

ПОТОК

Направить все ресурсы
в бесконечность моментов

Александр Коншин, Иви

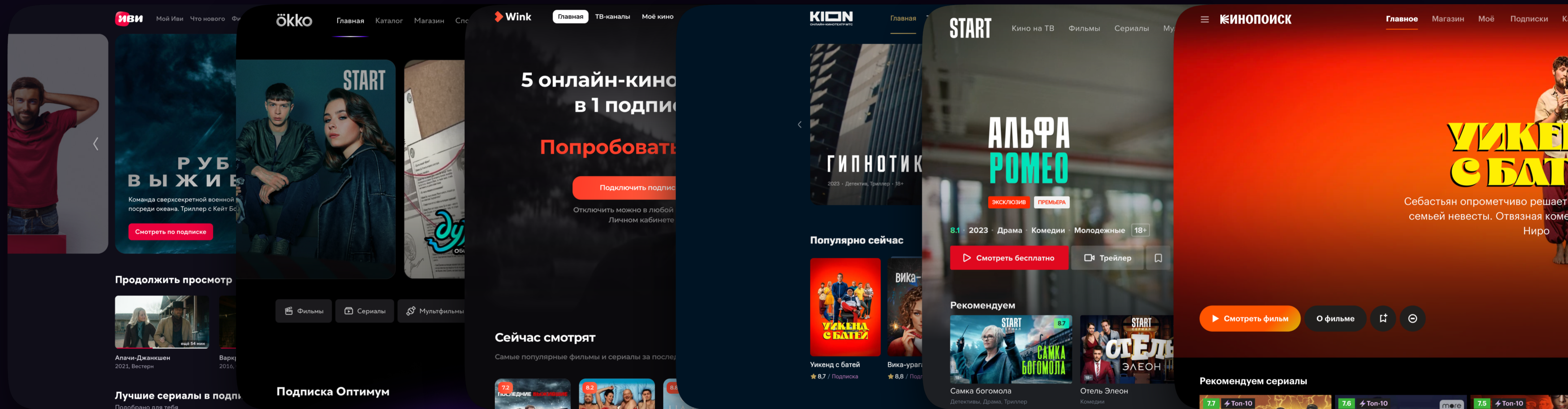


Зачем играть в TikTok

Проблема выбора



- Постер, мета, трейлер
- Пользователь устал, ушёл смотреть Reels



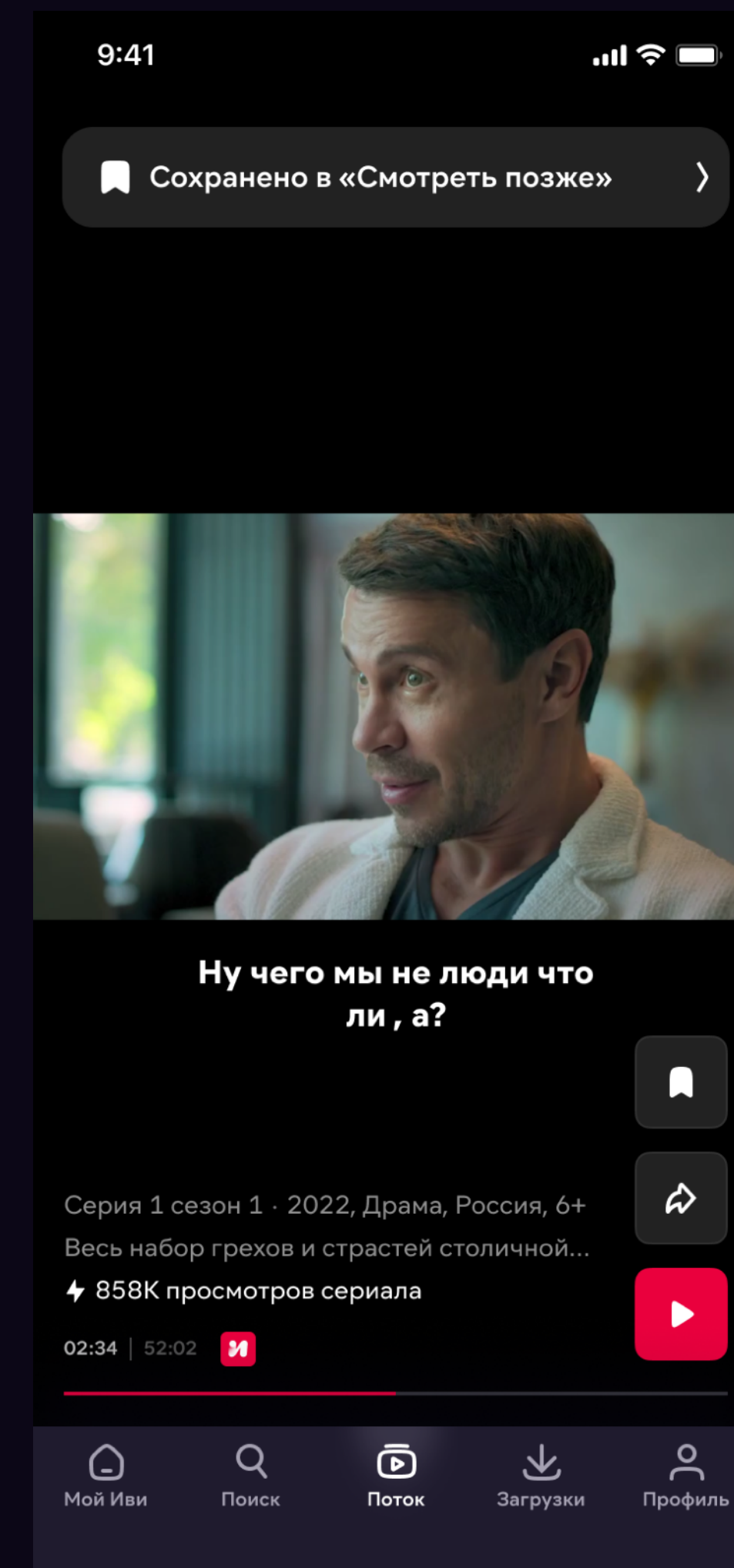
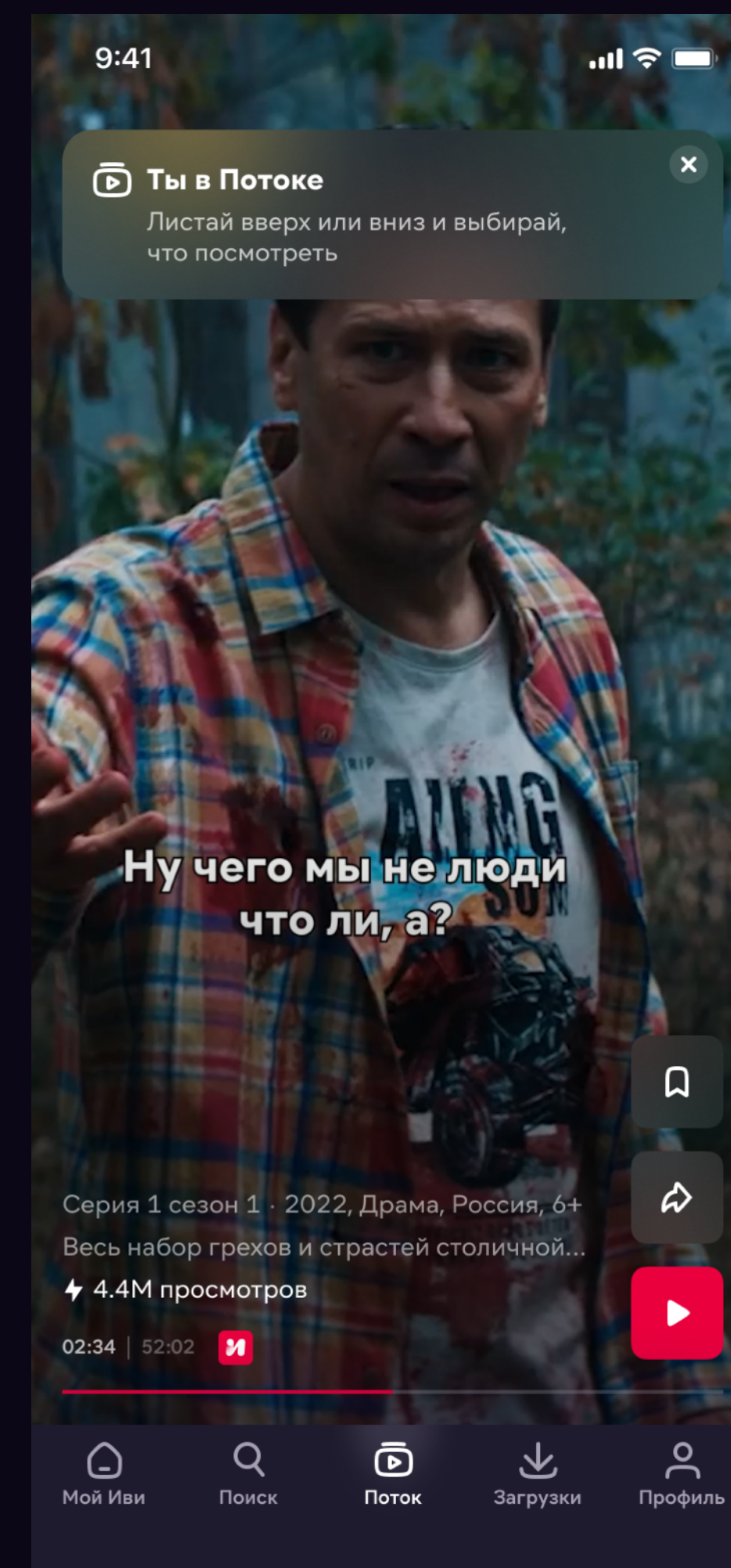


Изменить парадигму выбора

Выбирать в ленте и сразу смотреть

Требования к ленте

- бесконечная (> 4 000 тайтлов)
- по ~8-12 цепляющих сцен
- обязательны русские субтитры



Создавая Моменты



- Кадры — статичные картинки, из которых состоит видео.
- Шоты — это группы кадров от одной монтажной склейки до другой.
- Сцена — группа из одного или более шотов, объединённых по смыслу.



