https://doi.org/10.26898/0370-8799-2025-3-1
УДК 633.1:004.048

Тип статьи: оригинальная
Type of article: original

Классификация и картографирование сельскохозяйственных культур с использованием дистанционного зондирования и машинного обучения

(✉)**Каличкин В.К., Крылова О.С., Гарафутдинова Л.В.**
Сибирский федеральный научный центр агробиотехнологий Российской академии наук
Новосибирская область, р.п. Краснообск, Россия
(✉)e-mail: vk.kalichkin@gmail.com

Представлены результаты исследований по классификации и картографированию сельскохозяйственных культур с помощью машинного обучения (МО) с использованием данных дистанционного зондирования. Исследования проведены в 2022, 2023 гг. на территории землепользования хозяйств, расположенных в Новосибирской области. Использованы изображения, полученные с Sentinel-2 и Landsat 8-9. Признаками для обучения моделей МО выступали шесть спектральных полос и пять вегетационных индексов за разные даты вегетационных периодов. Применяли алгоритмы МО с контролируемым обучением: XGBoost, KNN, RF и SVM, а также нейронную сеть deep-FNN. Модели XGBoost, KNN и RF показали высокую точность классификации – 93–97% при разрешении 30 м/пиксель и 80–90% при разрешении 90 м/пиксель. Модель deep-FNN показала наименьшие результаты с точностью от 78 до 92% при разрешении 30 м/пиксель. Общее снижение точности на 8–12% при разрешении 90 м/пиксель в сравнении с разрешением 30 м/пиксель подчеркивает важность масштаба для эффективного распознавания культур. Также обучение моделей на объединенных данных спутников Sentinel-2 и Landsat 8-9 при разрешении 30 м/пиксель дало более высокие значения метрики F1-score, чем на данных отдельно по каждому из этих спутников. Различные метрики оценки (F1-score и ROC-AUC-score) подтвердили, что модель XGBoost была наиболее производительной и точной. Лучшая общая классификация достигнута для кукурузы, ячменя, однолетних и многолетних трав, а также залежи, с некоторым снижением точности для овса, гороха, вики и пшеницы мягкой озимой. Наименьшая точность отмечена при классификации картофеля, ярового рапса, пшеницы мягкой яровой и пара. Результаты исследований подчеркивают значимость выбора модели МО и масштаба разрешения спутниковых снимков для успешной классификации сельскохозяйственных культур.

Ключевые слова: классификация сельскохозяйственных культур, дистанционное зондирование, машинное обучение, картографирование

Crop classification and mapping using remote sensing and machine learning

(✉)**Kalichkin V.K., Krylova O.S., Garafutdinova L.V.**
Siberian Federal Scientific Centre of AgroBioTechnologies of the Russian Academy of Sciences
Krasnoobsk, Novosibirsk Region, Russia
(✉)e-mail: vk.kalichkin@gmail.com

Research results on crop classification and mapping using machine learning (ML) with remotely sensed data are presented. Studies were conducted in 2022, 2023 on the land use of the farms located in the Novosibirsk region. Sentinel-2 and Landsat 8-9 images were used. The features for training the ML models were six spectral bands and five vegetation indices for different dates of the growing seasons. XGBoost, KNN, RF and SVM supervised learning ML algorithms and deep-FNN neural network were applied. The XGBoost, KNN and RF models showed high classification accuracies of 93–97% at 30 m/pixel resolution and 80-90% at 90 m/pixel resolution. The deep-FNN model showed the lowest results with an accuracy of 78 to 92% at 30 m/pixel resolution. The overall 8–12% reduction in accuracy at 90 m/pixel resolution compared to 30 m/pixel resolution emphasizes the importance of scale for effective crop recognition. Also, training the models on the combined Sentinel-2 and Landsat 8-9 satellite data at 30 m/pixel resolution yielded higher values of the F1-score metric than on the data separately for each of these satellites. Various evaluation metrics (F1-score and ROC-AUC-score) confirmed that the XGBoost model was the best performing and most accurate. The best overall classification was achieved for corn, barley, annual and perennial grasses, and fallow, with some decrease in accuracy for oats, peas, vetch and soft winter wheat. The lowest accuracy was observed in the classification of potatoes, spring rape, soft spring wheat and fallow. The results of the research emphasize the importance of ML model selection and satellite imagery resolution scale for successful crop classification.

Keywords: crop classification, remote sensing, machine learning, mapping

Для цитирования: *Каличкин В.К., Крылова О.С., Гарафутдинова Л.В.* Классификация и картографирование сельскохозяйственных культур с использованием дистанционного зондирования и машинного обучения // Сибирский вестник сельскохозяйственной науки. 2025. Т. 55. № 3. С. 5–18. https://doi.org/10.26898/0370-8799-2025-3-1

For citation: Kalichkin V.K., Krylova O.S., Garafutdinova L.V. Crop classification and mapping using remote sensing and machine learning. *Sibirskii vestnik sel'skhozyastvennoi nayki = Siberian Herald of Agricultural Science*, 2025, vol. 55, no. 3, pp. 5–18. https://doi.org/10.26898/0370-8799-2025-3-1

Конфликт интересов
Авторы заявляют об отсутствии конфликта интересов.
Conflict of interest
The authors declare no conflict of interest.

INTRODUCTION

Spatial data in the form of large-scale crop maps at the level of a specific land use, municipal district or region are necessary for the development and management of government subsidy allocation and crop market planning. In addition, the growing demand for crop insurance has increased the practical requirements for crop mapping. Crop identification using traditional methods such as field surveys, although accurate, has limitations. Mapping a single crop plot takes considerable time and access to these plots can be difficult. Field surveys therefore remain an expensive approach to crop mapping, and regulatory agencies may not have the manpower to carry out such work on the ground. To meet this need, there is a need for online crop statistics at the plot level, including for tracking declared and actual crops sown. This can be achieved using remote sensing technologies and artificial intelligence methods. [1–3].

Since the early 1970s, classification of vegetation using remote sensing data has been used as an alternative to traditional field studies. However, it has not yet been possible to completely eliminate field surveys. They are necessary for quantitative assessment of the structure of agricultural crops, preparation of data and classification algorithms, and validation of models [4, 5]. Over the past two decades, the importance of research in crop mapping using remote sensing has continued to grow [6, 7]. The widespread popularity of satellite imagery over other platforms is due to its resolution and open access. Research has shown that multispectral remote sensing images obtained with appropriate sensors are suitable for crop mapping [8–11]. The second most widely used crop mapping technology is synthetic aperture radar (SAR) [12, 13]. The use of hyperspectral images remains uncommon, as does the use of terrestrial lidar scanner. For crop classification, a single sensor is usually used, as pre-processing the images from multiple sensors can be challenging. In global practice, images from the Sentinel and Landsat series of satellites are currently the most commonly used in this subject area [3, 14].

