Modelling a neural network for analysing the results of segmentation of satellite images

Mira Kaldarova1, Akerke Akanova1, Akgul Naizagarayeva2, Albina Kazanbayeva1, Nazira Ospanova3
1Department of Computer Science, S. Seifullin Kazakh Agrotechnical Research University, Astana, Republic of Kazakhstan
2Department of Information Systems, S. Seifullin Kazakh Agrotechnical Research University, Astana, Republic of Kazakhstan
3Faculty of Computer Science, Toraighyrov University, Pavlodar, Republic of Kazakhstan

ABSTRACT
The study’s relevance lies in addressing inaccuracies within satellite image segmentation, necessitating the development and implementation of neural network models for automated segmentation. The purpose of study is to develop a model of a neural network for training with data obtained from the segmentation of satellite images. The basis of the methodological approach in study is a combination of methods of system analysis of neural networks, which have had a substantial impact on the development of the computer vision industry, with an empirical study of the general principles of neural network modelling for the training on satellite images segmentation. In this study, the results were obtained, indicating that there is a fundamental possibility of developing and practical implementation of a neural network model to determine the quality of the obtained segmentation of images of agricultural fields. Satellite images of agricultural fields of the Republic of Kazakhstan are obtained, and segmentation of field images is performed using the developed neural network model for learning segmentation results. The practical importance of the results obtained in study lies in the possibility of their use in the development of functional models of neural networks for training the results of the segmentation of satellite images.

Keywords: Automatic image, Computer vision, Machine learning, Processing algorithms, Remote sensing

This is an open access article under the CC BY-SA license.

Corresponding Author:
Akerke Akanova
Department of Computer Science, S. Seifullin Kazakh Agrotechnical Research University
010011 Astana, Republic of Kazakhstan
Email:akanova.akerke@outlook.com

1. INTRODUCTION
This study addresses the challenge of developing and implementing neural network methods for satellite image segmentation that can meet reliability and quality standards. Despite advancements in algorithms, they often fall short, prompting the exploration of alternative segmentation approaches. Object segmentation through neural networks is a rapidly advancing field crucial for various applications, from agriculture to medical diagnostics. However, ensuring consistent quality and reliability remains a significant hurdle, necessitating continuous improvement in neural network modelling for segmentation training [1].

In turn, Rios et al. [2] conducted a joint study aimed at investigating the basics of modelling and control of neural networks. Researchers pay attention to the fact that the construction of neural network models for training the results of the segmentation of satellite images involves the creation of a mathematical model of the system, which is a differential equation that considers the key parameters of the dynamic system in real-time. Its solution allows for identifying the criteria for segmenting satellite images according to the specified parameters. Neupane et al. [3] aimed at investigating the principles of semantic segmentation of urban objects on satellite images based on deep learning, they note that the consistent development of deep
machine learning methods contributes to a shift in the paradigm of image systematisation from pixel-based methods and objects to semantic segmentation based on deep learning. According to researchers, such methods provide higher accuracy of satellite image processing compared to conventional ones.

For their part, Hua et al. [4] in a joint study aimed at investigating the general principles of segmentation of remote sensing images, pay attention to the fact that to train neural networks for high-resolution images, a substantial number of high-quality annotations on a pixel basis are needed. The introduction of special methods for controlling the task with a learning signal that considers special spatial terms is required to obtain them. A group of Guerin et al. [5] in a scientific paper aimed at investigating the semantic segmentation of satellite images, note that in recent years geodata, such as satellite and digital photographs, have become available in large quantities, and they are extremely rarely qualitatively segmented, which makes it difficult to use them. For this reason, the effective segmentation of satellite images is of great importance for creating opportunities for their subsequent practical use. Therewith, Yang et al. [6] in a study of bridge extraction algorithms based on deep learning and high-resolution satellite images concluded that the creation of a neural network model based on deep learning for the segmentation of satellite images contributes to the clear identification of objects in the resulting image.

