

III РСХБ



БИЗНЕС
МОСКВЫ



ALOLA 1

Задача №10 | РСХБ | ЛЦТ 26

Сканер российских вин для «Своё Вино»





О команде

- **Капитан:** Шмелев Александр Сергеевич, IT
- **Кол-во участников:** 2 человека
- **Краткое описание:**
 - команда появилась в день рождения капитана
 - встраиваемые системы, цифровое ТВ, компьютерное зрение, консультирование, разработка ПО
- **Город и регион:** Москва

КОМАНДА «ALOLA 1»

Краткое описание решения:

- «Что за вино?» распознаёт вино по фото и ведёт покупателя в каталог «Своё Вино».
- Для РСХБ это новый канал привлечения аудитории в момент выбора у полки.
- Сайт, Веб/PWA-приложение, Telegram-бот и Android с офлайн-режимом.

Уникальность решения:

- Распознавание одной бутылки или всей полки
- Добавление новых вин без переобучения модели
- Распознавание контрэтикетки по штрихкодам и QR
- Валидация датасета через реестр алкогольной продукции разрешенной к продаже в РФ
- Портал для разметки, подготовки тестовых датасетов, автоматизации бенчмарков, оценки моделей
- Интеграция с API vino-svoe.ru уже реализована. Поддерживается автоматическая синхронизация каталога и изображений.



Александр Шмелев

- Разработка, архитектура
- [@slmb1o](#)
- +79253271272
- Финалист конкурса управленцев „Лидеры России“, сезон 2023–2024



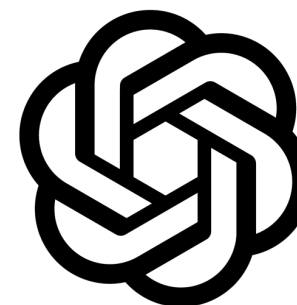
Ольга Шмелева

- Разметка и проверка данных
- [@olgansh15](#)
- +79166953115
- Финансовый аналитик



Qwen3.8-27B

- Re-rank
- DGX Spark
- Автоматический сбор датасета



ChatGPT Sol 5.6

- Графический дизайн
- Написание кода
- Автоматический сбор датасета



Claude Opus 5.5

- Написание кода
- Review кода

Краткая история команды:

01

Команда опирается на инженерный опыт капитана Александра Шмелева в разработке ПО, встраиваемых системах и компьютерном зрении. За плечами Александра МГТУ им. Н. Э. Баумана с отличием и MBA по стартап-менеджменту. Александр финалист конкурса «Лидеры России» 2023-2024.

В мае 2026 года он участвовал в команде, занявшей второе место на Lenta Tech Life Hack с решением для распознавания ценников по видео.

Почему вы выбрали именно эту задачу из предложенных на хакатоне?

02

Задача соединила наш опыт компьютерного зрения с работой над каталогом алкогольной продукции. Нам интересно превращать сложные технологии в понятный инструмент: человек фотографирует бутылку и получает информацию о вине. Здесь мы смогли применить наработки распознавания, обогащения и вычитки каталогов, автоматизируя поиск ошибок, дублей и недостающих сведений. Нас привлекла и возможность дальнейшего применения технологии в рознице: создание планограмм и контроль соответствия фактической выкладки планограмме по фотографиям полок.

С какими основными сложностями или вызовами вы столкнулись и как их преодолели?

03

Главную сложность создали почти одинаковые этикетки, блики и необычные ракурсы, а также неоднозначности в исходных данных. Мы проверили каталог, собрали тестовые фотографии из реальных условий и автоматизировали выделение бутылок и этикеток. Сложные случаи разбирали отдельно, а изменения сравнивали на воспроизводимых тестах. Для пополнения каталога выбрали поиск по визуальным признакам, который позволяет добавлять новые вина без переобучения модели.

