International Journal of Ecosystems and Ecology Science (IJEES) **Vol. 14 (4):** 147-166 (2024)
https://doi.org/10.31407/ijees https://doi.org/10.31407/ijees14.4

ADVANCES IN AGRICULTURAL IMAGE SEGMENTATION: FROM TRADITIONAL TECHNIQUES TO DEEP LEARNING APPROACHES

N. Gapon[1*], A. Puzerenko[1], M. Zhdanova[1,2], O. Safaryan[1], D. Rudoy[1]

[1]Don State Technical University, 1 Gagarin square, Rostov-on-Don, 344000, Russian Federation;*
[2]Research Laboratory "Modeling and Development of Intelligent Technical Systems of the AIC", 1, bld. 2 Gagarin square, Rostov-on-Don, 344000, Russian Federation;

[*]Corresponding Author N. Gapon, email: nikolay-rt@mail.ru;

Received September 2024; Accepted October 2024; Published November 2024;

DOI: https://doi.org/10.31407/ijees14.419

ABSTRACT

In recent years, image segmentation has been widely applied to solve problems in various fields. With the application of deep learning in machine vision, the excellent performance has been transferred to agricultural image processing by combining them with traditional methods. Segmentation methods have revolutionized the development of agricultural automation and are commonly used to analyze uniform sowing, crop type, identify pests, diseases, etc. This paper provides a review of the recent advances in traditional and deep learning-based methods for agricultural image segmentation. We present traditional methods that can effectively utilize the original image information and the high performance of deep learning-based methods. The review introduces key agricultural image datasets. The main metrics for evaluating the quality of image segmentation are presented, and the evaluation results of various classical and deep learning-based segmentation methods are presented.

Keywords: image segmentation, agricultural images, deep learning, neural networks, segmentation quality metrics.

INTRODUCTION

Agriculture has historically been one of the most labor-intensive areas of human activity. In the past, farmers manually cultivated fields, planted and grew crops, investing significant effort into each stage of the agricultural cycle. This process required not only physical strength, but also a deep knowledge of natural processes. However, with the development of technology, especially in the field of computer vision and artificial intelligence, agriculture has begun to be automated. Today, many tasks that previously required direct human intervention are performed by machines and computer systems. This allows for significant increases in productivity and reduction in labor costs. Machine vision-based approaches have become the driving force behind the advancement of the agricultural industry. Vision-based technologies are used in scenarios such as pest identification (Yuan et al., 2022), livestock behavioral traits (Qiao et al., 2019), etc., which can help reduce labor and time costs. An important part of this automation is the use of segmentation techniques, which allow for the identification, classification, and monitoring of various objects in the field, such as crops, soil, weeds, and pests. Segmentation is one of the main tasks of computer vision to achieve classification. It is an important component of computer vision applications. Unlike image prediction, segmentation generates pixel-level descriptions of objects embedded in their spatial information.

With the development of segmentation methods, they have been used to solve various agricultural problems, such as crop analysis (Kussul et al., 2017; Osadchy et al., 2024; Xhina et al., 2023), forest tree species labeling (Dechesne et al., 2017), weed segmentation (Zou et al., 2021), predictive agriculture (Anand et al., 2021), pest and disease identification (Kukusheva et al., 2024), etc. Segmentation methods have played an important role in modern agriculture, replacing traditional manual observation and measurement of phenotypic data. For example, segmentation methods are used to monitor the growth status of crops, predict canopy area and vegetation height, and determine fruit maturity (Utomowati et al., 2024). Machine vision plays a key role in agriculture by optimizing the sowing process and improving crop quality. By analyzing images, the system can determine the optimal harvest time to achieve maximum maturity. In the field of pest control, machine vision allows identifying pests by their shape, size, and texture, which facilitates more effective application of pesticides (Ilyushin and Afanaseva, 2020; Korotenko and Togusakov, 2024). However, working with images of agricultural land over large areas is challenging, as the heterogeneity of the soil and environment makes it difficult to obtain accurate information about the condition of plants and soil. Before deep learning was explored, the robustness of traditional segmentation methods in complex environments was greatly improved by color space transformation and color channel combination. For extreme lighting conditions and sharp shadow edges, vegetation index-based methods use separate color channels for pixel color and brightness, respectively (Riehle et al., 2020). Other methods use machine learning-based classification methods such as decision trees (Du et al., 2024; Gorelkina et al., 2024; Yang et al., 2015). Unlike traditional methods, deep learning automatically extracts features that are tailored to specific classification problems. This makes it effective in handling complex and diverse application scenarios. The successful application of deep learning in other fields has prompted its integration with traditional segmentation methods to address agricultural problems. Deep learning architectures such as VGG, FCN, U-Net, SegNet, DeepLab, and others have been widely used for segmentation. For segmentation tasks, U-Net is able to combine low-level and high-level features while preserving edge information and having a low computational burden. In segmenting high-resolution images of fruit tree branches, U-Net (Ronneberger et al., 2015) performs well using a modified cross-entropy loss (Rafikov et al., 2024). For segmenting complex images, DeepLabV3+ uses dilated convolution with a larger receptive field. According to pig counts, the MIoU value of the improved DeepLabV3+ is 74.62%, which is higher than other models (Bugubaeva et al., 2023). In addition to CNN architectures, generative adversarial networks (GANs) also perform well in semantic segmentation tasks. Discriminators in GANs can help learn the relationships between pixels (often ignored in CNN architectures), which can improve the performance and accuracy of the networks (Zhang et al., 2021). This paper aims to review existing methods for segmentation of agricultural images. Both classical methods and modern approaches based on deep learning are considered. This will allow us to identify the advantages and disadvantages of various approaches, as well as their applicability in agricultural production.

MATERIAL AND METHOD

Classical Segmentation Methods. Image segmentation allows dividing an image into different regions or objects, which is an important step in monitoring crop conditions, assessing crop quality, and identifying plant diseases. This section discusses both traditional segmentation methods and modern AI-based approaches that are used in agricultural practice. Threshold segmentation (Akhmetshin et al., 2023) is one of the most basic image processing methods, which is based on the selection of a certain threshold value that separates the image into background and objects. This method is applicable in situations where there is a clear difference in brightness between the segmented objects and the background (Figure 1).

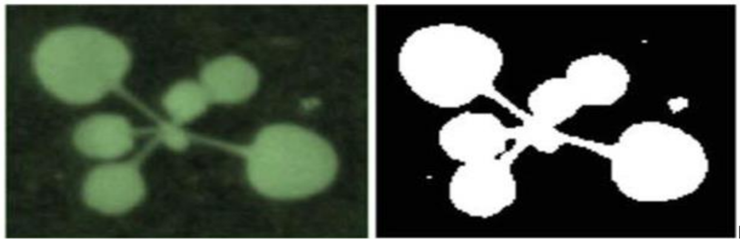

Figure 1. Threshold segmentation of images (Akhmetshin et al., 2023).

In agriculture, threshold segmentation is often used to isolate areas of vegetation or soil, which allows for quick and easy classification of large amounts of data. However, this method has limited applicability in conditions with uneven lighting or in the presence of complex textures in the image. Threshold segmentation clearly separates vegetation from the background, showing its effectiveness under simple conditions. This allows plants to be isolated for further analysis, although errors may occur when lighting or texture changes.
Edge detection methods such as the Canny operator and the Sobel operator (BenHajyoussef and Saidani, 2024) can detect sharp changes in pixel intensity, which is an indicator of the boundaries of objects in an image. These methods are widely used to detect plant contours, determine field boundaries, and other tasks related to the analysis of spatial structures. Edge detection is an important tool for analyzing complex images of agricultural fields, especially in conditions of high heterogeneity of objects in the image. The image shows the application of the Canny operator with different values of the σ parameter to detect the boundaries of objects (Figure 2). The original image shows a cat in its natural environment.

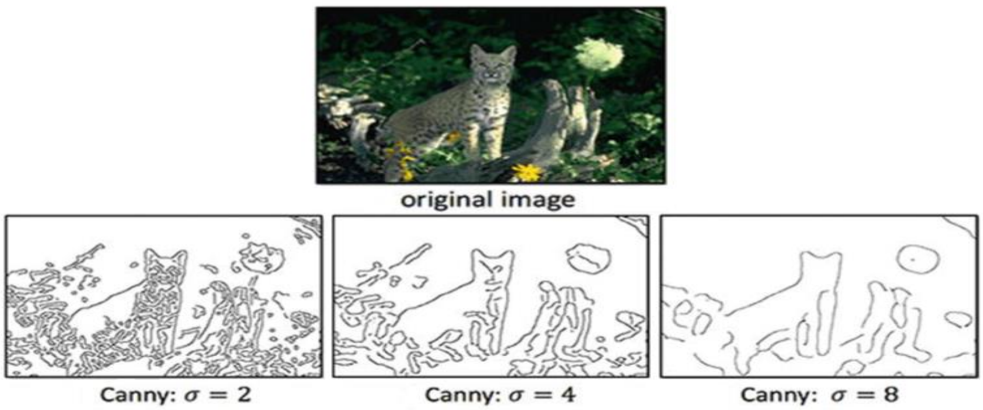

Figure 2. An example of using the Canny operator to highlight the boundaries of objects, (BenHajyoussef and Saidani, 2024).