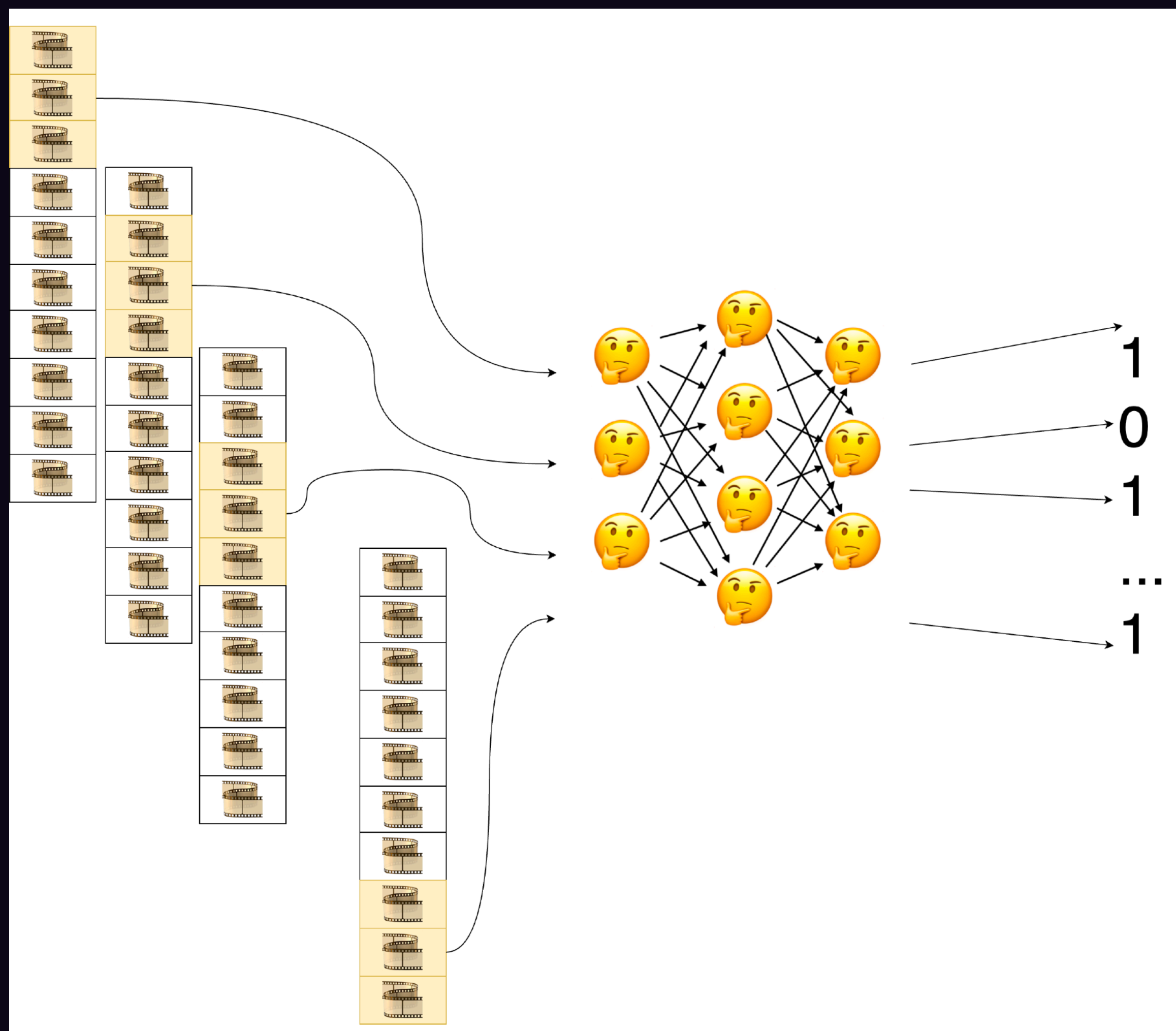
В Потокe используем сцены

Детекция шотов



- TransNetV2
- <https://arxiv.org/pdf/2008.04838.pdf>

Детекция целостных сцен

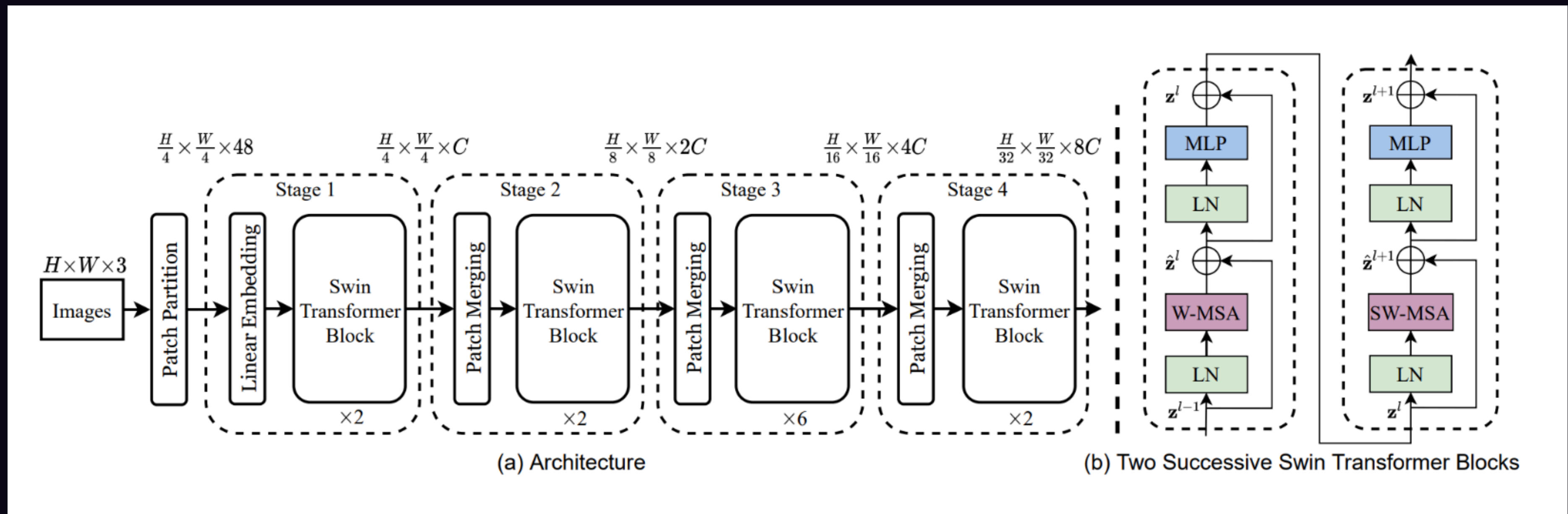


- Делим исходное видео на шоты
- Выделяем из шотов фиичи
- Двигаемся по шотам скользящим окном и последовательно подаём в сеть
- Выход — вероятность того, что шот является границей сцены
- Агрегируем предсказания каждого окна в финальные предикты



Детекция целостных сцен

Backbone

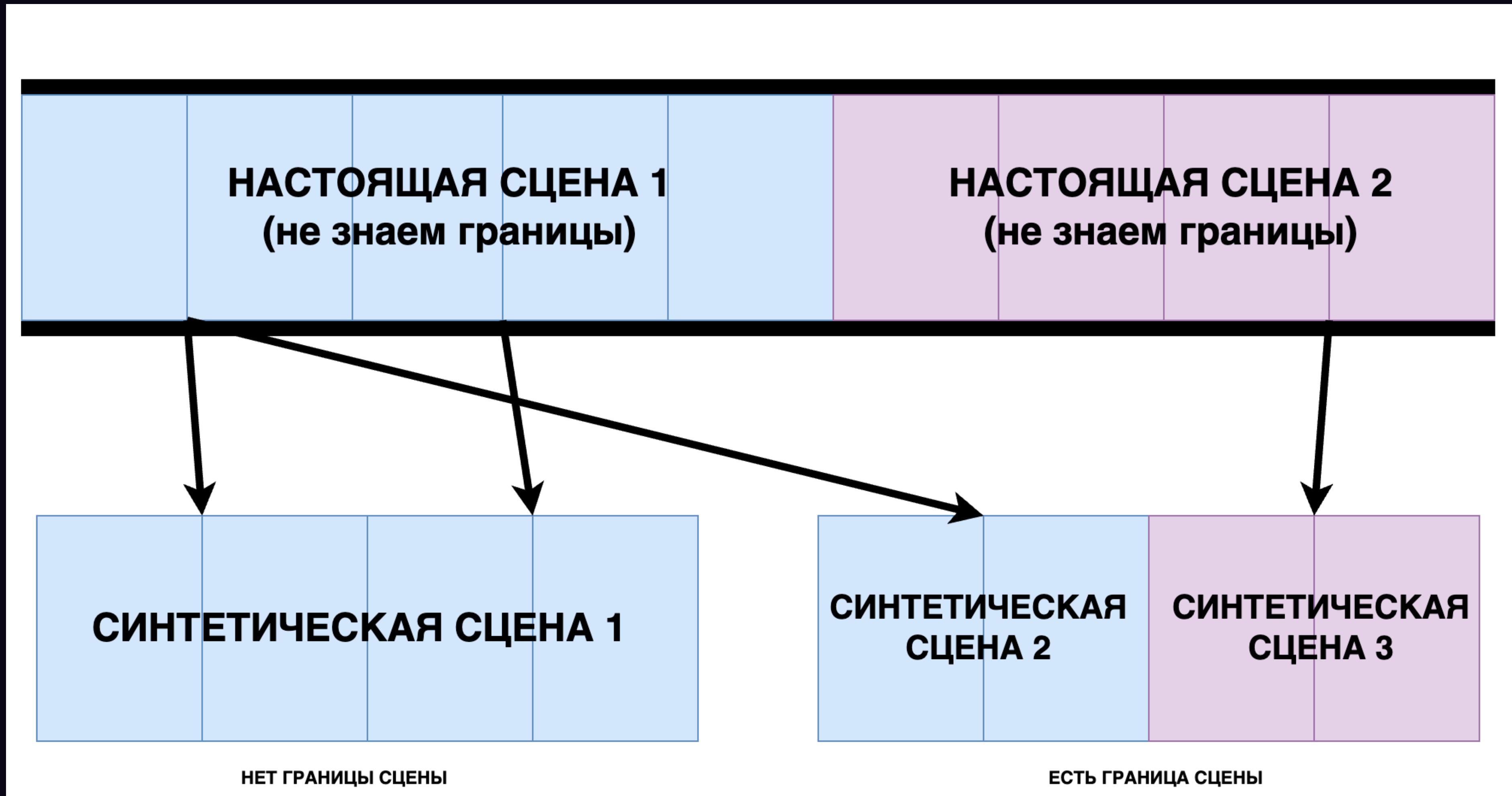


<https://github.com/microsoft/Swin-Transformer>



Детекция целостных сцен

На чём обучали?



