

В. Б. Мокін¹
Л. М. Скорина¹
Є. М. Крижановський¹
М. А. Гораш¹

ІНТЕЛЕКТУАЛЬНА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПОШУКУ НЕЗАРЕЄСТРОВАНИХ МІСЦЬ НАДХОДЖЕННЯ ЗАБРУДНЕНИХ ВОД У РІЧКУ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

¹ Вінницький національний технічний університет

Анотація

Запропоновано інтелектуальну інформаційну технологію пошуку незареєстрованих місць надходження забруднених вод у річку. Технологію реалізовано на основі запропонованого CIR-методу виділення вагомих ознак із використанням методів машинного навчання для ефективного і точного встановлення знаходження об'єктів.

Ключові слова: забруднення річки, стічні води, features-інжиніринг, розпізнавання зображень, машинне навчання, Keras, TensorFlow.

Abstract

The intellectual information technology of search of unregistered places of the receipt of contaminated water in the river is proposed. The technology is implemented on the basis of the proposed CIR-method for the allocation of features using the methods of machine learning for the efficient and accurate determination of the location of objects.

Keywords: river pollution, sewage, features-engineering, image recognition, machine learning, Keras, TensorFlow.

Вступ

На сьогоднішній день річки України страждають від надмірного забруднення. Джерелами такого забруднення є зареєстровані підприємства, що мають дозвіл на скидання вод й очисні споруди, поверхневий стік після опадів із сільськогосподарських угідь й міських територій, а також незареєстровані просторово-зосереджені забруднені води (ПЗЗВ) [1]. Знаходження місць потрапляння незареєстрованих ПЗЗВ у річку є необхідною задачею.

Якщо річка є джерелом питного централізованого водопостачання, то місця надходження до них ПЗЗВ мають бути точно встановленими. В Україні ведеться робота із актуалізації реєстру водотоків, але для встановлення точного місцезнаходження ПЗЗВ необхідно на основі супутникових знімків та векторних даних геопорталів знайти ділянки річки, де може мати місце факт скидання [1-3]. Це можна реалізувати використовуючи технології машинного навчання, такий як розпізнавання образів, класифікація об'єктів за набором зображень та ознак [2-6].

Результати дослідження

Поставлена задача може бути вирішена за умови використання відкритих векторних даних геопорталу OpenStreetMap (OSM) та супутникових даних Bing. Ці дані – загальнодоступні, дуже детальні й можуть бути використанні для системного аналізу за допомогою ArcGIS чи QGIS, а також бібліотек мов R, Python [2].

Пропонується CIR-метод для виділення ознак об'єктів із врахуванням причинно-наслідкових зв'язків та їх включення для формування вибірок даних [7].

Під час вирішення задачі пошуку незареєстрованих місць надходження забруднених вод у річку за векторними даними OpenStreetMap та супутниковими знімками Bing за CIR-методом встановлено вагомні ознаки, що наведено у табл. 1. Далі із застосуванням алгоритмів машинного навчання проводимо розпізнавання зображень, системний аналіз даних та класифікацію даних за

виділеними ознаками [7].

Таблиця 1 – CIR-метод виділення ознак

| Conditions / Умови (причини та умови, що сприяють появі об'єкта на зображеннях) | Impact / Впливи (наявність факторів підсилення впливу, які вирізняють цей об'єкт серед інших за певними критеріями) | Reactions / Наслідки (наслідки наявності об'єкта, які видно на зображеннях) |
|--|--|---|
| пониження рельєфу, що сприяє надходження ПЗЗВ у річку. | наявність споруд (підприємства, будинки приватного сектору) неподалік від берегової лінії. | - порізана берегова лінія; - густа берегова рослинність; - острівець водоростей та мілина у річці напроти місця витоку. |

Узагальнений алгоритм інтелектуальної інформаційної технології пошуку незареєстрованих місць надходження забруднених вод на ділянці річки [7] наведено у табл. 2.

