

Разработка архитектуры классификатора для оценки состояния объектов инфраструктуры с применением нейронных сетей

И.А. Абрамов, А.Ю. Камакин, Н.Д. Моисеев

Национальный исследовательский университет «Высшая школа экономики»

Геопространственный анализ – процесс изучения данных, содержащих географические компоненты, для выявления и изучения пространственных закономерностей. В наше время особое внимание уделяется его применению при планировании инфраструктуры городов [1]. Качественный анализ геопространственных данных позволяет выявлять проблемные зоны отдельных улиц и районов, что способствует эффективному распределению ресурсов при реализации проектов, направленных на улучшение качества жизни горожан.

С увеличением объема данных, содержащих географическую информацию, все больший приоритет отдается автоматизации их анализа, а с развитием глубокого обучения (deep learning), чаще встает вопрос о его применении в задачах геопространственного анализа. Геопространственный анализ активно использует методы автоматизированного сбора данных [2], что позволяет эффективно формировать датасеты для обучения моделей искусственного интеллекта.

В данной работе представлен программный комплекс, включающий в себя модуль сбора и подготовки изображений для формирования датасета, модуль классификаторов, а также модуль интерактивной веб-карты, демонстрирующей результат работы классификаторов. Карта предоставляет пользователю информацию о состоянии зданий, их назначении, а также иную информацию об особенностях различных районов города Москвы.

Модуль классификаторов состоит из комплексной архитектуры, состоящий из нескольких предобученных сверточных нейронных сетей, предназначенных для определения различных характеристик здания с изображения. Подход с использованием нескольких классификаторов позволяет получить более высокую точность при определении типа строения (жилое, коммерческое, промышленное и так далее) в условиях неравномерного распределения классов в датасете.

Помимо подготовки архитектуры, особое внимание уделяется исследованию возможности интеграции методов обработки естественного языка (NLP) и визуально-языковых моделей (VLM) для повышения точности классификации. Среди рассматриваемых моделей VLM, представлены PaliGemma 2, демонстрирующая высокую точность определения параметров объектов на изображении.

Исследование существующих архитектур классификаторов, также является одним из аспектов данной работы. Проведено исследование таких известных архитектур, как YOLO, RF-DETR, ResNet-50 и VGG19. Оценивались не только точность, в виде сравнения метрик Precision и Mean Average Precision (mAP), но и скорость обработки изображения. Данная оценка позволит выбрать наиболее подходящую модель, которая будет использоваться в нашем программном комплексе.

Литература

1. Kovacs-Györi A., Ristea A., Havas C., Mehaffy M., Hochmair H., Resch B., Juhasz L., Lehner A., Ramasubramanian L., Blaschke T. Opportunities and Challenges of Geospatial Analysis for Promoting Urban Livability in the Era of Big Data and Machine Learning // ISPRS International Journal of Geo-Information. 2020. Vol. 9, no. 12. P. 752. DOI: 10.3390/ijgi9120752.
2. Nekrasov H.A., Polivoda D.E., Prokofyeva E.N. The Geospatial Data Mining Concept Using Scrapping Technology // 2020 IEEE Moscow Workshop on Electronic and Networking Technologies, MWENT 2020, March 1-6, 2020, Moscow, Russia. IEEE, 2020. DOI: 10.1109/mwent47943.2020.9067406.