

# Исследование и реализация методов автоматического распознавания САРТСНА различных форматов на основе нейросетевых моделей

Студент группы 5.306М: Лаптев Александр Владимирович  
Научный руководитель: к.ф.-м.н., доцент Калачев Александр  
Викторович

11 июня 2025 г.

# Актуальность работы

Актуальность данной работы обусловлена как возрастающей сложностью CAPTCHA-систем, так и развитием инструментов, позволяющих преодолевать защитные механизмы web-ресурсов.

Анализ эффективности и разработка подходов для автоматизированного решения CAPTCHA могут применяться не только с точки зрения изучения устойчивости самих систем, но и в рамках исследования прикладного применения нейросетевых моделей в задачах распознавания информации в условиях ограничений.

# Цель работы

Целью работы является разработка и анализ комплексного подхода к автоматизации решения САПТСНА в различных форматах с использованием современных нейросетевых инструментов и API для распознавания.

# Задачи работы

Для достижения поставленной цели были сформулированы следующие задачи:

- 1 провести обзор существующих форматов CAPTCHA и методов их защиты;
- 2 разработать систему автоматического распознавания текстовых CAPTCHA с искажениями;
- 3 реализовать подход к решению графических CAPTCHA на основе методов компьютерного зрения и нейросетевых моделей;
- 4 построить решение для аудио CAPTCHA с использованием средств автоматического распознавания речи;
- 5 протестировать реализованные решения в реальных условиях, оценить точность распознавания и стабильность работы.

# Популярные форматы CAPTCHA

Проверочный код CAPTCHA – метод защиты, основанный на принципе аутентификации «вызов-ответ», предназначен для предотвращения различных автоматических действий путем выполнения пользователем простого теста, подтверждающего, что он человек, а не программа.

Наиболее популярными форматами CAPTCHA являются:

- 1 текстовый формат;
- 2 аудио формат;
- 3 графический формат.

# Пример САРТЧНА в текстовом формате



# Пример CAPTCHA в аудио формате

Press PLAY and enter the words you hear

PLAY



VERIFY





# Подходы к автоматизированному решению CAPTCHA

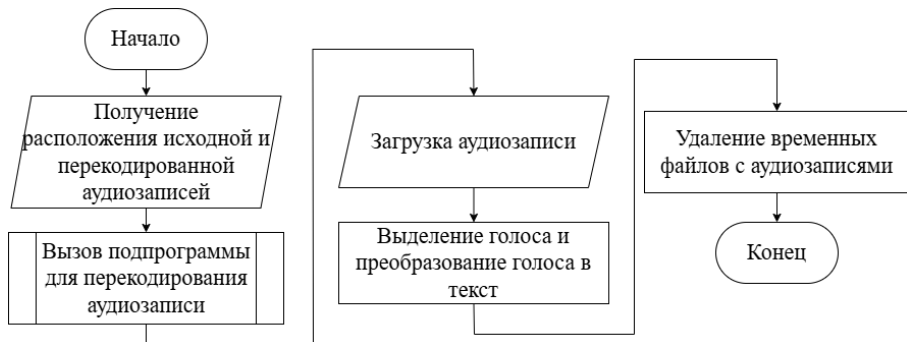
Подходы, которые использовались для автоматизации решения CAPTCHA в различных форматах:

- 1 аудио формат: облачный API с поддержкой продвинутых моделей автоматического распознавания речи (ASR);
- 2 текстовый формат: модель последовательного обучения (Seq2Seq) и алгоритмы шумоподавления на изображениях;
- 3 графический формат: одноэтапная модель для детекции объектов (YOLO) с поддержкой сегментации.

# Обработка аудиофайла САРТСНА

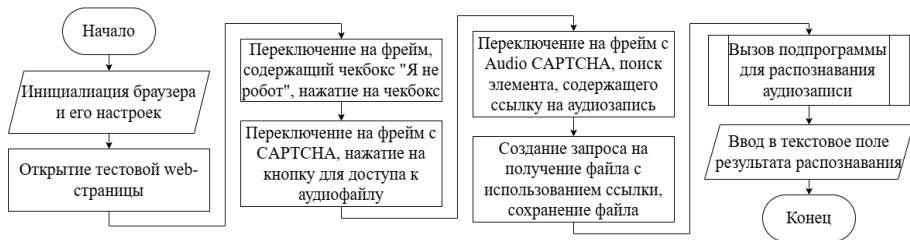
Процесс обработки аудиофайла состоит из нескольких этапов:

- 1 преобразование формата аудиофайла;
- 2 распознавание речи в аудиофайле;
- 3 сохранение результата распознавания.