Таблиця 2 – Алгоритм інтелектуальної інформаційної технології

| Етап | Алгоритм |
|--|--|
| 1. Збирання даних | - знайти векторні шари на доступних ГІС регіону, наприклад з OpenStreetMap; - використати відомі підходи до розпізнавання: об'єктний аналіз зображення; згорткові нейронні мережі та автоенкодері; метод нечітких k-середніх. |
| 2. Features-інжиніринг (виділення ознак) для формування моделі | - визначити ключові ознаки за допомогою CIR-методу на основі вибірки зареєстрованих місць надходження забруднених вод для автоматизованого виявлення природних (приток) та антропогенних (скидів стічних вод) місць впадіння в інших місцях річки; - провести автоматизований аналіз визначених на етапі 1 просторових даних та ідентифікувати виділені ознаки з використанням засобів опрацювання просторових даних ArcGIS, R, Python. |
| 3. Побудова моделі | - побудувати модель одним із методів машинного навчання, наприклад дерева рішень (бібліотеки xgboost, lightgbm) або нейронні мережі (Keras і Tensorflow), яка за виділеними ознаками навчиться виявляти місця, де може бути місце незареєстрованого скидання забруднення у річку. |
| 4. Ідентифікація ознак потенційних об'єктів | - ідентифікувати ознаки, відібрані на етапі 2, для незареєстрованих потенційних об'єктів із використанням засобів автоматизації, розроблених на етапах 1 і 2. |
| 5. Виявлення об'єктів | - здійснити класифікацію незареєстрованих місць забруднення поверхневих вод на заданій ділянці за даними етапу 4; - результат знайдених місць відобразити на карті, наприклад у QGIS чи ArcGIS; - здійснити перевірку знайдених місць із використанням БПЛА, відбору проб у воді, фото- і відеозйомки, у т.ч. після дощу. |

Висновки

Запропоновано інтелектуальну інформаційну технологію пошуку незареєстрованих місць надходження забруднених вод у річку з використанням методів машинного навчання. Запропоновано CIR-метод, що дозволяє виділити вагомі ознаки й точно визначити об'єкти незареєстрованого скидання забруднення. Використано об'єктний аналіз зображень та інші методи розпізнавання зображень і векторизації даних, просторовий аналіз для формування ознак і Machine learning для аналізу цих ознак, що підвищило ефективність встановлення місцезнаходження незареєстрованих джерел забруднення річки.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Наукові засади раціонального використання водних ресурсів України за басейновим принципом : монографія / За редакцією В. А. Сташук ; [В. А. Сташук, В. Б. Мокін, В. В. Гребінь, О. В. Чунарьов]. – Херсон : Грінь Д. С., 2014. – 320 с.
2. Системний аналіз та проектування ГІС : Електронний навчальний посібник / Є. М. Крижановський, В. Б. Мокін, А. Р. Яшолт, Л. М. Скорина. – Вінниця : ВНТУ, 2015. – 127 с. – Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/8960>.
3. Красовський Г. Я. Космічний моніторинг безпеки водних екосистем із застосуванням геоінформаційних технологій / Г. Я. Красовський ; Інститут проблем національної безпеки, Рада національної безпеки і оборони України. – К: Інтертехнологія, 2008. – 480 с.
4. Deep Learning [Electronic resource]. – Mode of access : <https://www.deeplearningbook.org/>.
5. Classification Methods for Remotely Sensed Data [Electronic resource]. – Mode of access : <https://www.taylorfrancis.com/books/9781420090741>.
6. Stroke Width Transform for Linear Structure Detection: Application to River and Road Extraction from High-Resolution Satellite Images [Electronic resource]. – Mode of access : https://link.springer.com/chapter/10.1007/978-3-319-59876-5_67.
7. Інформаційна технологія пошуку незареєстрованих місць надходження забруднених вод у річку за супутниковими і векторними даними геопорталів / В. Б. Мокін, Л. М. Скорина, Є. М. Крижановський, М. А. Гораш // Наукові праці Донецького національного технічного університету. Серія: «Інформатика, кібернетика та обчислювальна техніка». – 2018. – № 2 (27). – С. 30–35.

Мокін Віталій Борисович – д-р техн. наук, професор, завідувач кафедри системного аналізу, комп'ютерного моніторингу та інженерної графіки, Вінницький національний технічний університет, м. Вінниця, e-mail: vbmokin@gmail.com.

Скорина Любов Миколаївна – старший викладач кафедри системного аналізу, комп'ютерного моніторингу та інженерної графіки, Вінницький національний технічний університет, м. Вінниця.

Крижановський Євгеній Миколайович — канд. техн. наук, доцент кафедри системного аналізу, комп'ютерного моніторингу та інженерної графіки, Вінницький національний технічний університет, м. Вінниця.

Гораш Микола Анатолійович – аспірант кафедри системного аналізу, комп'ютерного моніторингу та інженерної графіки, Вінницький національний технічний університет, м. Вінниця.

Mokin Vitalii B. – Dr. Sc. (Eng.), Professor, Head of the Department of System Analysis, Computer Monitoring and Engineering Graphics, Vinnytsia National Technical University, Vinnytsia, e-mail: vbmokin@gmail.com.

Skoryna Liubov M. – Senior Lecturer at the Department of System Analysis, Computer Monitoring and Engineering Graphics, Vinnytsia National Technical University, Vinnytsia.

Kryzhanovskiy Yevhenii M. – Cand. Sc. (Eng), Associate Professor at the Department of System Analysis, Computer Monitoring and Engineering Graphics, Vinnytsia National Technical University, Vinnytsia.

Horash Mykola A. – Post-Graduate Student at the Department of System Analysis, Computer Monitoring and Engineering Graphics, Vinnytsia National Technical University, Vinnytsia.