Using the Canny operator, the contours of surrounding objects become clear and well distinguishable. As the σ value increases, the contours become smoother, which allows better identification of the main structures in the image, such as the boundaries of the object, which improves the accuracy of classification and analysis in agricultural applications. Region splitting and merging methods (Bins et al., 1996) aim to combine pixels into homogeneous regions based on their similar characteristics, such as intensity or texture. One of the most common methods is region growing, where the process starts with small "seeds" and gradually expands to include neighboring pixels if they meet a given similarity criterion. These methods are well suited for agricultural image segmentation (Figure 3), especially when analyzing images with a high degree of homogeneity, such as forest and agricultural areas.

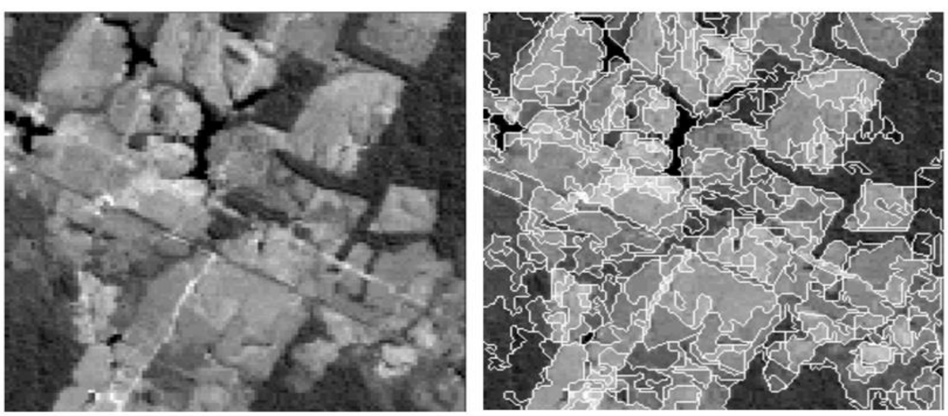

Figure 3. Example of image segmentation using the region split and merge method (Bins et al., 1996).

The region growth method has been successfully applied to segment images of Amazonia, distinguishing different land use types including forest and agricultural areas. The k-means clustering method (Reckling and Grosse, 2022) is an algorithm that divides an image into multiple clusters based on the similarity of pixels in a color space or feature space. This method is widely used to classify different types of vegetation or identify areas with different crop yields (Figure 4).

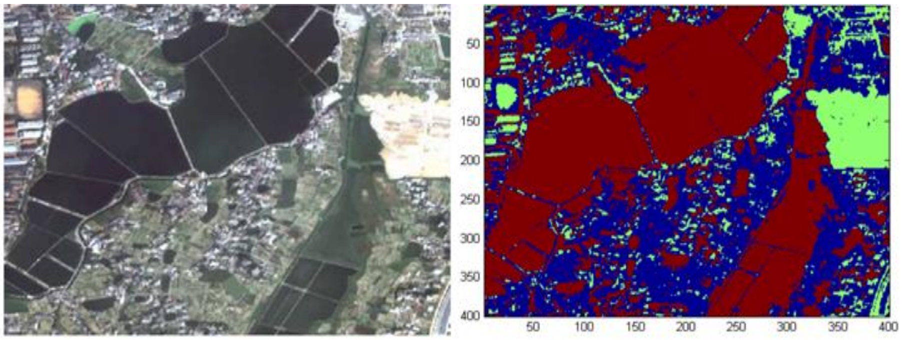

Figure 4. Example of image segmentation using k-means (Reckling and Grosse, 2022).

The main advantage of the method is its simplicity and efficiency in analyzing large amounts of data. However, the method requires a pre-selection of the number of clusters, which may limit its flexibility and adaptability.
The image is divided into several clusters, each representing a specific land use type, such as farmland, built-up area, and bare soil. The clustering using four clusters clearly separates the different land classes, significantly improving the classification accuracy and avoiding the mixed pixel problem.
The superpixel method (Li et al., 2020) is a technique that breaks an image into groups of adjacent pixels with similar characteristics, reducing the complexity of subsequent data processing. Superpixels improve the quality of segmentation by minimizing the number of elements to analyze, which increases the accuracy of subsequent steps such as identifying individual plants or detecting disease zones. This method is especially useful when it is important to preserve fine image detail (Figure 5).

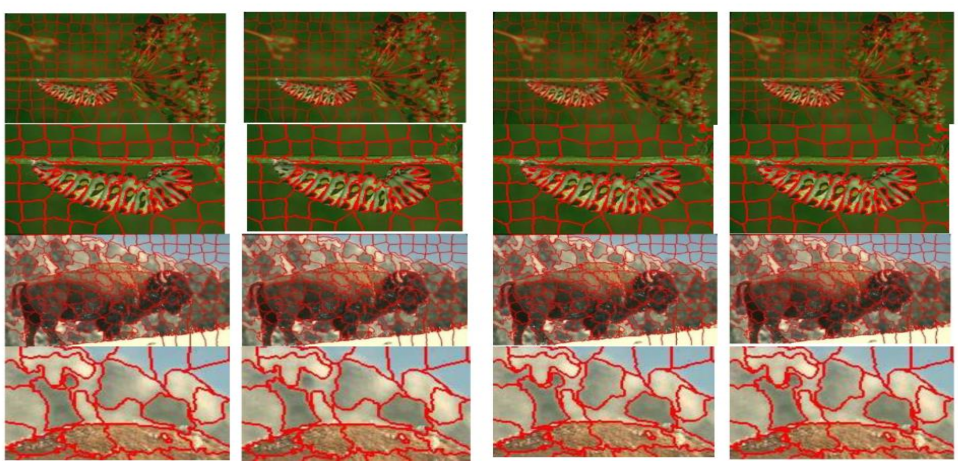

Figure 5. Example of image segmentation using superpixel method (Li et al., 2020).

Using the superpixel method, areas with uniform characteristics are highlighted in an image, allowing you to accurately outline the boundaries of objects such as leaves or animals in the background.
The watershed method (Xue et al., 2021) is based on the analogy with a geographical watershed, where the boundaries of segments are defined as lines dividing areas flowing to different minima in the image. This method is effective for segmenting complex images containing several objects with different intensities. In agriculture, the watershed method can be applied to separate areas with different crops or to identify heterogeneous areas within a single field (Figure 6).

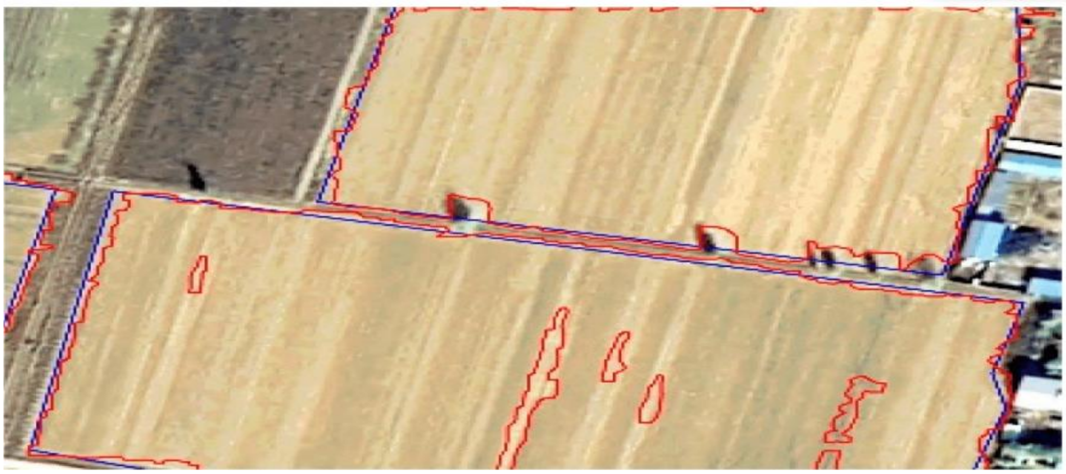

Figure 6. Example of image segmentation using watershed method (Xue et al., 2021).

