

Воронежский государственный университет

Проектная работа

на тему:

Оцифровка и анализ изображений и компьютерное зрение.

Работа с изображениями: Gimp, OCR. Работа с видеоматериалами: ELAN.

Подготовила студентка ВГУ

1 курс

Направление «Лингвистика»

Профиль «ТиМПИЯиК»

Заборовская Ксения Валерьевна

Преподаватель: Дони́на Ольга Валерьевна

Воронеж

2018

Мотивация.

Узнать все о современном направлении развития компьютерных технологий и научиться всему: от графического дизайна до создания субтитров.

Теория.

Оцифровка и анализ изображений.

Анализ изображений - это извлечение значимой информации из изображений; в основном из цифровых изображений с помощью методов цифровой обработки изображений. Задачи анализа изображений могут быть такими же простыми, как чтение штрих-кодированных тегов, или такими же сложными, как идентификация человека по его лицу.

Можно сказать, что в среднем анализ изображений сводится к извлечению значимой информации из изображений. Для каждой конкретной ситуации эта значимая информация может быть разной.

В анализе изображений как результат мы получаем вектор признаков. Компьютерное зрение решает более широкие задачи. В частности строятся модели. К примеру, по набору двумерных изображений можно построить трехмерную модель помещений. И есть еще одна смежная область — компьютерная графика, в которой генерируют изображение по модели.

Все это невозможно без использования знаний и алгоритмов из еще целого ряда областей. Таких, как распознавание образов и машинное обучение. В принципе можно сказать, что анализ изображений — это частный случай анализа данных, область искусственного интеллекта.

Для чего нужен анализ изображений

Существуют огромные архивы и коллекции изображений, и одной из самых главных задач является **индексирование и поиск** картинок. Что с ними можно делать? Самое простое — можно как-то по-умному построить **навигацию** по ним, классифицируя их по темам. Отдельная задача — это **поиск дубликатов** (например, когда вы приехали из отпуска и не хотите вручную искать повторяющиеся фотографии). Кроме того, поиск по дубликатам может помочь обнаружить незаконное использование вашей фотографии, которую вы однажды могли выложить в интернете. Отличная задача — **выбор лучшей фотографии**. С помощью алгоритма можно понять, какая картинка больше всего понравится пользователю. Например, если это портрет, лицо должно быть освещено, глаза открыты, изображение должно быть четким и т.д. В современных фотоаппаратах уже есть такая функция. Тоже задача поиска — **создание коллажей**, т.е. подбор фотографий, которые будут хорошо смотреться рядом.

Компьютерное зрение

Компьютерное зрение — это технология, с помощью которой машины могут находить, отслеживать, классифицировать и идентифицировать объекты, извлекая данные из изображений и анализируя полученную информацию.

Компьютерное зрение применяется для распознавания объектов, видеоаналитики, описания содержания изображений и видео, распознавания жестов и рукописного ввода, а также для интеллектуальной обработки изображений.

Проще говоря, сейчас ученые пытаются заставить компьютеры видеть, описывать и понимать все то, что мы видим.

OCR (Optical Character Recognition)

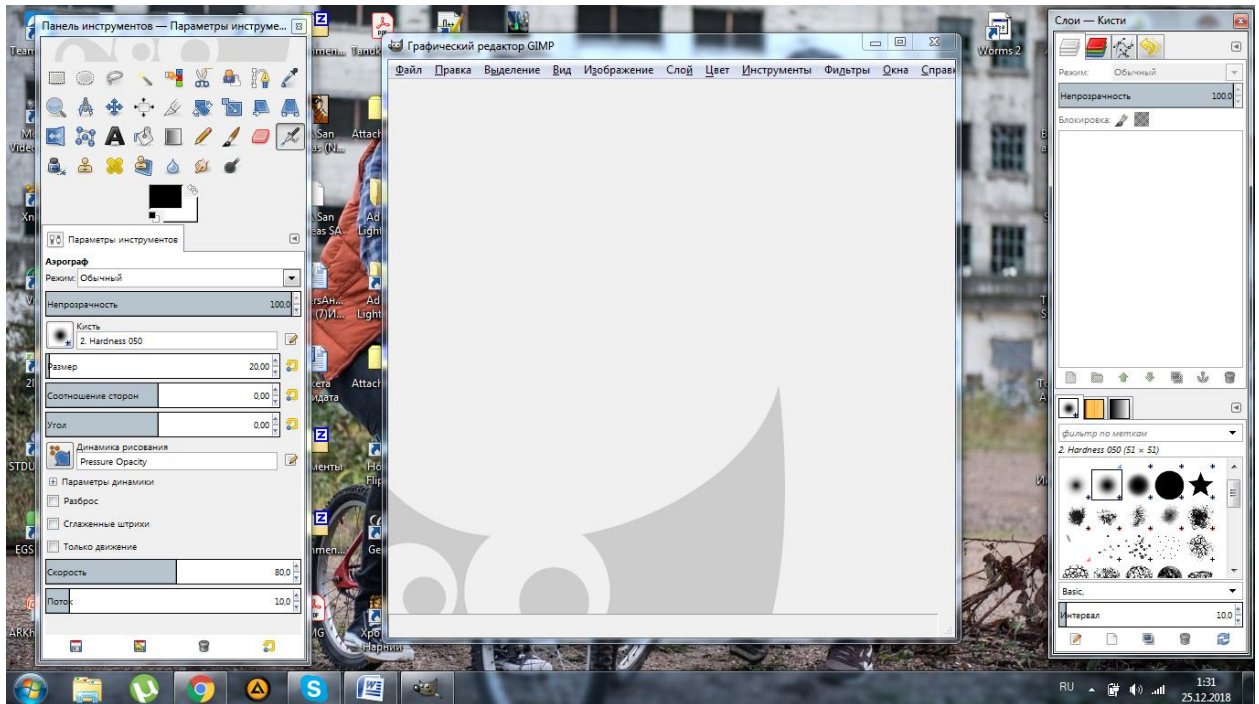
Еще одной очень интересной и полезной для студентов темой является OCR, с помощью которого можно запросто сделать из фотографии или скана учебника (или любого другого текста) Word-файл, который можно редактировать.

