

Technology of Selection and Recognition of Information Objects on Images of the Earth's Surface Based on Multi-Projection Analysis

Stepan Bilan, Vladyslav Hnatiienko, Oleh Ilarionov and Hanna Krasovska

Taras Shevchenko National University of Kyiv, Volodymyrska Street, 60, Kyiv, 01033, Ukraine

Abstract

The paper describes the technology for selecting and identifying objects on raster images of the earth's surface. A method for describing images using a set of projections obtained in different directions is presented. According to the proposed method, an initial scanning window with floating borders is formed. Projections are formed, the analysis of which made it possible to change the dimensions of the scanning window or to shift it in the appropriate direction until the object completely enters the window area. According to the generated projections, the geometric shape of the selected object is determined. In this case, the selection of edge pixels in the image is initially carried out, which makes it possible to improve the accuracy of identification. The method allows to increase the accuracy due to the formation of a larger number of projections in different directions. To do this, rotate the image itself at different angles is used. This approach allows you to determine the placement of the object and determine the presence of several objects on the window image in the case when in many projections the objects can be considered as one. The analysis of the density of distribution of ones pixels on the projections of the image of objects also made it possible to select several objects in the image. The use of threshold processing for the resulting projections made it possible to eliminate unnecessary areas that can be identified as a result of using the method. The method does not require large computational costs and is characterized by ease of analysis. Projection analysis allows feedback for effective correction of objects in the image and the image itself. Given the limited computing resources, the developed method is an attractive option for the implementation of systems for processing and recognizing satellite images.

Keywords ¹

Image, object selection, earth image, Radon transform, binary image projection, image edges

1. Introduction

Signals that are not informational for us for other people can carry a lot of information.

Currently, among all the tasks of image processing and recognition, one of the urgent tasks is the analysis of aerial photographs of the earth's surface. The analysis of images of the earth's surface is used in many areas of the national economy, such as: space, agriculture, military affairs, transport, geology, meteorology and others. One of the tasks in the analysis of images of the earth's surface is the task of selecting informative objects and their recognition. Particular attention is paid to improving the performance of operations and achieving real-time. This is relevant for aircraft that fly over the ground and carry out automatic analysis of the earth's surface. Various methods are used to solve this problem. [1-4]. A successful solution depends on the exact selection of the object in the image. Often the problem is assigning image elements to an object they do not belong to.

This distorts the result of recognition of the selected object in the image and leads to false solutions. In addition, images of selected objects, which have the same top view, but at the same time, have

International Scientific Symposium "Intelligent Solutions" (IntSol-2023), September 27-28, 2023, Kyiv-Uzhhorod, Ukraine

EMAIL: bstepan@ukr.net_(A.1); vladgnat1483@gmail.com_(A.2); oilarionov@gmail.com_(A.3); annavkrasovska@gmail.com_(A.4)

ORCID: 0000-0002-2978-5556; 0009-0000-2678-5158; 0000-0002-7435-3533; 0000-0003-1986-6130



© 2023 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

different heights of the object elements, have a significant impact on the recognition result. Existing methods for automatic selection of objects on the image of the earth's surface do not give the desired results. A particular challenge is to achieve high performance. Also, the solution of the problem depends on the choice of characteristic features, a correctly formed set of which gives high image quality. Since all objects in the image are viewed in a geometric projection from above, the recognition of such objects is considered as the recognition of geometric shapes and pictures on a two-dimensional plane. Methods aimed at solving such a problem use different approaches that do not use a single universal approach for all forms of images. Different characteristic features are used that require different measuring instruments and different computational models.

In this paper, we solve the problem of analyzing images of photographs of the earth's surface based on a multi-projection analysis of their binary representation. The multi-projection method allows you to automate the process of selecting local objects and form a homogeneous set of characteristic features that require unified algorithms for their processing.

