

**Геоинформационный набор данных для нейросетевого распознавания
источников загрязнения атмосферы**

Научный руководитель – Сарычев Дмитрий Владимирович

Мещерякова Анастасия Павловна

Студент (бакалавр)

Воронежский государственный университет, Воронеж, Россия

E-mail: nastya11.09.09.03@gmail.com

Загрязнение воздуха техногенными веществами является угрозой для экологической обстановки. Наиболее сильно это наблюдается в крупных городах. Для оперативной оптимизации состояния окружающей среды необходим соответствующий мониторинг [1]. Перспективным подходом к его выполнению является синтез технологий искусственного интеллекта (ИИ), дистанционного зондирования Земли (ДЗЗ) и геоинформационных систем (ГИС).

Цель работы заключалась в подготовке ГИС-набора данных для создания нейросетевой модели ИИ со способностью оперативного картографирования потенциальных источников загрязнения атмосферы по данным ДЗЗ. Методика работы состояла в: 1) дешифрировании целевых объектов на космических снимках высокого пространственного разрешения; 2) создании векторных представлений этих объектов – разметки; и 3) конвертации снимков и разметки в формат данных, необходимый для обучения и тестирования свёрточных нейронных сетей. Объектами дешифрирования выступали стационарные источники загрязнения атмосферы, видимые на спутниковых снимках высокого разрешения (рис. 1): устья дымовых труб и газоходов, жерла градирен, осуществляющие выход горячего воздуха (пара), и другие устройства для промышленных выбросов. В качестве модельных территорий взяты города Воронеж, Липецк и Тамбов – доступные автору для полевых обследований в случае необходимости подтверждения результатов дистанционного дешифрирования.

Для выбранных городов было осуществлено визуальное дешифрирование перечисленных потенциальных источников загрязнения атмосферы на космических снимках, предоставленных сервисами Google, Яндекс, ESRI и Bing. Векторная разметка выявленных объектов выполнялась в геоинформационной среде QGIS. Для каждого источника снимков формировался отдельный векторный слой разметки с классификацией выявленных целевых объектов по внешним характеристикам, размеру и степени их распознаваемости вследствие возможного наличия пара, дымки и других особенностей исходных изображений (рис. 2). Подготовленная разметка вместе с исходными снимками были сконвертированы в классы визуальных объектов PASCAL (формат данных Pattern Analysis, Statistical Modeling and Computational Learning, являющийся одним из стандартов для задач распознавания объектов методами машинного обучения).

По результатам дешифрирования для исследованных городов, в зависимости от источника, было выявлено следующее количество объектов: г. Воронеж – 217 (Яндекс), 222 (Google), 206 (ESRI), 183 (Bing); г. Липецк – 296 (Яндекс), 258 (Google), 268 (ESRI); г. Тамбов – 37 (Яндекс), 59 (Google), 38 (Bing). Преобладание стационарных источников загрязнения атмосферы в г. Липецк обусловлено его металлургической специализацией. Таким образом, результатом текущего этапа работ стала база данных, включившая 1784 образцов изображений целевых объектов по трём городам Центрального Черноземья. Результаты работы будут использованы для обучения и тестирования моделей ИИ с целью создания автоматической системы мониторинга источников загрязнения [2].

Источники и литература

- 1) Пасхина М. В. Разработка объектной структуры геоинформационных систем для задач экологического проектирования городской среды / М. В. Пасхина // Ярославский педагогический вестник. – 2011. – Т. 3, № 4. – С. 181-186.
- 2) Экологические риски аэротехногенного загрязнения промышленных городов Центрального Черноземья: монография / Под общей редакцией С.А. Куролала и О.В. Клепикова. – Воронеж: Издательство «Цифровая полиграфия», 2024. – 287 с.

Иллюстрации



Рис. : 1. Вид стационарного источника выбросов с земли и на фрагментах космических снимков на примере котельной Электростанции 1915 года, город Воронеж

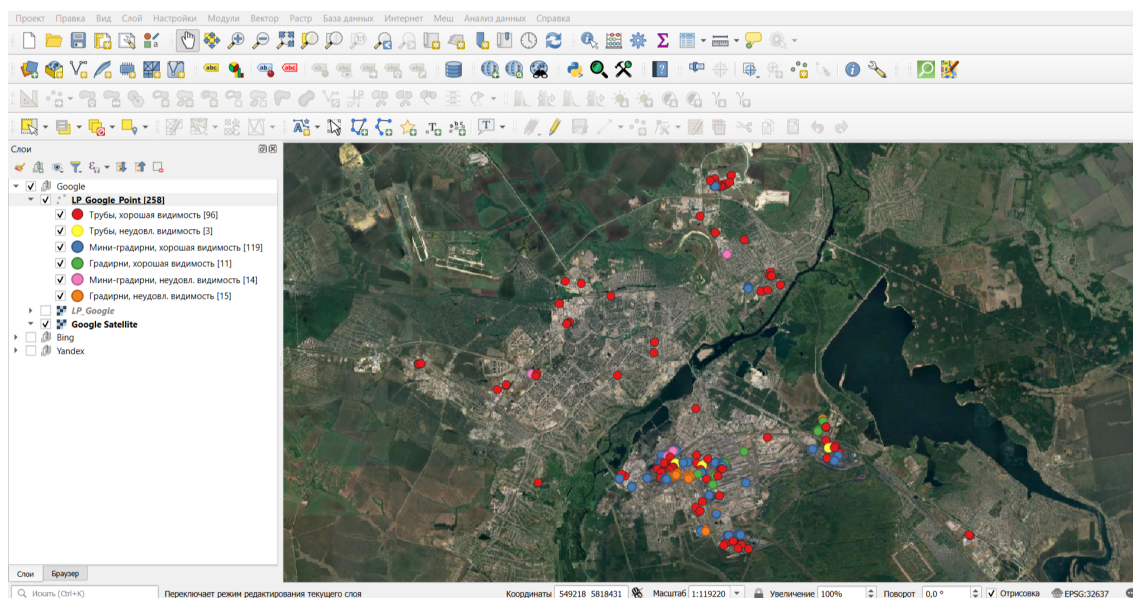


Рис. : 2. Рабочее пространство QGIS с выделенными в векторном слое потенциальными источниками загрязнения атмосферы