The watershed method is used to accurately delineate cropland boundaries, allowing for improved land use control and classification based on satellite imagery data.
Spectral segmentation (Moghaddam et al., 2020) uses the analysis of different spectral bands of an image to divide it into homogeneous regions. This method is especially useful when working with multispectral and hyperspectral images, which are often used in agriculture to monitor crop health and assess biophysical parameters (Figure 7).

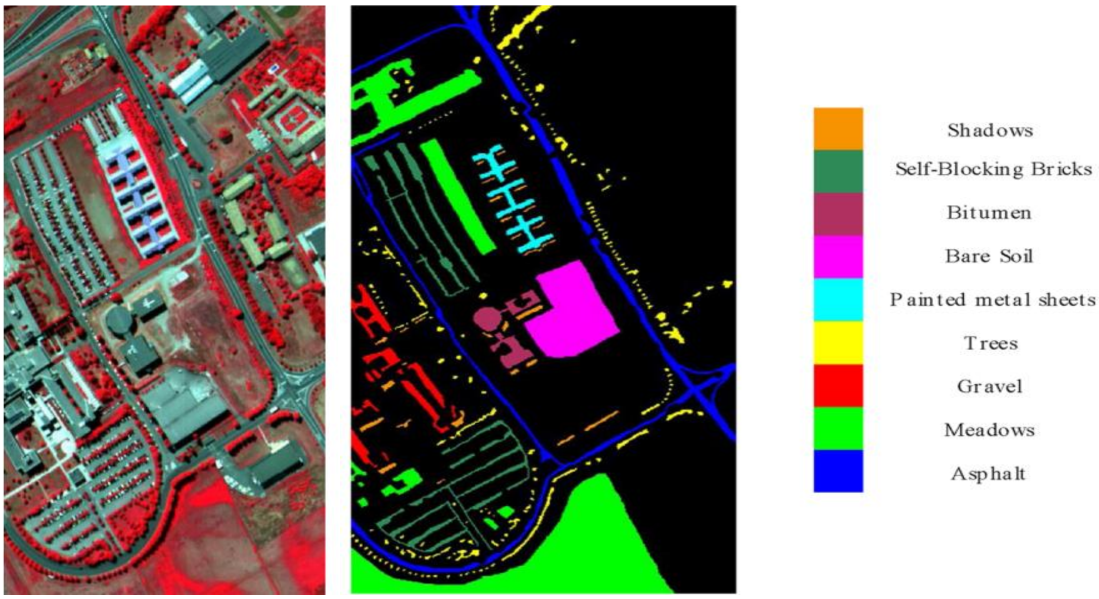

Figure 7. Example of image segmentation using spectral segmentation method (Moghaddam et al., 2020).

In the presented image, the hyperspectral data has been processed using spectral segmentation. As a result of the processing, different types of surfaces are distinguished, such as shadows, concrete blocks, bitumen, bare soil, painted metal sheets, trees, gravel, fields and asphalt. Each category is displayed with its own color, which allows for precise differentiation of different objects and surfaces in the image, ensuring high accuracy in monitoring and classifying areas.
The graph cut method (Belim and Belim, 2022) is based on representing an image as a graph, where pixels serve as vertices, and graph edges represent connections between these vertices, reflecting the similarity or difference between pixels.

Graph cut allows dividing an image into segments by minimizing a certain cost function, which makes this method especially useful for segmenting images with complex structures and heterogeneous areas (Figure 8).

Figure 8. Example of image segmentation using the graph cut method (Belim and Belim, 2022).

This approach can be effective in analyzing agricultural fields, where complex and heterogeneous structures are encountered.
The following images show an example of the use of the cut-based clustering method. The first image shows the original photograph of the road and its surrounding nature. The second image shows the result of applying the cut-based method to the graph, where the road section is highlighted in red. This demonstrates the ability of the method to accurately segment different areas of the image, separating the road from other sections.
Gabor filtering (Akhmetshin et al., 2024) is an image processing technique based on the use of a bank of Gabor filters to extract texture features. Each filter in the bank is tuned to a specific frequency and orientation, allowing textures corresponding to these parameters to be extracted. This approach is widely used for image segmentation based on texture characteristics, which is especially useful in agriculture when analyzing fields with different soil or vegetation types (Figure 9).

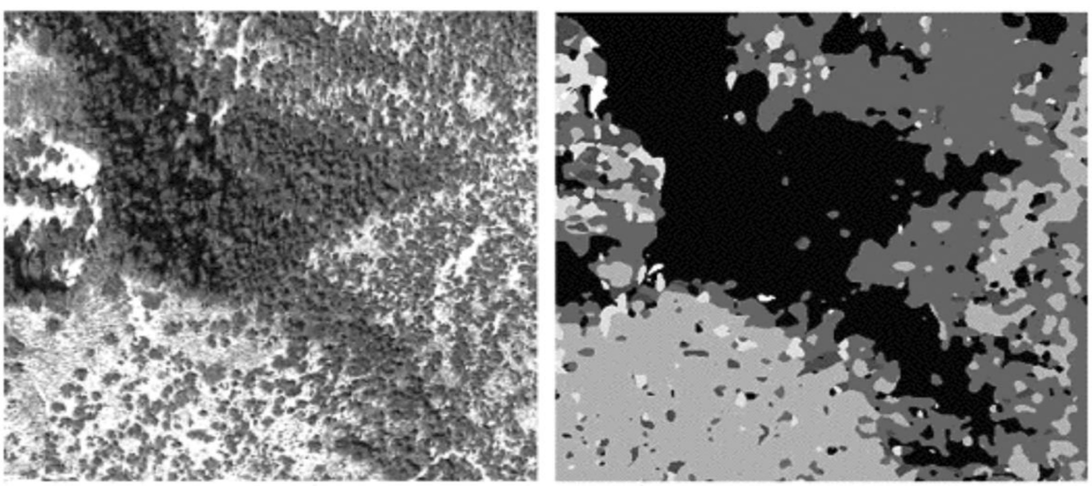

Figure 9. Application of Gabor filters to texture analysis of agricultural fields (Akhmetshin et al., 2024).

Gabor filters allow for precise extraction of texture features, such as different plant species or soil conditions.
The first image shows the original field, the second image shows the result of segmentation using Gabor filters, where different textures are clearly distinguished in individual segments, allowing for better analysis of the diversity and condition of the fields.
The Markov algorithm, or Markov Random Fields (MRF) model (Zheng et al., 2019), is used for image segmentation taking into account spatial dependencies between pixels (Figure 10).

a) b)

Figure 10. Result of satellite image segmentation using Markov algorithm (Zheng et al., 2019).

This method helps to improve the quality of segmentation, especially on noisy data or data with a high degree of texture heterogeneity, by using a probabilistic model that takes into account the influence of neighboring pixels on the classification of each individual pixel.

The image below shows the result of segmentation of a satellite image using the Markov algorithm. Image (a) shows the original satellite image of the study area, which contains agricultural fields, vegetation areas, and urban areas. Image (b) shows the result of segmentation, which identifies three main classes: agricultural fields (1), vegetation (2), and urban areas (3). The use of MRF allowed us to clearly separate the different land use types, despite the presence of heterogeneity and noise in the data, which makes this method especially useful for analyzing agricultural areas.

Segmentation Methods Using Deep Learning

Deep learning has become one of the most significant areas in image analysis, including segmentation tasks, in recent years. Unlike classical methods based on manual value adjustment, deep learning methods can automatically extract complex patterns and adapt to different conditions. In agriculture, where images can be diverse and complex, deep learning can achieve high accuracy and robustness in segmentation. This section discusses the main deep learning-based segmentation methods that have found application in agricultural field data analysis.

Convolutional neural networks (CNNs) (Kamilaris and Prenafeta-Boldú, 2018) are a powerful tool for solving computer vision problems, including image segmentation. Their key advantage is the ability to automatically extract hierarchically organized image features, ranging from simple elements to complex objects. In agriculture, CNNs are widely used to analyze and manage various aspects of agricultural fields, including crop segmentation, weed detection, and identification of damaged plant areas (Figure 11).

Figure 11. Field image segmentation using CNN (Kamilaris and Prenafeta-Boldú, 2018)

The image is segmented using CNN to mark areas of different vegetation. Labels (1) indicate healthy sugarcane plants, labels (2) indicate areas with soil, labels (3) identify areas where vegetation should be but are actually soil, and labels (4) indicate other unimportant image segments that do not belong to the main categories. This approach allows for accurate detection and analysis of problem areas in fields, enabling more efficient management of agricultural resources.