Оптическое распознавание символов (англ. **Optical Character Recognition –OCR**) – это технология, которая позволяет преобразовывать различные типы документов, такие как отсканированные документы, PDF-файлы или фото с цифровой камеры, в редактируемые форматы с возможностью поиска.

Практика.

Оцифровка изображения в GIMP.

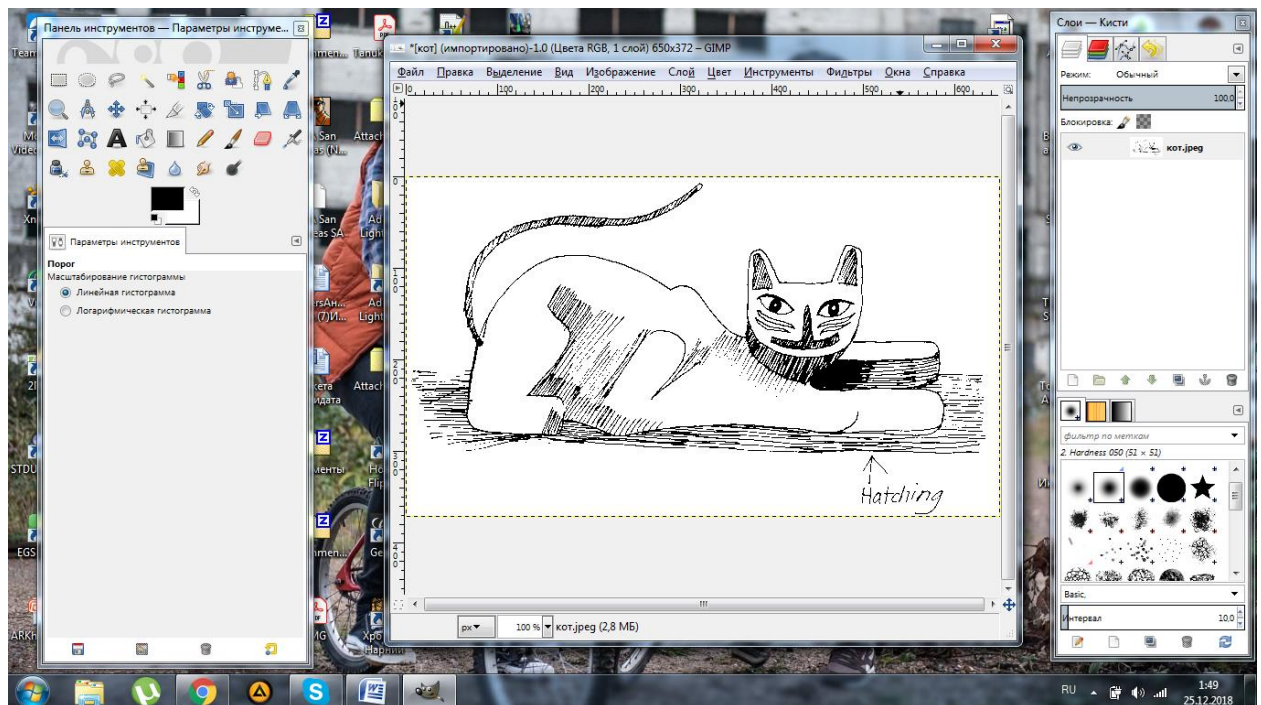
1. Качаем и устанавливаем GIMP по ссылке: <http://softcatalog.info/ru/programmy/gimp>
Не забудьте отказаться от всех предложение Яндекс.Браузера и Оперы!
2. Запускаем программу