Детекция целостных сцен

Пример работы



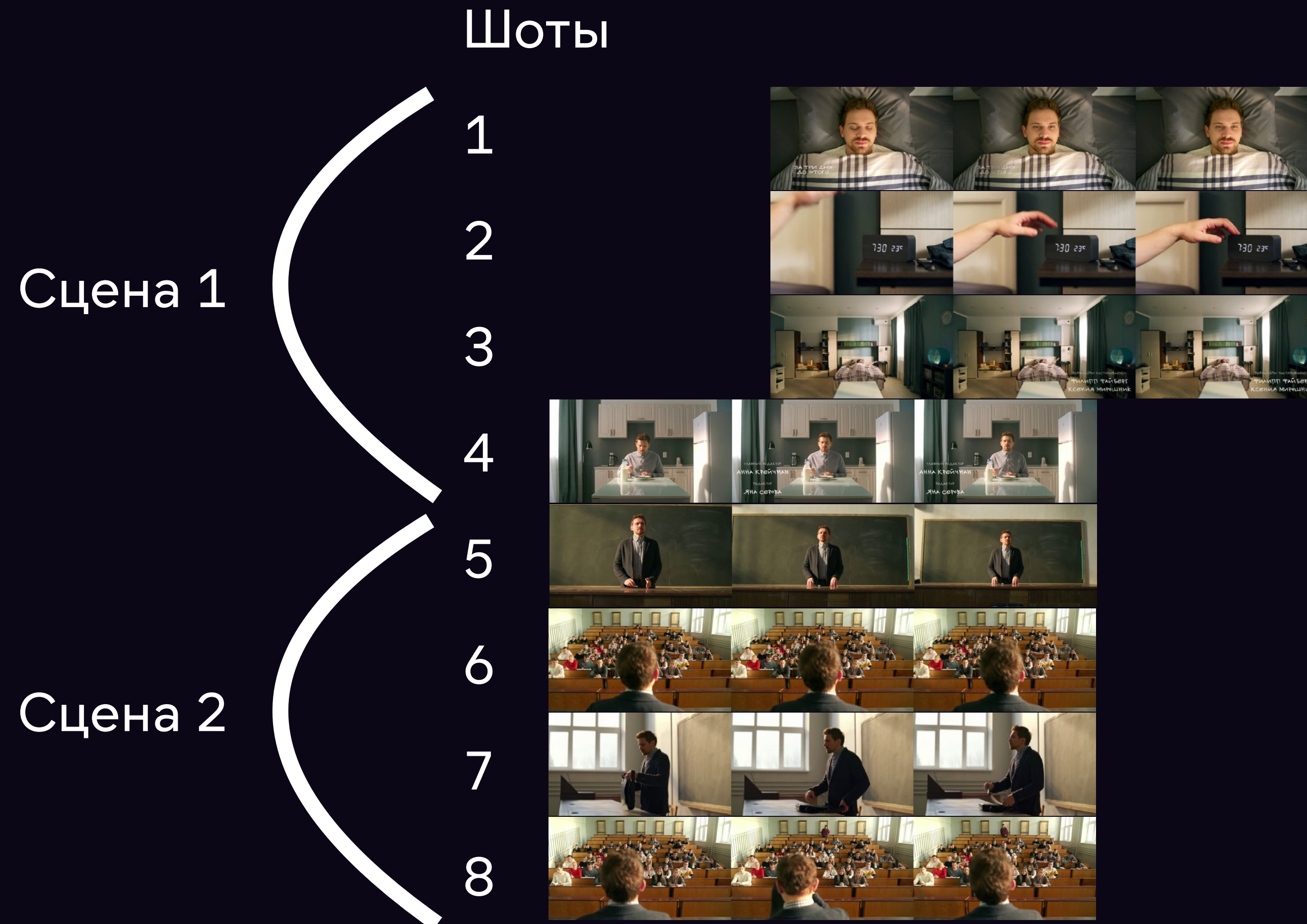
Шоты





Детекция целостных сцен

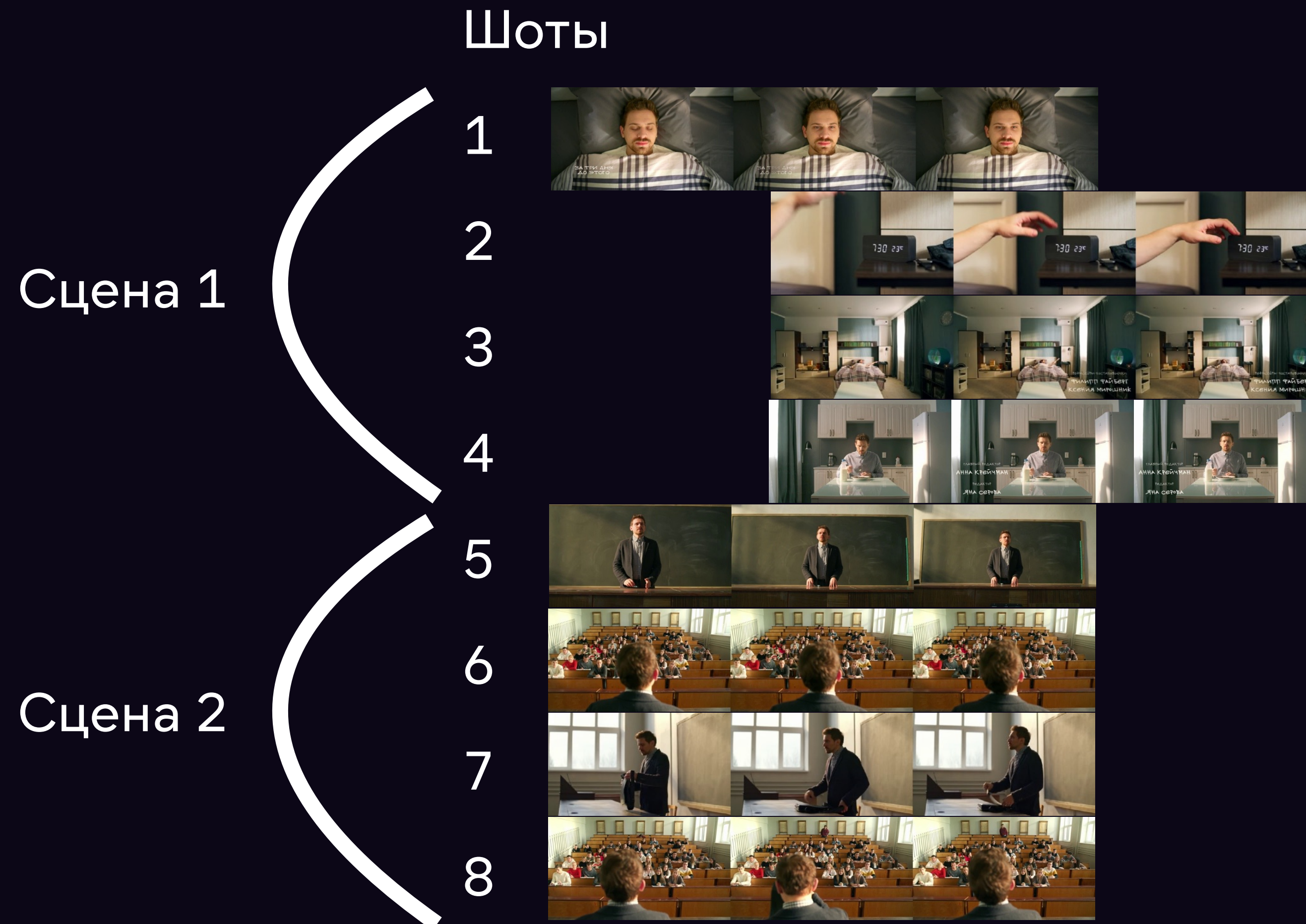
Пример работы



Ответ нейронки — 0

Детекция целостных сцен

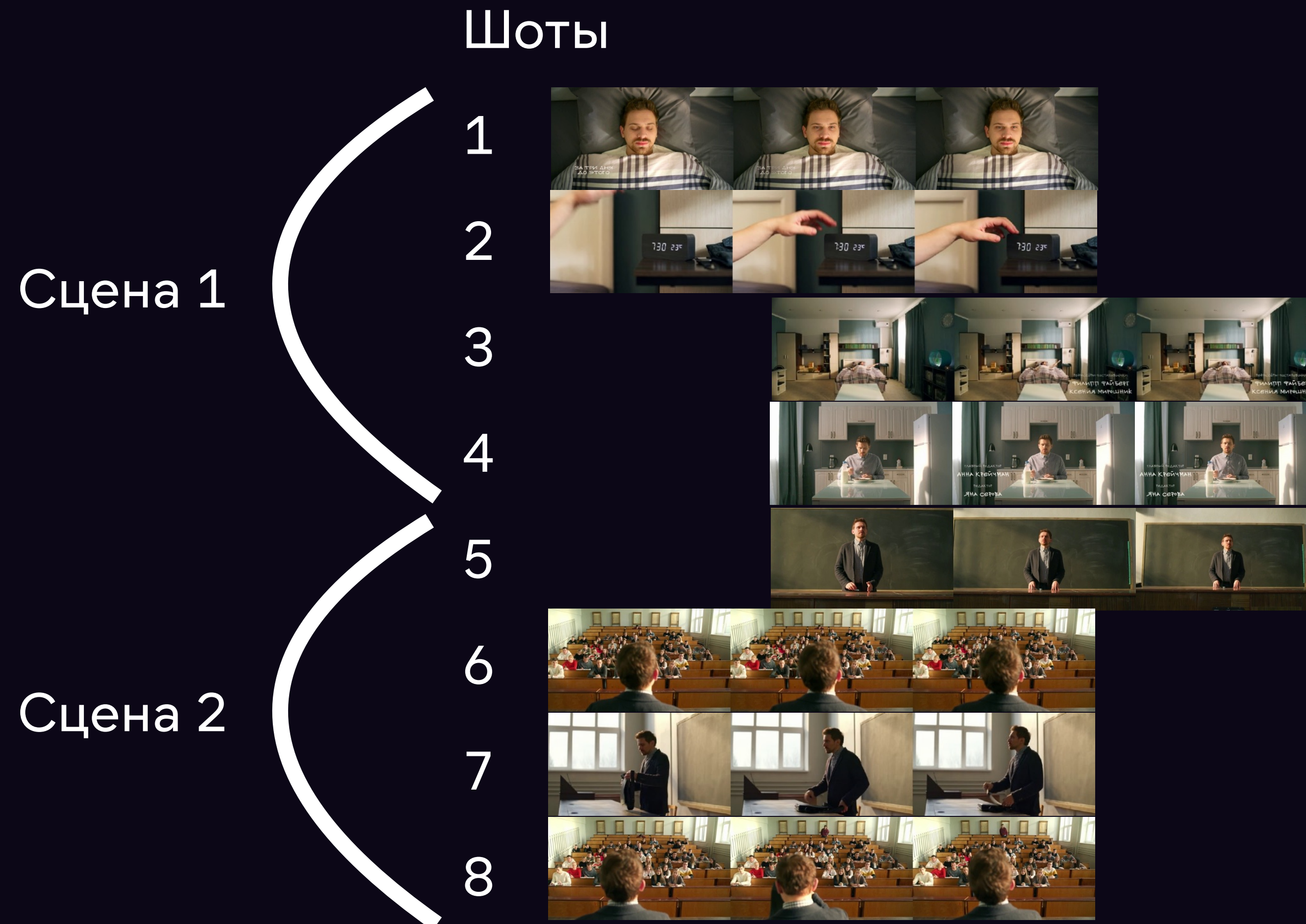
Пример работы



Ответ нейронки — 0

Детекция целостных сцен

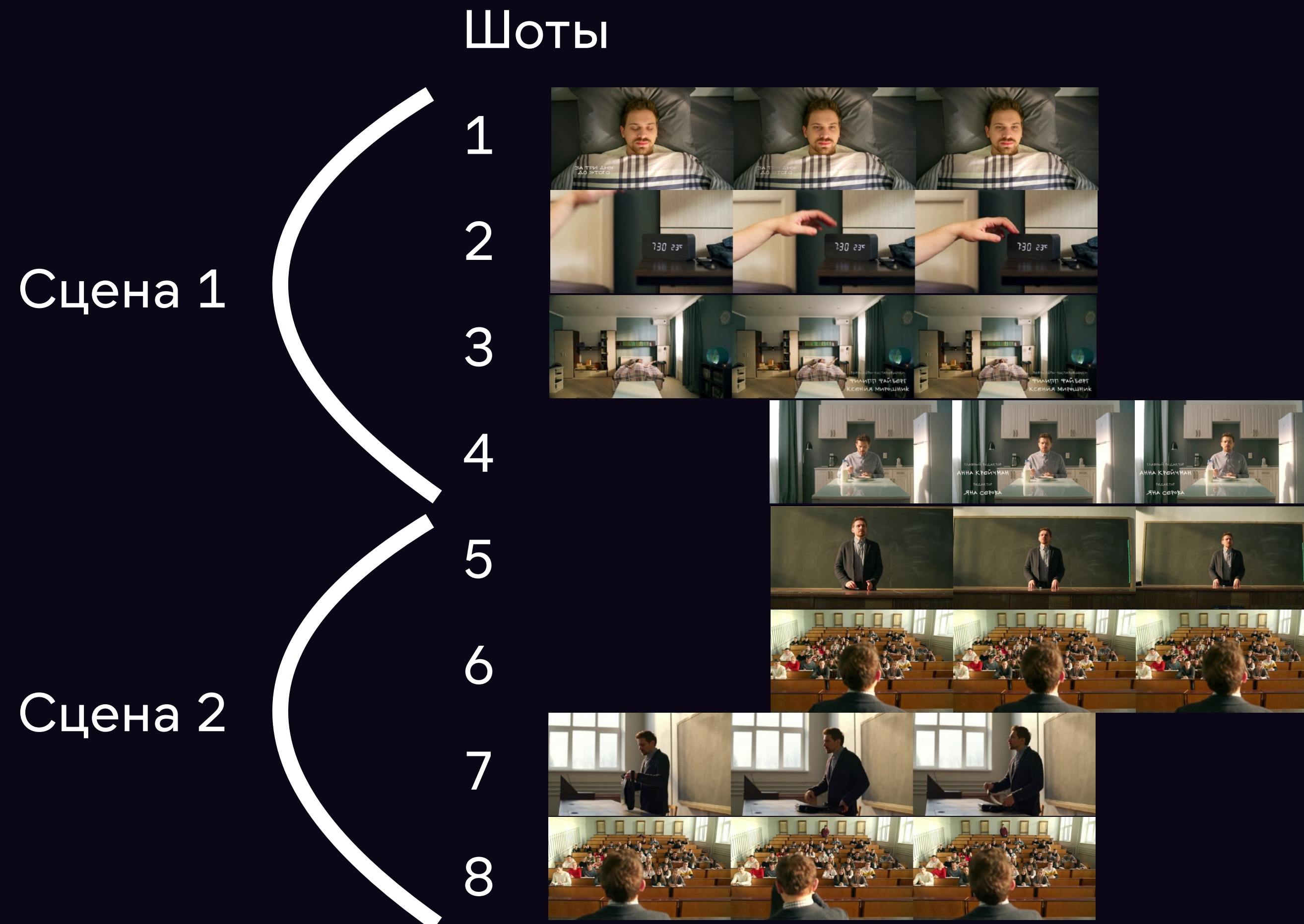
Пример работы



Ответ нейронки — 1

Детекция целостных сцен

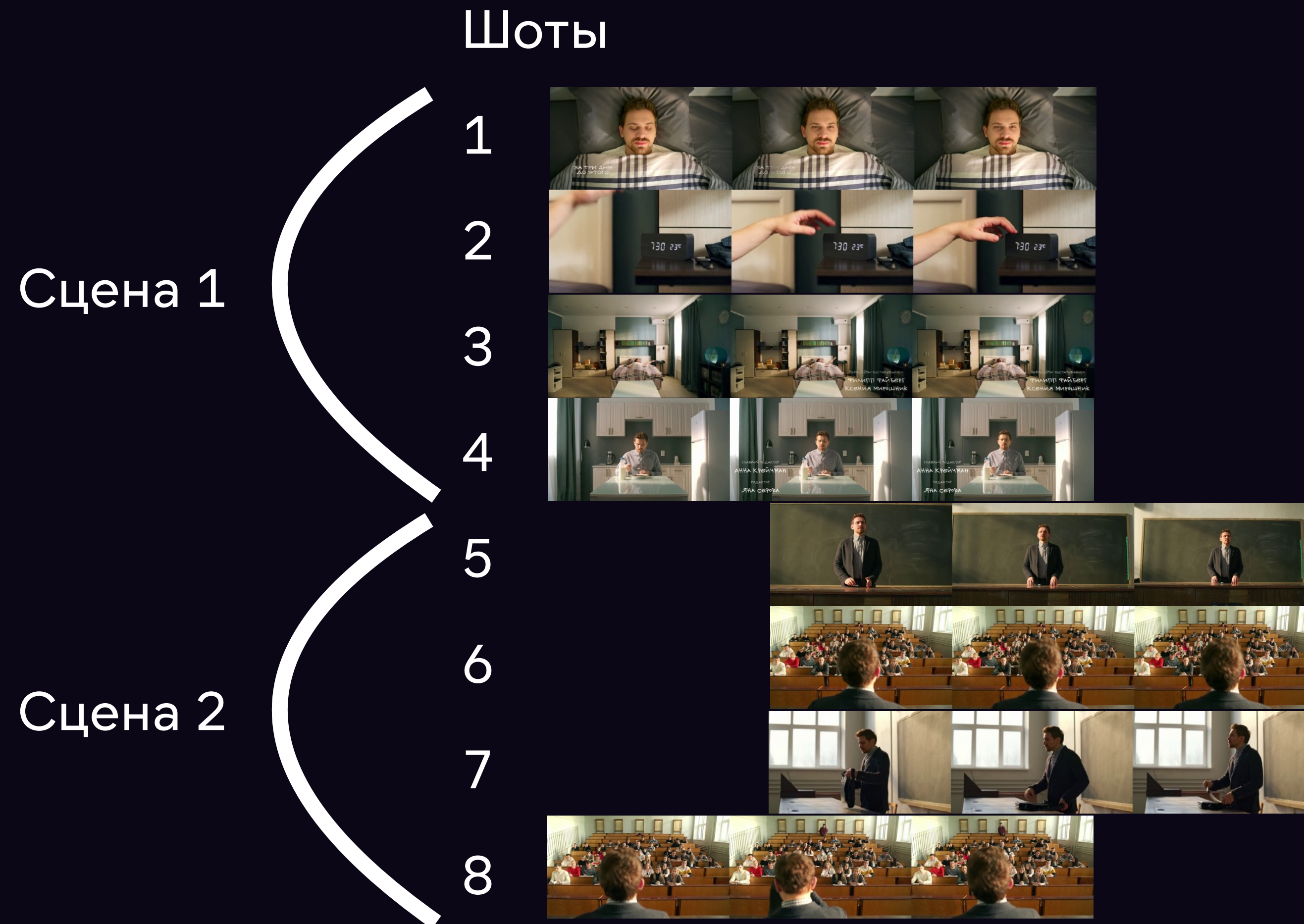
Пример работы



Ответ нейронки — 0

Детекция целостных сцен

Пример работы



Ответ нейронки — 0

Детекция целостных сцен

Пример работы



Ответ нейронки — 0

Детекция целостных сцен

Ограничения



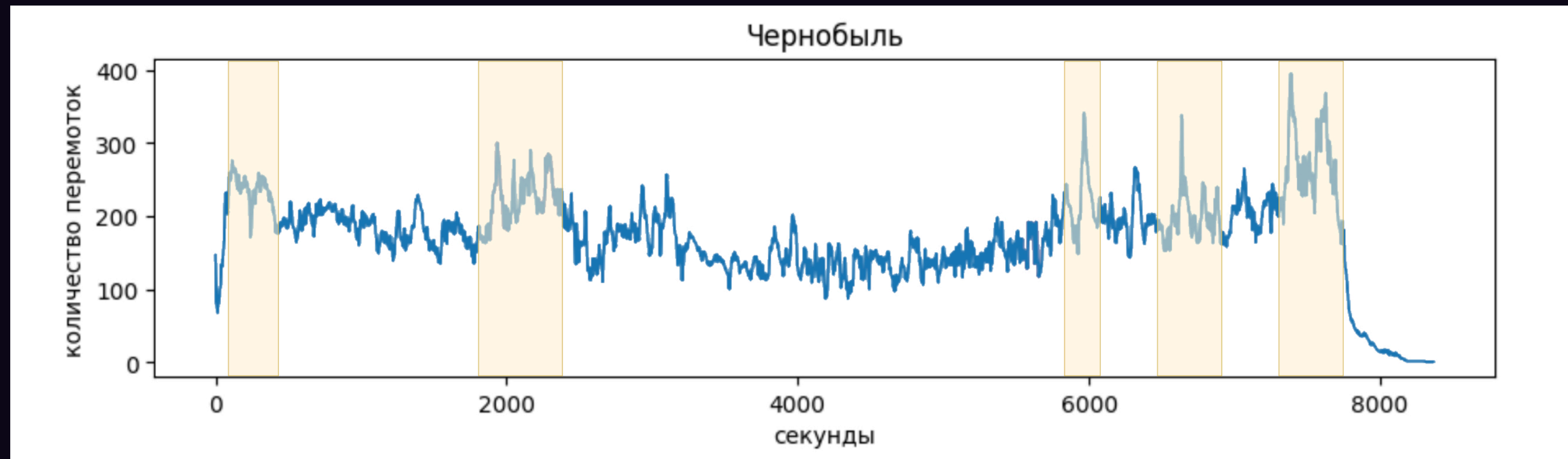