U-Net (Dimitrovski et al., 2024) is one of the most well-known architectures for image segmentation, originally developed for medical purposes but now widely used in various fields including agriculture. U-Net uses a symmetric structure with encoding and decoding parts, which allows it to restore image details at the pixel level and provide high segmentation accuracy (Figure 12).

Figure 12. Result of applying U-Net architecture (Dimitrovski et al., 2024) for image segmentation.

The images below show the results of applying the U-Net architecture to image segmentation. The first image shows the original image with road and building sections. After processing using U-Net (the second image), the road and building sections are clearly highlighted, which helps in accurately identifying the structure of agricultural land and creating maps needed for effective land management.

SegNet (Borodulin and Maksimenko, 2023) is an architecture specifically designed for semantic segmentation tasks. It is based on convolutional neural networks and uses an encoder-decoder structure to restore spatial information. This allows processing high-resolution images and obtaining detailed segmentation maps. In agriculture, SegNet is used to divide fields into areas with different types of soils or crops, as well as for the automatic analysis of large amounts of data, which is especially important when monitoring the condition of vast agricultural lands.

The first image shows a satellite photo of an area with various structures such as buildings and fields (Figure 13).

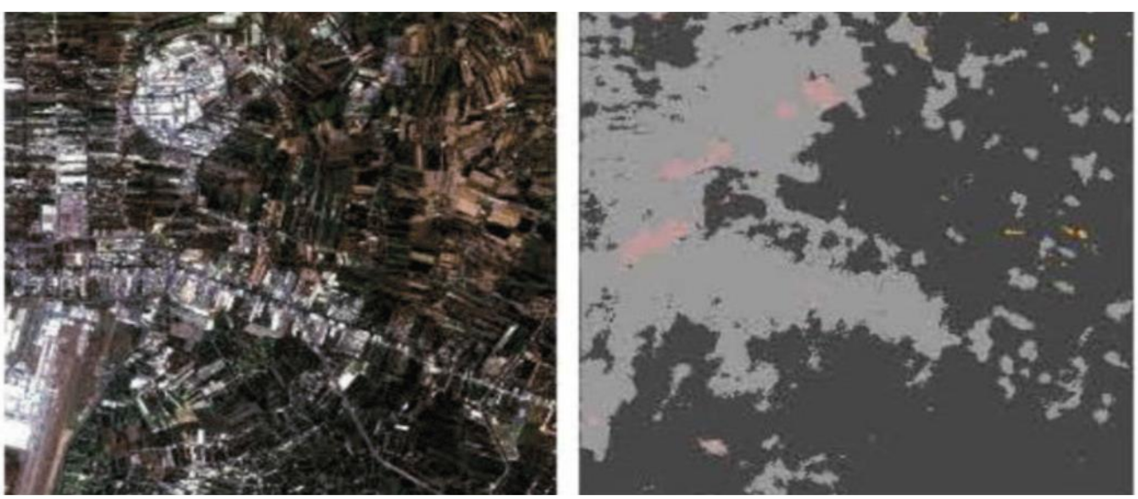

Figure 13. Result of applying SegNet (Borodulin and Maksimenko, 2023),
for semantic segmentation of satellite imagery of the territory.

The second image shows the result of segmentation performed using the SegNet architecture, where different land use types are highlighted. This allows for a clear separation of agricultural and non-agricultural zones, which is critical for effective agricultural land management and land use analysis.
DeepLab (Kradetskaya et al., 2024) is one of the most advanced architectures for image segmentation, incorporating Atrous Spatial Pyramid Pooling (ASPP) and Fully Convolutional Networks (FCN) technologies. These technologies allow the model to efficiently handle high-resolution images and accurately segment objects with significant changes in their scale. In agriculture, DeepLab is used to segment fields into different categories, such as crop types, weeds, and soils.
The first image shows a plant leaf with signs of disease (Figure 14).

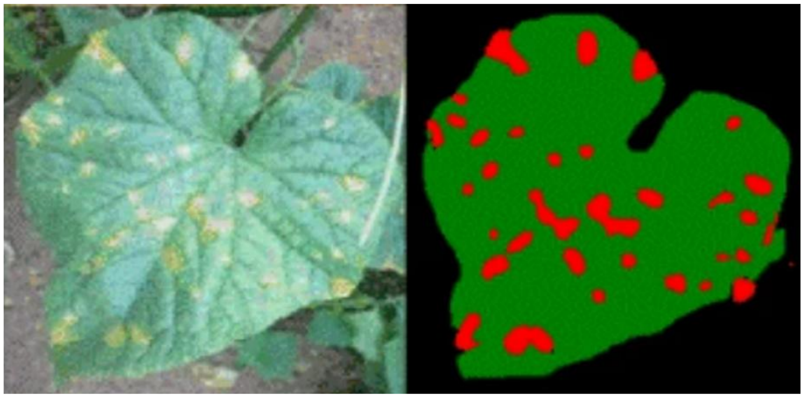

Figure 14. Example of image segmentation using DeepLab (Kradetskaya et al., 2024).

After processing using DeepLab (second image), the model successfully identified areas affected by the disease, which allows for effective monitoring of plant health and timely treatment measures.
Pyramid Scene Parsing Network (PSPNet) (Chen et al., 2021) uses the pyramid scene parsing method to handle objects of different scales in an image. This architecture allows for improved segmentation accuracy, especially in images containing many objects of varying sizes and shapes. In agriculture, PSPNet is used to analyze multi-layered crops, determine soil heterogeneity, and analyze complex landscape structures. The advantage of PSPNet is its ability to account for hierarchical image features, making it particularly effective when working with agricultural data where it is important to account for both large-scale and small-scale details (Figure 15).

Figure 15. An example of using PSPNet (Chen et al., 2021) to segment a bunch of grapes.

The image shows an example of using PSPNet to segment a bunch of grapes. The model first processes an image of a bunch of grapes and then extracts its contours, separating the bunch from the background. This approach significantly simplifies the task of analyzing and classifying the crop, ensuring high segmentation accuracy.
Mask R-CNN (Huang et al., 2024) is an extension of the Faster R-CNN architecture and adds the ability to accurately detect object boundaries at the pixel level. This method allows for both detection and segmentation tasks to be solved simultaneously, making it a particularly powerful tool for analyzing complex images. In agriculture, Mask R-CNN is used to accurately segment individual plants, identify their structural elements such as leaves or

fruits, and to detect the boundaries between different crops in an image. This method is particularly useful when there is a need to accurately analyze objects in high-resolution images (Figure 16).

Figure 16. Image segmentation using Mask R-CNN (Huang et al., 2024)

The first image shows bunches of grapes. After processing using Mask R-CNN (second image), the model highlights the bunches of grapes by accurately defining their boundaries. This helps in accurate analysis of the crop condition and further management of agricultural resources. Autoencoders (Mujkic et al., 2022) are a type of neural network that learns to reconstruct input data from its compressed representation. In segmentation tasks, autoencoders are used to extract hidden features that can then be used for clustering and segmentation of images. In agriculture, autoencoders are used to segment large amounts of data without clear labeling, which is especially important for analyzing large and diverse data sets, such as images of fields with different moisture levels or weed infestations. This method allows for efficient data analysis, revealing hidden patterns and anomalies, which contributes to more accurate management of agricultural processes. The image shows an autoencoder in action, reconstructing an image of an agricultural field and creating an anomaly map (Figure 17).

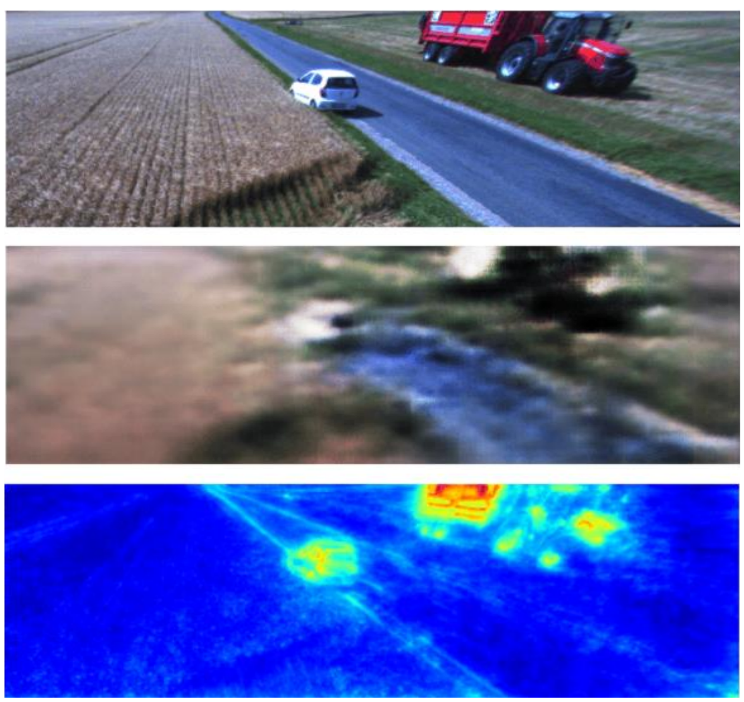

Figure 17. Operation of the autoencoder (Mujkic et al., 2022).