Machine learning (ML) and deep learning (DL) models have shown the ability to classify crops from medium and high-resolution satellite images with high accuracy[1] [15]. The algorithms and models used for crop mapping are varied and grouped into three categories:

– parametric ML classifiers (e.g. logistic regressions, hidden Markov models);

– nonparametric ML classifiers (e.g. support vector machine, decision tree, random forest, k-nearest neighbors, etc.);

– DL classifiers (neural networks of various modifications).

Parametric machine learning classifiers work by making certain assumptions about the statistical distribution of the data.

Nonparametric ML classifiers make no assumptions about the distribution of the data. The mathematical basis of DL allows neural networks to learn to recognize patterns in the data that is fed to them during the training process [16–18].

A review of the effectiveness of crop classification models using ML and DL based on remote sensing data conducted in the paper [19] showed that the most successful were random forest (RF), support vector machine (SVM), extreme gradient boosting (XGBoost) and neural networks (CNN, LSTM). These models achieved the highest overall accuracy in numerous studies (386 publications). However, the range of variation in the overall accuracy indicator was quite large. This indicates that the choice of the classifier and specific features significantly affects the accuracy of the crop mapping process [20, 21]. A possible solution to the problem of accuracy in crop classification may be to combine images from different satellites, which helps to reduce the time gap in the data. This technology is called multimodal remote sensing, and the classification efficiency of combined images from data [22] is superior to that of individual images.

The purpose of the study was to assess the accuracy of classification of agricultural crops by various models of ML and DL using remote sensing data and to carry out their mapping on the territory of the studied land use objects.

MATERIAL AND METHODS

Research on the classification and mapping of agricultural crops using remote sensing data was carried out in 2022 and 2023 on the land use territory of the EEP “Praktik” of the Novosibirsk State Agrarian University and the ES “Elitnaya” of the Siberian Federal Scientific Centre of Agro-BioTechnologies of the Russian Academy of Sciences, located in the Novosibirsk district of the Novosibirsk region (54°54'20.15"N, 82°52'20.06"E and 54°54'57"N, 82°57'6"E), with a total area of 950 ha. Classification and mapping were carried out for crops cultivated in farms: potatoes, soft spring wheat, soft winter wheat, corn, annual grasses, spring rape, barley, vetch, oats, peas, perennial grasses. The structure of land use of farms also includes complete fallow and unused lands (layland).

The work used images obtained from Sentinel-2 (Operational Land Imager (OLI)) and Landsat 8-9 (Thermal Infrared (TIRS) Collection 2), available for free download from Sentinelhub (https://apps.sentinel-hub. com) and the US Geological Survey (https:// earthexplorer.usgs.gov), respectively. Images with cloudiness below 75%, obtained from June 1 to August 31 were used. All satellite data were matched in accordance with atmospheric and radiometric correction. Scanned pictures of the images were converted to vector form and divided into objects in the form of a regular grid of 30 × 30 m (7657 data on the EEP “Praktik” object and 23,191 on the ES “Elitnaya” object) and 90 × 90 m (887 data on the EEP “Praktik”

[1]*Wu F., Maleki R., Oubara A., Gómez D., Eftekhari A., Yang G.* Machine Learning Approaches for Crop Identification from Remote Sensing Imagery: A Review // International Conference on Soft Computing and Pattern Recognition. Cham: Springer Nature Switzerland, 2023, pp. 325–336. DOI: 10.1007/978-3-031-27524-1_31.

object and 3081 on the ES “Elitnaya” object). Data containing information on spectral characteristics and vegetation indices were calculated using the “zonal statistics” function in Quantum GIS (QGIS).

Spectral bands from both sensor systems were used: blue, green, red, near infrared (NIR), shortwave infrared SWIR1 and SWIR2, as well as five vegetation indices: NDVI (Normalized Difference Vegetation Index), LAI (Leaf Area Index), NDMI (Normalized Difference Moisture Index), NDWI (Normalized Difference Water Index), GNDVI (Green Normalized Difference Vegetation Index).

A preliminary field survey was conducted to collect reliable data on crops. The boundaries of the plots were recorded digitally using SW maps in shp-file format and integrated into QGIS. Satellite images were also loaded into QGIS, their characteristics were developed, unnecessary stripes were eliminated, and vegetation indices were calculated. The values of the spectral characteristics of pixels from multispectral satellite images and vegetation indices were referenced in a single coordinate system 2505 Pulkovo 1942/Gauss-Kruger CM 87E to construct a data set. The first set consisted of the data from the Sentinel-2 satellite, the second – from Landsat 8-9 data, the third included a combination of the data from these satellites.

The input data for ML and DL were georeferenced contours of agricultural plots with crops placed on them, formed as classes of spatial objects using the "regular grid" function in QGIS. The classes were characterized by attribute characteristics of spectral bands and vegetation indices (11 features in total). Preliminary processing included data normalization, handling of missing values, and dividing the data into training and test sets in a ratio of 80 and 20%, respectively. The test sample was formed using the method of complete randomization. The correlation analysis was also conducted to identify a high correlation between the features and the recognizable classes. No high correlation was found, so all 11 features were included in the ML and DL models. In 2022, 12 crops (classes) were cultivated on the territory of the studied farms, of which two classes were present in both land use objects. In 2023, 11 crops (classes) were cultivated, of which two classes were also present in both objects. The distribution of the data by classes was unbalanced.

To build crop classification models, ML algorithms with supervised learning were used: Support Vector Machine – SVM, K-Nearest Neighbor – KNN, Random Forest – RF and XGBoost. ML algorithms were trained using a dataset containing pre-classified crop samples. Besides, to classify the data the deep feed-forward neural network – deep-FNN – with three open layers was used. This algorithm is designed to autonomously learn and extract hierarchical representations of data using interconnected layers of nodes. The deep-FNN neural network was trained over 20 epochs. Standard Python libraries for ML were used to build and train the models (numpy, pandas, skikit-learn, matplotlib, etc.).