- Качество зависит от размера окна
- Качество зависит от динамичности фильма
- Не всегда можно по одним визуальным фичам понять, где граница сцены



Ранжирование найденных сцен

По перемоткам

- Считаем количество перемоток НАЗАД для каждого 5-секундного интервала
- Вычисляем, сколько всего перемоток было сделано внутри границ каждой из сцен, нормируем результат на длину сцены

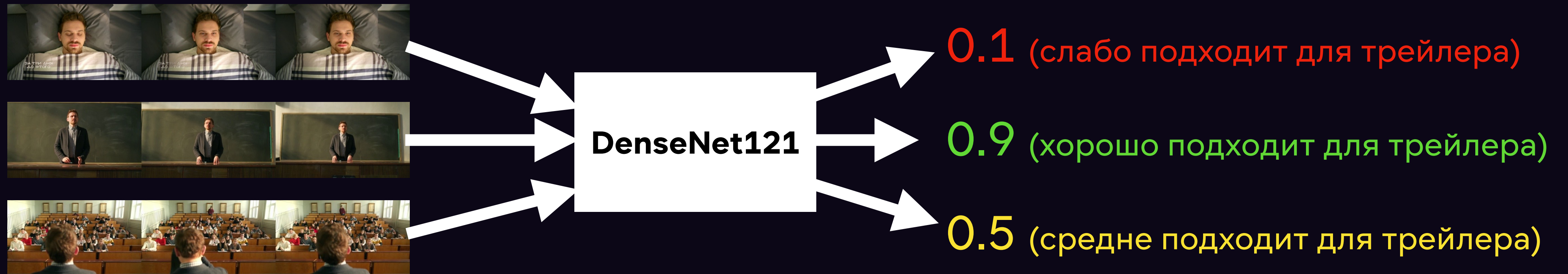




Ранжирование найденных сцен

По вероятности вхождения в трейлер

- Мы нарезали несколько тысяч единиц контента на шоты, и автоматически разметили, входил ли каждый из шотов в трейлер для фильма.
- На этих данных обучили модель бинарной классификации, которая выдаёт вероятность от 0 до 1



