

СОРЕВНОВАТЕЛЬНОЕ РАСПОЗНАВАНИЕ ТЕКСТА

Григорий Бартош, Александр Федотов, Михаил Ютман

ПОСТАНОВКА ЗАДАЧИ

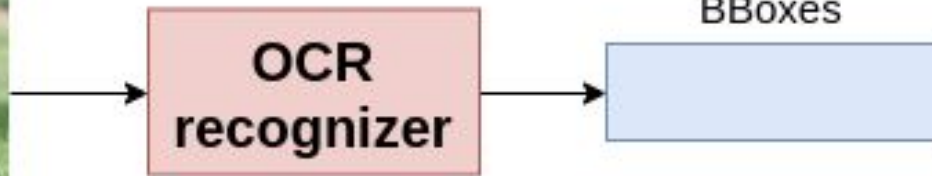
- Существуют картинки с текстом
- Хотим научиться качественно распознавать текст
- Optical Character Recognition – OCR



ОБЩАЯ СХЕМА

- OCR модели обучают на размеченных данных

Image with Text



ПРОБЛЕМЫ

- Проблема
 - OCR модели обучаются на размеченных данных
 - Таких данных мало
- Решение
 - Генерировать синтетические датасеты
- Проблема
 - Генерируются простые и смещённые данные
- Возможные решения
 - Эвристические ограничения
 - Соревновательная аугментация



ЦЕЛИ И ЗАДАЧИ

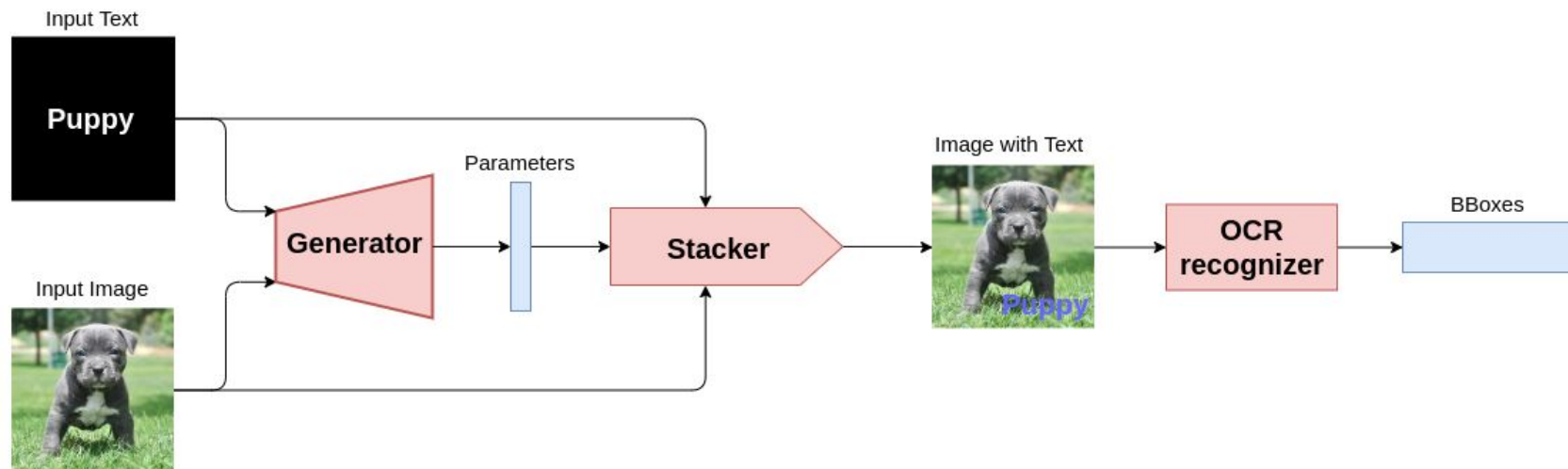
Цели

- Соревновательное обучение
 - Генерация изображений с наложенным текстом наиболее сложным образом
 - Распознавание текста