The selection of metrics for assessing classification accuracy was carried out taking into account a specific model. For the task of identification by SVM, KNN, RF and XGBoost algorithms with an unbalanced distribution of data by classes, the weighted avg F1-score was used, since this metric is a harmonic mean between precision and recall and allows taking into account not only the proportion of the correctly classified positive examples, but also the ability of the model to detect all positive examples, as well as the ROC-AUC-score metric. Categorical accuracy was used as an error metric to evaluate the performance of deep-FNN. To evaluate each class separately, the classification_report method from the scikit-learn library was used, which provides information for each class separately, as well as the average weighted F1-score values[2].

The models were validated by field recording of the placement of agricultural crops through-

[2]Dealing with unbalanced data. https://habr.com/ru/companies/otus/articles/769242/

out the entire territory of land use objects and comparing the results obtained with the classification results. Based on the results of the XGBoost model for the classification of agricultural crops and its validation, crop mapping was carried out in the QGIS software package.

RESULTS AND DISCUSSION

Crop mapping based on remote sensing remains challenging due to the high local variability and temporal dynamics of agricultural crops. These properties result in high intra-class variance of the spectral-temporal signal, which increases with the size and diversity of the study area, the number of crops to be recognized, and the variability of meteorological conditions. To address the problems associated with spectral-temporal variations, data with sufficient spatial resolution and high temporal density of observations are usually required, which can be combined in the used ML approaches with verified field data.

The peculiarity of our research was that optical remote sensing data with a resolution of 10 and 20 m per pixel of the Sentinel-2 satellite and 30 m per pixel of Landsat 8-9, expressed in a raster model, were transformed into a vector model using QGIS functions.

Using a vector model of data representation and grid construction, the spectral characteristics of pixels and the values of vegetation indices are combined in a single geodatabase of features in the form of tables, thereby creating the ability to work with ML algorithms and optimizing computing resources. To reduce classification errors due to the use of the images taken on the same day, the maximum possible number of images available for download during crop vegetation was used. In total, there were 26 such images for both satellites, years and land use objects.

At the first stage of the research, the accuracy of the ML algorithms was tested with all data sets in terms of years and land use objects (see Table 1). Almost all ML models showed fairly high performance and accuracy of crop classification when using the weighted average F1-score as a unit of measurement, but some features should be noted. Thus, the XGBoost model demonstrated the best results on all data sets in both years of research, while the classification accuracy was 94–97% at a scale of 30 m/pixel and 80–90% at a scale of 90 m/pixel.

The KNN model demonstrated approximately the same classification accuracy as the XGBoost model. A slightly lower classification accuracy was observed when implementing the RF model – 93–96% at a scale of 30 m/pixel and 80–88% at a scale of 90 m/pixel. The lowest classification accuracy rates were obtained when implementing the deep-FNN model, which varied within 78–92% at a scale of 30 m/pixel and within 56–79% at a scale of 90 m/pixel. The obtained results can be explained by the fact that ML algorithms can process tabular data types, while deep neural networks (DNNs) have difficulties processing such data. According to the authors of articles on machine learning on specialized Internet resources Habr (https://habr.com/ru/articles) and Kaggle (https://www.kaggle.com/), DNNs showed good results when working with images and text, but when processing tabular data, DNNs were inferior in classification accuracy to ML algorithms. The use of DL for analyzing tabular data remains understudied, and variants of ensemble ML models still dominate in most applications[3–5]. In this regard, classical ensemble approaches (XGBoost and RF) demonstrated higher classification accuracy than DNNs.

[3]Jaykumaran. Training 3D U-Net for Brain Tumor Segmentation (BraTS2023-GLI) Challenge. AVAILABLE AT: https://learnopencv.com/3d-u-net-brats. (accessed 24.10.2024).

[4]Learning from tabular data. AVAILABLE AT: https://habr.com/ru/articles/534186/ (accessed 29.10.2024).

[5]Data Science Trends on Kaggle! AVAILABLE AT: https://www.kaggle.com/code/shivamb/data-science-trends-on-kaggle (accessed 29.10.2024).

The F1-score metric values when training models on various Sentinel-2 and Landsat 8-9 satellite data at different resolutions, years of research and land use objects have unstable characteristics. This can be explained by the uneven number of observation days and the number of predictors formed depending on this number. In addition, the accuracy of the classification depended not only on the number of predictors, but also on the distribution of survey days over

Табл. 1. Результаты оценки точности моделей машинного обучения по метрике F1-score
Table 1. Results of assessing the accuracy of machine learning models using the F1-score metric

Model	EEP "Praktik"			ES "Elitnaya"		
	Sentinel-2	Landsat 8-9	Landsat + Sentinel	Sentinel-2	Landsat 8-9	Landsat + Sentinel
2022						
Scale 30 m/pixel						
KNN	0,95	0,92	0,95	0,96	0,96	0,96
SVM	0,92	0,87	0,91	0,94	0,93	0,95
RF	0,95	0,93	0,95	0,95	0,96	0,96
XGBoost	0,95	0,94	0,95	0,96	0,97	0,97
deep-FNN*	0,92	0,85	0,92	0,88	0,78	0,88
Scale 90 m/pixel						
KNN	0,81	0,82	0,77	0,86	0,83	0,85
SVM	0,76	0,75	0,71	0,82	0,77	0,82
RF	0,80	0,80	0,80	0,83	0,84	0,84
XGBoost	0,81	0,80	0,81	0,84	0,86	0,86
deep-FNN*	0,69	0,72	0,71	0,62	0,56	0,64
2023						
Scale 30 m/pixel						
KNN	0,97	0,96	0,97	0,96	0,94	0,96
SVM	0,96	0,91	0,96	0,94	0,86	0,94
RF	0,96	0,96	0,96	0,96	0,95	0,96
XGBoost	0,97	0,97	0,97	0,97	0,96	0,97
deep-FNN*	0,91	0,91	0,92	0,87	0,85	0,89
Scale 90 m/pixel						
KNN	0,89	0,85	0,89	0,85	0,80	0,86
SVM	0,81	0,74	0,81	0,85	0,72	0,82
RF	0,88	0,85	0,88	0,87	0,83	0,86
XGBoost	0,90	0,87	0,90	0,89	0,85	0,88
deep-FNN*	0,69	0,74	0,78	0,74	0,76	0,79

*For deep-FNN, the categorical_accuracy metric was used.

the vegetation period, namely: the more evenly the survey days are distributed in time, the higher the accuracy of the classification.

