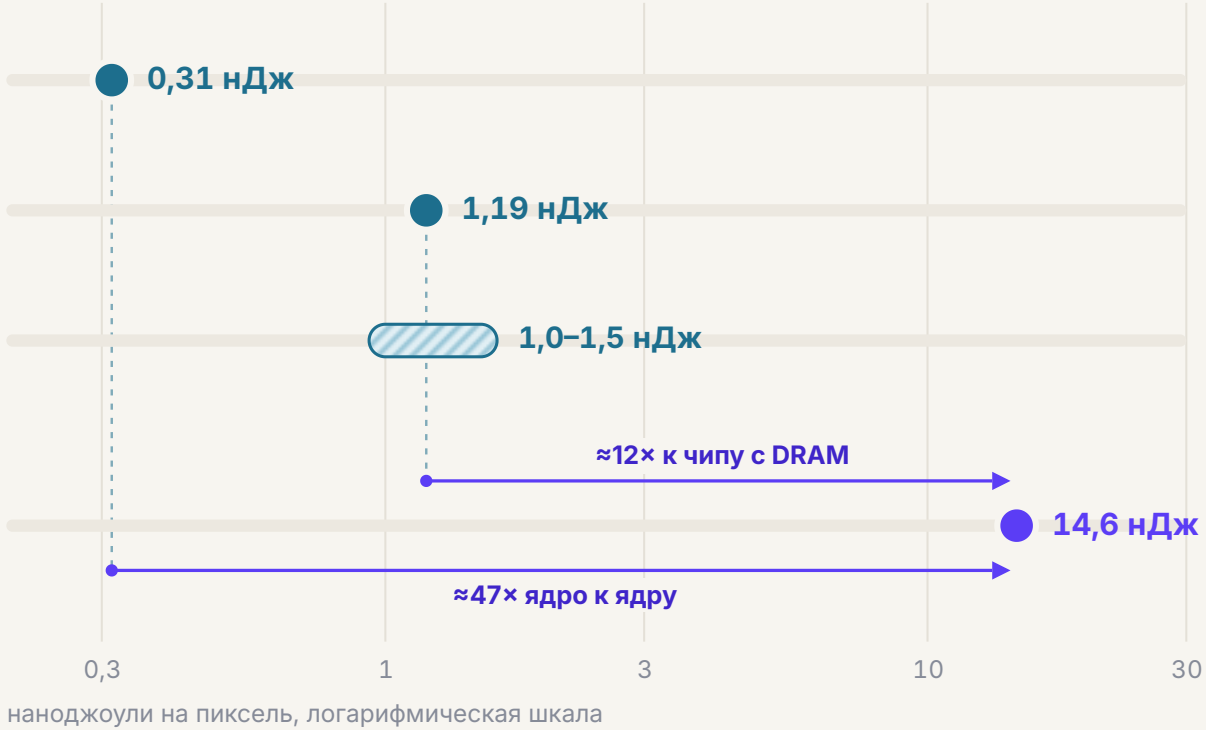
Чип декодера HEVC, ядро
40 нм, 4K30, измерено

Чип декодера HEVC с DRAM
40 нм, 4K30, измеренное ядро плюс смоделированная DRAM

Декодер VVC
28 нм, 4K60, прогноз

Нейросетевой видеодекодер (NVCA)
28 нм, 1080p25, только синтез; веса и DRAM не учтены

При 1080p30 это около 74 мВт против 908 мВт.



При 1080p30: около 74 мВт у чипа HEVC с DRAM против примерно 900 мВт у NVCA. Более старый техпроцесс у чипа HEVC, неучтённые веса и DRAM у NVCA: сравнение играет на руку нейросетевому проекту.

