А. В. Ужинский

Искусственный интеллект в науках о жизни

На протяжении нескольких лет в Лаборатории информационных технологий им. М. Г. Мещерякова реализуются проекты, связанные с применением искусственного интеллекта в науках о жизни.

С 2016 г. в ЛИТ развивается система управления данными проекта UNECE ICP Vegetation [1], участники которого собирают образцы мха и используют различные аналитические методы, в том числе и нейтронный активационный анализ на реакторе ИБР-2 ЛНФ ОИЯИ, чтобы получить данные по содержанию в воздухе тяжелых металлов, азота, стойких органических соединений и радионуклидов. Вобравшая в себя новые технологии и накопившая огромный массив данных, сегодня система переросла в интеллектуальную платформу. Например, мобильное приложение позволяет упростить процесс сбора и проверки данных, а для классификации изображений и прогнозирования загрязнений по данным дистанционного зондирования Земли активно используются модели глубокого обучения. Платформа построена на микросервисной архитектуре, что позволяет оперативно расширять ее функциональные возможности, автоматизировать операционные задачи и развивать блок аналитики.

В 2017 г. группа научных сотрудников ЛИТ выиграла грант РФФИ на разработку комплексной системы диагностирования болезней растений по изображениям и текстовому описанию.

Дроны с гиперспектральными камерами над полями, роботы с обычными камерами в теплицах, камеры высокого разрешения, установленные в вертикальных фермах, приложения в смартфоне, позволяющие по фотографии определить болезнь растений, — все это сегодня стало реальностью [2, 3]. Сейчас подобные технологии меняют сельское хозяйство так же, как в XX в. это делали тракторы и комбайны. Для того, чтобы прокормить растущее население Земли, необходимо к 2050 г. на 70 % увеличить производство продуктов питания. Использование различных устройств вместе с технологиями искусственного интеллекта позволяет увеличить урожайность и минимизировать потери

A. V. Uzhinskiy

Artificial Intelligence in Life Sciences

For several years, projects related to the application of artificial intelligence in life sciences have been implemented at the Meshcheryakov Laboratory of Information Technologies.

Since 2016, the data management system of the UNECE ICP Vegetation project [1] has been developed at MLIT. Its participants collect moss samples and employ various analytical methods, including neutron activation analysis at the IBR-2 reactor of FLNP JINR, to obtain data on the content of heavy metals, nitrogen, persistent organic pollutants and radionuclides in the air. Having incorporated novel technologies and accumulated a huge amount of data, the system has evolved into an intelligent platform. For example, the mobile application simplifies the data acquisition and verification process, while deep learning models are actively used to classify images and predict pollution from Earth remote sensing data. The platform is built on a microservice architecture, which enables to quickly enhance its functionality, to automate operational tasks and develop the analytics unit.

In 2017, a group of MLIT researchers won a grant of RFBR to elaborate a comprehensive system for diagnosing plant diseases using images and text descriptions.

Drones with hyperspectral cameras over fields, robots with conventional cameras in greenhouses, high-resolution cameras installed in vertical farms, smartphone applications that allow detecting plant diseases from photos — all this has become a reality today [2, 3]. Such technologies are transforming agriculture in the same way as tractors and combines did in the 20th century. To feed the growing population of the Earth, it is needed to increase food production by 70% by 2050. The use of different devices together with artificial intelligence technologies enables to boost yields and minimize losses in agriculture. The

в сельском хозяйстве. Наличие опытных специалистов в области нейронных сетей и машинного обучения, а также современные вычислительные ресурсы делают ЛИТ привлекательной для реализации совместных проектов в сельском хозяйстве.

Были разработаны платформа (pdd.jinr.ru) и мобильное приложение DoctorP (рис. 1), направленные на решение задач диагностирования и определения болезней домашних и сельскохозяйственных растений путем использования современных методов обработки графической и текстовой информации на высокопроизводительных вычислительных инфраструктурах [4]. В настоящее время в платформе есть модели для 30 сельскохозяйственных и декоративных культур: барбариса, винограда, вишни, голубики, клубники, кукурузы, огурцов, перца, пшеницы, смородины, томатов, хлопка, яблок, орхидей, роз и т. д. Общая модель для всех видов растений распознает 68 классов болезней и вредителей. В базе собрано свыше 6000 изображений. С начала 2023 г. платформой было обработано более 80 000 запросов пользователей. Для определения оптимальных подходов к обучению моделей команда разработчиков постоянно тестирует различные современные архитектуры нейронных сетей, политики аугментации данных и функции минимизации потерь. Все это позволяет добиться статистической точности моделей в 97 % [5]. Использовать интерфейс платформы могут все, начиная от специалистов агрохолдингов и заканчивая начинающими садоводами, для которых немаловажной особенностью программы являются рекомендации по лечению растений, верифицированные профессиональными агрономами.