However, it can be argued that higher classification accuracy for all ML models on both land use objects and in both years of research was obtained at a scale of 30 m/pixel compared to a scale of 90 m/pixel. Also, training of all models on combined data from Sentinel-2 and Landsat 8-9 satellites at a scale of 30 m/pixel showed more stable values of the F1-score metric. The results of the deep neural network with the FNN architecture (deep-FNN) were not included in the further analysis due to its low classification ability.

In ML, there are many metrics for assessing the accuracy and performance of models, each with its own advantages and disadvantages. However, a single universal metric that could optimally assess the quality of model performance in all cases has not yet been developed. In this regard, to assess the performance of models, it is not enough to use only one metric, but it is necessary to compare the values of two or more metrics that are most suitable for a given task.

The weighted average F1-score does not always allow for an unambiguous comparison of the performance of models due to its dependence on the threshold probability value. In this regard, in addition to the F1-score, this study uses the ROC-AUC-score metric, which allows for an adequate comparison of the performance of different models with an unbalanced data set (the higher the value of this metric, the higher the classification ability of the models). Thus, the highest values of the ROC-AUC-score metric were shown by the XGBoost method for both years of research for both objects using the combined data of the Sentinel-2 and Landsat 8-9 satellites at a scale of 30 m/pixel. In 2022, for the ES “Elitnaya” object, this indicator was 0.978, for the EEP “Praktik” object – 0.923, in 2023 – 0.972 and 0.979, respectively. The KNN method also showed close values of the ROC-AUC-score metric. In 2022, for the EEP “Praktik” object, it was only 0.1% less than the XGBoost method, in 2023 – 0.7%. At a scale of 90 m/pixel, the maximum ROC-AUC-score value was also shown by the XGBoost method. Thus, in 2023, for the EEP “Praktik” object, this metric is equal to 0.877, for the ES “Elitnaya” object – 0.895, in 2022, for the EEP “Praktik” object, the ROC-AUC-score is equal to 0.714, for the ES “Elitnaya” object – 0.896.

To visualize the accuracy assessment of ML models using ROC-AUC-score, a graph showing the results of crop classification when analyzing combined data (Sentinel-2 and Landsat 8-9) at a resolution of 30 m/pixel and 90 m/pixel for 2023 is presented as an example (see Fig. 1).

The ROC curve displays the ratio between the proportion of classes from the total number of feature carriers correctly classified as carrying the feature and the proportion of classes from the total number of classes not carrying the feature, incorrectly classified as carrying the feature. Typically, the ROC curve has a parabolic shape with varying curvature, and the higher the curvature, the better the model works. Fig. 1 shows that the ROC curve for the analysis of images at a scale of 30 m/pixel for the XGBoost method has better characteristics than all other ML methods for both the EEP “Praktik” object and the ES “Elitnaya” object. The worst ROC curve indicators are noted for the SVM method, the KNN and RF methods occupy an intermediate position. The same patterns were observed for the ROC curve when analyzing images at a scale of 90 m/pixel, but with reduced curvature. Taking into account the analysis of the model accuracy assessment by two metrics (F1-score and ROC-AUC-score), it can be concluded that the most productive and accurate at different scale values for both years of research and for both land use objects was the extreme gradient boosting method (XGBoost), the second place was taken by the k-nearest neighbors method (KNN), the third place was occupied by the random forest (RF). The lowest accuracy rates were shown by the support vector machine (SVM), i.e. it can be concluded that this algorithm is poorly suited for solving problems of this type. At the same time, the combined data from Sentinel-2 and Landsat 8-9 satellites at a scale of 30 m / pixel had a greater informative capacity for training models in crop classification problems.

The overall crop classification accuracy was determined using the output of the XGBoost model on the fused 30 m/pixel satellite data averaged over 2022, 2023 (see Table 2).

The highest overall accuracy for land use objects was obtained when classifying corn, barley, annual grasses, as well as perennial grasses and layland. The recognition accuracy of the above crops was 0.2–2.0% higher than the weighted average F1-score. For oats, peas, soft winter wheat and vetch, the overall classification accuracy was 0.5–2.0% lower than the weighted average F1-score. The lowest overall accuracy was obtained when classifying potatoes, spring rape, soft spring wheat and fallow.

Based on the results of classification of remote sensing images using the XGBoost model, land use maps of the EEP "Praktik" and the ES "Elitnaya" were created, containing 11 identified classes of agricultural crops, as well as layland and fallow for 2022 (a) and 2023 (b) (see Fig. 2).