Задачи

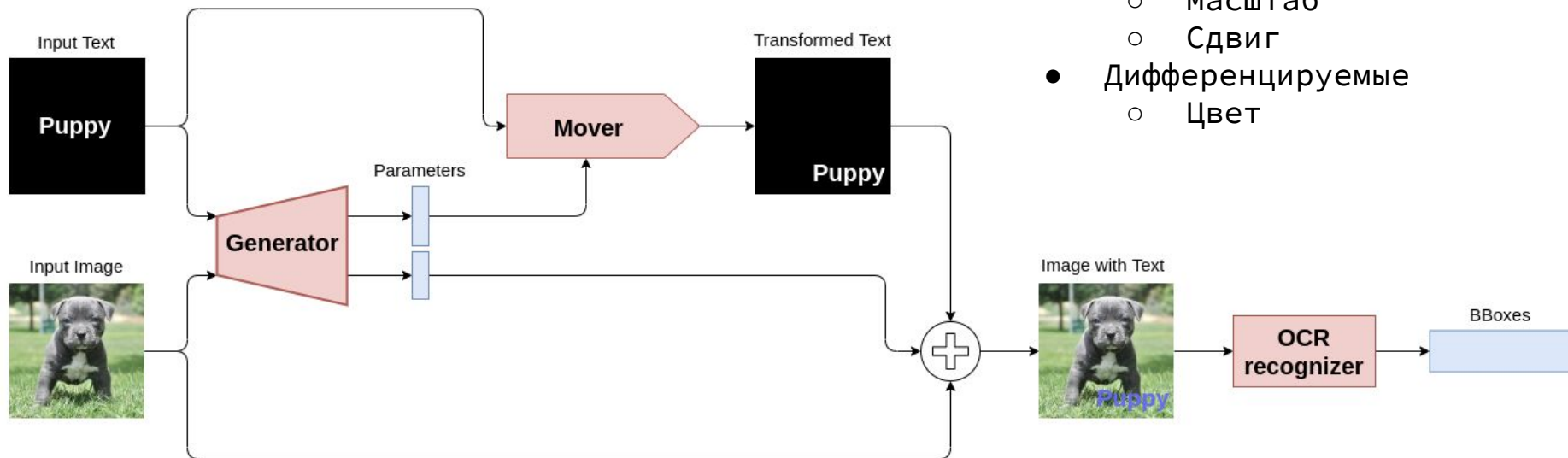
- Генерация параметров наложения
- Генерация наложения
- Распознавание текста