Нейросетевые модели pdd.jinr.ru используются в совместном проекте с НЦМУ «Агротехнологии будущего» на базе Тимирязевской академии для отслеживания влияния освещения на развитие растений. Исследования в данном направлении на основе оценок параметров развития, получаемых от модели, позволят подобрать оптимальные схемы по освещению, поливу, питанию и прочим аспектам выращивания сельскохозяйственных культур.

В другом проекте был реализован комбинированный алгоритм контроля салатов, выращиваемых на конвейерных лентах. На первом этапе для обнаружения объектов использовалась нейросетевая архитектура на базе YOLO, после чего производилась классификация салатов с использованием специализированных моделей. Такой подход позволил значительно сократить издержки на разметку данных, при этом точность классификации достигла 99 %.

В совместном проекте с компанией «Дока — Генные Технологии» разрабатываются механизмы

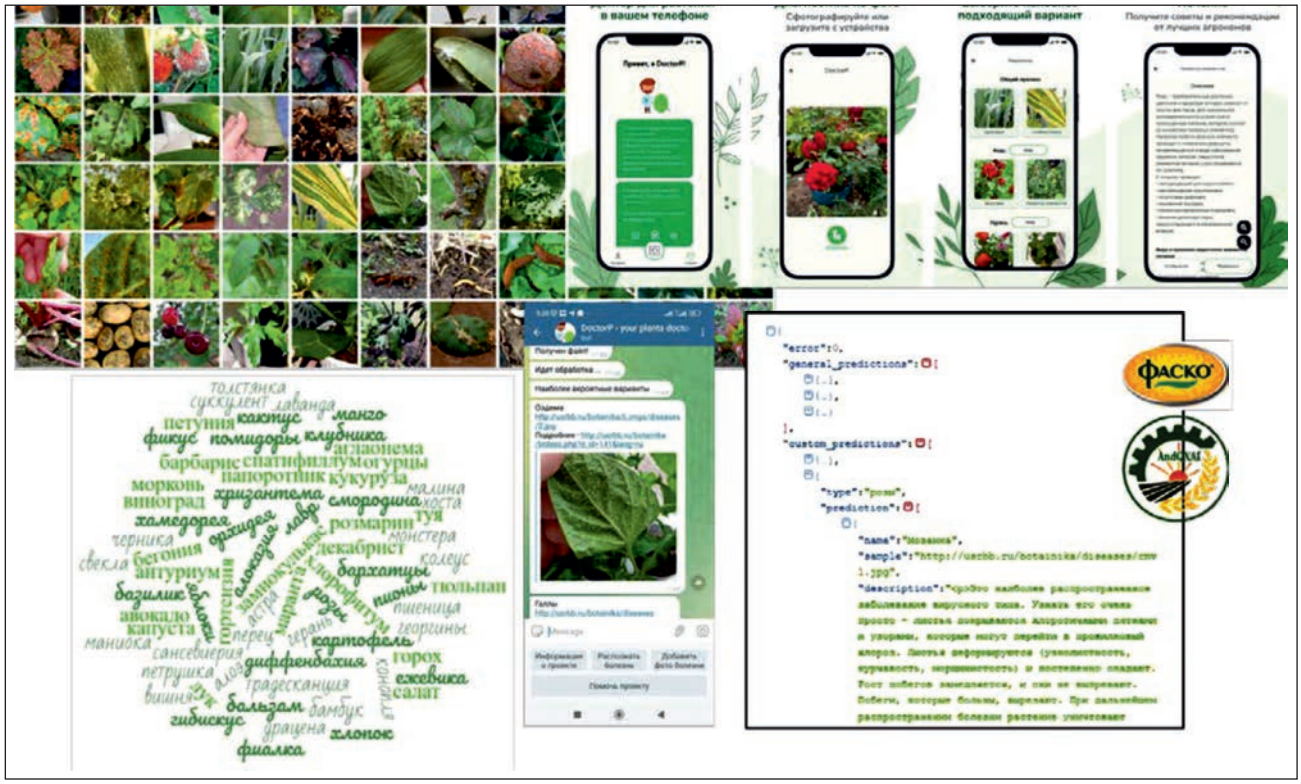

Рис. 1. Примеры интерфейсов платформы DoctorP

Fig. 1. Examples of DoctorP platform interfaces

Рис. 2. *a*) Изображения, полученные с камеры, установленной на сельхозтехнику, и обработанные моделью для определения болезней картофеля. *b*) Визуализация подобранного индекса, показывающая разницу между больными Y-вирусом (верхние изображения) и здоровыми (нижние изображения) растениями

