

ОТЗЫВ

рецензента о дипломной работе на соискание степени бакалавра студента Лучинского Владимира Дмитриевича

Рецензируемая работа по теме «Разработка системы распознавания городских объектов» посвящена изучению и практической реализации методов глубокого обучения для идентификации объектов городской застройки на изображениях.

Тема работы является актуальной поскольку распознавание объектов городской застройки является важным компонентом систем позиционирования и дополненной реальности на мобильных устройствах. Существующие методы глобального и локального позиционирования по сигналам спутников и сотовых станций, встроенные в мобильные устройства, имеют сравнительно невысокую точность особенно по азимуту. Кроме того они подвержены радиопомехам, экранированию и спуфингу. На практике это приводит к досадным ошибкам в определении стороны улицы, направления движения, размещения меток дополненной реальности. Визуальные методы позиционирования обеспечивают при наличии геотегов намного более высокую точность позиционирования и не зависят от внешних сигналов. Особенно важна точность визуальных методов распознавания для дополненной реальности, поскольку в отличие от других методов, ошибка положения объекта в пикселях не зависит от расстояния.

В работе проведён обзор предметной области, рассмотрены существующие решения Augmented.City и DELF, а также общий подход к поиску изображений с помощью глубокого обучения метрики; была выбрана архитектура нейросети на основе архитектур DeepLabv3+, ResNet и Inception. В работе реализована библиотека на C++, которая поддерживает распознавание городских объектов; семантическую сегментацию изображений; маскирование неустойчивых элементов изображения таких как пешеходы, автомобили, деревья. Также собран новый набор изображений городских сцен Санкт-Петербурга из 60 тыс. изображений, содержащих 892 класса. Проведено обучение распознающей сети. Проведены эксперименты, по результатам которых улучшение относительно существующей системы Augmented.City по точности топ-1 составило 38%, по точности топ-4 – 23%. Точность по сравнению с DELF оказалась хуже на 5%, но производительность выше в 100 раз. Время одного запроса на NVIDIA GTX 950M порядка 4.4 сек, что близко к требованиям реального времени.

В работе можно выделить следующие недостатки:

1. Работа написана в рамках соревновательного подхода к разработке и не учитывает реальные аспекты практического использования: не учитывается необходимость добавления новых классов без обучения, количество классов крайне мало по сравнению с количеством городских построек, не учитывается наличие типовых проектов, не учитываются данные GPS.
2. Хотелось бы видеть более подробное описание собранного набора данных по строениям Санкт-Петербурга.

Поставленная в работе задача в основном выполнена. Результаты имеют высокую практическую ценность.

Несмотря на замечания, учитывая сложность задачи, самостоятельность и объем проделанной работы, считаю, что работа Лучинского Владимира Дмитриевича заслуживает оценки - "ОТЛИЧНО".

к. ф.-м. н., ст. преп. кафедры информатики
математико-механического факультета СПбГУ
09.06.2020

Салищев С. И.

