

# AutoQA

РАЗРАБОТКА РОБОТА ДЛЯ  
АВТОМАТИЧЕСКОГО ТЕСТИРОВАНИЯ  
ВЕБ ПРИЛОЖЕНИЙ



Михаил Ющук  
DS Department Lead  
[m.yushchuk@quantumobile.com](mailto:m.yushchuk@quantumobile.com)





# МИХАИЛ ЮЩУК

## DS DEPARTMENT LEAD

- ✘ Руководжу командой из 15 data scientists
- ✘ Занимаюсь решением задач, которые ранее не решались
- ✘ Сотрудничаю с университетами для развития отрасли
- ✘ Пишу публикации для научных журналов
- ✘ Руководжу отделом R&D

# AutoQA

## АВТОМАТИЗАЦИЯ ТЕСТИРОВАНИЯ WEB ПРИЛОЖЕНИЙ

При текущем подходе в тестировании известны следующие недостатки:

- ✘ Ручное тестирование, долго и ненадежно.
- ✘ Автоматическое тестирование чувствительно к изменению кода веб сайта.
- ✘ Любой из подходов увеличивает время на разработку функционала и стоимость проекта на **30-50 %**

Мы решили разработать робота, который мог бы тестировать веб приложения с минимальным вовлечением людей. Мы назвали его SAM

# SAM. ИМИТАЦИОННОЕ ОБУЧЕНИЕ

## IMITATION LEARNING

Имея набор демонстраций, мы можем обучить SAM подражать выполненным действиям.

Особенности подхода:

- ✘ Чтобы успешно выполнить сценарий, робот должен повторить все шаги сценария без единой ошибки
- ✘ Робот должен следовать всем шагам на всех сценариях, у нас нет валидационных или тестовых выборок
- ✘ Робот не использует историю ввода и работает только с текущим состоянием экрана

Для обучение подражанию выбрали полносверточную нейронную сеть

# SAM



# SAM. ДЕМО

Successful  
and keys, Enter to send, Esc

or Failure in case of

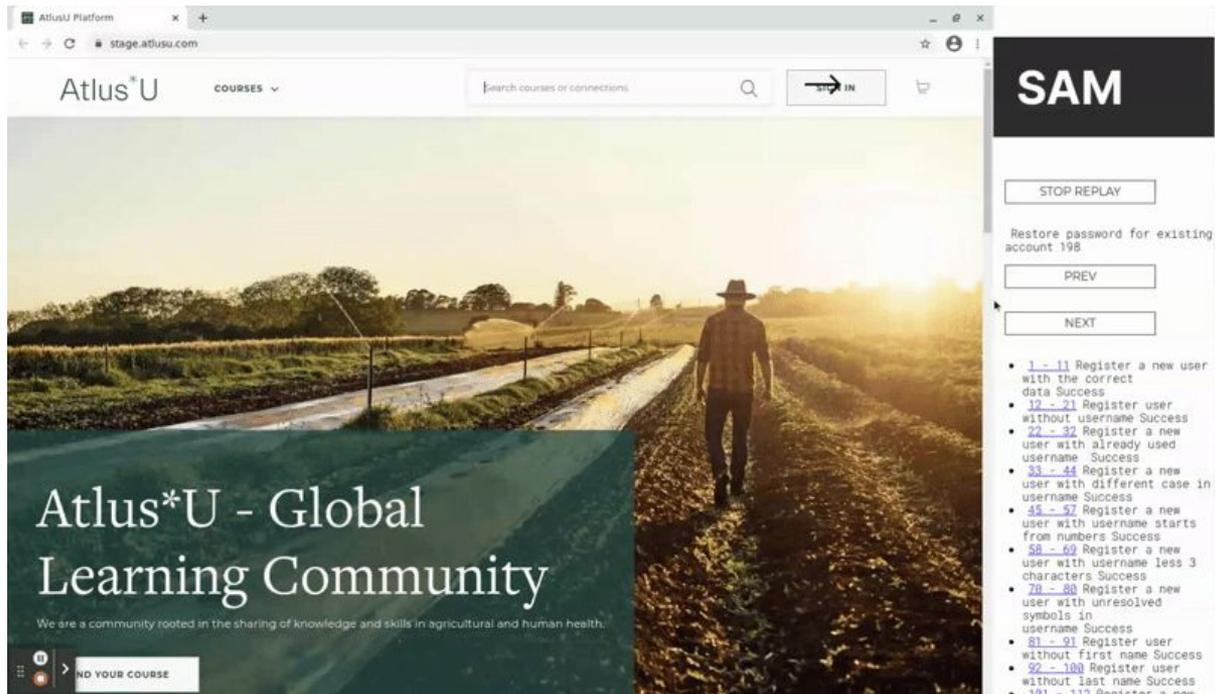
FAILURE

do not want to record this

Find out what model suggests

PREDICT

Связаться с технической поддержкой



## SAM. ДАННЫЕ

Датасет создан вручную с помощью SAM.