Bright areas on the map indicate deviations that require attention, which is important for field monitoring and problem prevention. Generative adversarial networks (GANs) (Zhuk, 2023) can be used to improve segmentation quality by generating realistic image samples that can then be used to train and improve models. GANs allow the generation of additional training examples, which is especially useful when there is limited data available for training. In agriculture, GANs are used to improve segmentation accuracy and improve field analysis models. For example, GANs can be used to generate additional data that helps improve image segmentation when analyzing complex agricultural landscapes where traditional methods may not be accurate enough. The image on the right (Figure 18) shows the segmentation maps generated by the GAN, which are used to improve the accuracy of the models.

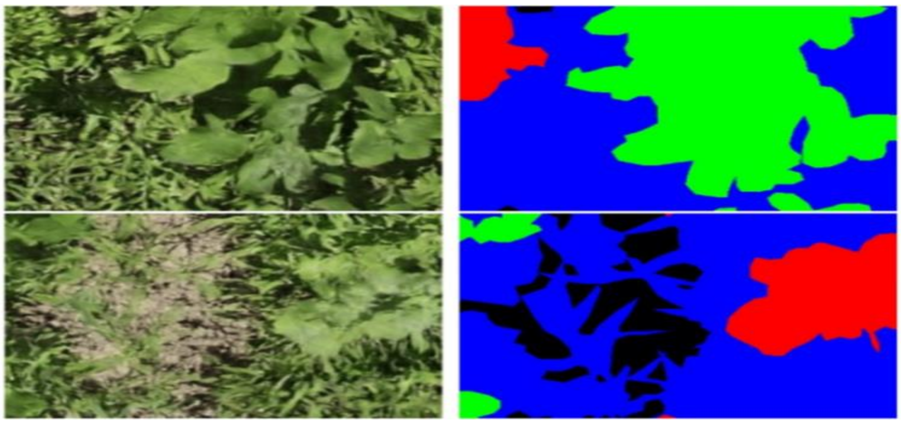

Figure 18. Example of using GAN network (Zhuk, 2023)

The images on the left show the corresponding original crop images. This approach helps improve segmentation models, especially in challenging scenarios such as dense or heterogeneous agricultural landscapes.
Convolutional models with graphical models (Abdullaev et al., 2023) are an approach that improves the segmentation results obtained by convolutional neural networks (CNNs) by using graphical models such as Conditional Random Fields (CRFs). CRFs allow for spatial dependencies and refinement of segment boundaries in an image. In agriculture, such combined models are used to more accurately segment fields where it is important to consider the relationships between objects, such as between different crops and weeds. This approach provides high accuracy and granularity of segmentation, which is critical when analyzing images of agricultural land.

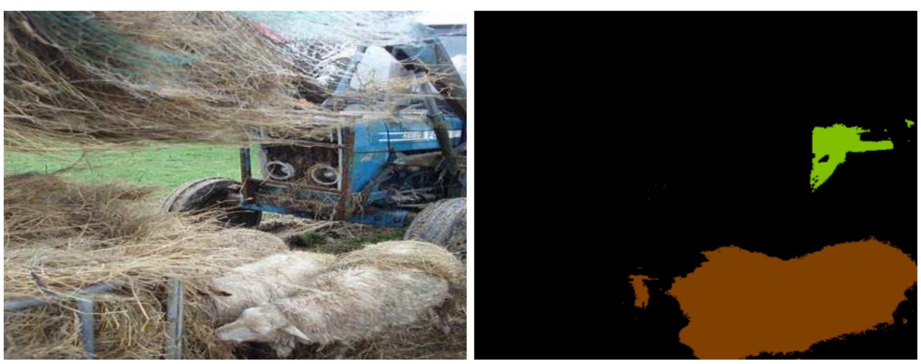

Figure 19. Segmentation result improved using CRF (Abdullaev et al., 2023)
Figure 19 shows an input image of an agricultural scene with animals that can be seen near a tractor. The segmentation result improved using CRF is shown in the second image, where the animals are highlighted in brown and some parts of the tractor are highlighted in green. This approach helps to more accurately identify and separate different objects in the image by taking into account their spatial relationships.
Long Short-Term Memory (LSTM) (Sagynbayeva et al., 2023) is a type of recurrent neural network (RNN) designed to handle data with temporal and spatial dependencies. In segmentation tasks, LSTMs are useful when it is

important to consider changes in objects or their characteristics over time. In agriculture, LSTM is used to analyze image sequences to track crop growth at different stages, predict changes in crop condition, and detect anomalies caused by climatic conditions. This provides a more accurate understanding of the dynamics of agricultural processes and helps optimize agricultural resource management (Figure 20).

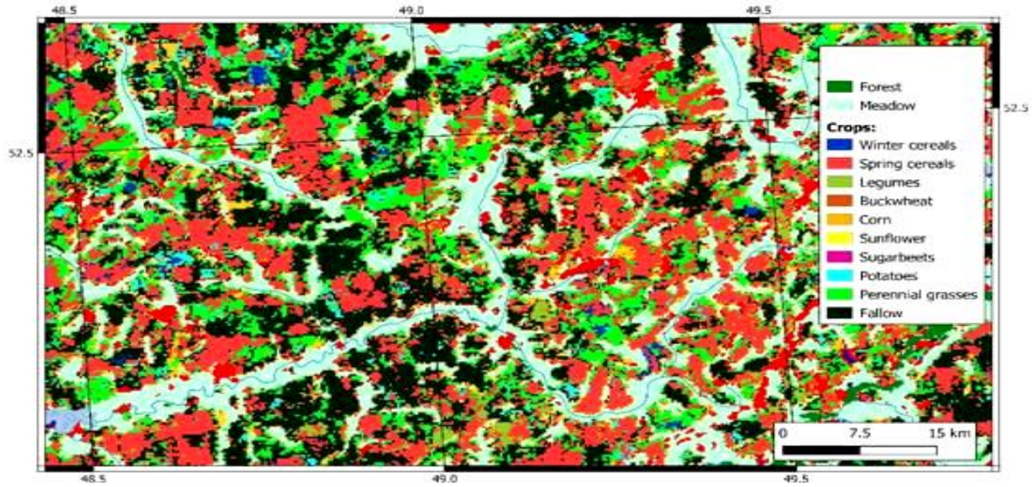

Figure 20. An example of using LSTM (Sagynbayeva et al., 2023) to analyze images of agricultural land.

The example presented uses LSTM to analyze sequential images of agricultural fields. The image shows how LSTM helps classify different crop types (e.g. winter grains, corn, sunflower) over large areas, taking into account temporal changes. This allows agronomists to more accurately predict crop yields and identify areas at risk, which significantly improves planning and management of agricultural processes.

Convolutional neural networks (CNNs) combined with Active Contour Models (ACMs) are a powerful tool for image segmentation, combining automatic feature extraction with accurate detection of object boundaries (Kharitonova, 2023). In this approach, CNNs identify key image features, such as plant texture or soil structure, while ACMs refine the boundaries, providing detailed segmentation. In agriculture, such models are used to accurately detect plant contours, determine crop boundaries, and identify pest infestations. This combination allows for high segmentation accuracy even in complex agricultural landscapes, enabling more efficient management of agricultural processes.