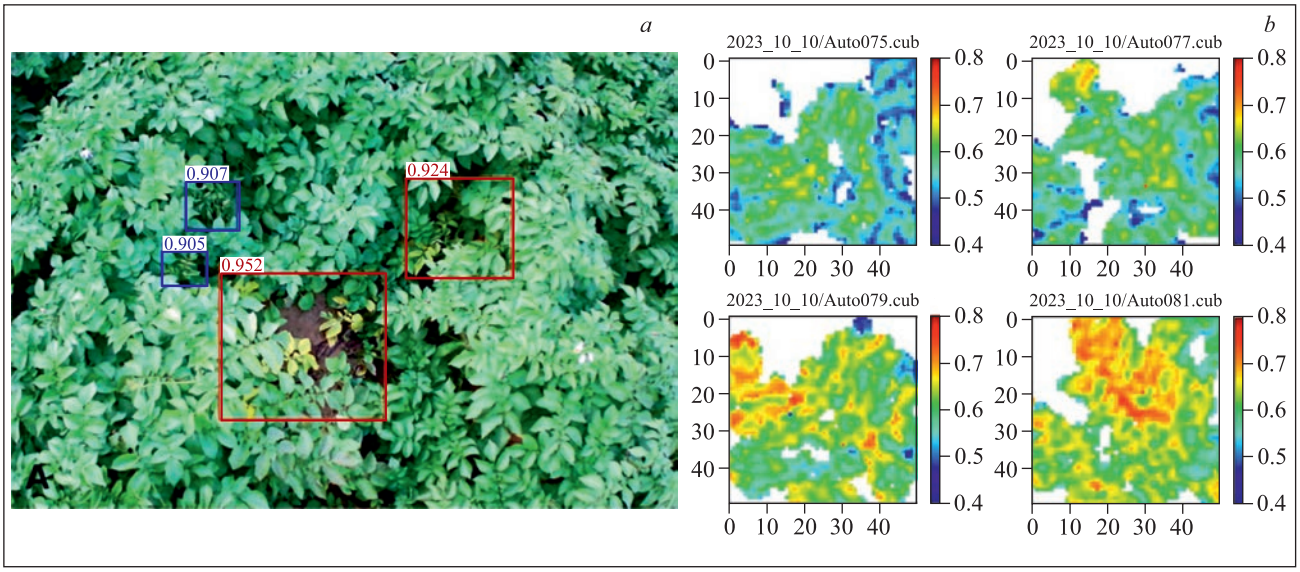

Fig. 2. *a*) Images obtained from a camera installed on agricultural equipment and processed by the model to define potato diseases. *b*) Visualization of the fitted index showing the difference between Y-virus diseased (top images) and healthy (bottom images) plants

presence of experienced specialists in the field of neural networks and machine learning, as well as advanced computing resources, makes MLIT attractive for the implementation of joint projects in agriculture.

A platform (pdd.jinr.ru) and a mobile application DoctorP (Fig. 1) were developed, which are aimed at solving the tasks of diagnosing and detecting the diseases of domesticated and agricultural plants using state-of-the-art methods of graphic and text information processing on high-performance computing infrastructures [4]. At present, the platform has models for 30 agricultural and ornamental crops: barberry, grapes, cherry, blueberry, strawberry, corn, cucumber, pepper, wheat, currant, tomato, cotton, apple, orchid, rose, etc. The general model for all plant species detects 68 classes of diseases and pests. The database contains over 6 000 images. Since the beginning of 2023, the platform has processed over 80 000 user requests. To define optimal approaches to training models, the development team steadily tests various advanced neural network architectures, data augmentation policies and loss minimization functions. All this enables to achieve a statistical accuracy of the models of 97% [5]. Everyone can use the platform interface, from specialists of agricultural holdings to novice gardeners, for whom an essential feature of the program is recommendations for plant treatment that are verified by professional agronomists.

The neural network models of pdd.jinr.ru are employed in a joint project with the World Class Research Centre “AgriTechnologies for the Future” on the basis of Timiryazev Academy to monitor the effect of lighting on plant development. Research in this area will make it possible to select optimal schemes for lighting, watering, nutrition and other aspects of growing crops relying on the estimates of development parameters obtained from the model.

A combined algorithm for monitoring lettuces grown on conveyor belts was implemented in another project. At the first stage, a neural network architecture on top of YOLO was utilized to detect objects, and afterwards lettuces were classified using specialized models. This approach enabled to significantly reduce the cost of data labeling, while the classification accuracy reached 99%.