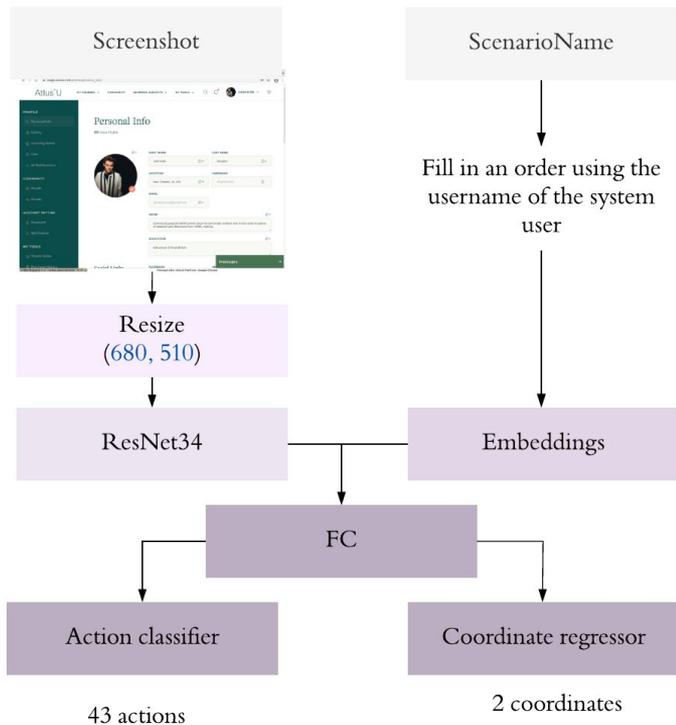
- ✘ **4520** снимков экрана
- ✘ **311** сценариев

Для каждого снимка даны целевые:

- ✘ Координаты клика
- ✘ Данные ввода

# СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

## АРХИТЕКТУРА



Модель принимает:

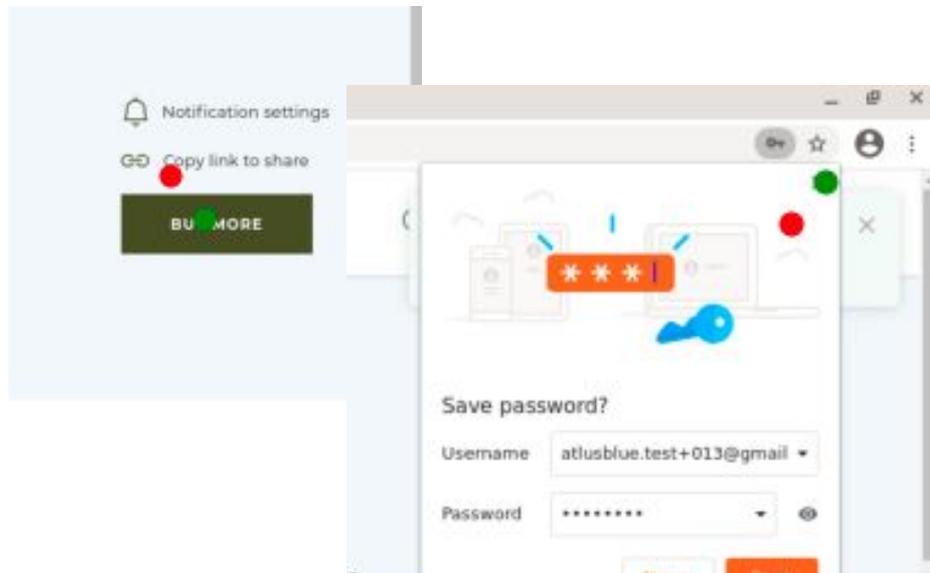
- ✘ Снимок экрана с пониженной дискретизацией в два раза
- ✘ Название сценария

И возвращает:

- ✘ Координаты клика (масштабированные значения X, Y)
- ✘ Классификацию действий (43 класса, напр. CLICK, SEND\_EMAIL, SEND\_PASSWORD, FINISH\_EPISODE и т.д.)

# СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

## РЕЗУЛЬТАТ РАБОТЫ



- ✘ Точность классификации действий: **96.90%**
- ✘ RMSE положения кликов: **93 px**. Клик часто находится рядом с целевым, но не попадает в необходимый элемент.
- ✘ Точность обнаружения в базовых UI элементов: icons ~35%, buttons ~75%, text fields ~90%

Было принято решение - помочь искать текст в кнопках с помощью модели CharGrid

# CHARGRID

ПОМОЖЕМ МОДЕЛИ НАЙТИ БУКВЫ

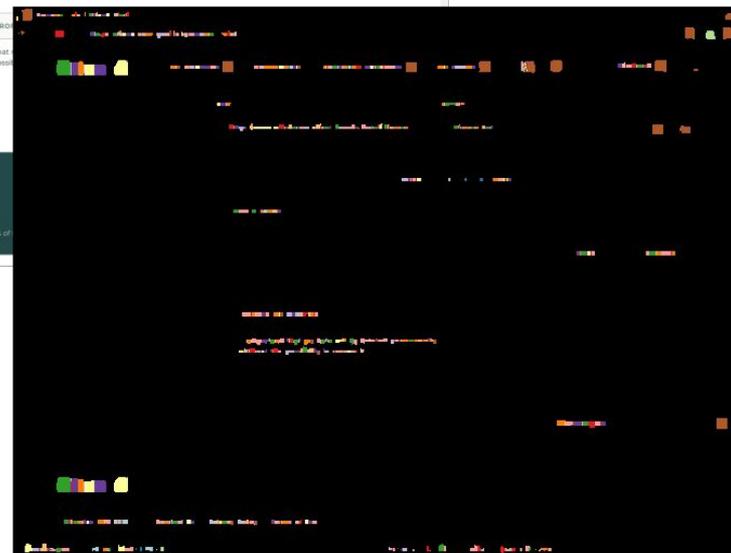
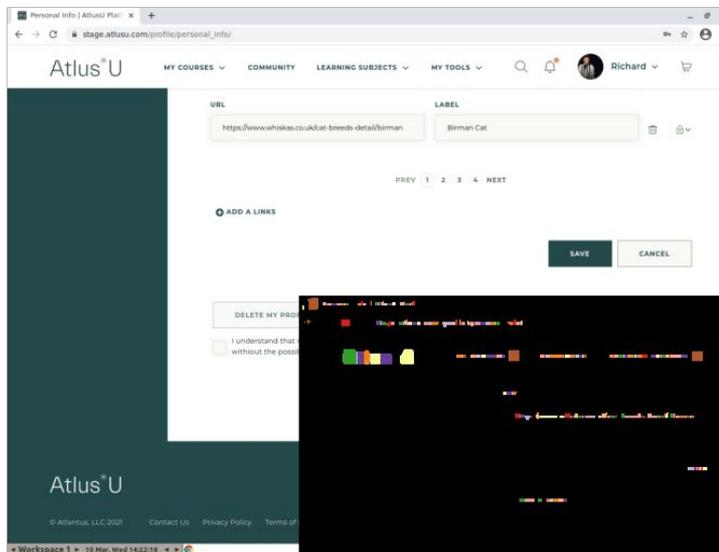
UNET для сегментации символов

Обучающий набор:

- ✘ Синтетические данные (110к картинок 64x56, символы: a-zA-Z0-9 + 41 UI элементов)

Результаты:

- ✘ Метрика IoU: 99.3%
- ✘ 5 FPS



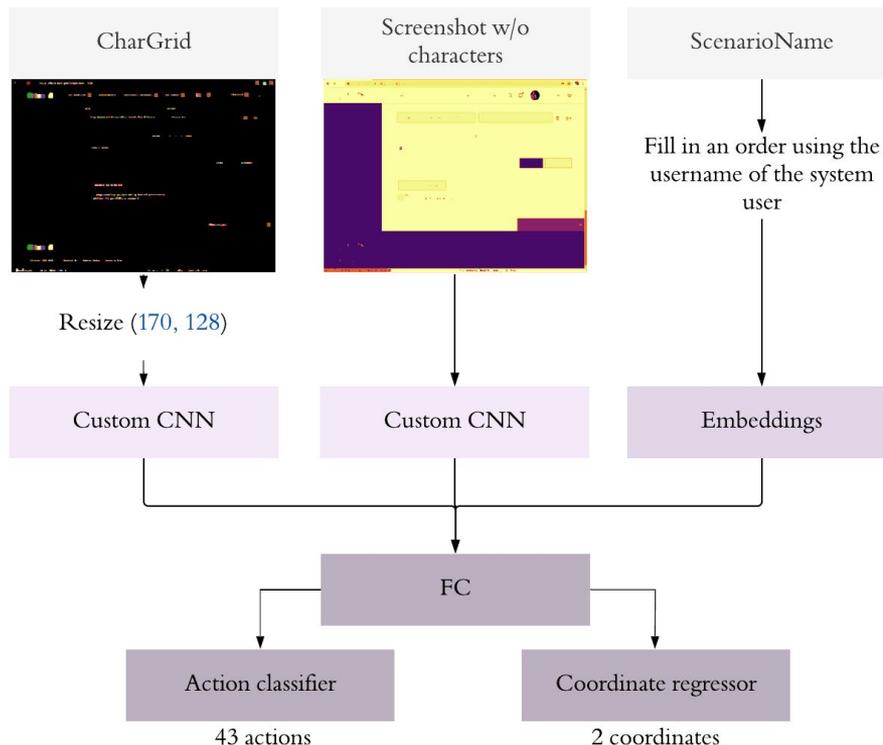
# СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ НА ОСНОВЕ CHARGRID

Модель принимает:

- ✘ Сегментацию символов (41 канал, 170x128)
- ✘ Снимок экрана без символов
- ✘ Название сценария

И возвращает:

- ✘ Координаты клика (масштабированные значения X, Y)
- ✘ Классификацию действий (43 класса, напр. CLICK, SEND\_EMAIL, SEND\_PASSWORD, FINISH\_EPISODE И Т.Д.)



# СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

## РЕЗУЛЬТАТЫ

✂ Точность классификации действий: ~~96.90%~~  
**99.91%**

✂ RMSE положения кликов: ~~93 px~~ **26 px**



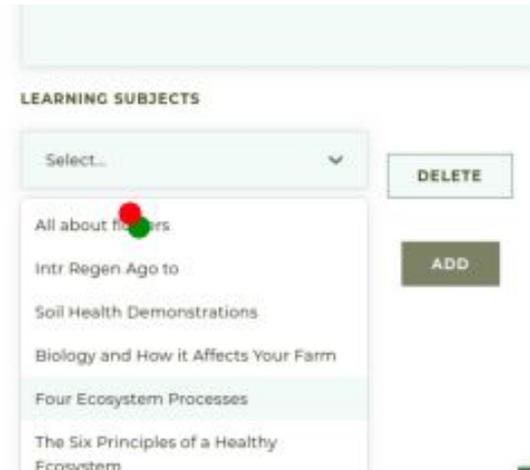
EMAIL

atiusblue.test+013@gmail.com

PASSWORD

SUBMIT

Forgot password? Sign Up



LEARNING SUBJECTS

Select...

DELETE

ADD

- All about farmers
- Intr Regen Ago to
- Soil Health Demonstrations
- Biology and How it Affects Your Farm
- Four Ecosystem Processes
- The Six Principles of a Healthy Ecosystem

По-прежнему, остается проблема с представлением UI элементов.

