

Решение кейса разметки видео- контента с использованием ИИ

Автоматизация процесса разметки видео-контента является ключевой задачей для снижения трудозатрат и повышения качества обработки данных. Наше решение позволит быстро и эффективно выделять ключевые объекты, события и сцены, а также проводить анализ тональности, символов и звукового сопровождения.

Терминология

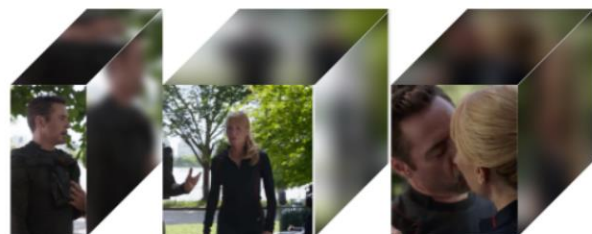
фильм



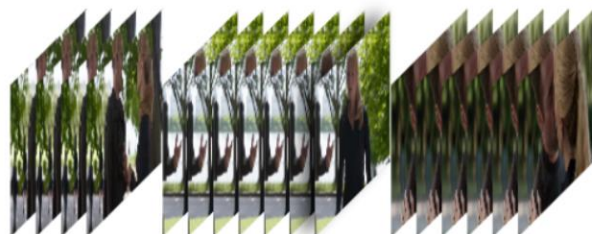
сцены



шоты



кадры



Обоснование выбора инструментария

Анализ объектов

	YOLOv8	SSD	Faster R-CNN
Маленькое потребление ресурсов	+	+	-
Точность детекции	+	-	+

Обработка изображений

	CV2	Scikit-Image	Mahotas
Производительность	+	-	-
Функциональность	+	+	-
Аппаратное ускорение	+	-	-

Классификация изображений

	Inception V3	ResNet-50	VGG-16
Аппаратные требования	+	+	-
Гибкость в доработке	+	+	-
Точность	+	-	-

Анализ окружения

	CLAP	Wav2Vec 2.0	AudioCLIP
Адаптивность	+	+	-
Аппаратные требования	+	-	+

Выделение сцен

	Агломеративная кластеризация	К-средние	DBSCAN
Устойчивость к шуму	+	-	+
Отсутствие необходимости в настройках	+	+	-

Выделение текста из аудио

	Speech Recording	PyAudio	Sounddevice
Простота использования	+	-	-
Поддержка новых форматов	+	-	-

Решённые технические задачи

- 1 Транскрибация аудио ✓
- 2 Описание событий по тексту ✓
- 3 Определение тональности звука ✓
- 4 Определение звукового окружения ✓
- 5 Определение объектов в видео ✓
- 6 Определение окружения объектов ✓
- 7 Определение движения объектов ✓
- 8 Определение масштабов объектов и Зон внимания в кадре ✓
- 9 Группировка шотов в сцены ✓
- 10 Выявление контекста сцены ✓
- 11 Реализация меток(18+/black) ✓
- 12 Поиск по размеченным видео ✓

Применение в бизнес-задачах



Интеграция в сервис модерации(автоматическое определение запрещенного/рекомендуемого контента)