2. Relative works

The problem of selecting and recognizing objects in images from images of the earth's surface has recently received much attention from specialists. Many works are devoted to the fractal analysis of images of the earth's surface [5, 6]. However, this approach requires significant resources. The methods mainly use reference images that are prepared in advance. In this situation, the resistance of informative features to various kinds of distortions is analyzed. The universal quality index is used as an indicator of durability [7]. Resistant to change are those areas for which the highest value of the universal quality index is provided. In [8], the image is divided into homogeneous segments of different shapes, which are determined by the characteristics of space and brightness, and the characteristics of the segments are compared with the reference segments, taking into account the value of the decision rule. Such methods are characterized by a large number of empirical parameters, which requires additional processing time. There are methods that use the gradient representation of the original image based on the Kirsch operator and automatically determine the clustering centers [9]. The use of the Kirsch operator limits the area of the analyzed neighborhood, which reduces the accuracy. Common methods that use artificial neural networks [10 - 12]. Such networks extract features and classify objects simultaneously. However, they cannot classify objects with different orientations, and also have low classification accuracy along with high performance. There are problems that are determined by the imposition of bounding boxes on the image. Therefore, methods based on the pyramid of attention are used [3, 13 - 15]. In such works, estimates are used, as well as a complex learning process, which affects the accuracy and speed of operations. In recent years, methods have been used that use the Radon transform and the hexagonal shape of the coverage to represent images [16]. Such methods allow you to select objects in the image with high accuracy based on the obtained Radon projections. However, they have a limitation due to the fact that modern computing systems use an orthogonal representation of graphic information. In this paper, we solve the problem of extracting informative objects based on two projections using an orthogonal representation of bitmaps.

3. Selection of objects based on multi-projection analysis

The exact selection of an object in a color image depends on the exact selection of its edges, which determine its geometric shape. To select edge pixels in an image, many well-known methods can be used [17 - 19], which give a different effect, but in general, edges are selected on all image objects.

Since it is very difficult to immediately determine the location of objects in the image in automatic mode, there is a need to search for methods without user intervention. One of these methods is the method using the Radon transform [16]. The Radon transform is used to solve many image processing problems in which there is a need to pre-select individual image elements. For a quadrangular mosaic (rectangular pixel shape), as a result of the Radon transformation, three projections are formed in the direction 0° , 45° and 90° . To solve most problems where there is a selection of objects of such a number of projections, it is enough. However, to improve the accuracy of the tasks being solved, six projections can be formed, obtained in the directions 0° , 30° , 60° , 90° , 120° and 150° . To obtain such projections without aliasing distortion, pixels with a hexagonal mosaic are used (pixels have the shape of an

equilateral hexagon), which can be modeled on a rectangular mosaic [16]. This made it possible to represent the image in a hexagonal mosaic. This representation of images and its advantages are described in the works [16. 20]. To select objects in images of the earth's surface, it is sufficient to analyze two projections in the directions 0° and 90° . However, this increases the time spent on solving the problem. You can also use one projection, but at the same time rotate the image by a given angle. The smaller the angle of rotation, the higher the accuracy of object selection. The selection of an object on the image of the earth's surface is carried out using the following sequence of operational steps.

1. An image of the earth's surface using aircraft or probes are formed.
2. With the help of the selected operator, edge pixels are selected in the image.
3. The number of selected pixels in the image by applying thresholding are reduced.
4. The resulting image is converted into a binary one, where the selected pixels correspond to the logical code "1", and the background pixels correspond to the logical code "0".
5. The resulting binary image is divided into regions of predetermined shapes and areas. As a rule, they are divided into four rectangular areas that have a common vertex in the center of the image.
6. For each selected area of the image, projections are formed in the direction of 0° and 90° .
7. The resulting projections are analyzed and, based on the analysis, a new shape of the selection window is determined.
8. The sixth and seventh steps are repeated until there is one selected object in the image in the window.
9. When the projections show that one object is selected in the generated window, the projections are analyzed and the geometric shape of the object is determined. The geometric shape of an object is determined by searching for the nearest reference set of projections.

Projection analysis can show that parts of object images are selected in the formed rectangular area. An example of such a situation in Figure 1 is shown.



Figure 1: An example of selecting individual parts of objects in an image

In this case, the system must decide in which direction to execute the algorithm. The selection window can move in the direction of the object, part of which fell into the field of the original window, or expand its border, which intersects the information object in the image. Both options are shown as an example in Figure 2. Resizing or moving the window is carried out until the projections show that the object is completely within the area of the analyzed window. In this case, the size of the object in the image may exceed the size of the analyzed window. Therefore, an attractive option is the option of expanding the boundaries (increasing the size of the window) towards the object, part of which fell into the original window. After the object is located in the field of the selection window, the analysis of the resulting projections shows which areas of the window need to be cut off. Thus, the window is first enlarged, and then reduced to the size of a rectangle covering the selected object. In accordance with this, the algorithm for generating the final window sizes consists of the following steps, which are expressed by the graph diagram of the algorithm (Fig. 3).