кадры шотов

Спойлеры



- Концовка — это спойлер
- Важный сюжетный поворот — это спойлер





Спойлеры

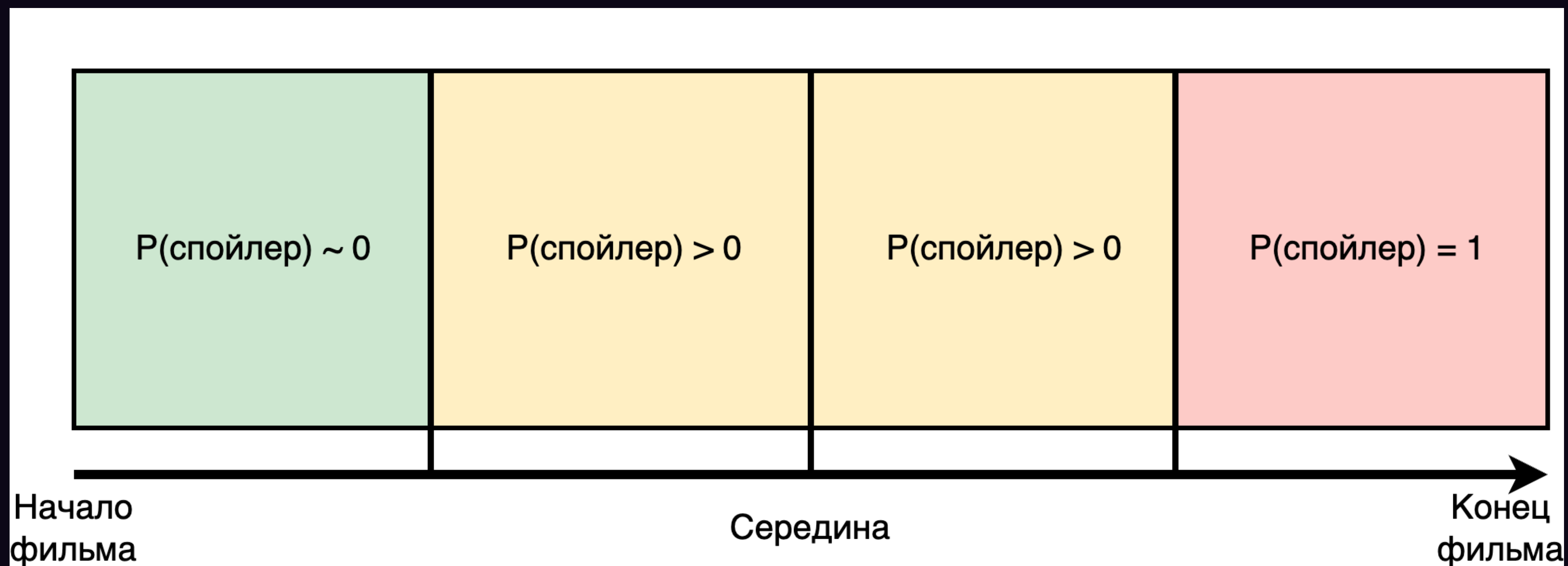
Что такое важный сюжетный поворот?

Литературный приём, который вносит радикальное изменение в направление или ожидаемый результат сюжета в художественном произведении.