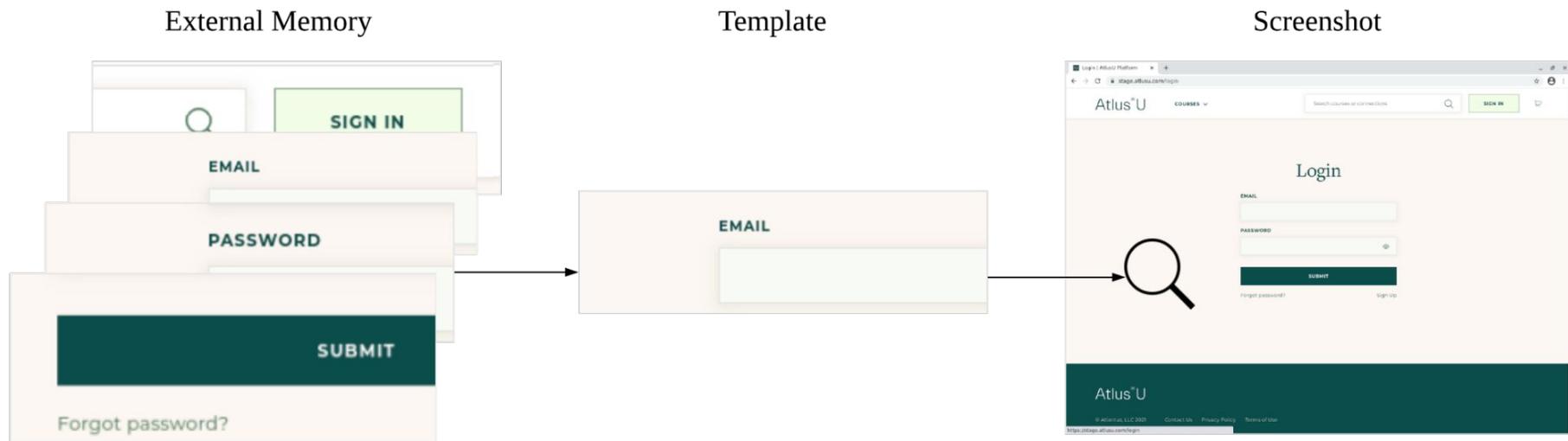
# ИТОГ ПЕРВОЙ ИТЕРАЦИИ

## ДАЛЬНЕЙШИЕ ДЕЙСТВИЯ

- ✘ При прохождении сценариев робот сталкивался с ситуациями где появлялись одинаковые скриншоты, но требовалось различное поведение. В таких случаях модель работала не корректно.
- ✘ Низкая точность предсказания клика по UI элементу веб страницы.
- ✘ Низкая скорость обработки картинок. 1 FPS
- ✘ Мы смогли завершить **5/311** сценариев успешно.

Наша модель не попадает в иконки, стало понятно, что данный подход не рабочий  
Решили выбрать иной подход - агенты с внешней памятью

# АГЕНТЫ С ВНЕШНЕЙ ПАМЯТЬЮ



End-To-End Memory Networks ([arXiv:1503.08895](https://arxiv.org/abs/1503.08895))

# ПОИСК ШАБЛОНА

## И ПЕРЕХОД К ГЛУБОКОМУ ОБУЧЕНИЮ

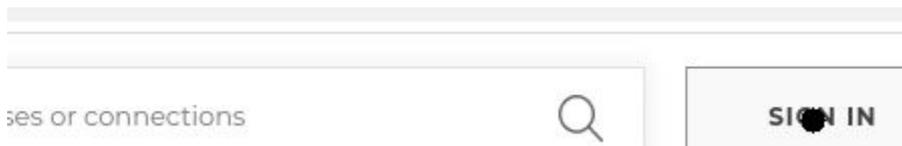
Используем алгоритм поиска шаблонов для нахождения элементов, по которым необходимо совершить клик

**Внешняя память:** фрагменты изображений в области каждого клика

**Цель:** находить фрагмент изображения в области клика

**Требования к модели:**

- ✘ Модель должна различать похожие элементы дизайна
- ✘ Модель должна быть устойчива к изменениям внешнего вида страницы
- ✘ Модель должна быть устойчива к смещению шаблона



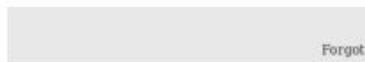
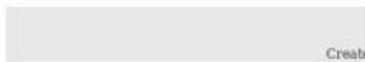
# ПОИСК ШАБЛОНА