In fact, the method works in such a way that a feedback between projections and window sizes is implemented. If there are situations when the projections do not show a separately selected object in the image, i.e. projection overlays are present, then the image is rotated by a certain discrete angle. After each discrete rotation, new projections are formed and analyzed to identify the object and determine its geometric shape. Instead of rotation, it is possible to form different projections in parallel, which greatly speeds up the image processing process. On Figure 4 are shown examples of projections obtained as a result of rotating the image of the selected area. Figure 4 shows that with a rotation angle of 30° , two objects are selected, the geometric shape of which can be determined. It is also possible to remove a

small number of pixels on the projections, which will allow you to select a separate object in the total set of selected pixels in the image.

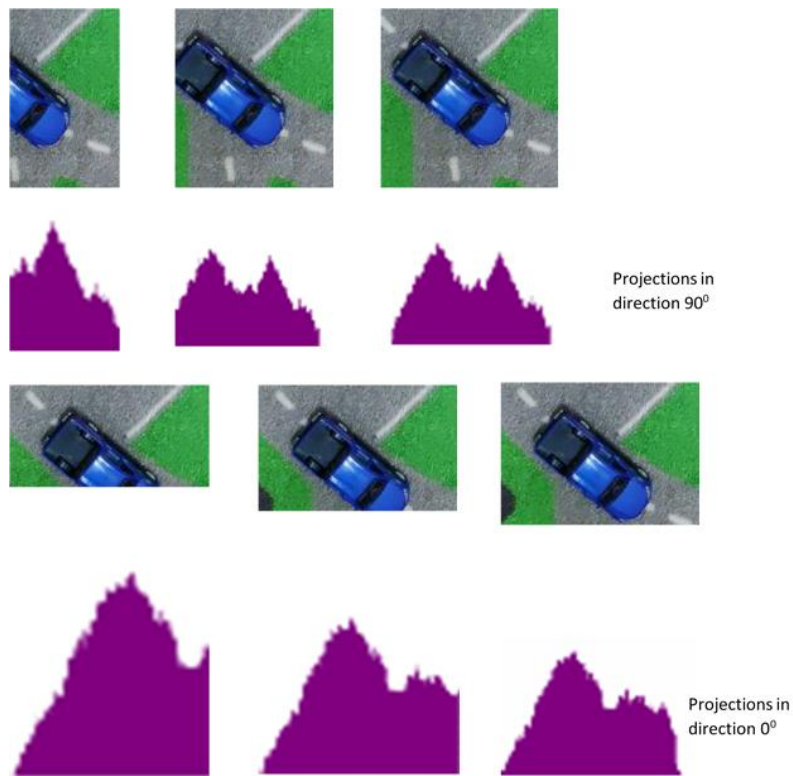


Figure 2: Examples of shifting and increasing the selection window towards the border that intersects the object in the image

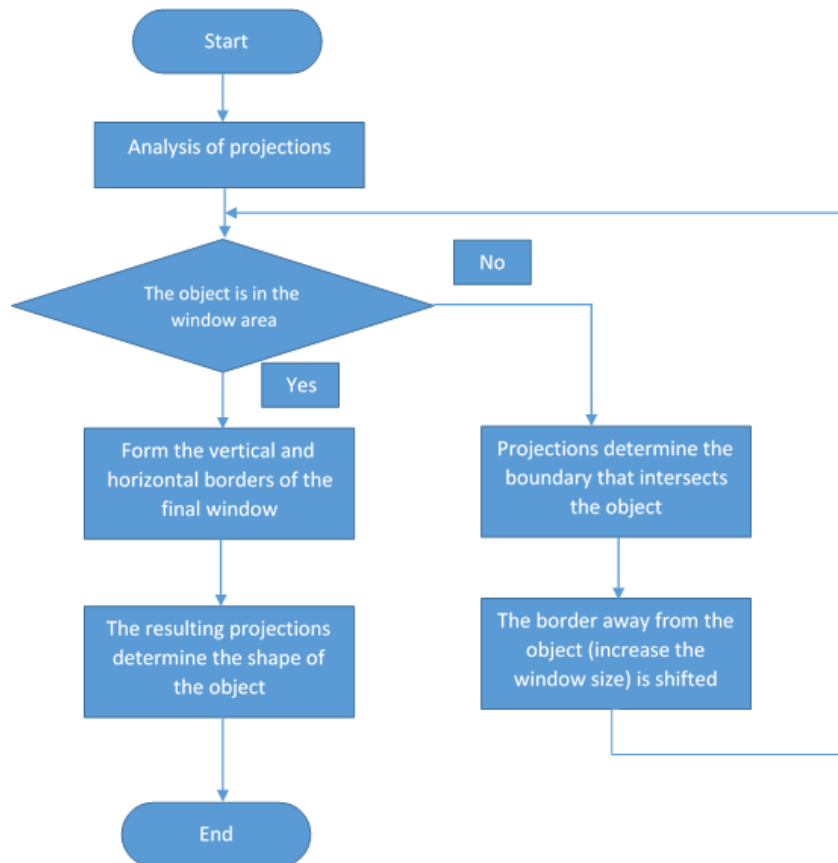


Figure 3: Graph diagram of the algorithm for searching for an object in the image