Спойлеры

Нехитрая логика



Вертикализация



Вертикализация

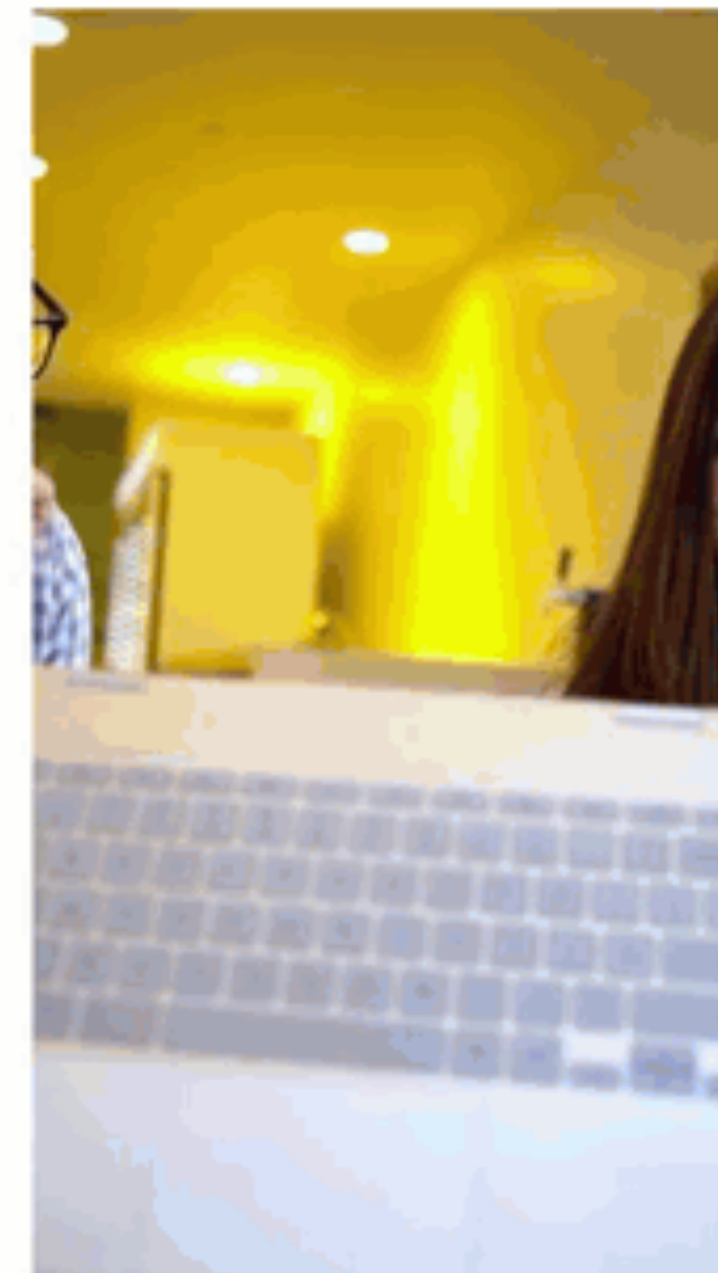
Google Autoflip



Original Video



Static Center Crop

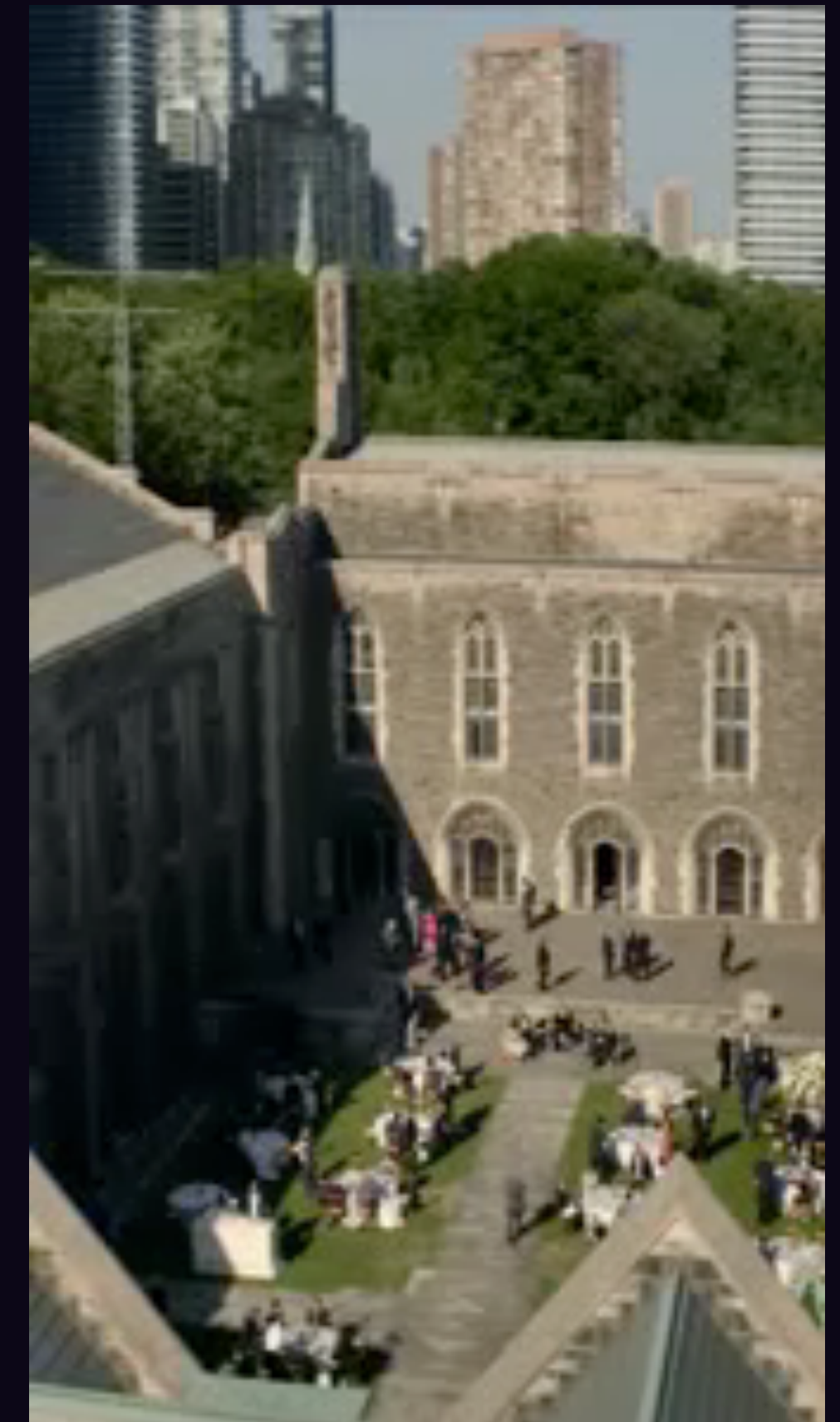


AutoFlip



Вертикализация

Быстрое самописное решение

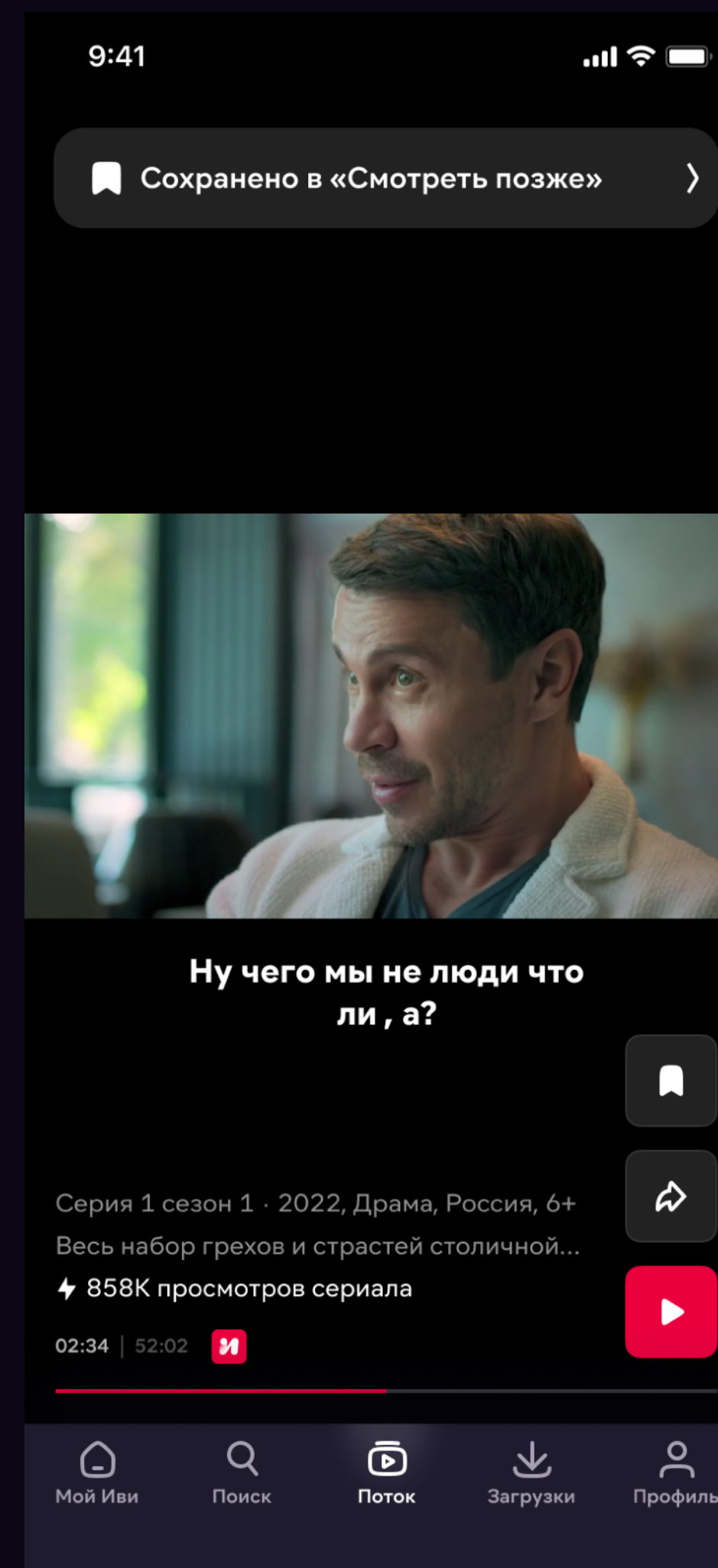


Вертикализация

И что решили?



- Откладываем разработку своего решения на потом
- Запускаемся на мобильных устройствах с горизонтальными моментами
- Редакция нарезает вертикальные руками в своём темпе

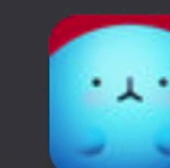


Субтитры

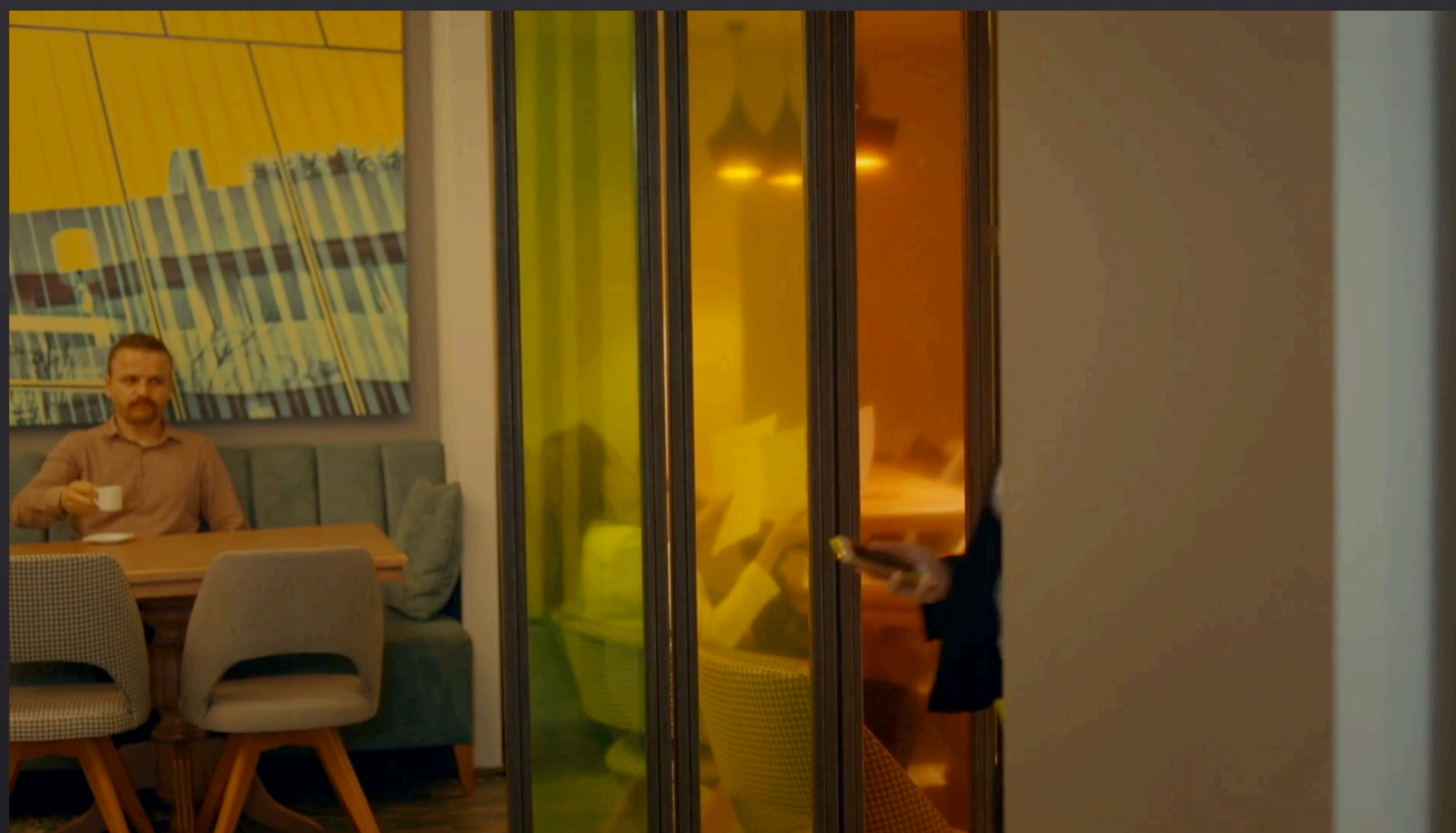


Мой Иви Что нового Фильмы Сериалы Мультфильмы TV+ Подписка

Поиск



Сериалы · Мелодрамы



ПОСТУЧИСЬ В МОЮ ДВЕРЬ

2020-2021 2 сезона 16+

Турция · Мелодрамы · Комедийные · Романтика

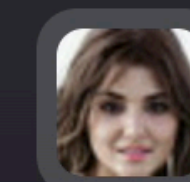
FullHD РУС РУС · Рус AI

95%

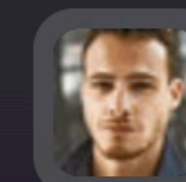
Тебе
подходит

8,7

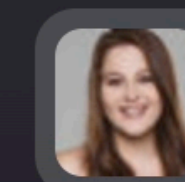
Рейтинг
Иви



Ханде Эрчел



Керем
Бюрсин



Эльчин
Афаджан

Великолепный дуэт Ханде Эрчел и Керема Бюрсина, сыгравших главные роли

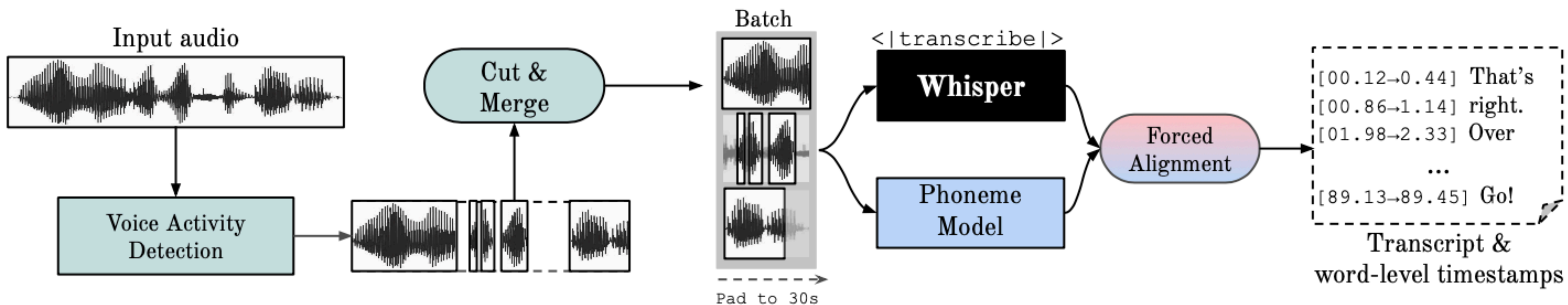
Трейлер



Субтитры WhisperX



<https://github.com/m-bain/whisperX>



Субтитры

Проблемы



- Редактор субтитров А. Захарова
Корректор А. Егорова
- Фондю любит тебя
- жми на колокольчик
- Подписывайтесь на наш канал
- [МУЗЫКА]
- [ШАГИ]