ОБЩАЯ АРХИТЕКТУРА



$$\max_G \min_{OCR} E_{x \sim p_{data}(x)} [Loss(OCR(G(x)))]$$

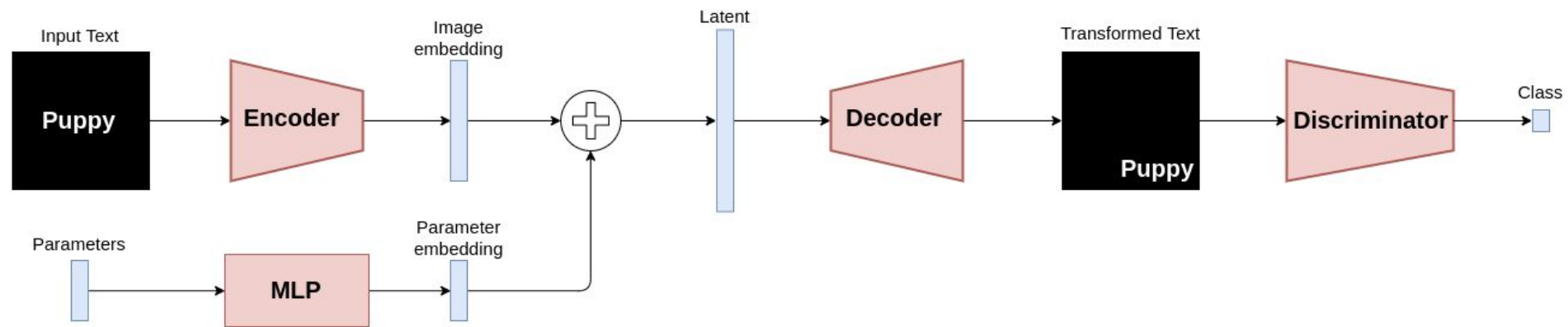
MOVER



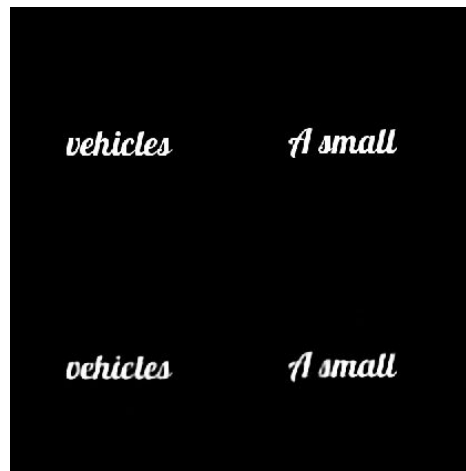
Параметры преобразования

- Не дифференцируемые
 - Поворот
 - Растяжение
 - Масштаб
 - Сдвиг
- Дифференцируемые
 - Цвет