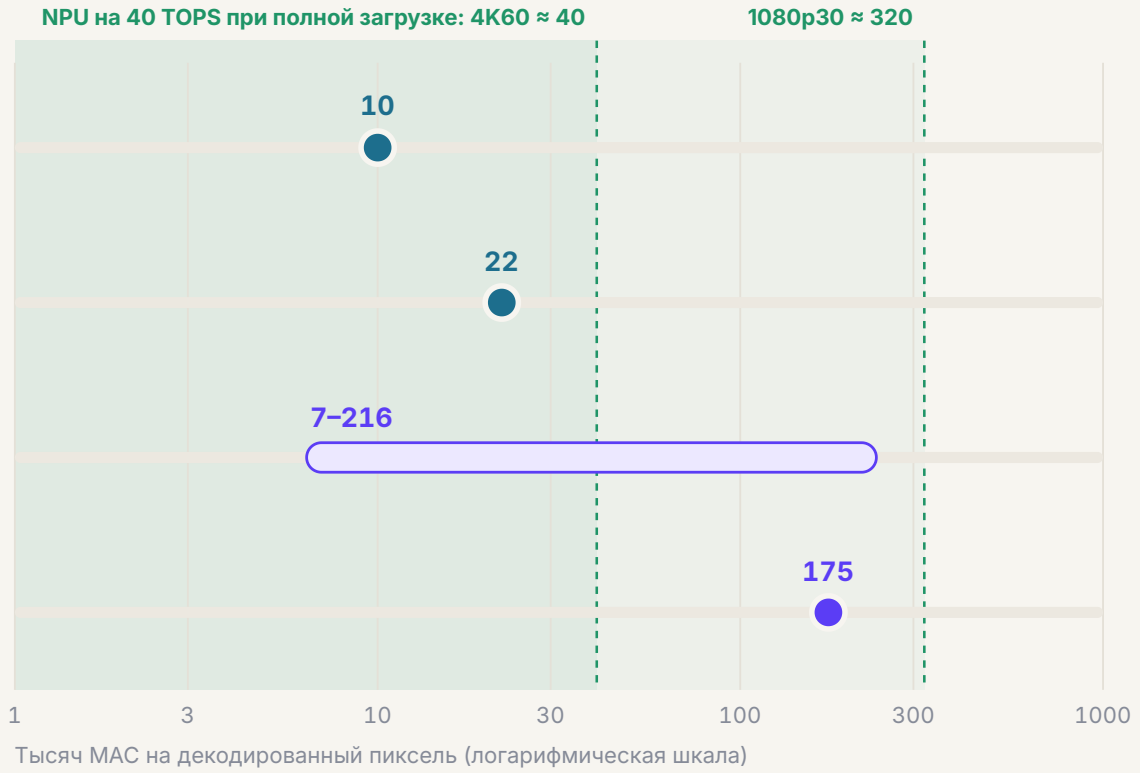
Что нейросетевой декодер требует от NPU

Нейросетевой внутрицикловой фильтр, очень низкая сложность
Исследование JVET · экономия битов ≈6 %
1,2 TOPS при 1080p30 · 10 при 4K60

Нейросетевой внутрицикловой фильтр, низкая сложность
Исследование JVET · ≈8 %
2,7 TOPS при 1080p30 · 22 при 4K60

Декодеры JPEG AI, от базового до высокого
12,5–27,9 % относительно VVC intra

DCVC-RT, целый нейросетевой видеодекодер
21 % относительно эталона VVC, условия статьи
22 TOPS при 1080p30 · 174 при 4K60



DCVC-RT нужно около 175 тыс. MAC на пиксель: около 22 TOPS при 1080p30 и 174 при 4K60, при полной загрузке **оценка**. Полностью загруженный NPU на 40 TOPS может позволить себе около 40 тыс. при 4K60.

Трафик памяти и побитовая точность — жёсткие ограничения

DRAM ДОМИНИРУЕТ

Сначала память, потом MAC

В чипе HEVC от MIT на DRAM приходится около трёх четвертей энергии системы: 0,88 из 1,19 нДж на пиксель. Нейросетевые декодеры переносят между кадрами большие карты признаков, и авторы DCVC-RT обнаружили, что скорость ограничивает ввод-вывод памяти, а не арифметика.