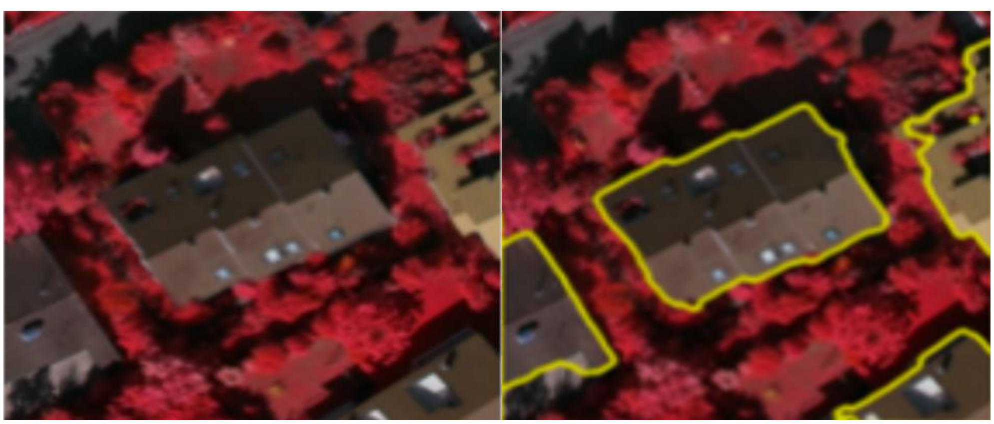

Figure 21. Example of image segmentation using CNN in combination with active contours (Kharitonova, 2023).

Figure 21 shows the segmentation process using CNN in combination with active contours. The original image is a satellite image of an area with buildings and vegetation. Using active contours allows for precise detection of the building boundaries, which improves the quality of subsequent data processing and increases the accuracy of the analysis.

RESULTS

Datasets for Agricultural Fields Analysis

Datasets play a key role in developing and testing image segmentation methods, especially in specific domains such as agriculture. High-quality datasets allow researchers to develop and evaluate models that can be effectively applied in real-world settings. This section reviews the main datasets used for agricultural field analysis, including both classic and modern datasets that are geared towards using deep learning methods.

Landsat and MODIS

One of the key sources of data for analyzing agricultural fields is satellite imagery, such as that from Landsat and MODIS (Roy et al., 2014). These datasets provide regular images of large areas of the Earth's surface, making them indispensable for monitoring land use, including agricultural areas. The images have limited resolution, but their advantage is the frequency of surveys and their global coverage. The data is used to analyze seasonal changes, estimate crop yields, and monitor land use changes.

PlantVillage

PlantVillage (Ismailova et al., 2024) is a state-of-the-art dataset for crop health analysis (Figure 22).

Figure 22. Examples of PlantVillage dataset

It includes thousands of images of various crops affected by diseases and is actively used to train classification and segmentation models. PlantVillage enables researchers to create and test algorithms that can effectively diagnose plant diseases and monitor their health at the level of individual leaves or fruits. In addition to images, the dataset contains metadata such as crop type, geographic location, and growing conditions, greatly expanding its application in agricultural research.

2020 IEEE GRSS Data Fusion Contest

Multispectral and hyperspectral datasets, such as those provided by the 2020 IEEE GRSS Data Fusion Contest (Schmitt et al., 2019), are more detailed datasets that include images of agricultural fields captured in more than 100 spectral bands. These data allow for in-depth analysis of plant and soil health, revealing subtle spectral differences that are not visible in regular RGB images. This level of detail is particularly useful for segmentation tasks related to distinguishing crop species and analyzing plant physiological health.

UAVid

The use of unmanned aerial vehicles (UAVs) in agriculture has opened up new opportunities for obtaining data with high spatial and temporal resolution.

The UAVid dataset (Lyu et al., 2020) includes high-resolution images (Figure 23) acquired from drones used to analyze both urban and agricultural landscapes.

Figure 23. Examples of UAVid dataset

This dataset is particularly valuable for segmentation and classification tasks, as the images include annotations that allow for the accurate identification and classification of various elements of agricultural fields, such as individual plants, beds, and areas with different soil characteristics.

Avo-AirDB

Avo-AirDB (Amraoui et al., 2022) (Figure 24) dataset is considered a rich source of information, especially for smart agriculture.

Figure 24. Examples of Avo-AirDB dataset

A dataset of 984 aerial photographs of avocado fruits is available to the public at a ground resolution of 2.7 cm per pixel. It was collected from over 113 hectares of an avocado farm in Morocco using a DJI Phantom 4 Pro UAV. It contains original aerial images and annotated images.

Sentinel-2

Temporal datasets, such as those from Sentinel-2 (Ilyushin and Martirosyan, 2024), are datasets that include regular observations of agricultural fields throughout the growing season. Sentinel-2 provides high-frequency, moderate-resolution images, making it particularly useful for monitoring crop growth dynamics, assessing the effectiveness of agricultural practices, and forecasting crop yields. Temporal data allows researchers to analyze seasonal changes and assess the impact of external factors such as climate change or disease.

Despite the availability of a large number of different datasets, researchers face a number of challenges when using them. One of the main problems is data heterogeneity due to differences in shooting conditions, image resolution, and annotation types. In addition, successful training of deep learning models requires a large amount of labeled data, which can be a labor-intensive and expensive process. It is also worth noting the need for data standardization and accessibility. Many existing datasets have unique formats and annotation schemes, which complicates their sharing and integration. Finally, data collected in one region or for one crop type may not be suitable for use in other settings, limiting its generality.

Evaluation Criteria for Segmentation Methods

There are several key metrics for assessing the quality of image segmentation, each reflecting different aspects of the accuracy and efficiency of the method. These metrics allow objective comparison of different algorithms and their applicability in different conditions, including agriculture.

Pixel Accuracy (PA) measures the proportion of pixels that were correctly classified by the model. The formula for calculating Pixel Accuracy is given below:

$$PA = \frac{\sum_{i=0}^{K} p_{ii}}{\sum_{i=0}^{K}\sum_{j=0}^{K} p_{ij}}$$

where p_{ii} — number of pixels correctly classified as class i;
p_{ij}— number of pixels belonging to class i but classified as class j.

2) Mean Pixel Accuracy (MPA) is the average accuracy across all classes, which allows to take into account the imbalance in the data between different classes:

$$MPA = \frac{1}{K+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k} p_{ij}}$$

where K — number of classes including background.

3) Intersection over Union (IoU), also known as the Jaccard index, is one of the most common metrics for assessing the quality of segmentation. It is calculated as the ratio of the area of intersection of the predicted segmentation region and the true region to the area of their union:

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

where A and B — true and predicted segmentation regions, respectively. The IoU value ranges from 0 to 1, where 1 corresponds to a complete match of the regions.

4) Mean Intersection over Union (mIoU) is the average IoU value across all classes:

$$mIoU = \frac{1}{K+1}\sum_{i=0}^{k} IoU_i$$

where IoU_i — IoU metric value for the i-th class.

5) Precision and Recall are classic metrics for assessing classification accuracy, including segmentation tasks. The formulas for calculating them are given below:

$$Precision = \frac{TP}{TP+FP}, \qquad Recall = \frac{TP}{TP+FN}$$

where TP — number of true positive predictions,
FP — number of false positive predictions,
FN — number of false negative predictions.

6) F1-Score is the harmonic mean of Precision and Recall, and is often used to evaluate models when both precision and recall are important to consider:

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision \cdot Recall}$$

There are several key metrics for assessing the quality of image segmentation, each reflecting different aspects of the accuracy and efficiency of the method. These metrics allow objective comparison of different algorithms and their applicability in different conditions, including agriculture.
Below is a table with the results of the evaluation of various classical segmentation methods used in agriculture.

Table 1. Results of evaluation of classical segmentation methods.

Metrics / Method	Pixel Accuracy	Mean Pixel Accuracy	IoU	mIoU	Precision and Recall	F1-Score
Threshold segmentation (Akhmetshin et al., 2023)	65%	0.60	0.45	0.40	60%, 55%	57%
Highlighting borders (BenHajyoussef and Saidani, 2024)	70%	0.65	0.50	0.45	65%, 60%	62%
Division and merger of regions (Bins et al., 1996)	72%	0.68	0.55	0.50	70%, 65%	67%
K-means clustering (Reckling and Grosse, 2022)	75%	0.70	0.60	0.55	75%, 70%	72%
Superpixels (Li et al., 2020)	78%	0.72	0.62	0.58	77%, 75%	76%
Watershed method (Xue et al., 2021)	73%	0.68	0.58	0.52	72%, 68%	70%
Spectral segmentation (Moghaddam et al., 2020)	80%	0.75	0.65	0.60	80%, 75%	77%
Clustering based on cuts (Belim and Belim, 2022)	76%	0.70	0.60	0.55	75%, 70%	72%
Gabor filtering (Akhmetshin et al., 2024)	74%	0.68	0.58	0.53	72%, 68%	70%
Markov's Algorithm (Zheng et al., 2019)	82%	0.78	0.68	0.63	80%, 78%	79%

This table allows you to compare the effectiveness of each method according to the main metrics for assessing the quality of segmentation.

Table 2. Evaluation results of segmentation methods using deep learning.