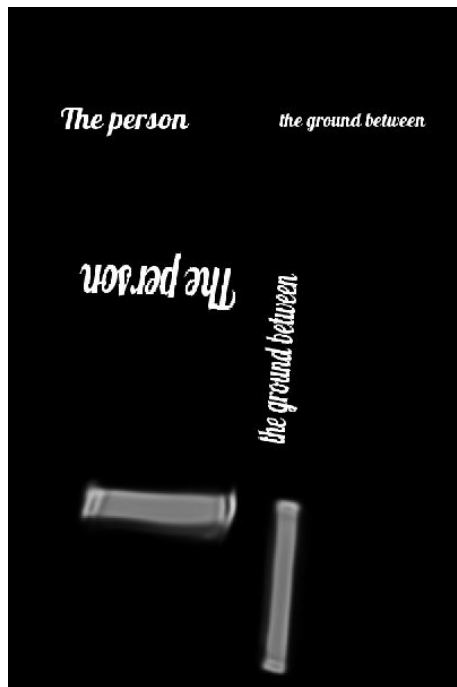
MOVER



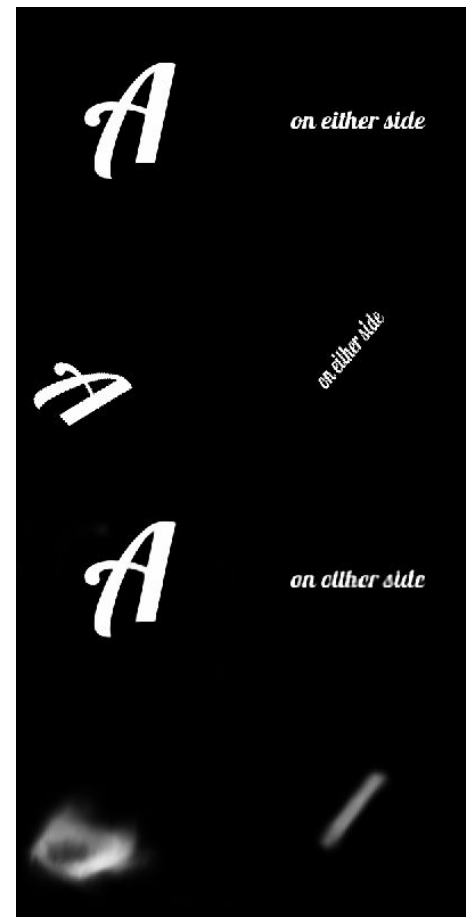
MOVER: РЕЗУЛЬТАТЫ



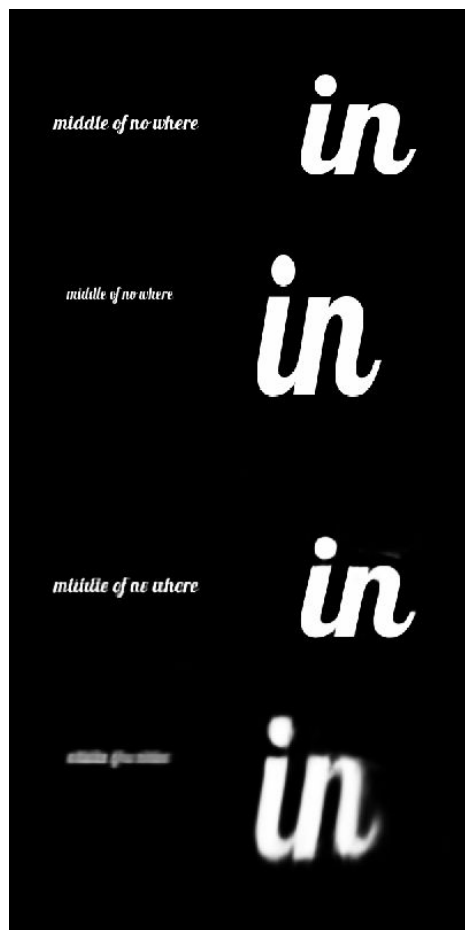
Автоэнкодер



Параметризованный автоэнкодер



Параллельное обучение



Без поворота

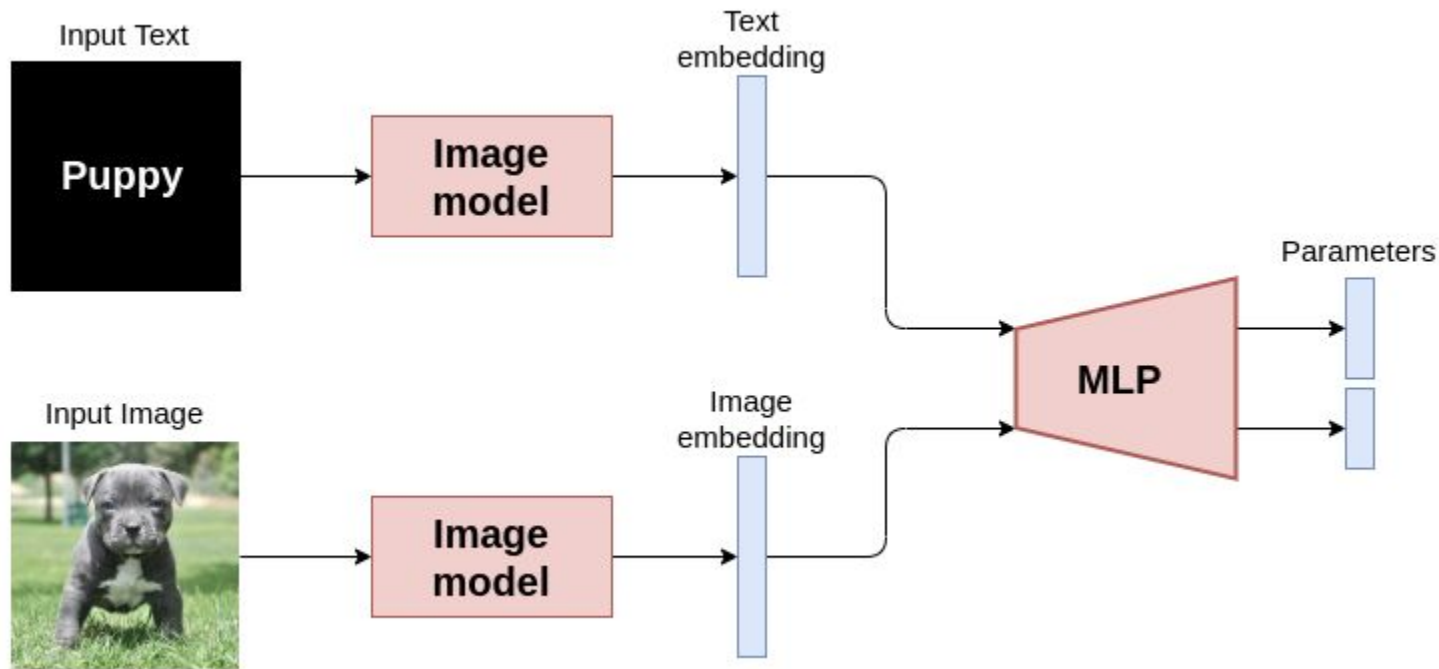


Новая функция ошибки

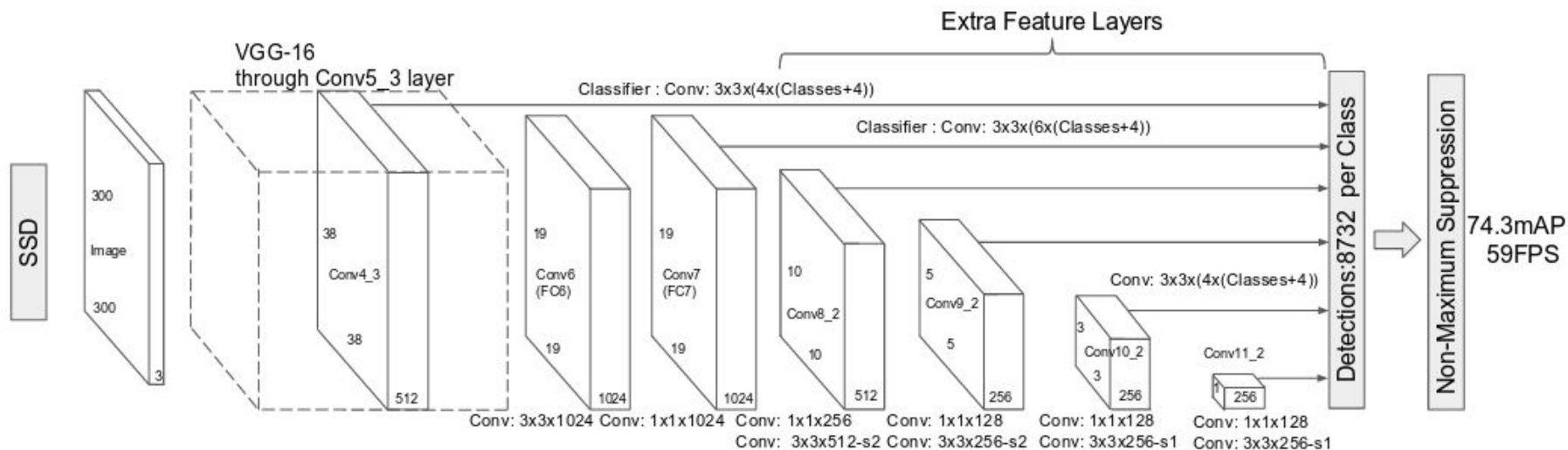


Дискриминатор

ГЕНЕРАТОР ПАРАМЕТРОВ



РАСПОЗНАВАНИЕ ТЕКСТА (OCR МОДЕЛЬ)



ИТОГИ И ВЫВОДЫ

Итоги

- Полностью реализованы пайплайны
- Исследован процесс обучение Mover'a
- Обучено распознавание текста на случайных примерах
- Не обучена полная модель, так как плохо обучен Mover

Выводы

- Преобразования отличные от изменения текстуры обучаются гораздо хуже