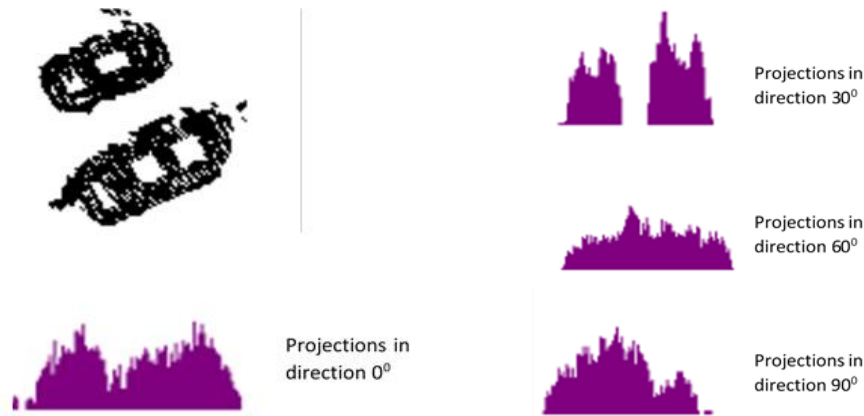


Figure 4: Examples of projections when rotating an image of a selected area by discrete rotation angles

4. A method for recognizing the geometric shape of an object based on the analysis of projections of its binary image

The formed projections of the selected object on the image allow you to determine its shape. Each projection is a set of numbers, each representing one straight line in the direction of the projection. For an $n \times m$ image, the projection can be given by the following model

$$P = \langle n_1, n_2, \dots, n_m \rangle,$$

where $n_i = \sum_j^m q_j$,

q_j – the value of the code of the j -th pixel in the corresponding i -th row of the image matrix.

The code value of each pixel can be either 0 or 1. In this case, the number n_i is equal to the number of ones in the corresponding i -th line of the image, which has the size of a rectangle covering the selected object. There are works that describe the forms of projections of various binary figures. For example, the projections of a circle filled inside a contour are the same. But at the same time, there are figures that can also give the same projections as circles. They may have voids inside the contour (Figure 5). The same can be said about other geometric shapes. You can eliminate such problems by using the outline representation of shapes (Figure 6). Especially if the outline is one pixel thick.

The search for straight lines from the resulting projections is determined by a large number in the sequence of numbers that determines the resulting projection. For example, if the projection is defined by a set of numbers $\langle 10, 2, 2, 2, 2, 2, 2, 2, 10 \rangle$, then it is safe to say that the edges of the object are limited to two segments ten pixels long. If this is a horizontal projection, then the resulting projections of the segments determine the top and bottom projections of objects. To determine the right and left boundaries, analyze the vertical projection or rotate the image by 90° . In this case, for a perfect rectangle, the projection will be determined by the following set of numbers $\langle 7, 2, 2, 2, 2, 2, 2, 2, 7 \rangle$.

Inclined line segments are determined by the ratio of sets of numbers in several projections. For example, there are two sets $P1 = \langle 1, 1, 1, 1, 1, 0, 0, 0, 0, 0 \rangle$ and $P2 = \langle 0, 0, 0, 1, 1, 1, 1, 1, 0, 0 \rangle$, then you can define a segment, as shown in Figure 7.

According to Figure 7 it can be seen that several line styles are possible according to the obtained projections. Therefore, to improve accuracy, it is better to rotate the image and analyze those projections that form the largest number on one of the projections. This means that the segment is perpendicular to the projection that forms the maximum number. The angle of rotation is fixed, which determines the angle of inclination of the segment. Because orthogonal tiling results in aliasing, the outline can be represented as a few pixels thick. In this case, allowable thickness limits are set for each projection, which can reduce the percentage of false description of the geometric shape, as well as reduce the time spent on processing. Before and during operation, the system is trained using the generated numerical sets that define projections and geometric shapes. For each geometric shape, two projections (two sets of numbers) are formed, which are determined by the ratios of the numbers in each numerical set.