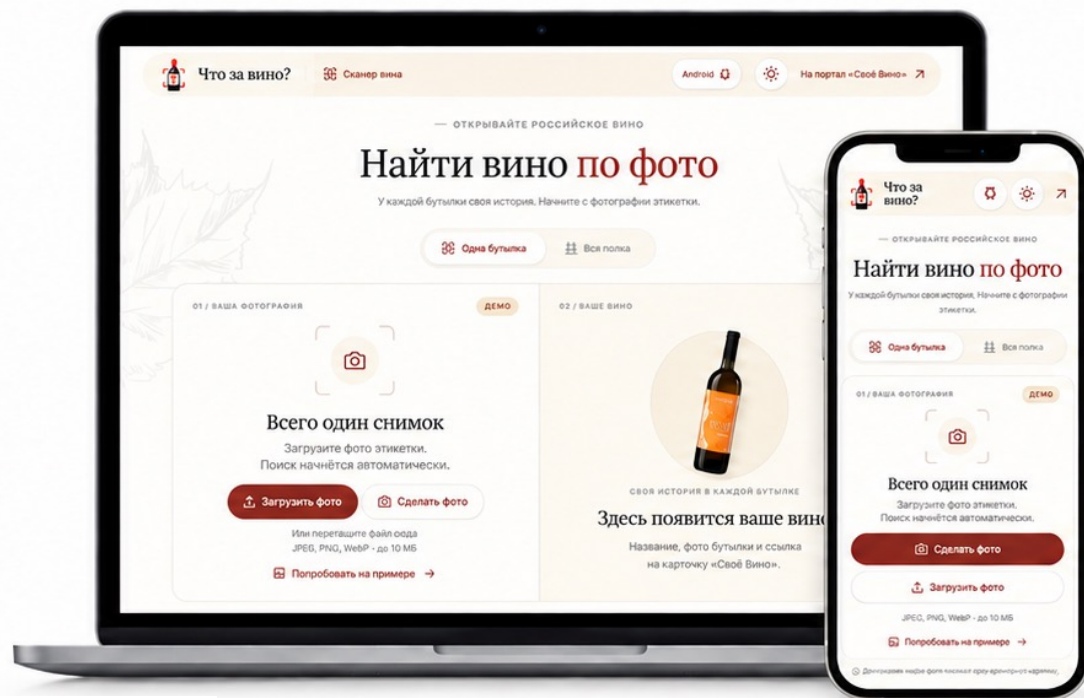
Техническая суть решения

- **Распознавание по нескольким сигналам.** SigLIP2 находит визуально похожие вина, VLM выделяет различия этикеток и уточняет выбор. Штрихкоды и QR быстро связывают фотографию с карточками каталога.
- **Чтение сложных штрихкодов.** Расширенный сервис использует OCR/VLM для чтения напечатанных цифр, когда обычного декодирования недостаточно.
- **Инструменты повышения качества.** Развитый сервис доразметки объединяет автоматическую подготовку изображений, проверку соответствия карточкам, разбор похожих вин и исправление ошибок.
- **Проверяемое развитие.** Benchmark harness сравнивает различные pipeline на единых наборах данных: измеряет точность и скорость, сохраняет результаты и помогает находить причины ошибок.

Маркетинговая суть решения

- **Повышение культуры потребления российского вина.** Помогаем пользователям разбираться в характеристиках, сортах и происхождении вина и делать осознанный выбор. Дальнейшее развитие — гастрономические сочетания с отечественными продуктами и истории о производителях, которые знакомят с винодельческими традициями и гастрономическим разнообразием регионов России.
- **Автоматизация для розницы и производителей.** Предлагаем развивать сервисы обогащения и проверки каталогов на базе наших наработок, контроля ценников, состояния полок и соблюдения планogramм. Ценность для бизнеса — сокращение ручной работы и быстрое выявление ошибок.
- **Совместные пилоты с РСХБ.** Объединяем опыт компьютерного зрения и работы с данными для создания edge-решений с обработкой фото и видео на камерах, мобильных устройствах и локальном оборудовании. Это снижает зависимость от интернет-соединения и позволяет оперативно выявлять отклонения — при контроле полок в магазинах, оценке качества продукции, мониторинге растений и объектов аквакультуры..

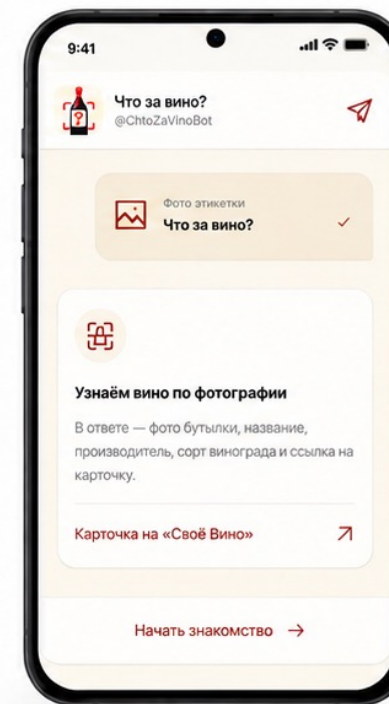
Продукты



+ PWA режим



Android App
Без интернета
Даже на парковке



Telegram Bot
Фото в чат. Вино в ответ.



WebApp
Desktop + Mobile

Качество распознавания



Mode	F1@1	F1@5
Без VLM re-rank	93%	99%
Qwen3.5 0.8B	94%	99%
Qwen3.5 9B	96%	99%

Для фото вин из каталога при 100% ответов $F1@k = R@k$.
Значения округлены.