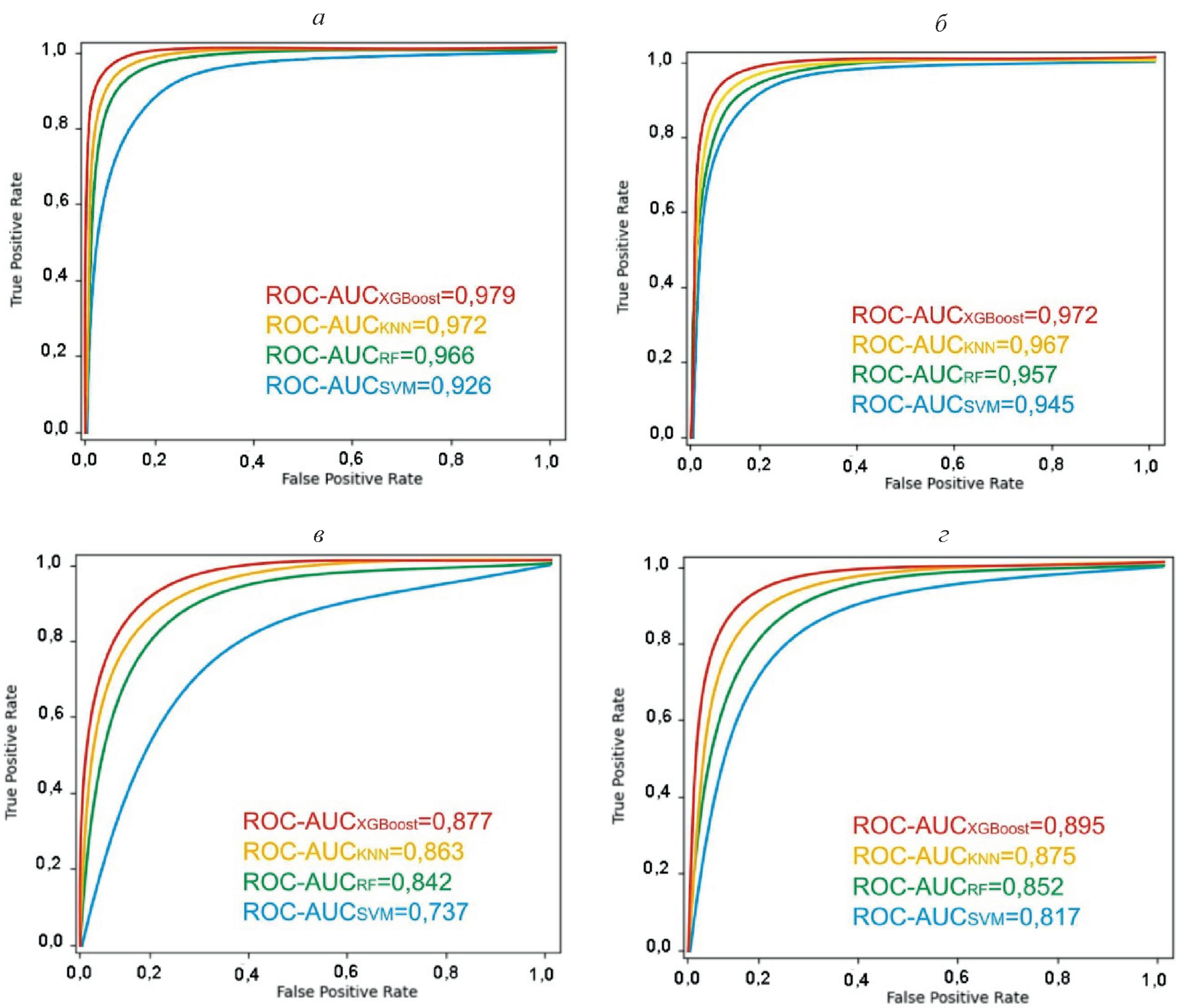

Рис. 1. ROC-кривые для разных методов машинного обучения. Масштаб 30 м/пиксель: *а* – УОХ «Практик»; *б* – ОС «Элитная»; масштаб 90 м/пиксель: *в* – УОХ «Практик», *г* – ОС «Элитная»

Fig. 1. ROC curves for different machine learning methods. Scale 30 m/pixel: *а* – EEP "Praktik", *б* – ES "Elitnaya"; scale 90 m/pixel: *в* – EEP "Praktik", *г* – ES "Elitnaya"

Mapping of agricultural crop species in the study areas was carried out using geoinformation technologies based on their classification (identification) using data analysis obtained from satellite images. ML models allow for efficient processing of large volumes of vector information from space images and contribute to increased classification accuracy. This opens up new opportunities for researchers and agronomists, allowing them to identify crop species and make informed agricultural decisions based on up-to-date information on their spatial localization.

Табл. 2. Общая точность классификации сельскохозяйственных культур моделью XGBoost

Table 2. Overall accuracy of crop classification by the XGBoost model

Class	Culture	Accuracy, %
1	Potato	83,5
2	Soft spring wheat	91,0
3	Layland	97,5
4	Soft winter wheat	95,0
5	Perennial grasses	96,7
6	Spring rape	88,0
7	Fallow	92,5
8	Corn	98,5
9	Annual grasses	98,0
10	Barley	98,5
11	Vetch	96,0
12	Oat	94,5
13	Peas	94,5
Weighted average F1-score		96,5

CONCLUSIONS

1. Evaluation of the ML models using the weighted average F1-score showed fairly high performance and accuracy of crop classification by the XGBoost, KNN, and RF algorithms, with the exception of the SVM method. The best results on all data sets in both years of research were demonstrated by the XGBoost model with a classification accuracy of 94 to 97% at a scale of 30 m/pixel and 80–90% at a scale of 90 m/pixel. The KNN model demonstrated approximately the same classification accuracy as the XGBoost model. A slightly lower classification accuracy was observed when implementing the RF model – 93–96% at a scale of 30 m/pixel and 80–88% at a scale of 90 m/pixel. The lowest classification accuracy rates were obtained when implementing the deep-FNN model, which varied within 78–92% at a scale of 30 m/pixel, and within 56–79% at a scale of 90 m/pixel.

2. The accuracy of crop classification on similar datasets was 8–12% lower at 90 m/pixel than at 30 m/pixel. This highlights the significant impact of scale on crop recognition accuracy. Also, training all models on combined Sentinel-2 and Landsat 8–9 data at 30 m/pixel showed higher F1-score values than on individual data from each of these satellites.

3. The analysis of the model accuracy assessment by two metrics (F1-score and ROC-AUC-score) showed that the most productive and accurate at different scale values for both years of research and for both land use objects was the extreme gradient boosting method (XGBoost), slightly inferior to it in accuracy were the k-nearest neighbors (KNN) and random forest (RF) methods. The lowest accuracy rates were shown by the support vector machine (SVM). The combined data from Sentinel-2 and