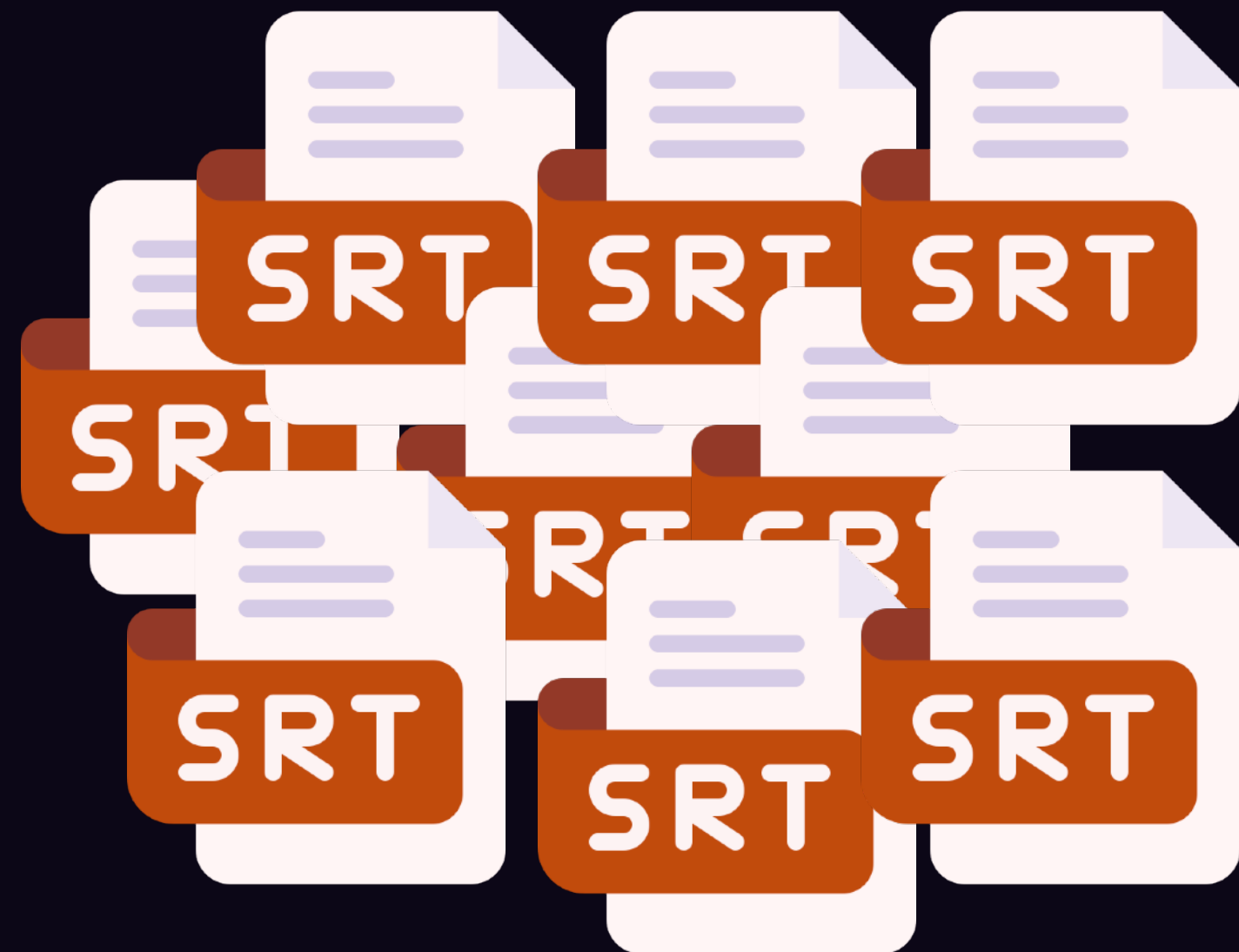




Субтитры

Оценка качества

- Собрали датасет из нескольких тысяч видео с субтитрами
- Определили набор метрик



```
| wer  
| wil  
| cer  
| time_segmentation_precision  
| time_segmentation_recall  
| time_segmentation_f1  
| time_segmentation_fbeta05  
| matched_wer  
| matched_recall  
| matched_precision  
| matched_f1  
| subtitle_nchars_distr_wasserstein  
| subtitle_nchars_distr_kl_div_forward  
| subtitle_nchars_distr_kl_div_reverse  
| subtitle_nchars_distr_kolmogorov_smirnov |
```



Субтитры

Метрики. Word Error Rate

$$WER = \frac{S + D + I}{N}$$

S — количество замен

D — количество удалений

I — количество вставок

N — количество слов в референсе

Диапазон — $[0 .. \infty]$

Идеал — 0



Субтитры

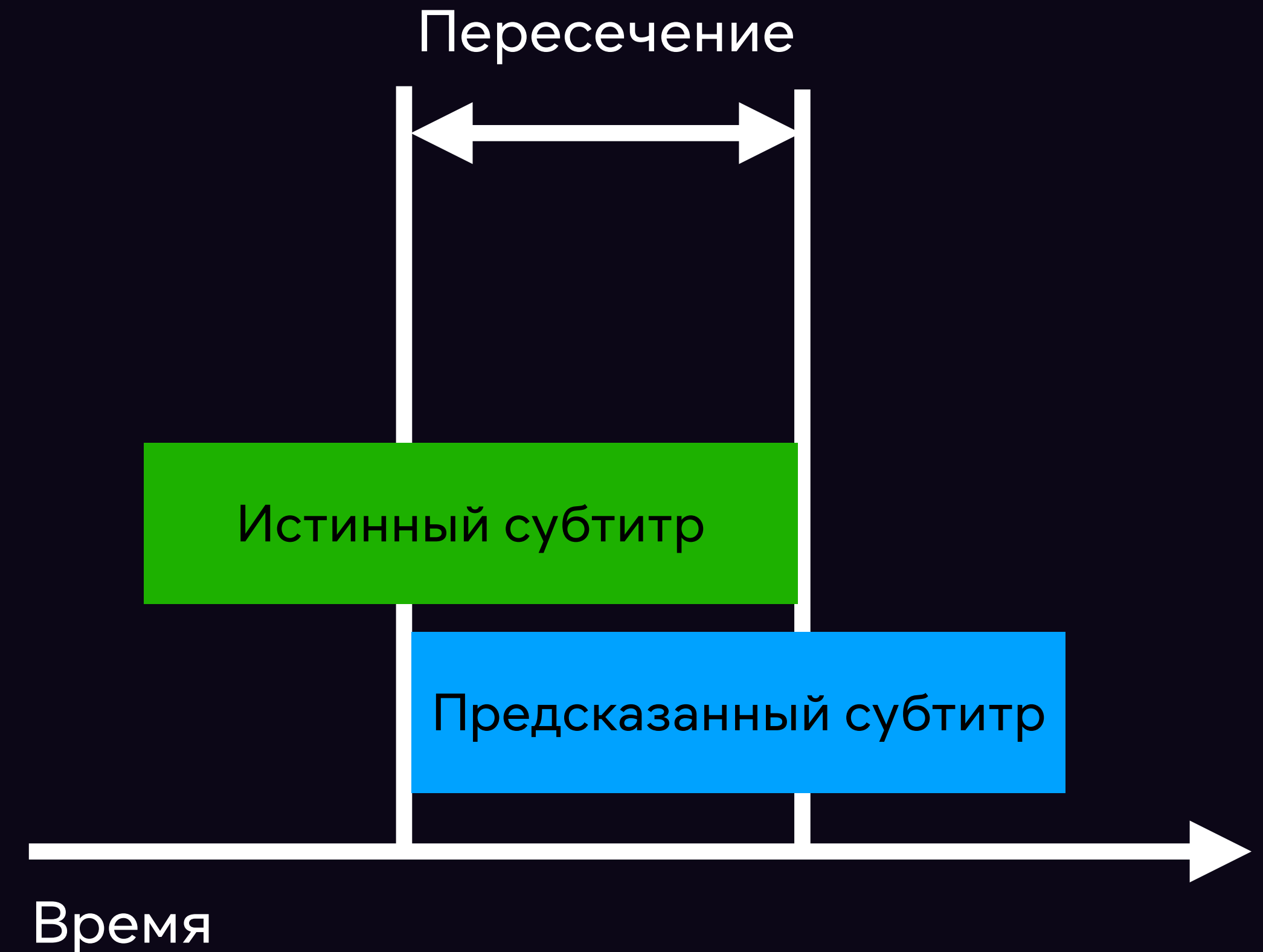
Метрики. F1, Precision, Recall

$$F_1 = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

$$\text{precision} = \frac{\text{пересечение}}{\text{длина предсказанного субтитра}}$$

$$\text{recall} = \frac{\text{пересечение}}{\text{длина истинного субтитра}}$$

Диапазон — $[0 .. 1]$

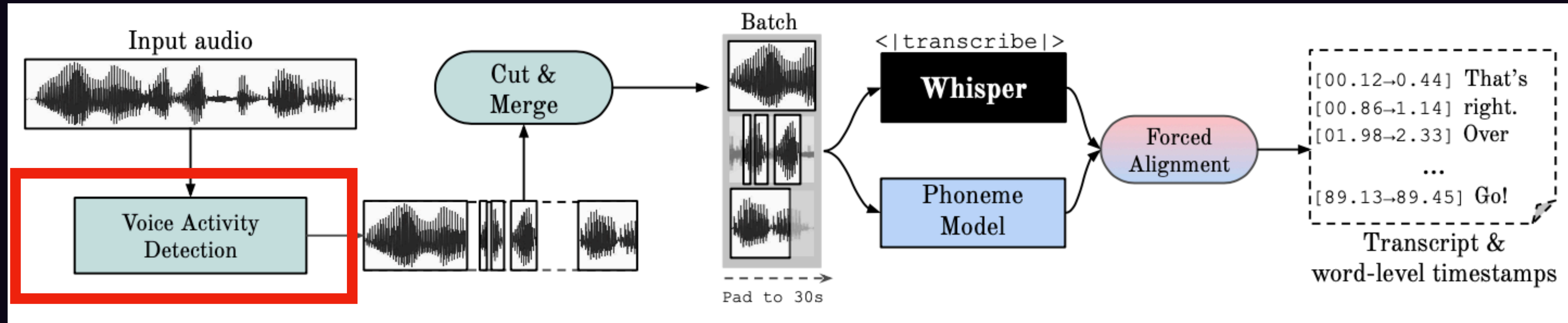




Субтитры

Тюнинг WhisperX

- Прикрутили ансамбль Voice Activity Detection (Pyannote + Speechbrain + Silero)

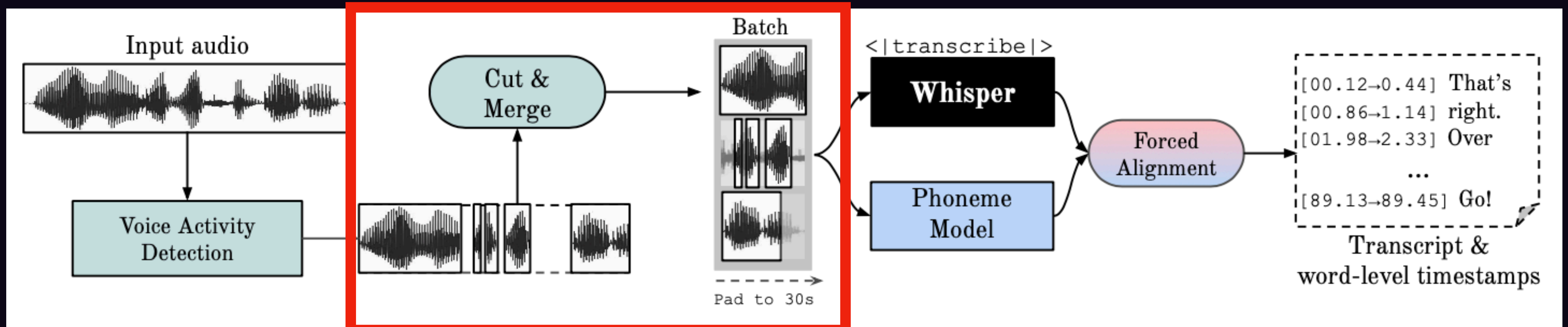




Субтитры

Тюнинг WhisperX

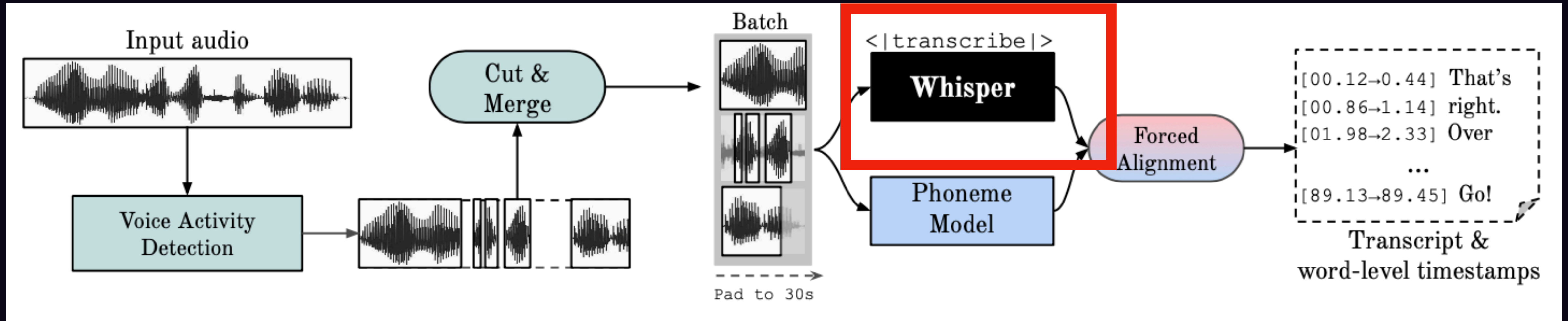
- Подтюнили группировку фрагментов при распознавании