Блок-схема алгоритма распознавания аудио САРТСНА.

# Тестирование решения для автоматизации решения аудио CAPTCHA



Блок-схема алгоритма решения аудио CAPTCHA.

# Подготовка датасета с текстовыми CAPTCHA

Для обучения модели был создан датасет из 100 000 изображений с текстовыми CAPTCHA, сгенерированными с использованием библиотеки `capcha` на Python. Датасет включает в себя следующие символы: ABCDEFGHIJKLMNOPQRSTUVWXYZ23456789.

Каждое изображение прошло этап предобработки, как показано на рисунке ниже:



а)



б)

Изображения CAPTCHA: а) – сгенерированное изображение, б) – результат обработки.

# Обучение модели для автоматизации решения текстовых CAPTCHA

Исходный датасет был случайным образом перемешан и разделен на три подмножества: обучающее, тестовое и валидационное в соотношении 80:10:10.

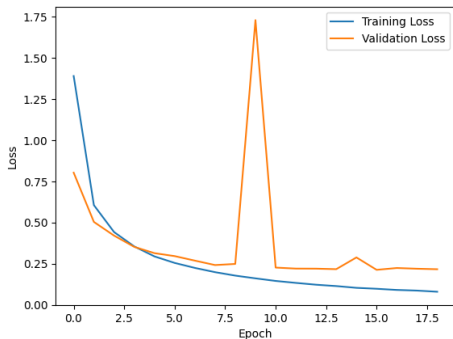
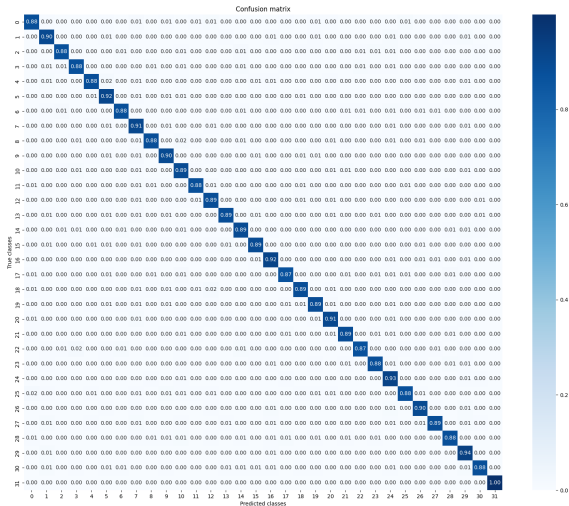


График изменения значений функции потерь в процессе обучения модели для решения текстовых CAPTCHA.

# Обучение модели для автоматизации решения текстовых CAPTCHA



Матрица ошибок обученной модели для решения текстовых  
CAPTCHA.

Точность распознавания моделью отдельных символов составила 0.9263.

Точность распознавания последовательностей различной длины представлена в таблице ниже.

Точность предсказаний для последовательностей различной длины.

Длина последовательности	Точность распознавания
4 символа	0.9305
5 символов	0.7450
6 символов	0.4575
7 символов	0.1915

## Подготовка датасета с графическими CAPTCHA



## Пример разметки изображения с тестовой графической CAPTCHA.



Набор классов, пути к выборкам и параметры конфигурации задаются в YAML-файле, который передается при обучении модели. Содержимое такого файла для данной модели:

---

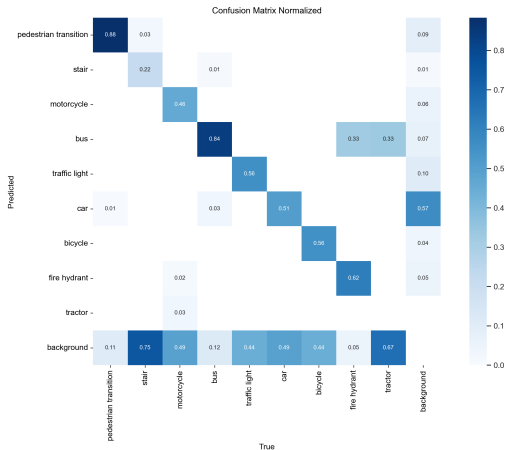
```
path: ../datasets/image_dataset
train: images/train
val: images/val

nc: 9  # Количество классов
names: ['pedestrian transition', 'stair', 'motorcycle',
  ↪ 'bus', 'traffic light', 'car', 'bicycle', 'fire
  ↪ hydrant', 'tractor']
```

---

Параметры конфигурации для обучения модели.

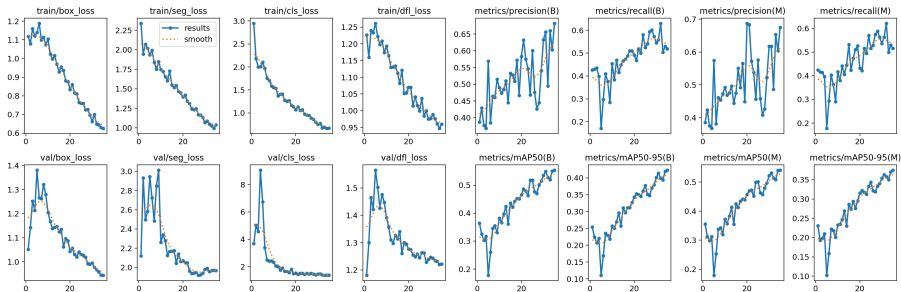
# Обучение модели для автоматизации решения графических CAPTCHA



Матрица ошибок для изображений валидационной выборки для модели YOLOv8.

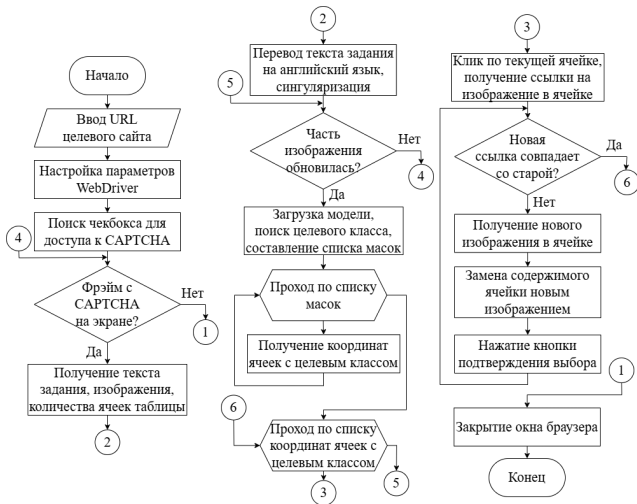
# Обучение модели для автоматизации решения графических CAPTCHA

Результаты обучения модели на основе YOLO отслеживались по ключевым метрикам (IoU, Precision, Recall, Loss), которые визуализировались автоматически. Примеры графиков с результатами обучения приведены ниже:



Изменение ключевых метрик в процессе обучения модели YOLOv8.

# Тестирование модели для автоматизации решения графических CAPTCHA



Блок-схема алгоритма решения графических CAPTCHA.

# Заключение

В результате выполненной работы были решены следующие задачи:

- 1 проведён обзор форматов CAPTCHA и существующих методов защиты от автоматических атак;
- 2 реализована система для распознавания CAPTCHA в текстовом формате на основе нейросетевой модели Sequence-to-Sequence;
- 3 создано решение для графических CAPTCHA с использованием модели YOLO, адаптированной для распознавания объектов на изображениях;
- 4 реализован подход к решению CAPTCHA в аудиоформате с использованием облачного API распознавания речи;
- 5 проведено тестирование всех компонентов системы в условиях, приближенных к реальным, с подтверждением их корректной и стабильной работы.

Перспективы дальнейших исследований включают:

- 1 расширение набора поддерживаемых типов CAPTCHA, включая более сложные динамические варианты;
- 2 оптимизацию времени обработки и точности распознавания;
- 3 исследование механизмов защиты CAPTCHA, устойчивых к современным методам автоматического анализа.