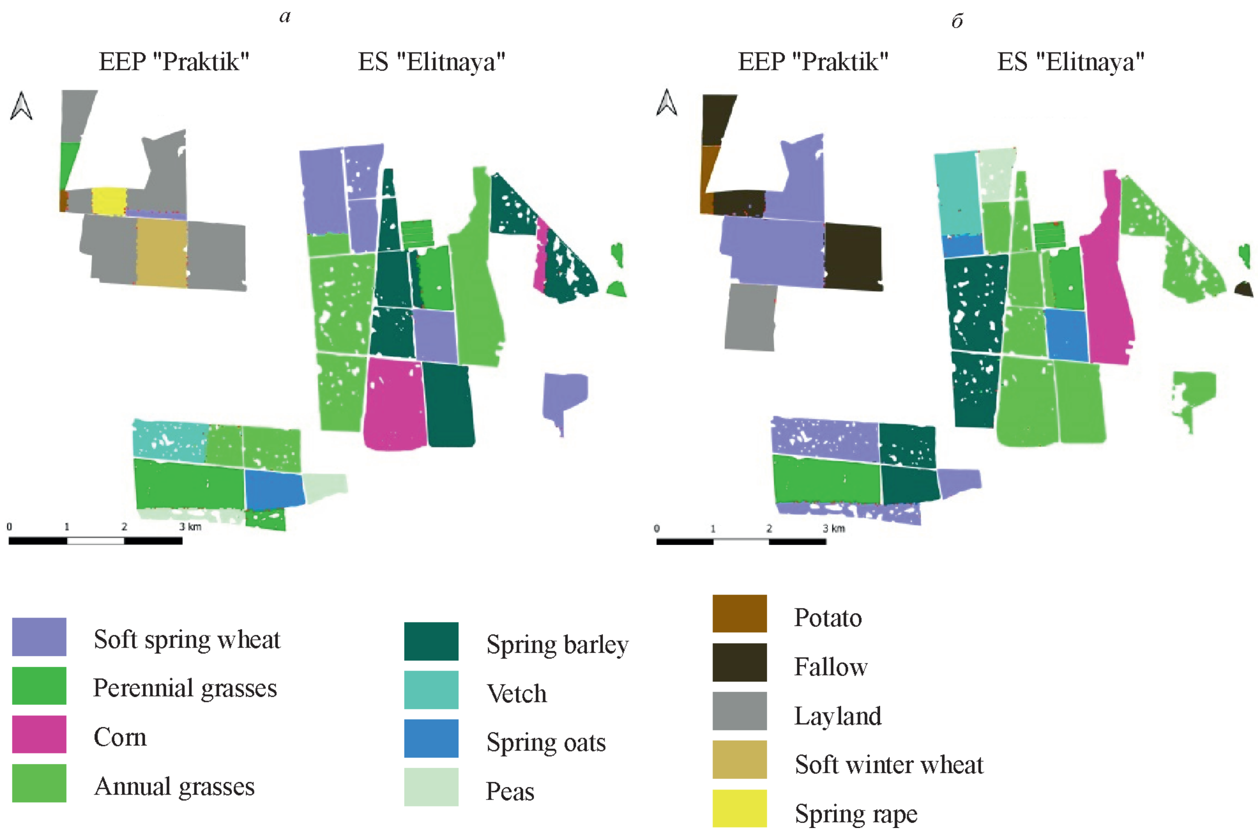

Рис. 2. Карты классификации сельскохозяйственных культур УОХ «Практик» и ОС «Элитная»
Fig. 2. Maps of classification of agricultural crops of the EEP "Praktik" and ES "Elitnaya"

Landsat 8-9 satellites at a scale of 30 m/pixel had a greater informative capacity for training models.

4. The highest overall accuracy (average for 2022, 2023) for land use objects was obtained for corn, barley, annual grasses, as well as for perennial grasses and layland when classifying the combined Sentinel-2 and Landsat 8-9 data at a resolution of 30 m/pixel using the XGBoost model. The recognition accuracy of the above crops was 0.2–2.0% higher than the weighted average F1-score. For oats, peas, soft winter wheat and vetch, the overall classification accuracy was 0.5–2.0% lower than the weighted average F1-score. The lowest overall accuracy was obtained when classifying potatoes, spring rape, soft spring wheat and fallow.

5. The use of remote sensing and ML allows the creation of crop placement maps in land use space, thereby reducing the labor costs of field geolocation and the creation of digital online statistics of cultivated crops. These maps can also be used to monitor crops and forecast yields.

СПИСОК ЛИТЕРАТУРЫ

1. *Плотников Д.Е., Хвостиков С.А., Барталев С.А.* Метод автоматического распознавания сельскохозяйственных культур на основе спутниковых данных и имитационной модели развития растений // Современные проблемы дистанционного зондирования Земли из космоса. 2018. Т. 15. № 4. С. 131–141.
2. *Якушев В.П., Захарян Ю.Г., Блохина С.Ю.* Состояние и перспективы использования дистанционного зондирования Земли в сельском хозяйстве // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 1. С. 287–294. DOI: 10.21046/2070-7401-2022-19-1-287-294.
3. *Илларионова Л.В., Степанов А.С., Фомина Е.А.* Распознавание и классификация посевов сельскохозяйственных культур Ха-

баровского края с использованием NDVI и EVI // Современные проблемы дистанционного зондирования Земли из космоса. 2023. Т. 20. № 2. С. 155. DOI: 10.21046/2070-7401-2023-20-2-155-165.

4. *Кононов В.М., Асадуллаев Р.Г., Кузьменко Н.И.* Алгоритм подготовки мультиспектральных спутниковых данных для задачи классификации сельскохозяйственных культур // Научный результат. Информационные технологии. 2020. Т. 5. № 2. С. 18–24. DOI: 10.18413/2518-1092-2020-5-2-0-3.
5. *Wu B., Zhang M., Zeng H., Tian F., Potgieter A.B., Qin X., Yan N., Chang S., Zhao Y., Dong Q., Boken V., Plotnikov D., Guo H., Wu F., Zhao H., Deronde B., Tits L., Loupian E.* Challenges and opportunities in remote sensing-based crop monitoring: a review // National Science Review. 2023. Vol. 10. N 4. P. 290. DOI: 10.1093/nsr/nwac290.
6. *Weiss M., Jacob F., Duveiller G.* Remote sensing for agricultural applications: A meta-review // Remote sensing of environment. 2020. Vol. 236. P. 111402. DOI: 10.1016/j.rse.2019.111402.
7. *Zeng Y., Hao D., Huete A., Dechant B., Berry J., Chen J.M., Joiner J., Frankenberg C., Bond-Lamberty B., Ryu Y., Xiao J., Asrar G.R., Chen M.* Optical vegetation indices for monitoring terrestrial ecosystems globally // Nature Reviews Earth & Environment. 2022. Vol. 3. N 7. P. 477–493. DOI: 10.1038/s43017-022-00298-5.
8. *Chakhar A., Ortega-Terol D., Hernández-López D., Ballesteros R., Ortega J.F., Moreno M.A.* Assessing the accuracy of multiple classification algorithms for crop classification using Landsat-8 and Sentinel-2 data // Remote sensing. 2020. Vol. 12. N 11. P. 1735. DOI: 10.3390/rs12111735.
9. *Song X.P., Huang W., Hansen M.C., Potapov P.* An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping // Science of Remote Sensing. 2021. Vol. 3. P. 100018. DOI: 10.1016/j.srs.2021.100018.
10. *Blickensdörfer L., Schwieder M., Pflugmacher D., Nendel C., Erasmi S., Hostert P.* Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany // Remote sensing of environment. 2022. Vol. 269. P. 112831. DOI: 10.1016/j.rse.2021.112831.
11. *Прохорец И.О., Степанов А.С.* Картографирование земель сельскохозяйственного назначения Хабаровского края методами машинного обучения с использованием изображений Sentinel-2 // Системы и средства информатики. 2024. Т. 34. № 1. С. 57–69. DOI: 10.14357/08696527240105.
12. *Дубровин К.Н., Степанов А.С., Верхотуров А.Л., Асеева Т.А.* Идентификация сельскохозяйственных культур с использованием радарных изображений // Информатика и автоматизация. 2022. Т. 21. № 2. С. 405–426. DOI: 10.15622/ia.21.2.7.
13. *Guo L., Zhao S., Gao J., Zhang H., Zou Y., Xiao X.* A novel workflow for crop type mapping with a time series of synthetic aperture radar and optical images in the Google Earth Engine // Remote Sensing. 2022. Vol. 14. N 21. P. 5458. DOI: 10.3390/rs14215458.
14. *Faqe I.G.R., Rasul A., Abdulla H.* Improving Crop Classification Accuracy with Integrated Sentinel-1 and Sentinel-2 Data: a Case Study of Barley and Wheat // Journal of Geovisualization and Spatial Analysis. 2023. Vol. 7. N 2. P. 22. DOI: 10.1007/s41651-023-00152-2.
15. *Воробьев Н.И., Пухальский Я.В., Астапова М.А., Сурин В.Г., Пищик В.Н.* Цифровая обработка фотометрических данных дистанционного зондирования полей озимой ржи // Аграрный вестник Урала. 2024. Т. 24. № 2. С. 152–162. DOI: 10.32417/1997-4868-2024-24-02-152-162.
16. *Асадуллаев Р.Г., Кузьменко Н.И.* Технология интеллектуального распознавания сельскохозяйственных культур нейронной сетью по мультиспектральным многовременным данным дистанционного зондирования Земли // Экономика. Информатика. 2022. Т. 49. № 1. С. 159–168. DOI: 10.52575/2687-0932-2022-49-1-159-168.
17. *Joshi A., Pradhan B., Gite S., Chakraborty S.* Remote-sensing data and deep-learning techniques in crop mapping and yield prediction:

A systematic review // Remote Sensing. 2023. Vol. 15. N 8. P. 2014. DOI: 10.3390/rs15082014.

18. *Токарев К.Е., Лебедь Н.И.* Мультиклассовое распознавание посевов сельскохозяйственных культур рекуррентной нейронной сетью глубокого обучения со сверточными слоями по цветным аэрофотоснимкам высокого разрешения // Международный сельскохозяйственный журнал. 2024. № 2. С. 192–195. DOI: 10.55186/25876740_2024_67_2_192.
19. *Alami M.M., Mansouri E., Imani Y., Bourja O., Lahlou O., Zennayi Y., Bourzeixet F., Houmma I.H., Hadria R.* Crop mapping using supervised machine learning and deep learning: A systematic literature review // International Journal of Remote Sensing. 2023. Vol. 44. N 8. P. 2717–2753. DOI: 10.1080/01431161.2023.2205984.
20. *Orynbaikyzy A., Gessner U., Mack B., Conrad C.* Crop Type Classification Using Fusion of Sentinel-1 and Sentinel-2 Data: Assessing the Impact of Feature Selection, Optical Data Availability, and Parcel Sizes on the Accuracies // *Remote Sensing*. 2020. Vol. 12. N 17. P. 2779. DOI: 10.3390/rs12172779.
21. *Barriere V., Claverie M., Schneider M., Lemoine G., d'Andrimont R.* Boosting crop classification by hierarchically fusing satellite, rotational, and contextual data // Remote Sensing of Environment. 2024. Vol. 305. P. 114110. DOI: 10.1016/j.rse.2024.114110.
22. *Karmakar P., Teng S.W., Murshed M., Pang S., Li Y., Lin H.* Crop monitoring by multimodal remote sensing: A review // Remote Sensing Applications: Society and Environment. 2024. Vol. 33. P. 101093. DOI: 10.1016/j.rsase.2023.101093.

REFERENCES

1. Plotnikov D.E., Hvostikov S.A., Bartalev S.A. Method for automated crop types mapping using remote sensing data and a plant growth simulation model. *Sovremenny`e problemy` distancionnogo zondirovaniya Zemli iz kosmosa = Current problems in remote sensing of the Earth from space*, 2018, vol. 15, no. 4. pp. 131–141. (In Russian).
2. Yakushev V.P., Zaxaryan Yu.G., Bloxina S.Yu. Current problems and prospects for the use of remote sensing of the Earth in agriculture. *Sovremenny`e problemy` distancionnogo zondirovaniya Zemli iz kosmosa = Current problems in remote sensing of the Earth from space*, 2022, vol. 19. no. 1. pp. 287–294. (In Russian). DOI: 10.21046/2070-7401-2022-19-1-287-294.
3. Illarionova L.V., Stepanov A.S., Fomina E.A. Recognition and classification of agricultural crops in Khabarovsk region using NDVI and EVI. *Sovremenny`e problemy` distancionnogo zondirovaniya Zemli iz kosmosa = Current problems in remote sensing of the Earth from space*, 2023, vol. 20, no. 2, p. 155. (In Russian). DOI: 10.21046/2070-7401-2023-20-2-155-165.
4. Kononov V.M., Asadullaev R.G., Kuz`menko N.I. Algorithm of multi-spectral satellite data preparation for agricultural crop classification. *Nauchny`j rezul`tat. Informacionny`e texnologii = Research result. Information technologies*, 2020, vol. 5, no. 2, pp. 18–24. (In Russian). DOI: 10.18413/2518-1092-2020-5-2-0-3.
5. Wu B., Zhang M., Zeng H., Tian F., Potgieter A.B., Qin X., Yan N., Chang S., Zhao Y., Dong Q., Boken V., Plotnikov D., Guo H., Wu F., Zhao H., Deronde B., Tits L., Loupian E. Challenges and opportunities in remote sensing-based crop monitoring: a review. *National Science Review,* 2023, vol. 10, no. 4, p. 290. DOI: 10.1093/nsr/nwac290.
6. Weiss M., Jacob F., Duveiller G. Remote sensing for agricultural applications: A meta-review. *Remote sensing of environment*, 2020, vol. 236, p. 111402. DOI: 10.1016/j.rse.2019.111402.
7. Zeng Y., Hao D., Huete A., Dechant B., Berry J., Chen J.M., Joiner J., Frankenberg C., Bond-Lamberty B., Ryu Y., Xiao J., Asrar G.R., Chen M. Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nature Reviews Earth & Environment*, 2022, vol. 3, no. 7, p. 477–493. DOI: 10.1038/s43017-022-00298-5.
8. Chakhar A., Ortega-Terol D., Hernández-López D., Ballesteros R., Ortega J.F., Moreno M.A. Assessing the accuracy of multiple classification algorithms for crop classification using Landsat-8 and Sentinel-2 data. *Remote*

sensing, 2020, vol. 12, no. 11, p. 1735. DOI: 10.3390/rs12111735.