ТОЧНОСТЬ

Каждый декодер должен вычислить те же биты

Энтропийное декодирование ломается, если два устройства вычисляют вероятность по-разному. Решения: целочисленные модели (DCVC-RT), гиперприор с общим масштабом (MLVC), 16-битный целочисленный инференс (SADL в JVET), точки проверки соответствия (JPEG AI).

Для команды разработчиков чипа это снова дисциплина целочисленных преобразований HEVC: стандарт должен точно определять арифметику.

Разделение, которое пошло в серию: кодек на жёсткой логике плюс NPU



> Canghai от Tencent: развёрнуто более 100 000 чипов первого поколения; второе поколение добавляет H.266 и суперразрешение на кристалле. Chips&Media и Allegro DVT продают IP кодеков и нейросетевые видео-IP бок о бок.

Что бы я сделал — по ролям

Команды SoC и ТВ-чипов

- Декодеры на жёсткой логике для каждого стандарта, который нужно воспроизводить, — AV2 уже сейчас
- NPU под улучшение в разрешении дисплея, без лишнего круга через DRAM
- SEI нейросетевых постфильтров и детерминированные целочисленные режимы
- H.267 как начало проектирования в 2029 году или позже

Поставщики IP-блоков

- Декодеры AV2 и VVC уже сейчас
- Ускорители постфильтров и сжатие буфера кадров
- Инструменты проверки соответствия для нейросетевой обработки

Видеоплатформы

- Нейросетевая предобработка на стороне кодера окупается первой
- Суперразрешение на клиенте там, где позволяют устройства
- Сквозные нейросетевые кодеки — только в замкнутых системах, которые вы контролируете

Где нейросетевые инструменты можно лицензировать

Стандарт	Нейросетевое содержание	Статус, ноябрь 2026
H.267 (рабочее название)	Открыто: кандидаты — и классические (ECM), и нейросетевые инструменты	Оценка ответов — январь 2027 года; версия 1 — конец 2029 года
VVC и HEVC с VSEI	Нейросетевой постфильтр описан в SEI; для декодеров необязателен	Действует; расширен в октябре 2025 года
AV2	Классика, без роялти	Версия 1.0.0 — 28 мая 2026 года; YouTube в вебе — с начала 2027 года (план)
JPEG AI	Сквозной обучаемый кодек изображений	Часть 1 — 2025 год; поддержки по умолчанию в продуктах не найдено
IEEE 1857.11	Нейросетевое кодирование изображений, рассчитанное на NPU	Опубликован в декабре 2024 года
Video Coding for Machines	Кодирование для машинного анализа, а не для просмотра	Проект международного стандарта, начало 2026 года
ML-расширения Opus	Необязательные и обратно совместимые	В libopus 1.5 и 1.6

645

патентных семейств ByteDance в работе над стандартом после VVC (GreyB, 2025)

$0,50 · $1,00–1,50

ставка пула VVC за телефон и за 4K-телевизор, Регион 1

$0

роялти за AV1 и AV2 от участников AOMedia

КТО ЗАРАБАТЫВАЕТ

Деньги идут вслед за стандартами и платформами

ЗАРАБАТЫВАЮТ СЕЙЧАС

Платформы

Оставляют экономию себе

Нейросетевой даунскейлер Netflix; суперразрешение YouTube и Meta; Microsoft открыла код MLVC под лицензией MIT

ЗАРАБАТЫВАЮТ СЕЙЧАС

Производители устройств и чипов

Продают отличия от конкурентов

ISP камер; ТВ-процессоры апскейлинга; собственные чипы кодеков Canghai от Tencent

ЗАРАБАТЫВАЮТ СЕЙЧАС

Поставщики IP-блоков

Лицензируют каждый новый стандарт

Заказчики декодера AV2 у Chips&Media, Allegro DVT и VeriSilicon — в течение трёх недель после выхода спецификации

ЗАРАБАТЫВАЮТ СЕЙЧАС

Производители GPU

Привязывают к своей платформе

DLSS 4 в 250+ играх и приложениях; SDK нейросетевых текстур в бета-версии

С КАЖДОГО УСТРОЙСТВА

Патентообладатели и пулы

Получают отчисления с устройства

VVC: $0,50 за телефон, $1,00–1,50 за 4K-телевизор; инструменты H.267 — после 2029 года

ВЫХОД ЧЕРЕЗ ПРОДАЖУ

Стартапы нейросетевых кодеков

Куплены

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

> Три самых известных стартапа нейросетевых кодеков — WaveOne, iSIZE и Deep Render — купили Apple, Sony и InterDigital.

