

Worksheet интеграции SAM 2

Ответьте на восемь вопросов перед интеграцией SAM 2 в видео-продукт.

1. Источник промпта — откуда приходит клик?

- ☐ Человек в петле (аналитик, врач, монтажёр)
- ☐ Handoff детектора (bounding box YOLO / RF-DETR)
- ☐ Open-vocabulary детектор (Grounding DINO / Florence-2)
- ☐ Vision-language агент (Gemini / Claude / GPT)

2. Размер чекпойнта — качество vs стоимость

- ☐ Hiera-Tiny (38.9M) — edge / Jetson / экономия
- ☐ Hiera-Small (46M) — edge с чуть лучшим качеством
- ☐ Hiera-Base+ (80.8M) — облачный sweet spot (\$/качество)
- ☐ Hiera-Large (224.4M) — accuracy-bound: VFX / медицина

3. Target деплоя

- ☐ Облако (single-stream propagation)
- ☐ Облако (multi-stream, batched)
- ☐ Edge — Jetson Orin Nano / NX / AGX
- ☐ On-device (ноутбук / desktop с TensorRT FP16)

4. Типичная длина видео и cadence ресета памяти

- ☐ Короткие клипы (≤ 300 кадров): ресет не нужен
- ☐ Средние (300–2,000 кадров): ресет каждые 300 кадров
- ☐ Длинные ($> 2,000$): SAM2Long или chunked propagation

5. Multi-object scope

- ☐ Один объект на клип — propagation достаточно
- ☐ Несколько (2–5) — отдельные propagation на объект
- ☐ Много (> 5) — рассмотрите SAM 3 (concept segmentation)

6. Occlusion-стратегия

- ☐ Порог occlusion score (рекомендуем: 0.7)
- ☐ Не выдавать маску выше порога
- ☐ Выводить low-confidence кадры на ручной просмотр

7. Аудит лицензий

- ☐ Apache 2.0 на весах — коммерция OK
- ☐ CC BY-SA на SA-V — copyleft на fine-tune'ax
- ☐ При fine-tune фиксировать data lineage до публикации

8. Fallback при отказе SAM 2

- ☐ Залипание на distractor: память SAM2.1++
- ☐ Long-video drift: SAM2Long или chunked re-prompt
- ☐ Тонкие края (волосы): handoff к ручному roto
- ☐ Multi-instance concept: эскалировать в SAM 3