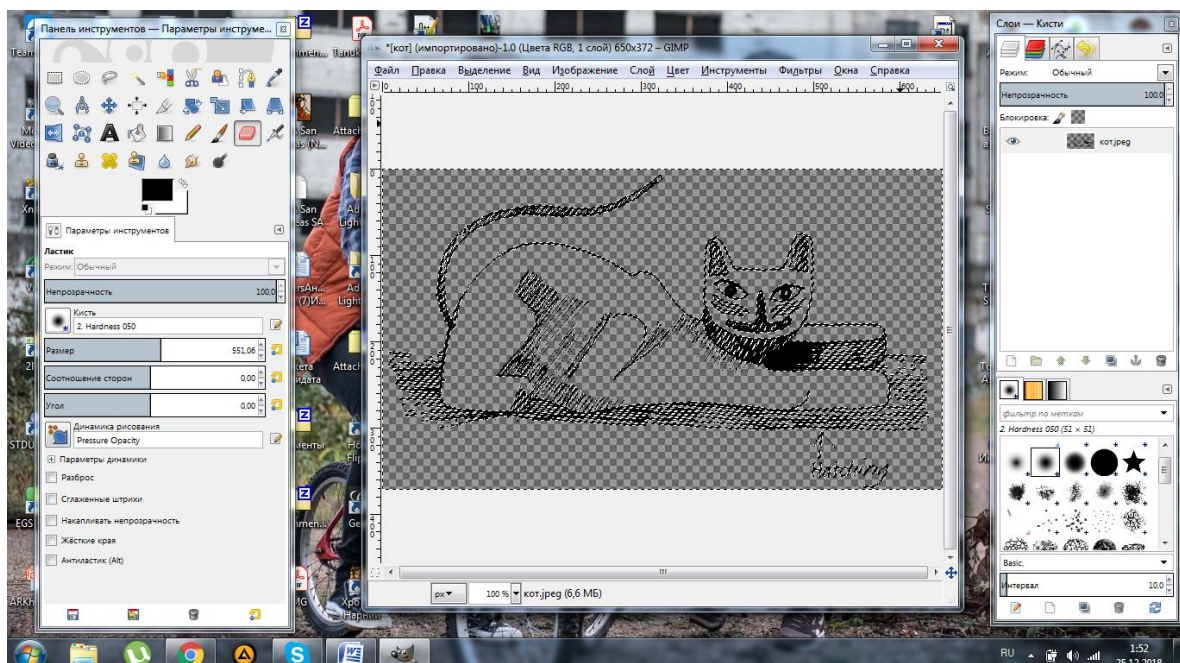
3. В поисковике вбиваем что-то вроде «pen drawing» и ищем сфотографированный рисунок в карандаше или ручке, желательно без цвета. Фон пусть будет некрасивым желтым или серым, а вокруг самого рисунка могут находиться ненужные предметы по типу ручек или карандашей. Я нашла котика.



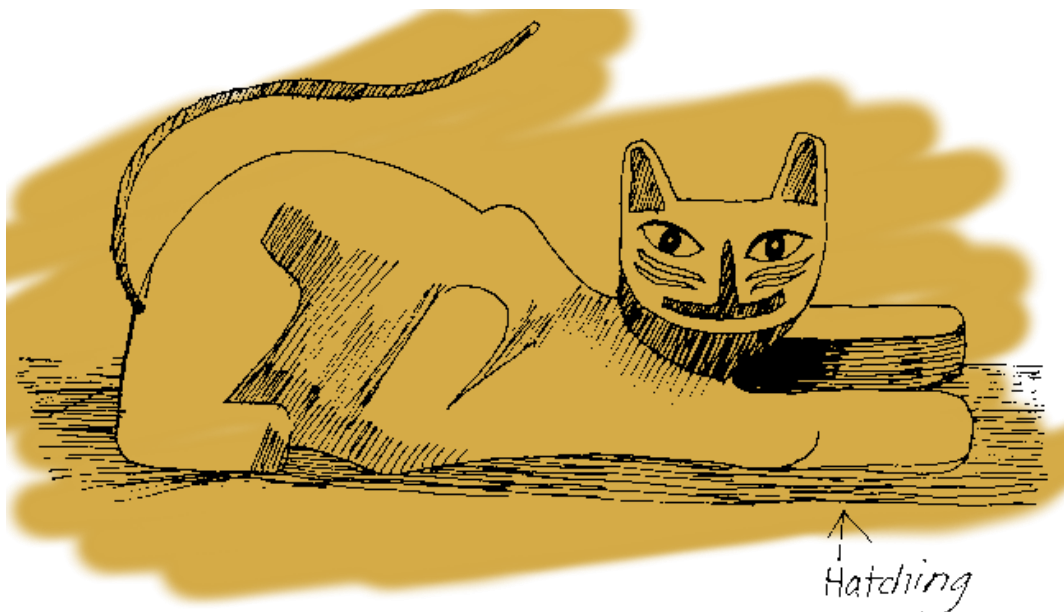
4. Загружаем фото в GIMP. Нажимаем на верхней панели инструментов Цвет – Порог - ок (Colors-Threshold). После всего этого наше фото должно стать больше похоже на оцифрованное изображение:



5. Далее мы кликаем на сам слой изображения (правая панель) и выбираем «Создать альфа-канал» («Add Alpha Channel»).
6. Нажимаем на «Выделение по цвету» в левой панели инструментов (или просто нажимаем сочетание клавиш Shift+O) и кликаем по цвету контура. После этого в верхней панели выбираем Выделение – Инвертировать. Таким образом мы выделили белый фон.
7. Берем «Ластик» (Shift+E) и проводим им по всему изображению, не боясь за сам рисунок. После всех этих манипуляций он должен выглядеть как-то так:



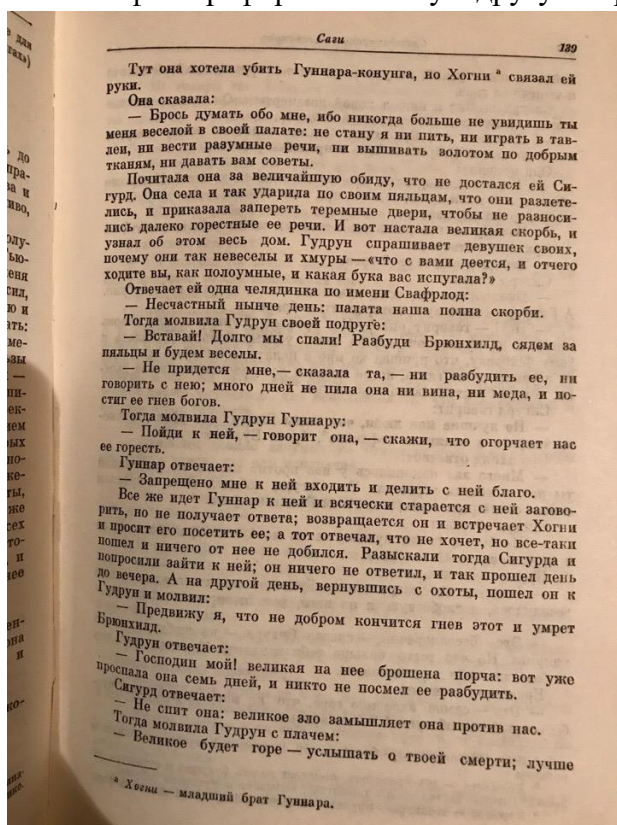
8. Отменяем выделение во вкладке «Выделение» и экспортируем наше произведение (Shift+Ctrl+E) в формате png. Готово! Теперь у вас есть оцифрованный контур, который можно поместить поверх другого изображения или раскрасить (для этого нужно создать новый слой и поместить его ниже контура).



Работа с OCR

В 2018 году существует немало способов конвертирования текста из формата pdf или jpeg в редактируемые word-документы. Даже больше, сейчас это можно сделать даже онлайн! Мы опробуем это сделать на примере сайта <https://finereaderonline.com>

За основу возьмем случайную фотографию из учебника по зарубежной литературе (вы можете сфотографировать любую другую страницу любой другой книги):