Одна бутылка и группа: матчинг с каталогом

Фото бутылки или полки. Результат через wine_slug связан с каталогом «Своё Вино»

Одна бутылка



Фото бутылки

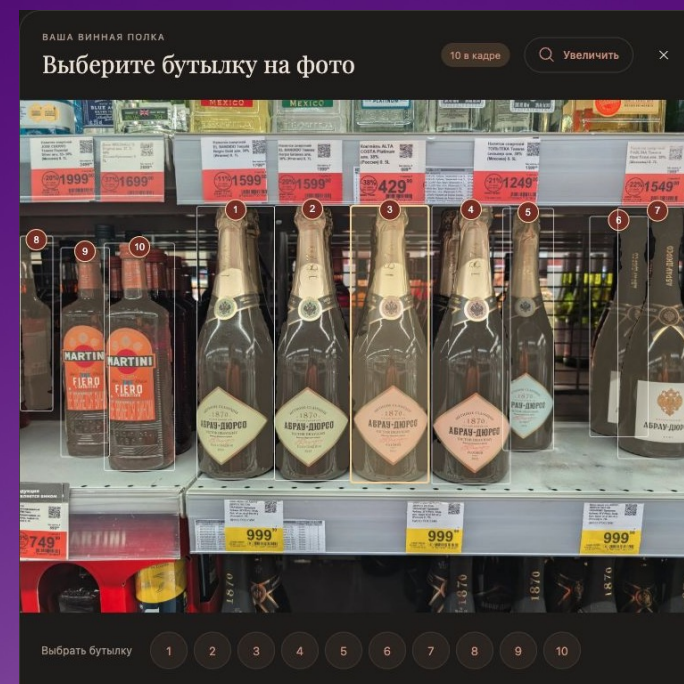


Эталон каталога

«Балаклава Chardonnay Brut»