The main purpose of this study is to evaluate the possibilities of creating a neural network model for training data determined during the segmentation of satellite images. The main objectives of this study are: the analysis of algorithms for processing and converting data obtained during the segmentation of agricultural field images and the development of a neural network model to determine the quality of the obtained segmentation of these images. The relevance of the problem under study, as already mentioned above, is caused by a substantial number of inaccuracies in the data obtained through the segmentation of satellite images and the associated need to create and then implement neural network models for training automatic segmentation of satellite image results.

2. MATERIALS AND METHODS

This study employed a methodological approach combining system analysis of neural networks with empirical observation of satellite images from fields in Kazakhstan. The empirical investigation involved segmenting these images through experimental deep learning neural network modelling. The theoretical framework drew from an analysis of various papers addressing issues in neural network modelling for satellite image segmentation. System analysis of neural network usage experiences greatly influenced computer vision enhancement, leading to the identification of specific deep neural network architectures. This facilitated understanding the fundamental principles governing the functioning of these networks, including construction features and accuracy parameters. The empirical study focused on developing neural networks for semantic segmentation of satellite images, enabling precise classification and division of objects within the images. Semantic classification included a sequence of programme operations from three main steps: i) initially, all classes present in the processed image are selected; ii) the next step is segmentation/detection, which allows identifying not only the entities of classes but also information about the spatial location of all the determined classes; and iii) in the last step, an accurate conclusion is obtained. Separate labels are created for each specific pixel. These labels indicate that a certain pixel belongs to the established entity of the class.

The study introduces a segmentation process for satellite images of the fields in the Republic of Kazakhstan, obtained from remote satellites. Unlike previous studies focused on Pokrovka village in the North Kazakh region, this research broadens its scope to cover various regions of the country. The segmentation process involves several sequential operations, categorizing images into seven distinct classes: urban land, agricultural land, pastures, forests, rivers and lakes, wastelands, and unknown areas (e.g., cloud, fog). Utilizing a convolutional neural network (CNN) with the advanced DeepLabV3 plus structure, the model operates on high-resolution images with predictions matching the input data size of 2,448×2,448 pixels. To optimize computational costs without significantly compromising accuracy, the input image size is adjusted to 1,000×1,000 pixels before prediction. Additionally, the server model is packaged into an executable file for ease of deployment, and a graphical user interface using the QT framework enhances user interaction and usability.

In this neural network, the GrabCut image segmentation method was used, which is based on graph sections. Starting from the user-defined bounding box, which is located around the segmented object, this algorithm evaluates the distribution of the colour gamut of the object under study, based on a mixed Gaussian model. This is necessary to build a random field based on pixel labels with a function that highlights certain areas marked with identical labels and run optimisation based on a graph section to output values. Due to the fact that this estimate will have a higher degree of accuracy compared to the initial one obtained from the bounding box, if such a need arises, this procedure can be repeated until convergence is achieved.

Modelling a neural network for analysing the results of … (Mira Kaldarova)
3. RESULTS

It is worth giving some examples of standard deep neural network architectures that have made a substantial contribution to the field of computer vision since they are often used as the basis of semantic segmentation systems:

i) AlexNet: CNN architecture created in Toronto, which won the ImageNet 2012 contest with an accuracy of 84.6%. It consists of five layers: 3 ultra-precise layers using the rectified linear unit (ReLU) activation function for non-linearity, one layer of subdiscretisation, and one layer of reset. It is an ultra-high precision neural network [7].

ii) Visual geometry group (VGG)-16: this neural network model was developed at Oxford University and won the ImageNet 2013 contest with 92.7% accuracy. It uses a stack of convolution layers with small receptive fields in the first layers instead of multiple layers with large receptive fields. It is an improved version of AlexNet, in which large filters (size 11 and 5 in the first and second ultra-precise layer, respectively) were replaced with several 3x3 filters alternating with each other. The VGG16 network was trained for several weeks, using NVIDIA TITAN BLACK graphic cards [8].