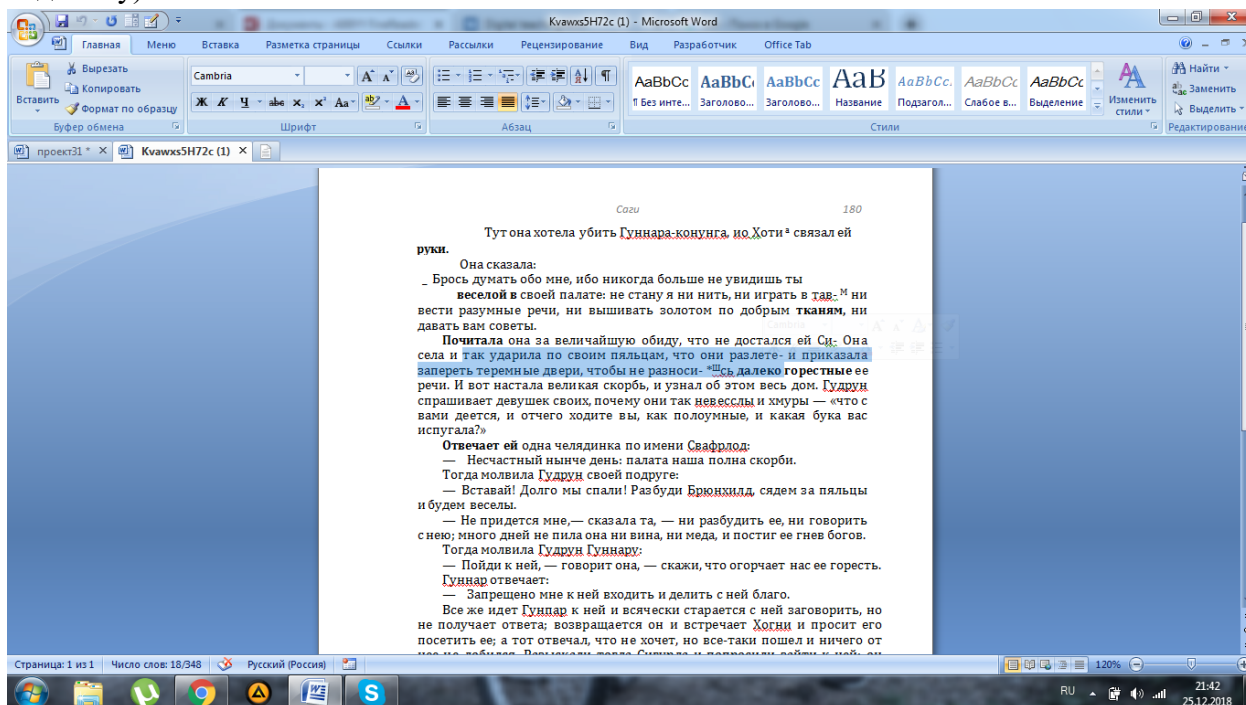


Заходим на сайт и нажимаем «Распознавание». Загружаем файл, указываем язык текста и конечный формат документа.

The screenshot shows the ABBYY FineReader Online web interface. At the top, there's a header with the logo and three tabs: "Распознавание" (Recognition), "Документы" (Documents), and "Магазин" (Store). Below the header, the "Распознавание" section is active. It contains three numbered steps: 1. "Загрузите файлы для распознавания" (Upload files for recognition) with a "Загрузить файлы" button and a note about file size and format. 2. "Выберите язык документа" (Select document language) with a dropdown menu set to "Русский" and a link to "Переключить на шрифты с фактурой [?]". 3. "Выберите формат сохранения документа" (Select document save format) with a grid of icons for various formats: .docx (Microsoft Word), .xlsx (Microsoft Excel), .rtf (Rich Text Format), .txt (Plain Text), .pptx (Microsoft PowerPoint), .odt (Open Document), .pdf (PDF), .fb2 (Fiction Book), and .epub (Electronic Publication). The .docx format is highlighted.

Нас сразу же перенаправляют на страницу с уже готовым документом в нужном нам формате, который остается только скачать.

На выходе у нас должен появиться обычный файл, который спокойно можно редактировать! (Единственный минус этого сайта — он предоставляет лишь 10 страниц для в месяц, так что если хотите оцифровать побольше — будьте добры оплатить платную подписку)