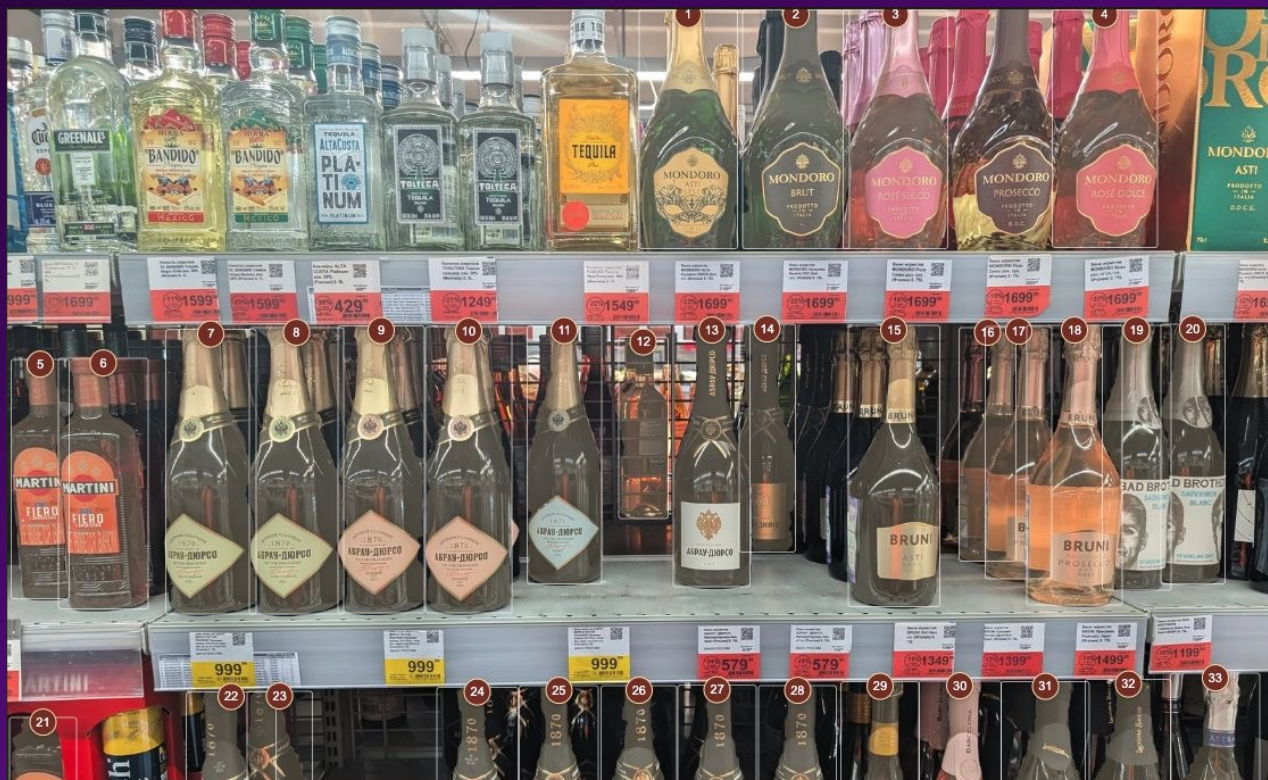
Фото и эталон одного SKU

Группа на полке

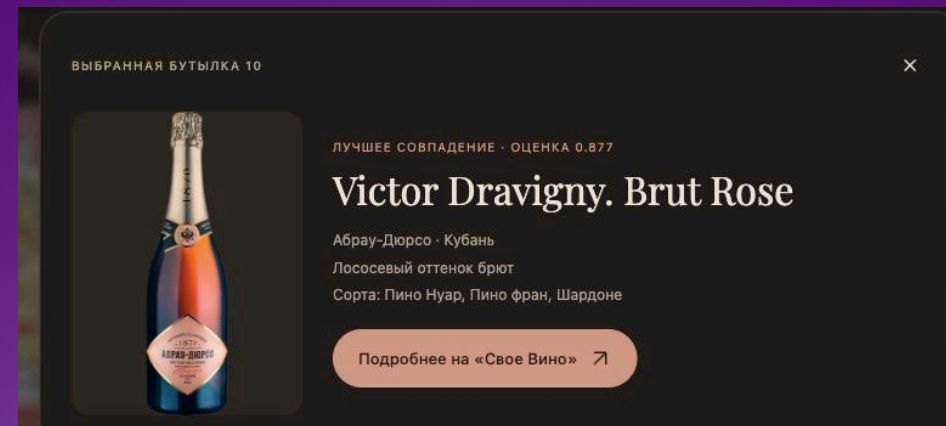


Области бутылок и готовые совпадения.
Выбор открывает результат без нового запроса.

Один снимок всей полки



Все найденные бутылки на фото



Выбор сразу открывает результат

Достоинства решения



- **Новые вина без переобучения**

Для расширения каталога достаточно добавить эталонные изображения и обновить поисковый индекс.

- **Несколько сигналов для распознавания**

Визуальное сходство, различия этикеток, штрихкоды и QR помогают уточнять выбор среди похожих вин.

- **Меньше ручной работы**

Автоматизируем подготовку изображений и извлечение признаков; спорные случаи проверяем в сервисе разметки.

- **Измеримое качество.**

Сравниваем модели и pipeline на единых тестовых наборах по точности и скорости, разбираем причины ошибок.

- **Доступность в разных условиях**

Web (включая PWA режим), Telegram бот и приложение Android; мобильное приложение поддерживает распознавание без интернета.

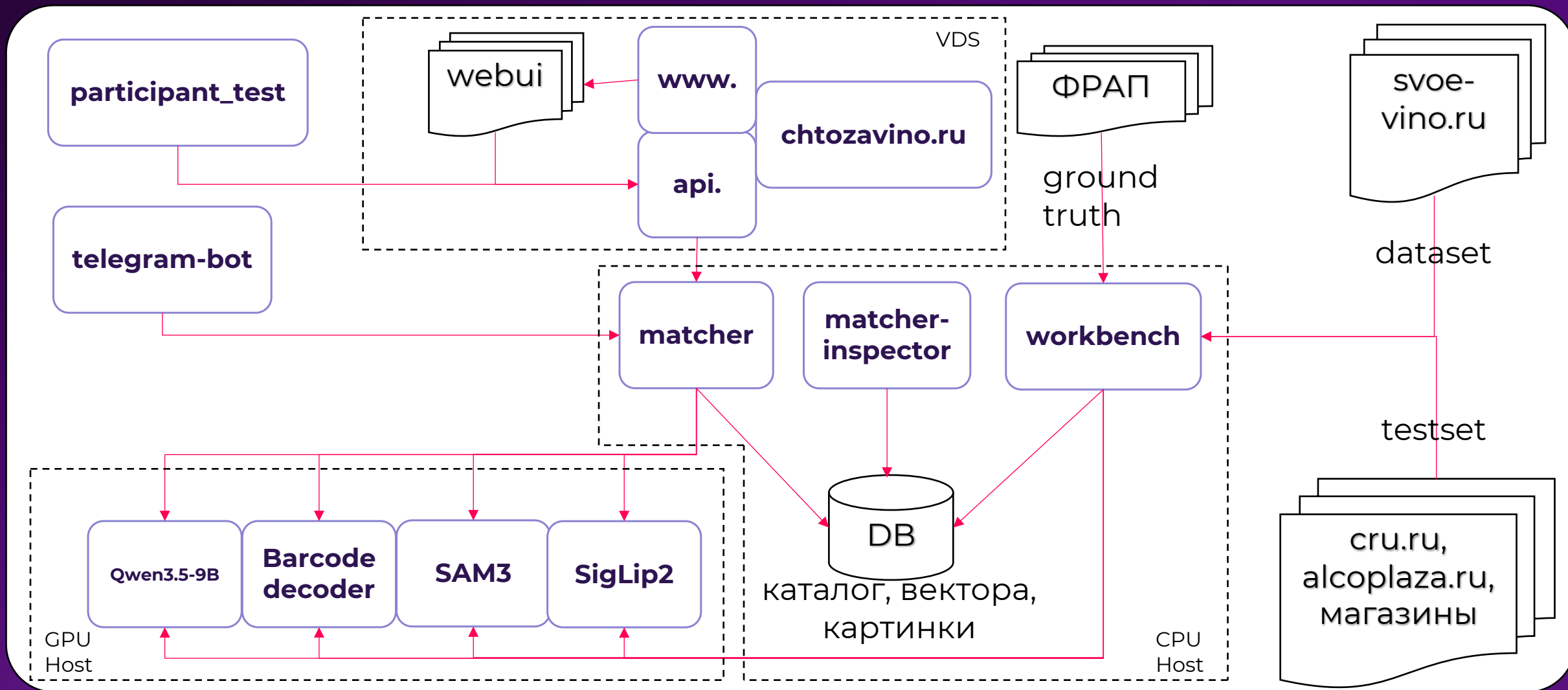
- **Наличие интеграции с API vino-svoe.ru**

