

AI PartsFinder: интеллектуальный поиск запчастей по фото

Решение для оптимизации
логистики и продаж в
компании «Интерполе»

Евгений Зорин, ML Engineer



Проблема, с которой мы сталкиваемся: Поиск запчастей — узкое место

Растущая номенклатура

Более 44 000 наименований запчастей, база постоянно растет, усложняя идентификацию.

Сложность для клиентов

Клиенты не всегда могут описать нужную деталь, что приводит к потере времени и отказу от покупки.

Нагрузка на персонал

Менеджерам и складским сотрудникам сложно идентифицировать детали "на словах", особенно новичкам.

⊗ Итог: замедление логистики, потери заказов, рост издержек.

AI PartsFinder: для кого это решение?



Клиенты

Фермеры и механизаторы смогут быстро найти нужную деталь прямо в поле, избегая простоя техники.



Менеджеры по продажам

Сократят количество ошибок при оформлении заказов и ускорят обслуживание клиентов.



Складские сотрудники

Легко определяют детали при приемке и сортировке, особенно новые сотрудники.

Наше решение: AI PartsFinder в действии

1

Загрузка фото

Клиент или сотрудник делает снимок необходимой детали.

2

Интеллектуальный поиск

Модель AI выделяет деталь на изображении и сравнивает ее с нашей базой данных.

3

Топ-5 кандидатов

Система выдает 3-5 наиболее похожих вариантов деталей с артикулами.

4

Быстрый заказ

Возможность немедленного заказа или уточнения информации с менеджером.



Результат: значительное сокращение отказов и ускорение обработки запросов.

Архитектура MVP: как это работает



- **Извлечение признаков (ResNet50):** Создание уникальных векторов для каждой детали.
- **База векторов (FAISS):** Быстрый поиск ближайших соседей среди десятков тысяч запчастей.
- **Веб-интерфейс (Flask + Tailwind):** Интуитивно понятный интерфейс для пользователей и менеджеров.

Почему embedding-поиск?

| | | |
|--------------------------------|--|--|
| Классификация (40 000 классов) | Высокая точность на фиксированной базе. | Огромные ресурсы на обучение, переобучение при добавлении новых деталей. |
| Embedding (ResNet + FAISS) | Легко расширять базу, работает с десятками тысяч классов, быстрый поиск, высокая масштабируемость. | Незначительно ниже точность на самых сложных случаях. |

✓ Выбор: embedding-поиск — наиболее гибкое, масштабируемое и перспективное решение для «Интерполе».

Как работает embedding-поиск?



Обучение/Подготовка модели

Используем (или дообучаем) ResNet50 для извлечения высококачественных векторных представлений (эмбеддингов) изображений.



Формирование векторной базы

Для всех 40 000+ деталей «Интерполе» сохраняем их уникальные эмбеддинги в FAISS.



Запрос от пользователя

Клиент делает фото детали, и AI PartsFinder извлекает ее эмбеддинг.



Поиск ближайших соседей (kNN)

Сравниваем полученный эмбеддинг с нашей векторной базой, находя топ-N наиболее похожих запчастей.

Этот подход обеспечивает быстрый, точный и масштабируемый поиск, даже для огромных каталогов.

Почему ResNet50: Основа нашего embedding-поиска

Преимущества ResNet50

- **Доказанная эффективность:** Одна из лучших архитектур для извлечения признаков.
- **Предобученность:** Уже "знает", как распознавать объекты, благодаря обучению на ImageNet.
- **Адаптивность:** Легко модифицируется для использования в качестве feature extractor.
- **Быстрый инференс:** Подходит для производственных систем.

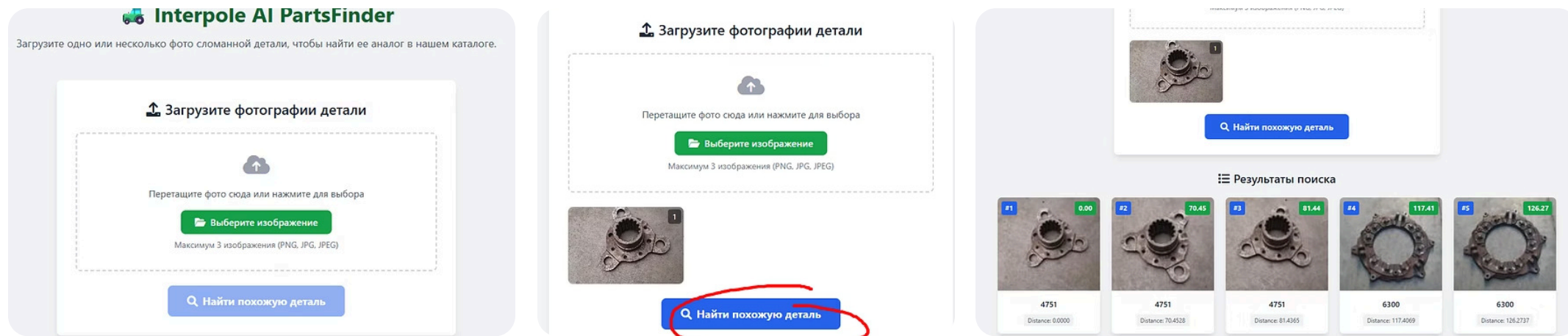
Как формируются эмбединги

ResNet50, изначально предназначенный для классификации (с 1000 выходных нейронов), модифицируется путем **удаления последнего слоя**.

В результате мы получаем **вектор размерности 2048** — это уникальный "отпечаток" изображения, его эмбединг.