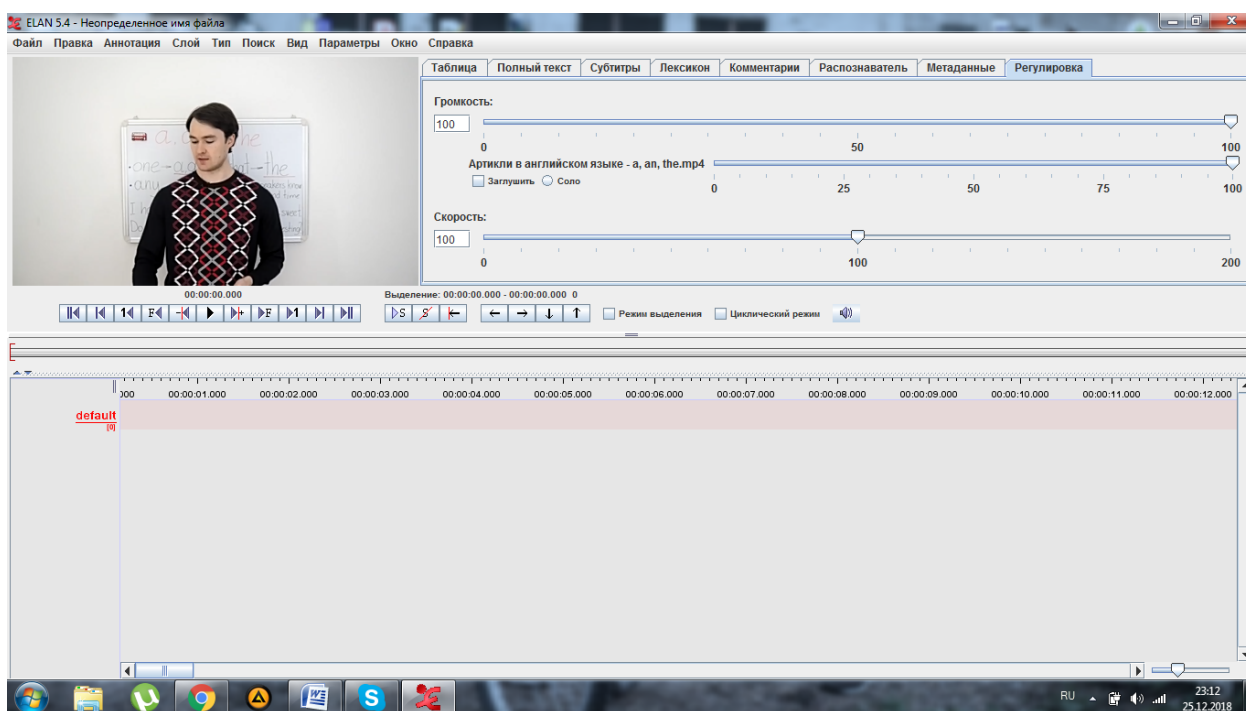
Работа с ELAN

ELAN – профессиональная программа для создания комплексных аннотаций для видео и аудио. Аннотацией может быть что угодно: предложение, фраза, перевод, комментарий или описание. Аннотации могут быть созданы на нескольких слоях, которые называются уровнями (ярусами). Эти уровни могут быть иерархически связаны друг с другом. Аннотации могут быть либо привязаны ко времени файла, либо могут относиться к уже существующим аннотациям.

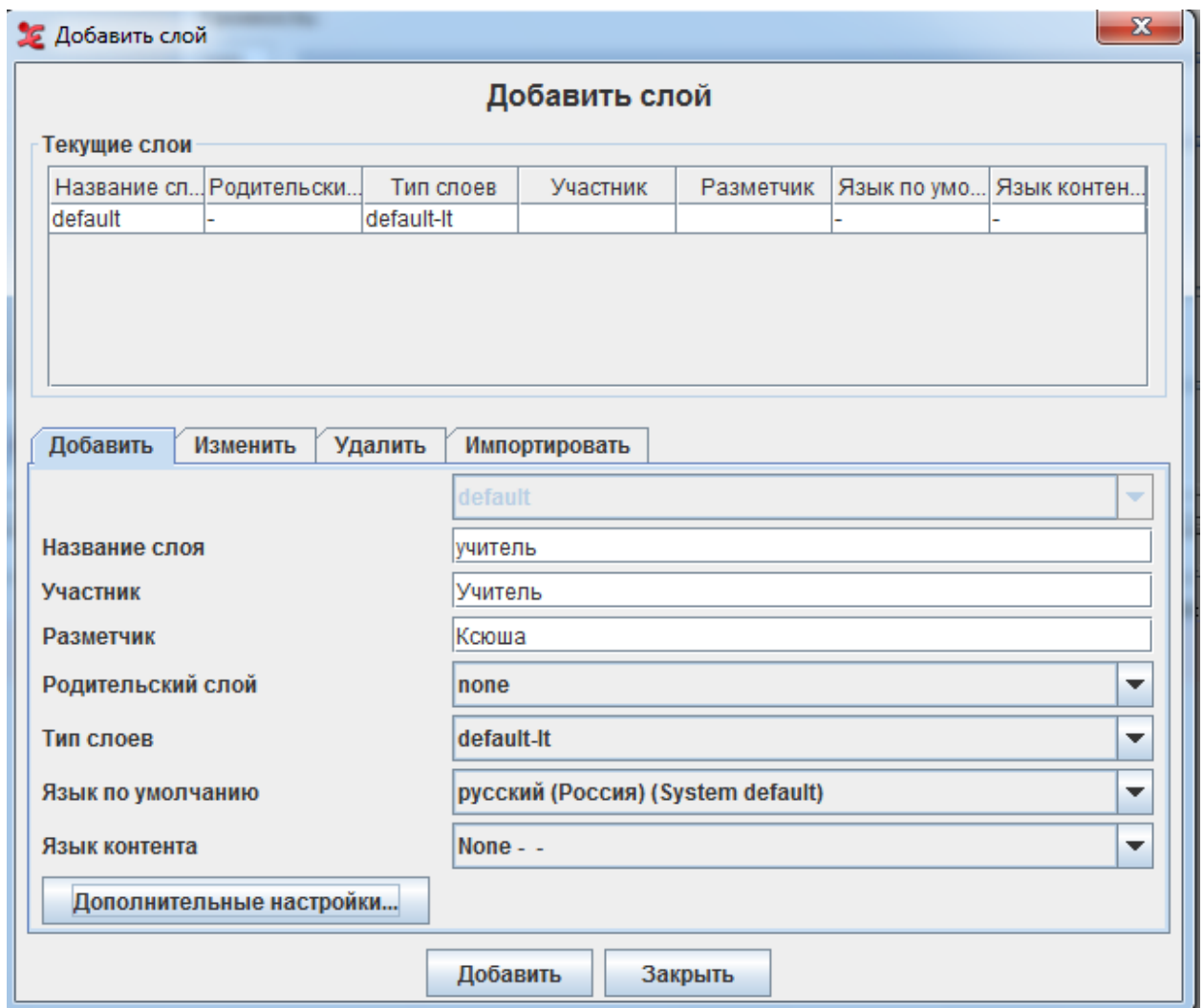
Начинаем работать! Для начала качаем программу: <https://tla.mpi.nl/tools/tla-tools/elan/download/>