Metrics / Method	Pixel Accuracy	Mean Pixel Accuracy	IoU	mIoU	Precision and Recall	F1-Score
CNN (Kamilaris and Prenafeta-Boldú, 2018)	88.5%	0.72	0.68	0.74	85.0%, 84.5%	84.31%
U-Net (Dimitrovski et al., 2024)	92.0%	0.73	0.68	0.74	95.0%, 94.9%	94.96%
SegNet (Borodulin and Maksimenko, 2023)	89.8%	0.71	0.71	0.72	87.8%, 87,8%	87.85%
DeepLab (Kradetskaya et al., 2024)	96.0%	0.94	0.94	0.81	97.0%, 96.9%	96.98%
PSPNet (Chen et al., 2021)	97.0%	0.97	0.97	0.77	98.0%, 97.5%	97.01%
Mask R-CNN (Huang et al., 2024)	93.0%	0.75	0.75	0.75	95.0%, 94.0%	94.52%
Autoencoder (Mujkic et al., 2022)	88.0%	0.81	0.74	0.73	87.8%, 87.5%	83.0%
GAN (Zhuk, 2023)	87.5%	0.70	0.74	0.73	85.0%, 84.0%	84.5%
CNN + CRF (Abdullaev et al., 2023)	89.0%	0.71	0.71	0.60	85.0%, 83.5%	84.0%
LSTM (Sagynbayeva et al., 2023)	91.0%	0.82	0.90	0.85	95.1%, 94.5%	94.1%
CNN + Active Contours (Kharitonova, 2023)	90.0%	0.77	0.75	0.75	89.0%, 88.5%	88.8%

The following table presents the evaluation results of state-of-the-art segmentation methods based on deep neural networks. These methods have shown high performance in challenging agricultural image segmentation tasks.

CONCLUSION

- In conclusion, it can be noted that image segmentation technologies are already deeply integrated into modern agricultural management methods. The variety of methods, from simple threshold approaches to complex deep learning models, reflects the desire of researchers and developers to find the most effective tools for analyzing agricultural data. Each of the segmentation methods considered offers its own unique advantages, allowing to solve a wide range of problems related to crop monitoring, field boundary detection, and weed and pest identification. This review highlights the importance of further development and adaptation of these technologies to changing agricultural conditions to ensure sustainable and efficient resource management in the future.
- A review of image segmentation methods in agriculture showed that classical methods, although easy to implement, have limited effectiveness in complex agricultural landscapes. Their accuracy rates, as can be seen from Table 1, are at the level of 65–82% according to the Pixel Accuracy metric.
- Modern methods based on deep neural networks demonstrate significantly higher results. According to Table 2, models such as U-Net, DeepLab, and PSPNet achieve accuracy of over 90%, which significantly exceeds classical approaches.
- Thus, deep learning methods provide high accuracy and adaptability in segmenting agricultural images, making them preferable for practical applications. However, their effective use requires a large amount of labeled data and significant computing resources.

Acknowledgements. The work was carried out within the framework of the project "Mathematical modeling and algorithms for modeling plant growth based on an automated mapping system" (FZNE-2024-0006).

REFERENCES

1. Abdullaev I, Prodanova N, Bhaskar KA, Lydia EL, Kadry S, Kim J, (2023). Task offloading and resource allocation in iot based mobile edge computing using deep learning, Computers, Materials & Continua 76(2), 1463-1477. https://doi.org/10.32604/cmc.2023.038417;
2. Akhmetshin E, Nemtsev A, Shichiyakh R, Shakhov D, Dedkova I, (2024). Evolutionary algorithm with deep learning based fall detection on Internet of Things environment, Fusion: Practice and Applications 14(2), 132-145. http://dx.doi.org/10.54216/FPA.140211;
3. Akhmetshin E, Sultanova S, Anupama CSS, Kumar KV, Lydia EL, (2023). Surveillance video-based object detection by feature extraction and classification using deep learning architecture. In: Intelligent Data Engineering and Analytics (Ed. by V. Bhateja, F. Carroll, J.M.R.S. Tavares, S.S. Sengar and P. Peer), pp. 369-378. Springer, Singapore. http://dx.doi.org/10.1007/978-981-99-6706-3_32;
4. Amraoui KE, Lghoul M, Ezzaki A, Masmoudi L, Hadri M, Elbelrhiti H, Simo AA, (2022). Avo-AirDB: An avocado UAV Database for agricultural image segmentation and classification, Data in Brief 45, 108738. http://dx.doi.org/10.1016/j.dib.2022.108738;
5. Anand T, Sinha S, Mandal M, Chamola V, Yu F, (2021). AgriSegNet: Deep aerial semantic segmentation framework for IoT-assisted precision agriculture, IEEE Sensors Journal 21(16), 17581-17590. http://dx.doi.org/10.1109/JSEN.2021.3071290;
6. Belim SV, Belim SY, (2022). Images segmentation based on cutting the graph into communities, Algorithms 15(9), 312. http://dx.doi.org/10.3390/a15090312;
7. BenHajyoussef A, Saidani A, (2024). Recent advances on image edge detection. In: Digital Image Processing - Latest Advances and Applications (Ed. by F. Cuevas, P.L. Mazzeo and A. Bruno). IntechOpen. http://dx.doi.org/10.5772/intechopen.1003763;
8. Bins LSA, Fonseca LG, Erthal GJ, Ii FM, (1996). Satellite imagery segmentation: A region growing approach, Simpósio Brasileiro de Sensoriamento Remoto 8(1996), 677-680;

9. Borodulin D, Maksimenko A, (2023). Enhancing reforestation efforts: A comprehensive analysis of modern approaches to cultivate planting material for forest crops, Journal of Global Innovations in Agricultural Sciences 11, 499-506. https://doi.org/10.22194/JGIAS/11.1180;
10. Bugubaeva A, Chashkov V, Mamikhin S, Kuprijanov A, Kuanyshbaev S, Nugmanov A, Bulaev A, Shcheglov A, Manakhov D, Zharlygasov Z, Isakaev Y, Uxikbayeva M, Badawy W, Joldassov A, (2023). Assessment of the state of vegetation cover of recultivated dumps of uranium deposits in Northern Kazakhstan, Brazilian Journal of Biology 83, e279616. https://doi.org/10.1590/1519-6984.279616;
11. Chen S, Song Y, Su J, Fang Y, Shen L, Mi Z, Su B, (2021). Segmentation of field grape bunches via an improved pyramid scene parsing network, International Journal of Agricultural and Biological Engineering 14(6), 185-194. http://dx.doi.org/10.25165/j.ijabe.20211406.6903;
12. Dechesne C, Mallet C, Le Bris A, Gouet-Brunet V, (2017). Semantic segmentation of forest stands of pure species as a global optimization problem, ISPRS Annals of the Photogrammetry Remote Sensing and Spatial Information Sences IV-1/W1, 141-148. http://dx.doi.org/10.5194/isprs-annals-IV-1-W1-141-2017;
13. Dimitrovski I, Spasev V, Loshkovska S, Kitanovski I, (2024). U-Net ensemble for enhanced semantic segmentation in remote sensing imagery, Remote Sensing 16(12), 2077. http://dx.doi.org/10.3390/rs16122077;
14. Du S, Hashim N, Kamarudin S, (2024). Unveiling the technology acceptance model: A systematic review of evolution, impact across industries and future directions, Pakistan Journal of Life and Social Sciences 22(1), 1220-1240. https://doi.org/10.57239/PJLSS-2024-22.1.0083;
15. Gorelkina AK, Timoshchuk IV, Mikhaylova ES and Neverov EN, (2024). Monitoring the state of waterbodies and taking into account anthropogenic impact, International Journal of Agriculture and Biosciences 13(3), 444-448. https://doi.org/10.47278/journal.ijab/2024.143;
16. Huang X, Peng D, Qi H, Zhou L, Zhang C, (2024). Detection and instance segmentation of grape clusters in orchard environments using an improved mask R-CNN model, Agriculture 14(6), 918. http://dx.doi.org/10.3390/agriculture14060918;
17. Ilyushin Y, Afanaseva O, (2020). Modeling of a spatial distributed management system of a preliminary hydro-cleaning gasoline steam column. In: International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM, 2020-August (2.1), pp. 531-538. STEF92 Technology, Sofia. https://doi.org/10.5593/sgem2020/2.1/s08.068;
18. Ilyushin Y, Martirosyan A, (2024). The development of the oderberg electrolyzer electromagnetic field’s state monitoring system, Scientific Reports 14, 3501. https://doi.org/10.1038/s41598-024-52002-w;
19. Ismailova E, Shemshura O, Sadanova A, Baimakhanova G, Turlybayeva Z, Kuldybayev N, Yelubayeva A, Kopzhassarov B, Issina Zh, Temreshev I, Beknazarova Z, (2024). Monitoring studies of the occurrence of fire blight pathogen in Kazakhstan and identification of antagonistic microorganisms suppressing its development, Brazilian Journal of Biology 84, e285493. https://doi.org/10.1590/1519-6984.285493;
20. Kamilaris A, Prenafeta-Boldú FX, (2018). A review of the use of convolutional neural networks in agriculture, The Journal of Agricultural Science 156(3), 312-322. http://dx.doi.org/10.1017/S0021859618000436;
21. Kharitonova YuS, (2023). Legal means of providing the principle of transparency of the artificial intelligence, Journal of Digital Technologies and Law 1(2), 337-358. https://doi.org/10.21202/jdtl.2023.14;
22. Korotenko V, Togusakov O, (2024). Ecological consciousness and sustainable development: Bridging environmental education with green economy practices, International Journal of Ecosystems and Ecology Science 14(4), 17-24;
23. Kukusheva A, Kakezhanova Z, Uakhitov Z, Sarbasov A, Nasiyev B, (2024). Role of tillage and bio-humus in the restoration of degraded pastures of Agropyrum pectiniforme in the steppe zone of northeast Kazakhstan under changing climatic conditions, International Journal of Agriculture and Biosciences 13(4), 560-573. https://doi.org/10.47278/journal.ijab/2024.161;
24. Kradetskaya O, Dzhazina D, Kairzhanov Y, Dashkevich S, Chilimova I, Utebayev M, Shelayeva T, (2024). Amino acid composition of proteins and technological indicators of grain quality of genotypes and varieties of spring soft wheat in northern Kazakhstan, Journal of Global Innovations in Agricultural Sciences 12, 197-206. https://doi.org/10.22194/JGIAS/12.12211;
25. Kussul N, Lavreniuk M, Skakun S, Shelestov A, (2017). Deep learning classification of land cover and crop types using remote sensing data, IEEE Geoscience and Remote Sensing Letters 14(5), 778-782. http://dx.doi.org/10.1109/LGRS.2017.2681128;