Решение готово к пилоту

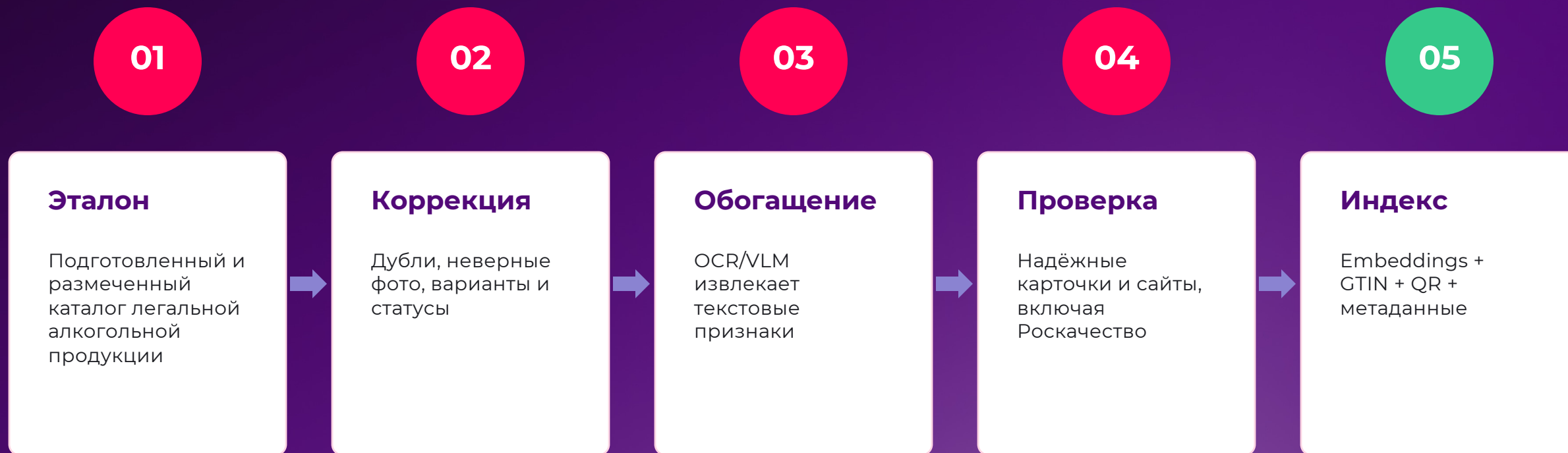
- **Наличие API для интеграции с другими сервисами**

API позволяет встроить распознавание в сервисы партнёра (например, для планограммы), а отдельные компоненты — развивать и заменять независимо.

Архитектура



Подготовка данных до модели



OCR — источник признаков, а не автономная истина.

Несколько эталонов одного вина

Фронт, оборот и каталожное фото ведут к одной карточке товара



Каталожное фото
"Свое вино"



Фронт



Оборот

**AGORA
Yachting
Cabernet Sauvignon**



**Одна идентичность
в каталоге**

Внутри view берём
лучший cosine среди
эталонов вина.