Изображения похожих объектов будут иметь **близкие векторы** в этом многомерном пространстве, что позволяет точно находить совпадения.

AI PartsFinder: демонстрация MVP



Слева — фото, загруженное клиентом или сотрудником. Справа — найденные кандидаты из нашей базы данных, отсортированные по степени схожести.

Экономическая эффективность: осязаемые выгоды

2-3x



Сокращение отказов

AI PartsFinder снизит отказы от заказов, повышая удовлетворенность клиентов.

Рост скорости

Ускорение обработки запросов менеджерами, что приведет к увеличению производительности отдела продаж.

Минимум ошибок на складе

Оптимизация работы склада за счет снижения количества ошибок новичков при сортировке.

Увеличение выручки

Прямой рост выручки компании благодаря сохранению ранее потерянных заказов.

Наша команда: движущая сила проекта



Agile подход - ключ к успешному внедрению, так как позволяет команде быстро реагировать на изменения в проекте на всех его этапах и эффективно адаптироваться. Гибкая методология Agile дает возможность постоянно пересматривать и корректировать планы, чтобы соответствовать текущим потребностям бизнеса. Это помогает минимизировать риски и обеспечивает своевременную и качественную реализацию проекта. Agile-подход фокусируется на ценностях, а не на жестких процессах, что позволяет команде быстро реагировать на изменения и адаптироваться к новым условиям.

План реализации



Месяц 1: Данные и Модель

Сбор и предобработка обучающих данных. Разработка и тестирование первой (baseline) версии AI-модели поиска запчастей.



Месяц 2: MVP Интерфейс

Создание минимально жизнеспособного продукта (MVP) веб-интерфейса для демонстрации функциональности поиска по фото.



Месяц 3: Тестирование

Интенсивное тестирование MVP на реальных данных на складе и с участием ключевых клиентов. Сбор обратной связи.



Месяц 4: Пилотный Запуск

Пилотный запуск AI PartsFinder в одном из ключевых регионов для оценки эффективности в боевых условиях.



Месяц 5: Масштабирование

Доработка и оптимизация системы на основе результатов пилота, планирование и начало масштабирования на все регионы.

Риски и пути решения

1

Риск 1: Точность Embedding-поиска

Проблема: Эмбединги могут путать визуально похожие детали, например, шестерни разного диаметра или подшипники.

Решение:

- Комбинировать embedding-поиск с уточняющей моделью (ensemble).
- Использовать YOLO для детекции и обрезки фона.
- Применять методы кластеризации и центроиды (FAISS + centroid vectors).

2

Риск 2: Разнородность и качество фото

Проблема: Фотографии от клиентов/сотрудников имеют разное качество, ракурс и освещение, что может приводить к ошибкам.

Решение:

- Обязательная предобработка: автообрезка, нормализация яркости/контраста.
- Дообучение модели на «грязных» данных, приближённых к реальным фото.
- Применение аугментаций (blur, rotations, lighting shift) при тренировке.

3

Риск 3: Скорость поиска в большой базе

Проблема: Рост базы данных (>50 000 изображений) может привести к увеличению времени отклика и задержкам при поиске.

Решение:

- FAISS с GPU-ускорением (IVF, HNSW индексы).
- Внедрение кэширования популярных запросов.
- Многоуровневый поиск: быстрый грубый поиск → уточняющий поиск.

Стоимость проекта

1

Разработка MVP

≈ 1–2 млн руб.

(1 месяц, команда из 5 специалистов)

2

Инфраструктура

≈ 500 тыс. руб. в год

(Серверы, хранилище данных)

3

Поддержка и Дообучение

≈ 300–600 тыс. руб. в год

(Обслуживание системы, улучшение AI-модели)



Дальнейшее развитие

Дообучение ResNet50 на собственных данных

Полная база фотографий в размере 1.5 млн. фото позволяет выполнить дообучение нейросети, увеличив ее узкую специализацию и повысив точность.

Предобработка фото (YOLO8 nano)

Предварительная детекция детали, центрирование и обрезка фона обеспечит повышение качества входных данных, что, в свою очередь, оптимизирует производительность приложения без увеличения нагрузки.

Аугментации при обучении

Применение поворотов, шумов, изменение яркости/контраста позволит повысить устойчивость эмбедингов к достаточным или не качественным входным данным.

Интеграция с ERP/CRM

Бесшовная связь с существующими системами компании для автоматизации процессов заказа, учета и управления запасами.

Автоматическая классификация

Идентификация и категоризация новых деталей при поступлении на склад, сокращая ручной труд и повышая точность учета.

Мобильное приложение

Удобный доступ к поиску и управлению заказами с любого устройства для клиентов и сотрудников склада для еще большей оптимизации работы.

Расширение функционала AI PartsFinder планируемое в дальнейших версиях, для максимальной эффективности и удобства использования.

Заключение: трансформация бизнеса с AI PartsFinder

AI PartsFinder — это не просто инструмент, это катализатор роста и оптимизации для вашей компании.



Сокращение отказов

Снижение количества потерянных заказов в **2-3 раза** за счет быстрого и точного поиска.



Рост лояльности

Повышение удовлетворенности клиентов благодаря быстрому обслуживанию и отсутствию ошибок при подборе.



Ускорение работы

Значительное ускорение процессов в отделе продаж и на складе, повышая общую производительность.



Основа для автоматизации

Создание прочного фундамента для дальнейшей цифровизации и автоматизации всей логистической цепочки.

Внедряя AI PartsFinder, вы инвестируете в будущее своего бизнеса, обеспечивая его устойчивость и конкурентоспособность.