26. Li C, Guo B, Wang G, Zheng Y, Liu Y, He W, (2020). NICE: Superpixel segmentation using non-iterative clustering with efficiency, Applied Sciences 10(12), 4415. http://dx.doi.org/10.3390/app10124415;
27. Lyu Y, Vosselman G, Xia GS, Yilmaz A, Yang MY, (2020). UAVid: A semantic segmentation dataset for UAV imagery, ISPRS Journal of Photogrammetry and Remote Sensing 165, 108-119. http://dx.doi.org/10.1016/j.isprsjprs.2020.05.009;
28. Moghaddam SHA, Mokhtarzade M, Beirami BA, (2020). A feature extraction method based on spectral segmentation and integration of hyperspectral images, International Journal of Applied Earth Observation and Geoinformation 89, 102097. http://dx.doi.org/10.1016/j.jag.2020.102097;
29. Mujkic E, Philipsen MP, Moeslund TB, Christiansen MP, Ravn O, (2022). Anomaly detection for agricultural vehicles using autoencoders, Sensors 22(10), 3608. http://dx.doi.org/10.3390/s22103608;
30. Osadchy E, Abdullayev I, Bakhvalov S, Klochko E, Tagibova A, (2024). Jellyfish search algorithm based feature selection with optimal deep learning for predicting financial crises in the economy and society, Fusion: Practice and Applications 14(2), 186-198. http://dx.doi.org/10.54216/FPA.140215;
31. Qiao Y, Truman M, Sukkarieh S, (2019). Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming, Computers and Electronics in Agriculture 165, 104958. http://dx.doi.org/10.1016/j.compag.2019.104958;
32. Rafikov T, Zhumatayeva Z, Mukaliyev Z, Zhildikbayeva A, (2024). Evaluating land degradation in East Kazakhstan using NDVI and Landsat data, International Journal of Design & Nature and Ecodynamics 19(5), 1677-1686. https://doi.org/10.18280/ijdne.190521;
33. Reckling M, Grosse M, (2022). On-farm research to diversify organic farming systems, Organic Farming 8(1), 1-2. https://doi.org/10.12924/of2022.08010001;
34. Riehle D, Reiser D, Griepentrog H, (2020). Robust index-based semantic plant/background segmentation for RGB-images, Computers and Electronics in Agriculture 169, 105201. http://dx.doi.org/10.1016/j.compag.2019.105201;
35. Ronneberger O, Fischer P, Brox T, (2015). U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (Ed. by N. Navab, J. Hornegger, W. Wells and A. Frangi), pp. 234-241. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28;
36. Roy DP, Wulder MA, Loveland TR, Woodcock CE, Allen RG, Anderson MC, Helder D, Irons JR, Johnson DM, Kennedy R, Scambos TA, Schaaf CB, Schott JR, Sheng Y, Vermote EF, Belward AS, Bindschadler R, Cohen WB, Gao F, Hipple JD, Hostert P, Huntington J, Justice CO, Kilic A, Kovalskyy V, Lee ZP, Lymburner L, Masek JG, McCorkel J, Shuai Y, Trezza R, Vogelmann J, Wynne RH, Zhu Z, (2014). Landsat-8: Science and product vision for terrestrial global change research, Remote Sensing of Environment 145, 154-172. http://dx.doi.org/10.1016/j.rse.2014.02.001;
37. Sagynbayeva A, Mambetov B, Kalachev A, Boranbay Z, Borissenko Y, Baigazakova Z, Turlybekov K, (2023). Assessing pine forest restoration post-fire using NDVI and GIS technologies, International Journal of Design & Nature and Ecodynamics 18(6), 1459-1467. https://doi.org/10.18280/ijdne.180620;
38. Schmitt M, Hughes L, Ghamisi P, Yokoya N, Hänsch R, (2019). 2020 IEEE GRSS Data Fusion Contest, IEEE Dataport. https://dx.doi.org/10.21227/rha7-m332;
39. Utomowati R, Suranto S, Suntoro S, Muryani C, (2024). Assessment of food supply ecosystem services and water supply ecosystem services to optimise sustainable land use planning in Samin watershed, Central Java, Indonesia, International Journal of Sustainable Development and Planning 19(9), 3577-3586. https://doi.org/10.18280/ijsdp.190925
40. Xhina E, Bilo I, Ktona A, Paparisto A, Liçi O, Qirinxhi X, Haxhiu D, (2023). Machine learning methods for understanding biodiversity and its conservation, International Journal of Ecosystems and Ecology Science 13(3), 39-44. https://doi.org/10.31407/ijees13.306;
41. Xue Y, Zhao J, Zhang M, (2021). A watershed-segmentation-based improved algorithm for extracting cultivated land boundaries, Remote Sensing 13(5), 939. http://dx.doi.org/10.3390/rs13050939;
42. Yang W, Wang S, Zhao X, Zhang J, Feng J, (2015). Greenness identification based on HSV decision tree, Information Processing in Agriculture 2(3-4), 149-160. http://dx.doi.org/10.1016/j.inpa.2015.07.003;
43. Yuan Y, Chen L, Wu H, Li L, (2022). Advanced agricultural disease image recognition technologies: A review, Information Processing in Agriculture 9(1), 48-59. http://dx.doi.org/10.1016/j.inpa.2021.01.003;
44. Zhang J, Rao Y, Man C, Jiang Z, Li S, (2021). Identification of cucumber leaf diseases using deep learning and small sample size for agricultural Internet of Things, International Journal of Distributed Sensor Networks 17(4), 1-13. http://dx.doi.org/10.1177/15501477211007407;

45. Zheng C, Pan X, Chen X, Yang X, Xin X, Su L, (2019). An object-based Markov random field model with anisotropic penalty for semantic segmentation of high spatial resolution remote sensing imagery, Remote Sensing 11(23), 2878. http://dx.doi.org/10.3390/rs11232878;
46. Zhuk A, (2023). Artificial intelligence impact on the environment: Hidden ecological costs and ethical-legal issues, Journal of Digital Technologies and Law 1(4), 932-954. https://doi.org/10.21202/jdtl.2023.40;
47. Zou K, Chen X, Wang Y, Zhang C, Zhang F, (2021). A modified U-Net with a specific data argumentation method for semantic segmentation of weed images in the field, Computers and Electronics in Agriculture 187, 106242. http://dx.doi.org/10.1016/j.compag.2021.106242;