iii) GoogLeNet: this network from Google became the winner of the ImageNet 2014 contest with an accuracy of 93.3%. It consists of 22 layers and a new building block of neural networks, called the initial module. The initial module consists of: a network connection layer, a subdiscretisation layer, a large-size convolution layer, and a small-size convolution layer. GoogLeNet is a kind of ultra-high precision neural network based on the Inception architecture. The use of initial modules allows making a choice between several sizes of ultra-high precision filters in each block. The initial network places these modules on top of each other, from time to time combining layers with the maximum pool in increments of 2, which allows for halving the grid resolution (GoogLeNet, 2022).

iv) ResNet: this model is from Microsoft, which won the ImageNet 2016 contest with 96.4% accuracy. It is well known for its depth (152 layers) and the introduction of residual blocks. Residual blocks solve the problem of learning a really deep architecture by introducing connections with missing identifiers so that layers can copy their input data to the next level.

Satellite image segmentation involves dividing images into pixel groups representing specific objects while determining object types [9]-[11]. This process ensures homogeneous areas within images, with key factors like textural, spectral characteristics, shape, size, and context influencing segmentation quality. Neural networks, functioning as classifiers, analyze pixel values and perform tasks like object recognition and semantic segmentation. Deep learning techniques require images of sufficient resolution for effective training and verification, enabling recognition of specific patterns. For tasks like windfall detection, images covering tens of meters with resolutions of 256x256 pixels are typically adequate for accurate recognition. The augmentation process in deep learning reduces overfitting risk and determines the number of images used in training [12]. It involves various image transformations like color adjustment and zooming. Conventional machine learning methods, lacking specialized feature extraction algorithms, tend to perform worse since they don’t consider neighboring pixels. In the study, satellite images of North Kazakhstan fields were used, acquired via remote satellite survey as shown in Figure 1.

![Satellite image of the fields of the Republic of Kazakhstan](image1.png)

Figure 1. Satellite image of the fields of the Republic of Kazakhstan [13]
After processing satellite images, the developed neural network model was utilized to obtain the segmentation of field images and assess the final quality parameters of the segmentation. The segmentation methods employed enable the differentiation of classes corresponding to various types of objects, both natural and anthropogenic. Typically, empirical verification confirms any information regarding the number of classes and their probabilistic characteristics, without the need for classification of training samples [14]. Semantic segmentation is the process of assigning a semantic label to coherent areas of an image, which may encompass pixels, sub-pixels, super-pixels, or specific image fragments. Pixel-by-pixel segmentation involves assigning a label to each pixel in high-resolution images or assigning class membership to lower-resolution images where individual objects cannot be adequately represented. Various parametric and nonparametric classifiers, including artificial neural networks (ANNs), support vector machines (SVMs), decision tree classifiers, and expert systems, are employed for pixel-by-pixel segmentation [15]. In this study, two primary objectives were outlined: the processing and transformation of data obtained from the segmentation of agricultural field images, and the development of a neural network model to evaluate the quality of the resulting segmentation of these images.

A number of features of the proposed solutions contribute to obtaining advantages. In particular, in this neural network, the GrabCut image segmentation method was used, which is based on graph sections, which is why this estimate will have a higher degree of accuracy. Therewith, if such a need arises, this procedure can be repeated until convergence is achieved. The advantages of this study should be considered the positive dynamics of increasing the accuracy and reducing error parameters. In addition, a holistic neural network model was developed for training as shown in Figure 2.

In the course of training, the dropout layer is sequentially enabled. Figure 3 shows the approximate results of neural network training. According to the results of the neural network training, the error was 0.25780 (Figure 3(a)), the accuracy was 0.8213 (Figure 3(b)). Figure 4 shows the features of satellite image processing and the choice of segmentation methods and sequences.