Мы попробуем сделать субтитры к видео-уроку, так как это может пригодиться вам в будущем. Для этого заходим на YouTube и ищем там любой небольшой урок на любую тему. Чтобы скачать видео, в адресной строке перед «youtube» пишем «ss» и качаем с открывшегося сайта.

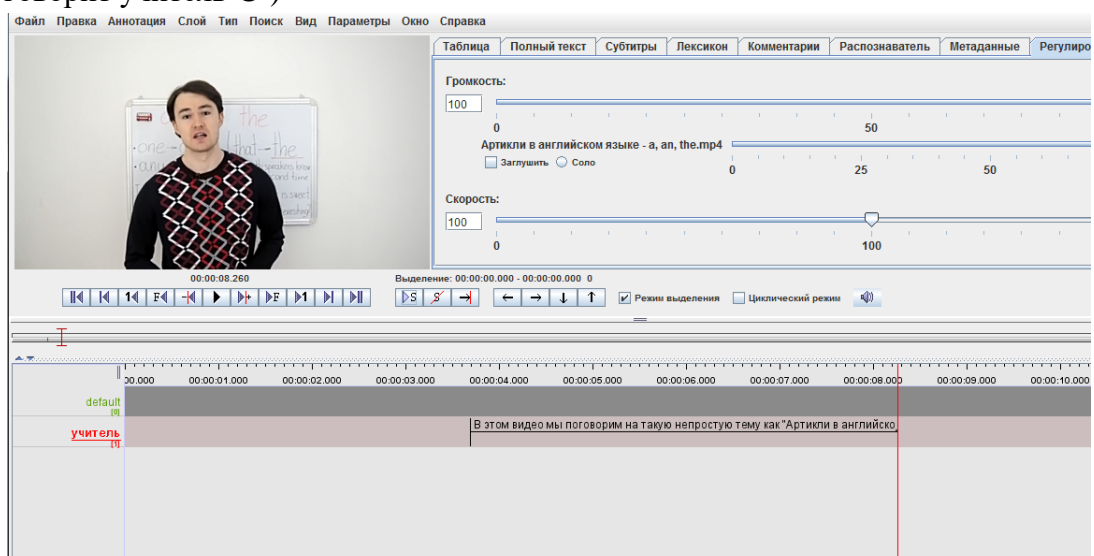
Запускаем ELAN и нажимаем Ctrl+N, чтобы создать новый файл. В открывшемся окне выбираем скачанное видео и нажимаем «ОК».



Далее добавим новый слой, на котором будут находиться наши субтитры. Выбираем Слой – Добавить слой... Название слоя и участник слоя могут быть одинаковыми, в разметчиках указываем свое имя. В дополнительных настройках можно выбрать цвет аннотаций, цвет фона и шрифт.