## ДАННЫЕ

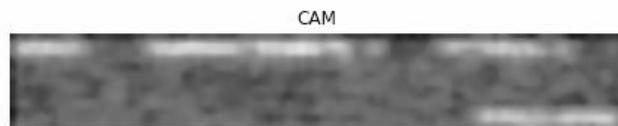
False



False



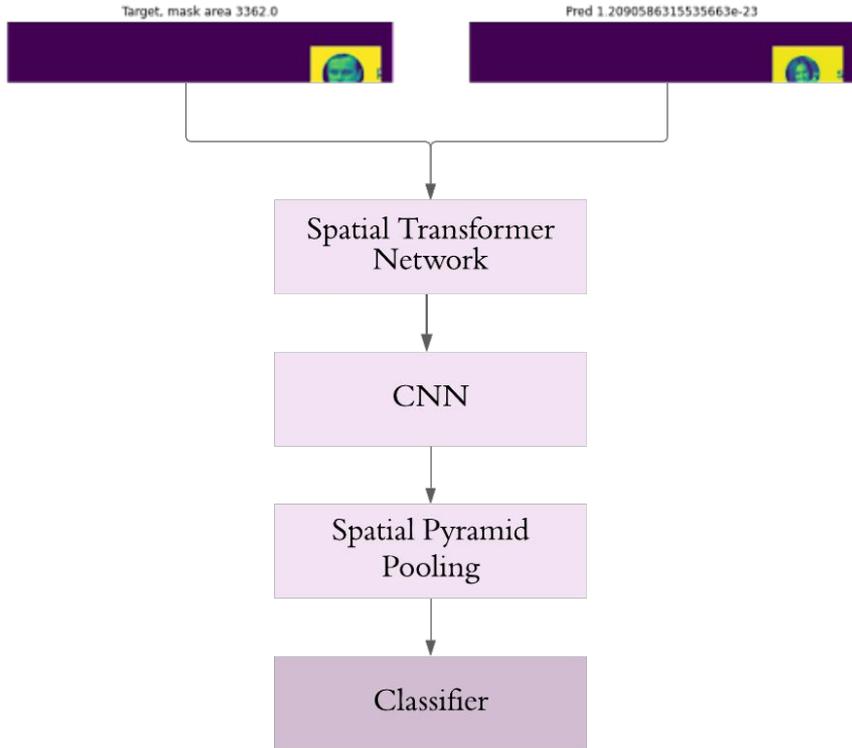
True



- ✘ **Обучающая выборка:** набор скриншотов веб страниц, метки для которых созданы при помощи сверточной модели. Датасеты ([RoboFlow Website Screenshots](#), [Phish-Iris Dataset](#)). Всего 20000 картинок.
- ✘ **Тестовая выборка:** изображения и метки, собранные с использованием SAM. 4000 картинок

# MATCHINGNET

## АРХИТЕКТУРА



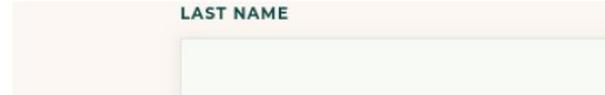
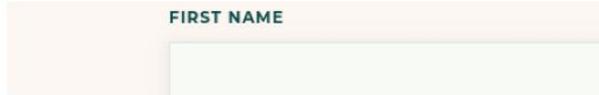
MATCHINGNET - сеть, разработанная нами

- ✘ Мы начинаем работу с простой CNN, которая принимает на вход два снимка. Полученная точность 56%
- ✘ При добавлении сложных примеров в обучающую выборку точность увеличивается до 60%
- ✘ После изменения архитектуры сети на сиамскую сеть точность возросла до 74%
- ✘ После добавления регуляризации [Temperature scaling loss](#) точность составляет 90%
- ✘ С добавлением [STN](#) и [SPP](#) финальная точность составила 95%
- ✘ Мы смогли завершить 220/311 сценариев успешно.

# MATCHINGNET

## ПРИМЕРЫ ОШИБОК

### Похожие элементы



### Полностью одинаковые элементы под маской



95% ТОЧНОСТЬ

300 FPS



# ИТОГ

МЫ НАШЛИ ПОДХОД, РЕШАЮЩИЙ  
ДАННУЮ ЗАДАЧУ

- ✘ Создан сервис для сбора и разметки действий на веб страницах - SAM
- ✘ Обучена сверточная нейронная сеть для классификации действий и регрессии координат клика, представлена модель для детектирования символов на веб страницах
- ✘ Обучен CharGrid, который сейчас применяется в других проектах компании для распознавания текста
- ✘ Обучена модель поиска шаблонов MatchingNet, которая показала лучший результат
- ✘ Количество успешно выполненных сценариев 220/311

# ЦЕЛЬ РОБОТА - УПРОСТИТЬ ЖИЗНЬ

Научить робота делать нашу ежедневную работу в сети по команде

AutoQA первый шаг на пути

# Q&A

More information at <https://quantumobile.com/>