When training data is unavailable, employing data clustering algorithms is a potent method for processing satellite images. Clustering categorizes objects into distinct clusters based on similarity, although it requires specifying the cluster count and struggles with complex shapes [1]. In contrast, deep learning architectures, such as fully convolutional networks (FCNs), enhance computational performance and accuracy by increasing network depth. FCNs, leveraging CNNs, enable end-to-end segmentation of images, yielding higher accuracy than traditional methods, especially in satellite image analysis [3]. Ongoing refinements of FCN architecture are crucial to address specific challenges in satellite image segmentation. The key experiments that must be conducted involve further exploration of deep learning techniques in satellite image segmentation, particularly focusing on refining existing architectures and developing novel approaches to address specific challenges, such as handling complex shapes and improving segmentation accuracy. Investigating the transferability of deep learning models trained on one geographic region to others could provide valuable insights into the generalizability of these approaches across different contexts.

The significance of deep learning methodologies like CNNs and FCNs in semantic segmentation tasks, particularly in satellite image analysis, is highlighted. These methods show promise in improving the
precision and speed of image processing. The discussion emphasizes the need for ongoing research and refinement of deep learning architectures to advance satellite image analysis. The results underscore the effectiveness of deep neural networks in addressing semantic segmentation challenges, suggesting avenues for future research to explore refining existing models and developing novel approaches. Investigating the transferability of deep learning models across geographic regions could offer valuable insights into their applicability in varied contexts.

Figure 3. Neural network training results (a) the error and (b) the accuracy

Figure 4. Features of satellite image processing and selection of segmentation methods and algorithms [1]
4. DISCUSSION

The practical application of deep learning neural network models for satellite image segmentation differs significantly from conventional pixel segmentation methods. In this approach, all images processed for training and verification must possess the requisite resolution to facilitate subsequent analysis effectively. Deep learning algorithms excel in discerning patterns within objects or geographic areas, necessitating a diverse array of image sizes to capture predetermined characteristics accurately. The developed neural network model aimed at assessing the quality of agricultural field image segmentation in the Republic of Kazakhstan has showcased remarkable accuracy and minimal error parameters.

Kislov et al. [12] considered the problems of extending deep learning approaches for segmentation of forest disturbances on very high-resolution satellite images. According to the researchers, accurate detection of forest disturbances is a challenging problem in the field of environmental monitoring. It is noted that progress in this issue has been made due to the rapid development of remote sensing devices based on neural networks with complex architectures and deep learning principles. Thus, in the works of Arora et al. [16], Kroner et al. [17] jointly considered a number of problematic aspects of road segmentation based on satellite images. The authors note that modeling of neural networks for training satellite image segmentation allows obtaining remotely sensed images with higher resolution and the ability to cover a larger geographical region. The obtained conclusions do not contradict the results of this study, at the same time, the issue of problems with the accuracy of the obtained images remains open.

In turn, Liu and Cao [18], Chang and Leung [19] having considered the general principles of constructing an improved two-scale residual network for ultra-high-resolution images, noted that in recent years considerable experience has been accumulated in the practical application of neural networks for single-image super-resolution (SISR) tasks when segmenting satellite images. According to researchers, modelling neural networks for learning the results of segmentation of satellite images provides wide opportunities in improving the quality of resolution of images obtained at a considerable distance. This study agrees with the researchers’ conclusions, however, the need to achieve high accuracy of segmentation of satellite images is notable.

Wei et al. [20] discuss challenges in assessing image quality without a reference, proposing a meta-learning approach based on optimization. They highlight the importance of this issue in remote satellite sensing within computer vision. Recent advancements in deep learning have enabled successful image quality assessment from satellite imagery, reflecting the growing ubiquity of digital images in various domains. Huang et al. [21] explore object recognition principles in satellite image segmentation using deep learning. They emphasize two crucial stages in remote object recognition: feature extraction and classifier development. Neural network models effectively address these tasks, as validated by their study. However, they stress the need to tailor feature extraction and classifier development to specific conditions in each case.