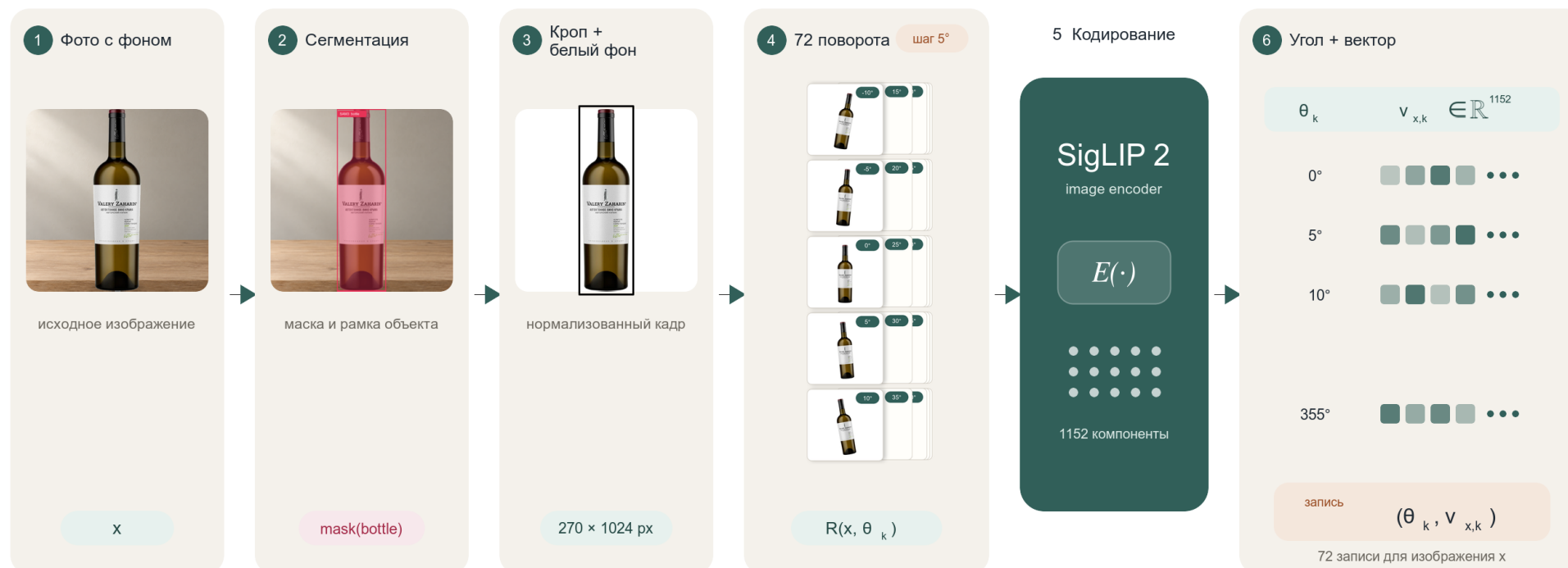
Новый ракурс добавляем в индекс без обучения нового класса

Подготовка фотографий

Подготовка фото → повороты → SigLIP 2 → векторы

$$\theta_k = 5^\circ \cdot k, k = 0, \dots, 71$$

$$v_{x,k} = E(R(x, \theta_k)) \in \mathbb{R}^{1152}$$



- SigLIP 2: только RGB, без alpha-канала → прозрачность заменяем белым фоном
- Ракурс эталона $\neq$ ракурс кандидата → косинусная близость: 1.00 → до 0.75 (−0.25)

* Все выводы — по результатам тестов.

Stress-test: реальные фото, а не витрина

**3 354**

уникальных изображения в собранном наборе

2 228

запросов в фиксированном benchmark snapshot

Сложные случаи

Блики · наклон · кроп · полка · похожие этикетки

Источники: Яндекс.Картинки, DuckDuckGo, iRecommend, Otvovik, alcoplaza, cru, cigarpro и другие сайты.

Как выбрать товар на сложном фоне

Сначала выбираем основную упаковку, затем сравниваем её с каталогом



**В обоих проверенных кадрах
selector выбрал банку**

Сигналы выбора

Контакт с рукой **0,40**

Площадь объекта **0,22**

Положение в кадре **0,14**

Остальные признаки **0,24**

Нет приоритета класса «бутылка».
Учитываем резкость, уверенность детектора и
другие признаки.

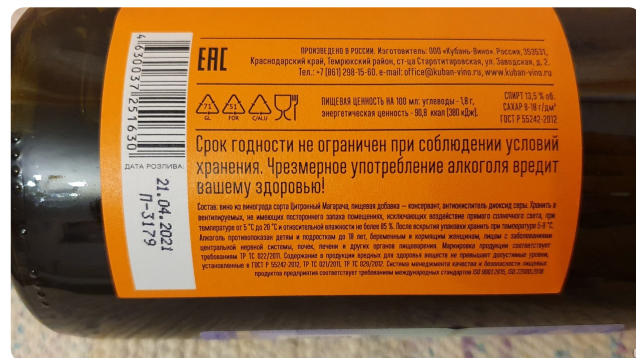
Основной путь: ответ без лишних моделей



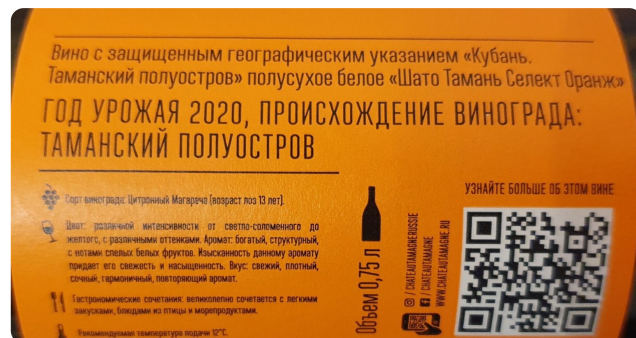
Экспрес-путь завершается сразу после декодирования штрихкода

VLM не работает на каждом запросе. Он включается только для похожих этикеток.