9. Song X.P., Huang W., Hansen M.C., Potapov P. An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. *Science of Remote Sensing*, 2021, vol. 3, p. 100018. DOI: 10.1016/j.srs.2021.100018.
10. Blickensdörfer L., Schwieder M., Pflugmacher D., Nendel C., Erasmi S., Hostert P. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote sensing of environment*, 2022, vol. 269, p. 112831. DOI: 10.1016/j.rse.2021.112831.
11. Proxorecz I.O., Stepanov A.S. Mapping of the Khabarovsk region arable lands by machine learning using Sentinel-2 images. *Sistemy` i sredstva informatiki = Systems and means of informatics*, 2024, vol. 34, no. 1, pp. 57–69. (In Rusian). DOI: 10.14357/08696527240105.
12. Dubrovin K.N., Stepanov A.S., Verxoturov A.L., Aseeva T.A. Crop identification using radar images. *Informatika i avtomatizaciya = Informatics and Automation*, 2022, vol. 21, no. 2, pp. 405–426. (In Russian). DOI: 10.15622/ia.21.2.7.
13. Guo L., Zhao S., Gao J., Zhang H., Zou Y., Xiao X. A novel workflow for crop type mapping with a time series of synthetic aperture radar and optical images in the Google Earth Engine. *Remote Sensing*, 2022, vol. 14, no. 21, p. 5458 DOI: 10.3390/rs14215458.
14. Faqe I.G.R., Rasul A., Abdulla H. Improving Crop Classification Ac-curacy with Integrated Sentinel-1 and Sentinel-2 Data: a Case Study of Barley and Wheat. *Journal of Geovisualization and Spatial Analysi,* 2023, vol. 7, no. 2, p. 22. DOI: 10.1007/s41651-023-00152-2.
15. Vorob`ev N.I., Puxal`skij Ya.V., Astapova M.A., Surin V.G., Pishhik V.N. Digital processing of photometric data of remote sensing of winter rye fields. *Agrarny`j vestnik Urala = Agrarian Bulletin of the Urals*, 2024, vol. 24, no. 2, pp. 152–162. (In Russian). DOI: 10.32417/1997-4868-2024-24-02-152-162.
16. Asadullaev R.G., Kuz`menko N.I. Technology of intelligent agricultural crops recognition by neural network based on multispectral multitemporal Earth remote sensing data. *E`konomika. Informatika = Economics. Information technologies,* 2022, vol. 49, no. 1, pp. 159–168. (In Russian). DOI: 10.52575/2687-0932-2022-49-1-159-168.
17. Joshi A., Pradhan B., Gite S., Chakraborty S. Remote-sensing data and deep-learning techniques in crop mapping and yield prediction: A systematic review. *Remote Sensing*, 2023, vol. 15, no. 8, p. 2014. DOI: 10.3390/rs15082014.
18. Tokarev K.E., Lebed` N.I. Multiclass recognition of agricultural crops by a recurrent deep learning neural network with convolutional layers based on high-resolution color aerial photographs. *Mezhdunarodny`j sel`skoxozyajstvenny`j zhurnal = International agricultural journal*, 2024, no. 2, pp. 192–195. (In Russian). DOI: 10 .55186/25876740_2024_67_2_192.
19. Alami M.M., Mansouri E., Imani Y., Bourja O., Lahlou O., Zennayi Y., Bourzeixet F., Houmma I.H., Hadria R. Crop mapping using supervised machine learning and deep learning: A systematic literature review. *International Journal of Remote Sensing,* 2023, vol. 44, no. 8, pp. 2717–2753. DOI: 10.1080/01431161.2023.2205984.
20. Orynbaikyzy A., Gessner U., Mack B., Conrad C. Crop Type Classification Using Fusion of Sentinel-1 and Sentinel-2 Data: Assessing the Impact of Feature Selection, Optical Data Availability, and Parcel Sizes on the Accuracies. *Remote Sensing,* 2020, vol. 12, no. 17, p. 2779. DOI: 10.3390/rs12172779.
21. Barriere V., Claverie M., Schneider M., Lemoine G., d'Andrimont R. Boosting crop classification by hierarchically fusing satellite, rotational, and contextual data. *Remote Sensing of Environment*, 2024, vol. 305, p. 114110 DOI: 10.1016/j.rse.2024.114110.
22. Karmakar P., Teng S.W., Murshed M., Pang S., Li Y., Lin H. Crop monitoring by multimodal remote sensing: A review. *Remote Sensing Applications: Society and Environment*, 2024, vol. 33, p. 101093 DOI: 10.1016/j.rsase.2023.101093.

ИНФОРМАЦИЯ ОБ АВТОРАХ

(✉)**Каличкин В.К.,** главный научный сотрудник, доктор сельскохозяйственных наук, профессор; ORCID 0000-0002-7765-3451**,** SPIN-код 1587-4615**; адрес для переписки:** Россия, 630501, Новосибирская область, р.п. Краснообск, а/я 463; e-mail: v.kalichkin@gmail.ru

Крылова О.С., младший научный сотрудник, аспирант; SPIN-код 9571-1912

Гарафутдинова Л.В., младший научный сотрудник; ORCID 0000-0001-5498-6539, SPIN-код 1924-2317

AUTHOR INFORMATION

(✉)**Vladimir K. Kalichkin,** Head Researcher, Doctor of Science in Agriculture, Professor; ORCID 0000-0002-7765-3451, SPIN-code 1587-4615; **address:** PO Box 463, Krasnoobsk, Novosibirsk Region, 630501, Russia; e-mail: v.kalichkin@gmail.ru

Olga S. Krylova, Junior Researcher, Post-graduate Student; SPIN-code 9571-1912

Lyudmila V. Garafutdinova, Junior Researcher; ORCID 0000-0001-5498-6539, SPIN-code 1924-2317

Дата поступления статьи / Received by the editors 29.11.2024
Дата принятия к публикации / Accepted for publication 20.12.2024
Дата публикации / Published 15.04.2025