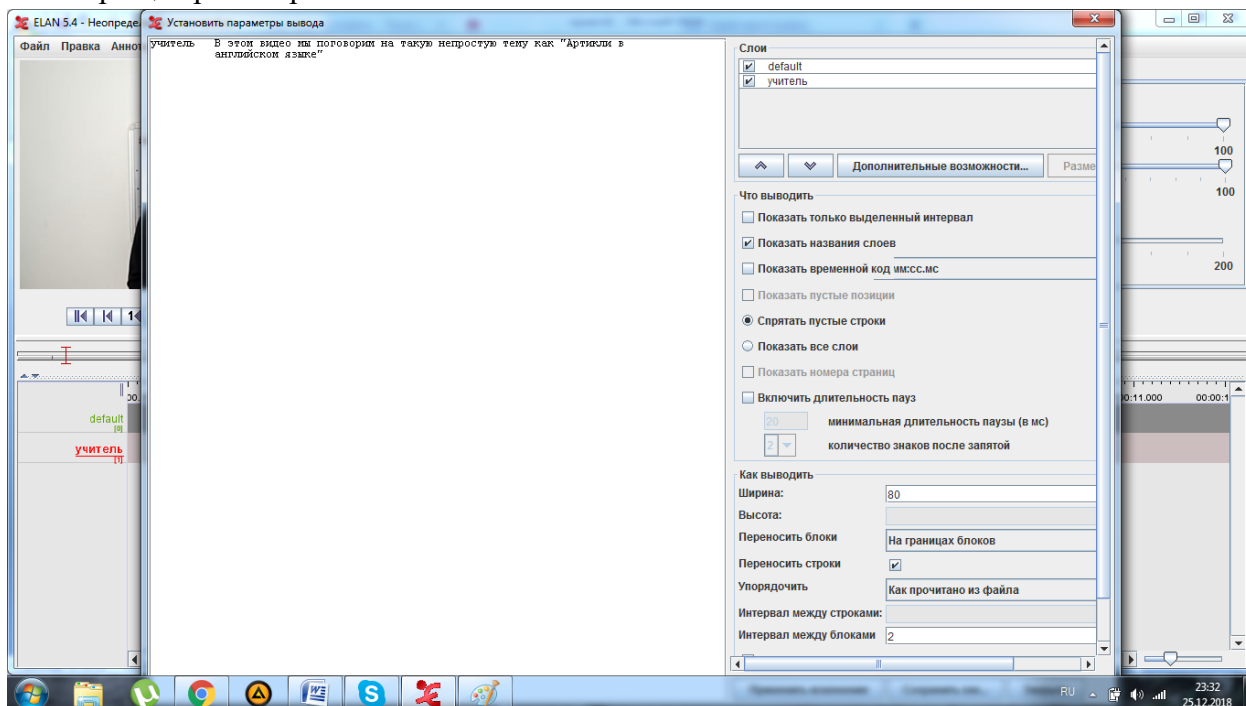
Теперь, имея этот слой, мы можем переключиться на него, дважды щелкнув по названию. После этого мы выбираем определенный участок, который мы хотим затранскрибировать, выделяя его мышкой (drag and drop). Дважды кликнув на выделенное поле нашего слоя откроется окно, в котором мы можем печатать что угодно (но в нашем случае то, что говорит учитель 😊)



Таким образом можно сделать субтитры ко всему видео.

Далее надо его экспортировать. Нажимаем Файл – Экспортировать как... -

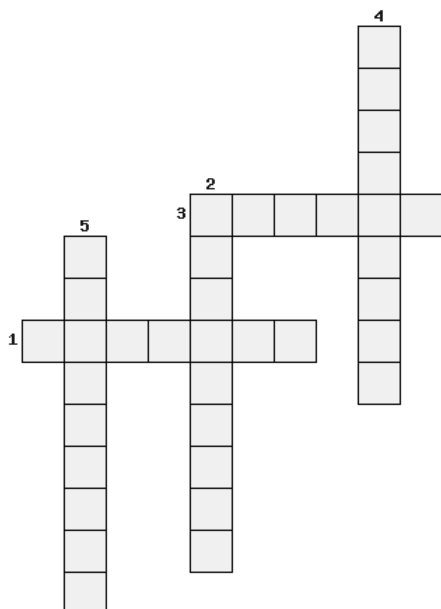
Многострочный текст. Желательно включить функцию «Показать временной код», чтобы было проще ориентироваться.



Сохраняем txt-файл, который позже можно перекодировать в srt и наслаждаться полноценными субтитрами поверх видео!

Закрепление материала.

1. Пройдите забавный тест на Quizizz! <https://join.quizizz.com> (код викторины **837461**)
2. Оцифруйте рукописные материалы из своих лекций так, чтобы их легко было распечатать!
3. Разгадайте кроссворд:



1. Что распознается в OCR?
2. Что можно создавать в ELAN?
3. Извлечение значимой информации из изображения
4. Повторяющиеся фотографии, которые можно искать с помощью анализа изображения
5. Перевод документов в электронный вид.

4. Соотнесите название программ/сайтов с их задачами:

ELAN	оцифровка изображений
GIMP	перевод документов в word-формат
Finereaderonline	создание аннотаций

5. Подготовьте развернутую презентацию на тему «Компьютерное зрение: сейчас и в перспективе»

Используемые материалы:

<https://tla.mpi.nl/tools/tla-tools/elan/elan-description/>

<https://habr.com/company/yandex/blog/203136/>

<https://habr.com/company/yandex/blog/251161/>