Экспресс-путь



При успешном декодировании поиск по эмбедингам не требуется



EAN-13 на фотографии 24 Мп распознаётся примерно за 0,2 с

** MacBook Pro M2 Max · медиана 10 запусков*

Re-rank по опроснику различий

Подготовка данных offline (автоматическое)

1. Кластеризация датасета по схожести



2. Извлечение признаков с этикеток



VLM

texts:
"30/70",
"750ml"
marks: "red
sticker at
bottom left"



VLM

texts:
"50/50",
"750ml"
marks: "brown
sticker at
bottom left"



VLM

texts:
"70/30",
"750ml"
marks: "magenta
sticker at
bottom left"

Qwen3.5-9B

Qwen3.8-27B
или Qwen3.8-Max

3. Поиск различий

LLM

Поиск уникальных
различий и вывод правил

If the label's bottom-left square shows 30/70, it is Card A; if it shows 50/50 (printed as 8/8), it is Card B; if it shows 70/30, it is Card C.



Online сопоставление (matching)

4. Re-rank

VLM анализирует каждого кандидата из кластера и отвечает на вопрос о различиях этикеток. На основе ответов система уточняет порядок кандидатов в выдаче.



Производительность и ресурсы

GX10 (DGX Spark): 250 тестовых фото, 4 одновременных запроса

Режим	Задержка ¹ , с / фото	Серверные часы на 1 000 фото ²
Без VLM	2,853	0,1981
Qwen3.5 0,8B	4,140	0,2875
Qwen3.5 9B	4,570	0,3174

Подготовка каталога, секунд / вино

GX10 (DGX Spark)	4,09
RTX4090 48 ГБ	3,40

Шаг поворота эталона 5°: 72 вектора (0–355°).

55 вызовов VLM на 250 фото.
Вызов только при наличии схожих вин.

Токены Qwen3.5-9B LLM/VLM на операцию

Этап	Единица	Вход	Выход
Индекс: SAM3, SigLIP2 (шаг 5°)	вино	0	0
Этикетка (VLM):	карточка	≈ 1 403	≈ 365
Правила (LLM):	кластер	≈ 2 892	≈ 4 023*
Матчинг (VLM):	вызов VLM	≈ 1 818	≈ 20

Средние отдельных замеров. Матчинг при 1536 px.
* Выход для правил включает токены рассуждений.

¹ 9B: полный прогон 289 с. 0,8B: 55 вызовов и общий этап до VLM..
² часов сервера / 1000 фото = средняя задержка, с × 1000 / (4 × 3600). Цена = часы × тариф.
Обе модели Qwen3.5 используют NVFP4 на DGX Spark.

Что готовим заранее и что работает на телефоне

ВНЕ УСТРОЙСТВА (заранее)

Workbench · desktop GPU

- 1 **Каталог**
проверка карточек, GTIN и QR
- 2 **Эталонные изображения**
DIS · crop · белый RGB
- 3 **Векторный индекс**
SigLIP2 Base/16 224 на GPU

Результат: готовый пакет для приложения

MODEL PACK

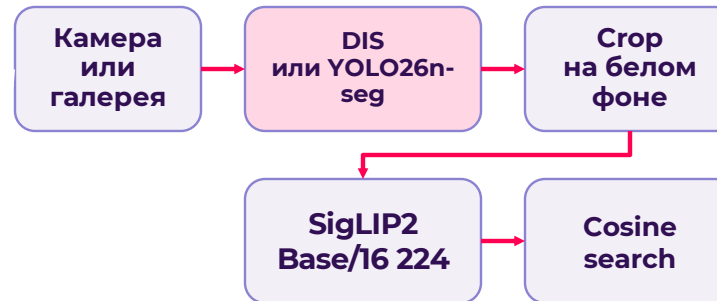
DIS
YOLO26n-seg
SigLIP2

вектора
метаданные ·
штрихкоды
превью

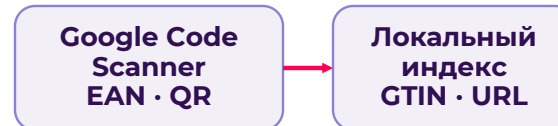
НА УСТРОЙСТВЕ

Android 9+ · обработка фото без интернета

ФОТО



ШТРИХКОД или QR



Результат и история остаются на телефоне



Готовая интеграция с API vino-svoe.ru

Автоматический импорт каталога через API

Уже реализовано	Возможности
Загрузка каталога	Карточки вин и оригинальные фотографии из API
Сверка изменений	Новые, исчезнувшие и возвращённые вина
Контроль обновлений	Сравнение полей и фото; выбор при конфликтах

Новые вина и изменения не нужно переносить вручную

Дополнительные функции

Уже работает

В текущих приложениях

Приложение Android

Работает без интернета, и готово к публикации в Google Play и RuStore

Обратная связь

Подтверждение правильного матчинга и ещё три варианта ответа в Telegram.