In turn, Wang et al. [22], Ganakwar [23] in a joint study of the use of neural networks to improve image parameters for ultra-high resolution of uncontrolled remote sensing, note that an effective solution to the problem of creating ultra-high resolution requires testing a variety of approaches based on training pairs of images with resolutions of varying degrees of clarity. Therewith, neural networks provide high efficiency in the process of creating remote images and their segmentation by training the results of segmentation of this kind [24], [25]. The conclusions of the researchers are fully confirmed by the results obtained in this study. Notably, the effectiveness of the segmentation of remote images should be monitored in almost every case. The analysis of the results derived from this study, juxtaposed with findings from related research on neural network utilization for satellite image segmentation training, underscores their fundamental alignment with the key parameters assessed for comparative analysis. This alignment serves as compelling evidence of the robust scientific credibility and validity of the study's outcomes, affirming their potential for subsequent practical application in the development and implementation of neural network models trained on segmented satellite image data.

5. CONCLUSION

Processing and transformation of data obtained during the segmentation of images of agricultural fields in Kazakhstan allowed the development of a neural network model to determine the quality of the resulting image segmentation. This approach allowed obtaining a substantial increase in the accuracy of segmentation, in contrast to general methods, the practical application of which assumed the presence of a typical data set. This provides substantial prospects in the field of the practical application of neural networks for learning the results of satellite images in the future.

The advantage of this study is the achievement of a higher degree of data accuracy when segmenting satellite images. This is due to the automatization of the data processing obtained remotely from a satellite through the use of a deep learning neural network model. According to the results of the analysis of the neural network training, the error was 0.25780, the accuracy was 0.8213. This determines substantial
prospects for the subsequent practical application of neural network models for learning the results of the segmentation of satellite images when segmenting images of agricultural fields in Kazakhstan. The disadvantages of this study are the presence of probabilities of erroneous allocation of classes of objects during the segmentation of satellite images. Semantic segmentation involves assigning a semantic label (also known as a class) to each coherent area of the image. The prospects for further research in this area are determined by the possibility of improving the accuracy and quality of segmentation of images obtained by remote satellite imagery. This is of great importance in almost all fields of modern economic activity, in which there is a need to obtain accurate information about the state of natural and technological objects through the use of satellite photography of the terrain.

REFERENCES


BIOGRAPHIES OF AUTHORS

Mira Kaldarova is a Doctoral Student at the Department of Computer Science, S. Seifullin Kazakh Agrotechnical Research University. Her scientific interests are neural networks, machine learning, and automatic segmentation of satellite images. She can be contacted at email: mira_kaldarova@outlook.com.

Akerke Akanova is a Ph.D., Senior Lecturer at the Department of Computer Science, S. Seifullin Kazakh Agrotechnical Research University. Her research interests include improving satellite image segmentation with neural network models. She can be contacted at email: akanovaakerke65@gmail.com.

Akgul Naizagarayeva is a Researcher at the Department of Information Systems, S. Seifullin Kazakh Agrotechnical Research University. The processing and transformation of data obtained during image segmentation is a field of scientific interest. She can be contacted at email: Naizagarayeva91@proton.me.

Albina Kazanbayeva is a Ph.D., Senior Lecturer at the Department of Computer Science, S. Seifullin Kazakh Agrotechnical Research University. Her scientific interest lies in machine learning techniques for satellite image segmentation. She can be contacted at email: KazanbayevaAlbina@hotmail.com.

Nazira Ospanova is a Ph.D., Associate Professor at the Faculty of Computer Science, Toraighyrov University. Her scientific interest lies in the impact of neural networks on remote sensing data accuracy. She can be contacted at email: nazira.ospanova@proton.me.