Субтитры Тюнинг WhisperX

- Подобрали гиперпараметры Whisper





Субтитры

Тюнинг WhisperX

- Добавили фильтрацию матерных слов и галлюцинаций по словарю
- Сделали ограничение на длину одного субтитра
- Добавили паддинги после длинных фраз, чтобы зритель успевал прочесть строчку



Субтитры

Результаты



Снизилось количество ошибок **WER 21% → 9%**

Повысилась точность локализаций фраз **F1 69% → 89%**

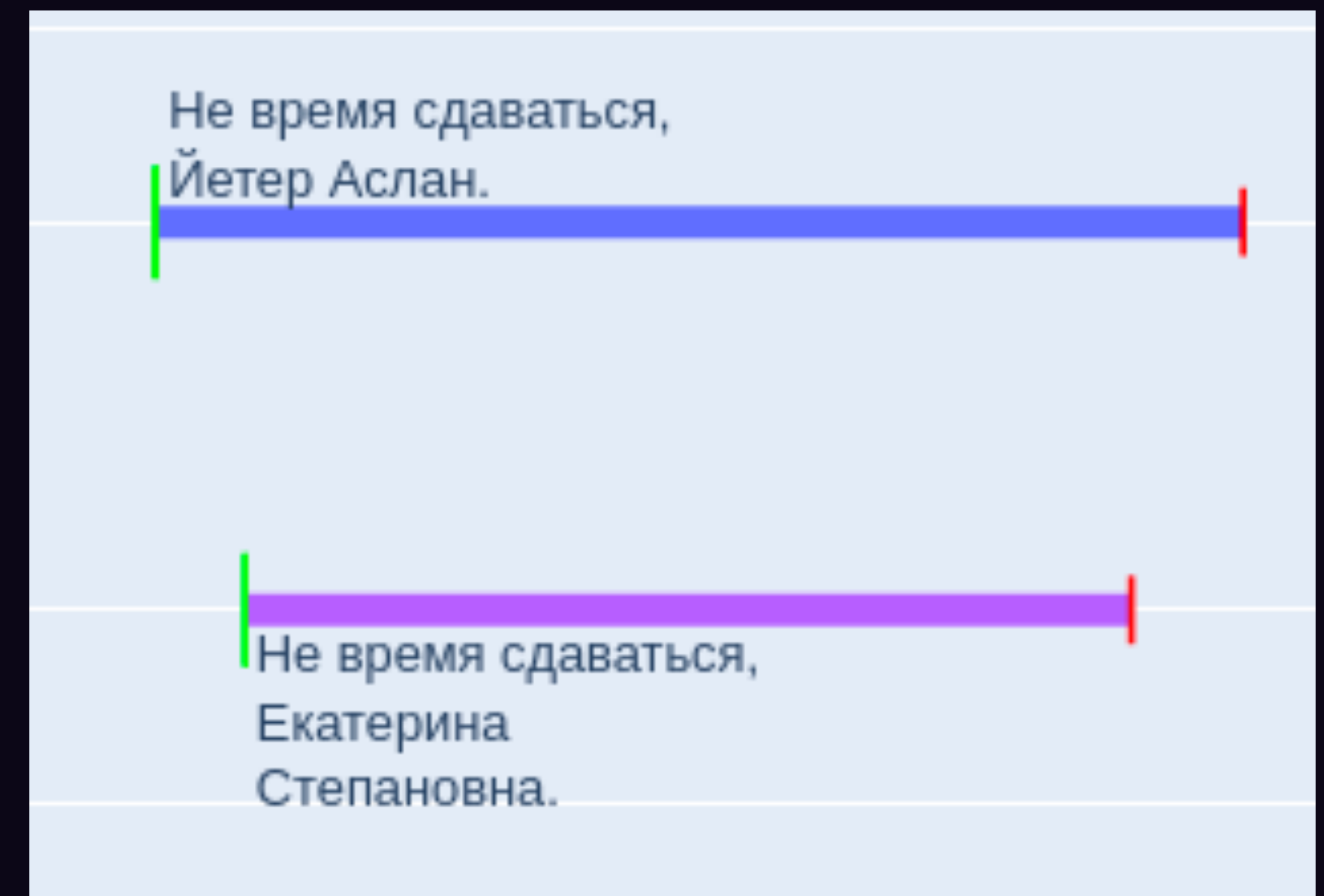
Снизилось количество галлюцинаций



Субтитры

Оставшиеся проблемы

- Имена собственные.
Решение — прикрутить NER, завести базу имён персонажей и матчить с распознанными.
- Изредка галлюцинирует.
Решение — добавлять новые галлюцинации в словарь.





Люди против машин

Редакторская часть

- Люди пока выбирают моменты лучше
- Ручная вертикализация
- Производительность ~ 1 000 моментов в месяц





Модерация

Оставшиеся проблемы

- Неинтересные моменты
- Обнаженка, кровь и прочий 18+
- Спойлеры
- Кривая обрезка сцен
- Ошибки в субтитрах













Модерация




← **Друзья** [Иви](#) [Google](#) [Yandex](#)

Комедийные Мелодрамы Зарубежные

2.2.0 | 2023-11-09 15:18

 3:19 - 3:34	 4:59 - 5:22	 6:21 - 6:39	 9:02 - 9:20	 10:39 - 10:56	 11:28 - 11:45
 16:06 - 16:26	 17:25 - 17:45	 18:01 - 18:18	 18:18 - 18:41		



Что-то не так?

- Содержит более одной смысловой сцены ⓘ
- Обрезается речь в начале или конце ⓘ
- Смысловая сцена не закончена
- Не интересный
- Содержит 18+ ⓘ

ОДОБРИТЬ **ОТКЛОНИТЬ**

0:00 0:01 0:02 0:03 0:04 0:05 0:06 0:07 0:08 0:09 0:10 0:11 0:12 0:13 0:14 0:15 0:16 0:17 0:18 0:19 0:20 0:21 0:22 0:23 0:24 0:25 0:26 0:27 0:28 0:29 0:30 0:31 0:32 0:33 0:34

3:19 - 3:33



Запуск

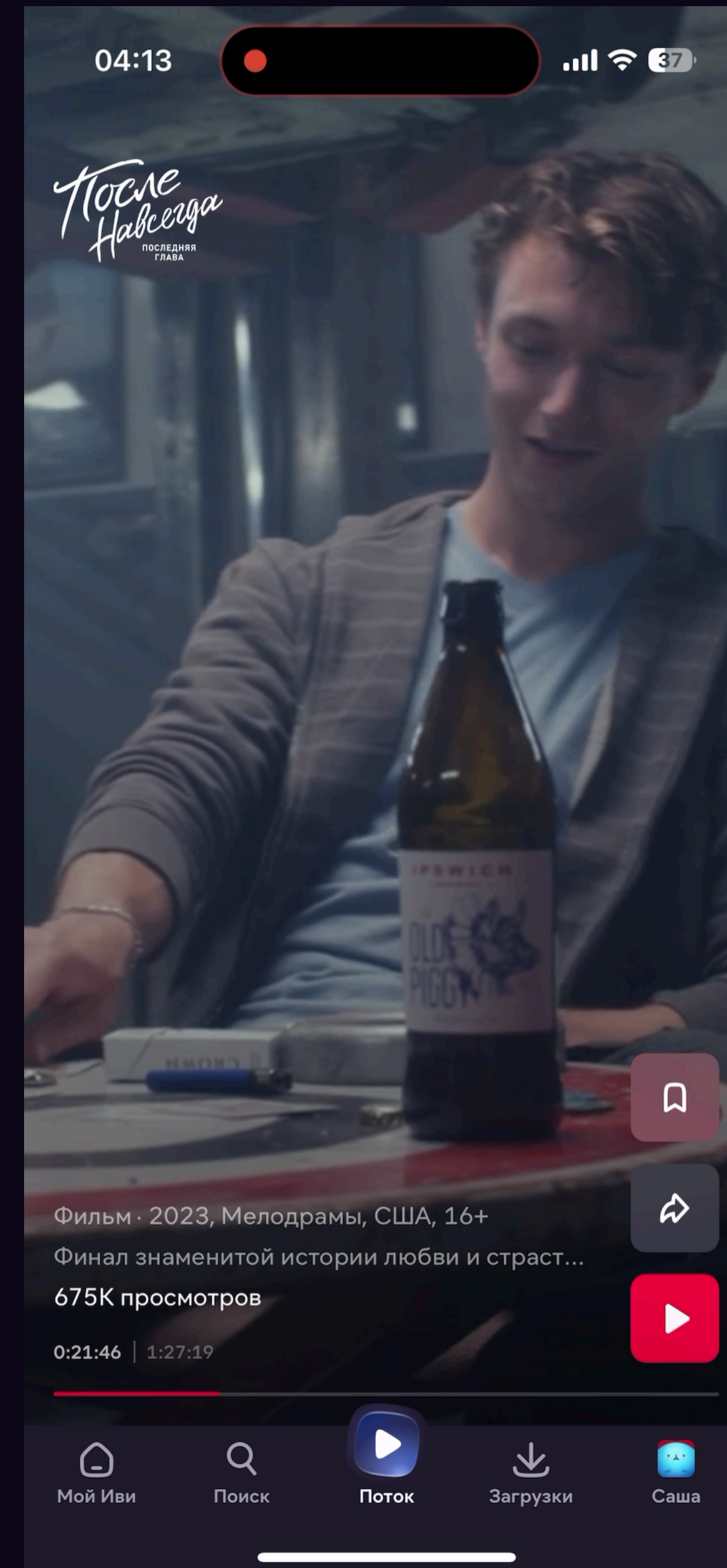
10 команд, 6 месяцев работы

- Дизайнеры
- Back-end
- Encoder
- Computer Vision
- Аналитики
- Рекомендательная система
- iOS
- Android
- SmartTV
- Редакция

Запуск

Промежуточные итоги

- Платформы: iOS, Android, LG WebOS, Samsung Tizen
- 3 700 тайтлов в Поток
- 15 000 горизонтальных и 3 700 вертикальных моментов в ленте



Планы



- Второй подход к вертикализации видео в CV
- Дообучение моделей и тюнинг пайплайнов на статистике просмотров и фидбэке модерации

Спасибо за внимание!

Поток (iOS, Android, smart TV):



<https://www.ivi.ru/flow>

Александр Коншин:



<https://sashulyak.github.io>