Определение подходящей рекламы на основе разметки

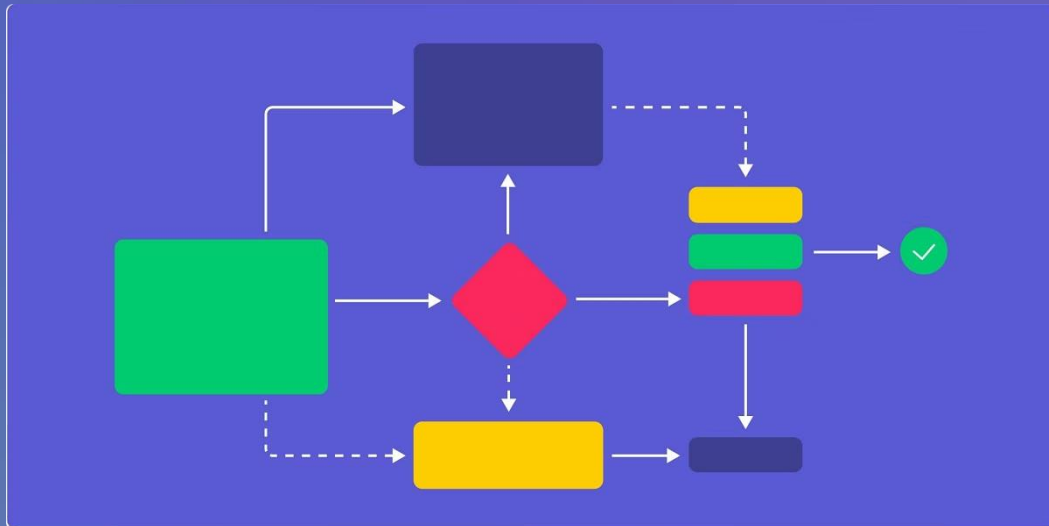


Поиск по размеченным видео



Генерация клипов на основе ключевых моментов и зон привлечения внимания

Алгоритм работы программы



Загрузка видео

1

Пользователь загружает видео на платформу

Компьютерное зрение

2

Нейросетевые модели анализируют видео

Преобразование видео

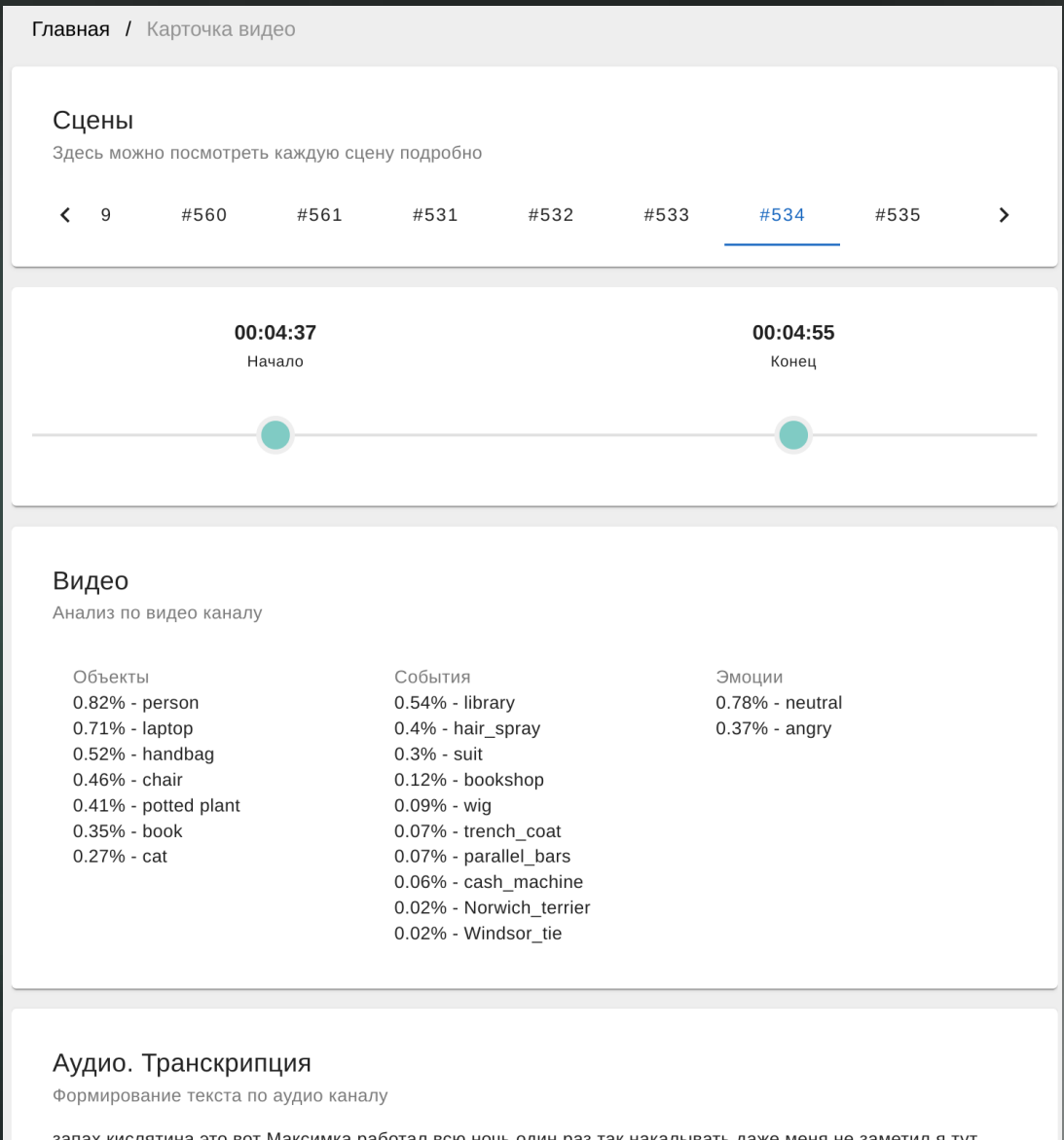
3

Система проводит обработку и группировку видео и аудио

Вывод результатов

4

Пользователь получает размеченные данные



Интерфейс приложения

Видео

Загруженное видео для анализа

Аналитика

Результаты выделения объектов, сцен и событий

Визуализации

Представление разметки в виде удобного описания

Поиск

Возможность найти видео на основе введенного текста или фото

Перспективы развития

Задачи на реализацию

Реализовать:

- Добавление данных в словари меток
- Поиск по фото
- Определение динамики видео

Релиз

Создание полностью готовой версии приложения для разметки видеоконтента

Квантование

Провести квантование весов моделей для уменьшения их размерности

Масштабируемость

Наращивание вычислительных мощностей для обработки больших объемов видео



Наша команда

Шинковой
Данил

shinkovoy2003
@mail.ru



Design, ML

Плаксин
Константин

kos.iaya@mail.ru



NLP

Горшенин
Алексей

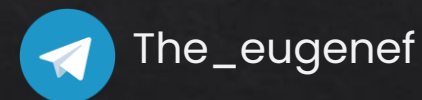
augomgtu
@gmail.com



ML, DS

Франчук
Евгений

rodneyhalo
@gmail.com



Fullstack, ML

Король
Алексей



Fullstack