Шесть направлений и риск в каждом

Направление	Кто покупает	Почему сейчас	Главный риск
Нейросетевая предобработка на стороне кодера	Стриминговые платформы, платформы пользовательского видео, вещатели	Не нужно менять устройства; доказано в Netflix	Крупные платформы делают это сами
IP и настройка улучшения на стороне дисплея	Производители чипов для ТВ, приставок и телефонов	NPU уже есть в устройствах	Коммодитизация; собственные команды
Нейросетевые кодеки для замкнутых систем	Поставщики корпоративных коммуникаций, промышленных решений и телеприсутствия	MLVC доказывает, что это работает, и его код открыт	Нужно контролировать оба конца; фрагментация NPU
Микросхемы или IP: кодек плюс NPU	Производители SoC	Начинается цикл AV2	Сильные действующие игроки
Оценка, проверка соответствия и детерминированный инференс	Все, кто выпускает нейросетевые инструменты	Побитовая точность — жёсткое требование	Малый рынок, если не в пакете
Вклад в стандартизацию H.267	Лицензиары, крупные платформы	Ответы оцениваются в январе 2027 года	Долго, дорого, без гарантий

Делайте скидку на заявления без четырёх уточнений

Заявление	Что говорят факты
«Нейросетевые кодеки уже обходят VVC»	В лаборатории. DCVC-RT экономит 21 % относительно эталонного кодера VVC; его декодер стоит около 175 тыс. MAC на пиксель.
«ИИ-кодеки заменят H.26x и AV1 через несколько лет»	Не в воспроизведении: версия 1 H.267 — в конце 2029 года, чипы — около 2031 года оценка. Кодек, который приходит сейчас, AV2, — классический.
«MLVC сокращает полосу на 88 %»	Относительно H.264 при 360p. Относительно аппаратного HEVC: 75,5 % при 360p и 65,4 % при 540p, на собственном тестовом наборе Microsoft.
«Нейросетевые текстуры экономят 85–96 % видеопамяти»	В демосценах. Вышедшая игра выиграла около 30 %, и только на отдельных ассетах.
«JPEG AI — стандарт ISO, значит, камеры и браузеры его примут»	Поддержки по умолчанию в продуктах не найдено. Веб принял JPEG XL — классический формат.

Спрашивайте об опорном кодеке, метрике, разрешении и стоимости декодирования. Заявление, в котором нет всех четырёх, — результат исследования, а не продукт.

За чем следить и что остаётся открытым

ОРИЕНТИРЫ

Январь 2027	Оценка ответов по H.267: сколько опираются на нейросетевые инструменты и какой ценой на пиксель
Начало 2027	Первые потоки AV2 на YouTube в вебе; декодеры AV2 в анонсированных SoC
Второе полугодие 2027	Первая тестовая модель H.267
2027	MLVC за пределами одноранговых звонков Teams
2027–2028	Первый продукт, использующий JPEG AI по умолчанию
Конец 2029	Версия 1 H.267 запускает цикл разработки микросхем

ОТКРЫТЫЕ ВОПРОСЫ

01 Точность в масштабе: побитово точное целочисленное декодирование на NPU разных производителей — ценой какой потери сжатия?

02 Сначала память, потом MAC: трафик DRAM на пиксель при нейросетевой постобработке

03 Стоимость на пиксель, при которой нейросетевой инструмент в H.267 перестаёт окупаться

04 Какие замкнутые системы, кроме звонков, первыми перейдут на сквозные нейросетевые кодеки

05 Как измерять достоверность, когда декодер генерирует детали

Эти направления мы можем исследовать вместе — через технический обмен и техническое сотрудничество.

i@tanshuai.com · tanshuai.com