Интерактивные демо

Доступно на сайте для ограниченного числа примеров

Гастрономические пары

Блюда региональных кухонь России.

Этикетка понятным языком

Объяснение сортов и терминов.

Паспорт вкуса

Характеристики и ароматы вина.

Путеводитель по линейке

Знакомство с коллекциями одного производителя.

Идеи развития

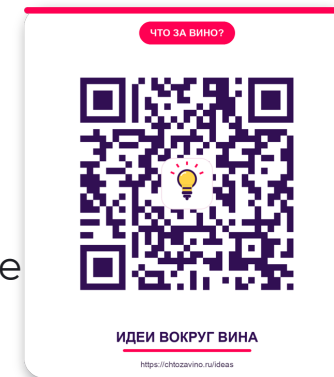
Макеты будущих сервисов

Российские продукты к выбранному блюду в магазине рядом. Подбирает отечественные продукты из ассортимента магазина (GPS + API ритейла)

Поиск похожих вин по вкусу

Контроль планogramм и обнаружение аномалий в выкладке (недокомплект, отсутствие, несоответствие ценников)

Аудиоистории о вине к застолью (LLM генерирует рассказ)



Подробнее: <https://chtozavino.ru/ideas>

Распознавание по фото и переход в каталог «Своё Вино»

Как проверяем

01 Согласование

Выборка вин и реальные фото.
Критерии приёмки до начала теста.

02 Проверка сценариев

Одна бутылка, полка, контрэтикетка.
Отдельно проверяем вина вне каталога.

03 Ограниченный запуск

Тестовая витрина с API распознавания.
Учёт переходов в карточки «Своё Вино».

От РСХБ Владелец пилота, тестовый контур
и согласованная выборка вин.

Что измеряем

Качество

Топ-1, ложные совпадения.
Для полки отдельно:
обнаружение и верный матчинг.

Скорость

p50 / p95 времени ответа
при 1 и 4 запросах.

Пользовательский эффект

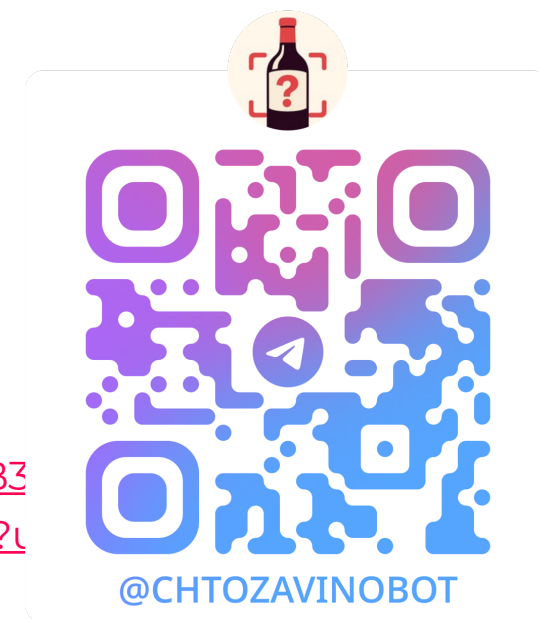
Доля сессий с переходом
в карточку найденного вина.

Результат пилота

Отчёт и решение о расширении.

Ссылки

- Сайт «Что за вино?» - <https://chtozavino.ru/>
- Телеграм-бот «Что за вино?» - <https://t.me/ChtoZaVinoBot>
- Android приложение - <https://chtozavino.ru/android>
- Документация: <https://github.com/s1mb1o/svov-vino-lab-dev#readme>
- Прототип: <https://chtozavino.ru/hackaton>
- GitHub repo: <https://github.com/s1mb1o/svov-vino-lab-dev> tag: hackaton-snapshot-2909
- Google Drive: <https://drive.google.com/drive/u/0/folders/1YfZ9ghYz8yJpznCIEiqgKho64EB3>
- ML модели: https://drive.google.com/drive/folders/1k_suN_6i6jaaS_tiLsm5xUypy5vh4HQ?u



Приложение. Проверка

Для доступа к /v1/eval/predict используется Basic Auth:

- Имя пользователя: evaluator
- Пароль: 00f9be74a3537041291d62d4d9c4198b1becf6385f557f1556c4c965c1417580

Запуск проверочного скрипта:

```
MATCHER_BASIC_PASSWORD="00f9be74a3537041291d62d4d9c4198b1becf6385f557f1556c4c965c1417580"
./participant_test.sh \
  --images-dir ./queries \
  --manifest ./queries.tsv \
  --endpoint "https://evaluator:${MATCHER_BASIC_PASSWORD}@chtozavino.ru/v1/eval/predict" \
  --output ./predictions.jsonl
unset MATCHER_BASIC_PASSWORD
```

Спасибо за
ваше
внимание !