Mechanisms for detecting potato diseases are elaborated in a joint project with Doka-Gene Technologies. Since one of the company’s major goals is to grow high-quality seed, disease detection is crucial, especially in the early stages. Studies are conducted in several directions at once (Fig. 2). To detect problems in fields, software and hardware complexes are employed, which are installed on agricultural equipment and shoot with high-resolution cameras operating in the visible spectrum. To search for patterns that allow detecting diseased plants even before visible symptoms emerge, studies using hyperspectral images are carried out. Different versions of neural network models for image classification involving one-shot learning ap-

определения различных заболеваний картофеля. Поскольку одна из основных задач компании — выращивание высококачественного семенного материала, определение болезней, особенно на ранних стадиях, имеет критически важное значение. Исследования ведутся сразу в нескольких направлениях (рис. 2). Для определения проблем на полях используются программно-аппаратные комплексы, устанавливаемые на санитарные машины и производящие съемку на камеры высокого разрешения, работающие в видимом спектре. Для поиска закономерностей, позволяющих определять больные растения еще до появления видимых симптомов, проводятся исследования с использованием гиперспектральных изображений. Тестируются как различные варианты нейросетевых моделей для классификации изображений с использованием one-shot learning подходов, так и статистические алгоритмы для работы на уровне пикселей.

Еще одно направление исследований — автоматизация анализа состояния растений в тепличных комплексах. Основной задачей проекта является упрощение рутинных операций агрономов и предоставление им удобных инструментов учета и мониторинга. В ходе реализации разрабатывается автономная роботизированная платформа, способная передвигаться по различным типам поверхностей, фиксировать интересующие показатели и производить съемку на высоте более 4 м. В теплицах основной проблемой являются отслеживание, локализация и идентификация вредителей на ранних стадиях их проявления, что требует разработки нейросетевых моделей и алгоритмов для классификации, сегментации, обнаружения и отслеживания объектов.

Платформа продолжает развиваться. База данных пополняется снимками пользователей и тем самым улучшает точность моделей. В перспективе в платформу будут добавлены возможности обработки видеопотока, модели для определения нехватки основных элементов (азот, фосфор, кальций, железо и т. д.), а также средства формирования рекомендаций по выращиванию и отслеживанию развития наиболее востребованных сельскохозяйственных культур.

---

proaches, as well as statistical algorithms for working at the pixel level, are tested.

Another area of research is the automation of analysis of the condition of plants in greenhouse complexes. The main objective of the project is to simplify the routine operations of agronomists and provide them with convenient accounting and monitoring tools. During the implementation, an autonomous robotic platform capable of moving on various types of surfaces, recording indicators of interest and shooting at a height of more than four meters is developed. The main challenge in greenhouses is tracking, localizing and detecting pests in the early stages of their manifestation, which entails the elaboration of neural network models and algorithms for object classification, segmentation, detection, and tracking.

The platform continues to develop. The database is supplemented with user images and thereby enhances the accuracy of the models. In the future, the ability to process video streams, models for defining the shortage of basic elements (nitrogen, phosphorus, calcium, iron, etc.), as well as tools for generating recommendations for growing and monitoring the development of the most popular agricultural crops, will be added to the platform.

Список литературы / References

1. *Ужинский А. В.* Современные информационные технологии в экологическом мониторинге // Новости ОИЯИ. 2021. №4. С. 32–35;

Uzhinskiy A. Modern Information Technologies in Environmental Monitoring // JINR News. 2021. No. 4. P. 32–35.

2. *Uzhinskiy A.* Advanced Technologies and Artificial Intelligence in Agriculture // Appl. Math. 2023. V. 3. P. 799–813; https://doi.org/10.3390/appliedmath3040043.

3. *Ужинский А. В.* Искусственный интеллект в сельском хозяйстве // Открытые системы. 2023. №3. С. 20–23;

Uzhinskiy A. Artificial Intelligence in Agriculture // Open Systems. 2023. No. 3. P. 20–23 (in Russian).

4. *Ужинский А. В.* Искусственный интеллект против болезней растений // Открытые системы. 2022. №3. С. 29–31;

Uzhinskiy A. Artificial Intelligence against Plant Diseases // Open Systems. 2022. No. 3. P. 29–31 (in Russian).

5. *Uzhinskiy A., Ososkov G., Goncharov P., Nechaevskiy A., Smetanin A.* One-Shot Learning with Triplet Loss for Vegetation Classification Tasks // Comput. Opt. 2021. V. 45. P. 608–614; https://doi.org/10.18287/2412-6179-CO-856.