However, the resulting initial numerical set is quite difficult to process. Therefore, to simplify the processing of sets of numbers, thresholding is used, with the help of which the numbers in each set are

sequentially reduced until zero values appear. An example of such processing for the initial set of numbers $\langle 10, 2, 2, 2, 2, 5, 2, 2, 2, 3, 2, 2, 2, 2, 9 \rangle$ is carried out in time steps, presented in Table 1.

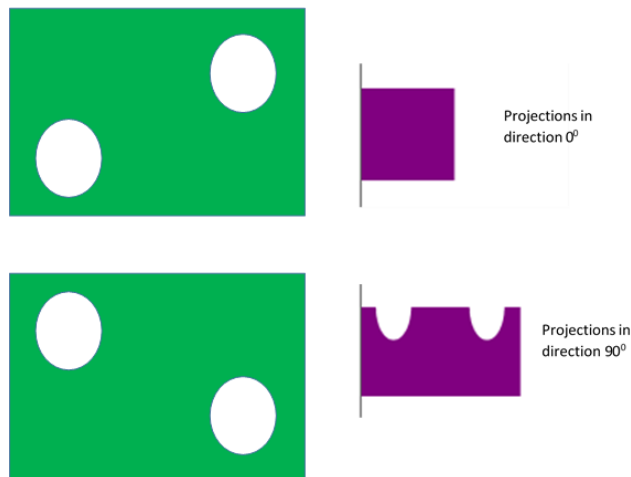


Figure 5: Examples of figures with the same projections, but different geometric shapes

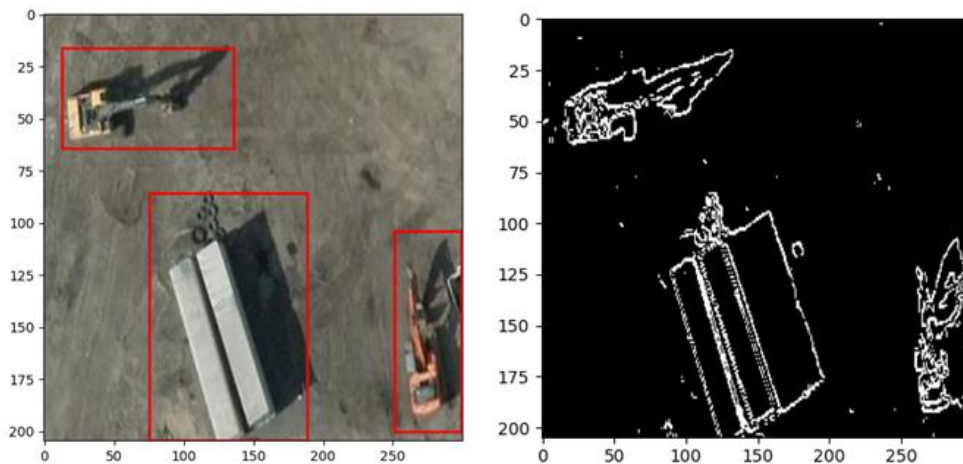


Figure 6: An example of contour representation of figures and their geometric projections

											P_2
				1							1
					1						1
						1					1
							1				1
								1			1
											0
											0
											0
P_1	0	0	0	1	1	1	1	1	0	0	

Figure 7: An example of determining the shape of a segment by two projections

The table shows that perpendicular to this projection there are three lines located in the first, sixth and fifteenth row (column). If this projection is formed in the horizontal direction, then these line segments are parallel to the horizontal axis. In this example, the number 3 can be regarded as noise and, within the confidence interval, it is related to the contour thickness. The numbers in the far right column of Table 1 indicate the number of pixels that were removed by thresholding. According to the results of processing the numerical sequence of the vertical projection, segments of vertical lines are determined. For an example of a vertical projection that forms an array of numbers $\langle 7, 2, 3, 3, 3, 3, 3, 2, 6 \rangle$, the image in fig. 8 is shown.

Table 1
Single projection thresholding example

Processing step	Array of numbers	Memory of steps
0	10, 2, 2, 2, 2, 5, 2, 2, 2, 3, 2, 2, 2, 2, 9	
1	9, 1, 1, 1, 1, 4, 1, 1, 1, 2, 1, 1, 1, 1, 8	
2	8, 0, 0, 0, 0, 3, 0, 0, 0, 1, 0, 0, 0, 0, 7	2
3	7, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 6	3
4	6, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 5	
5	5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4	4
	
9	1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0	9

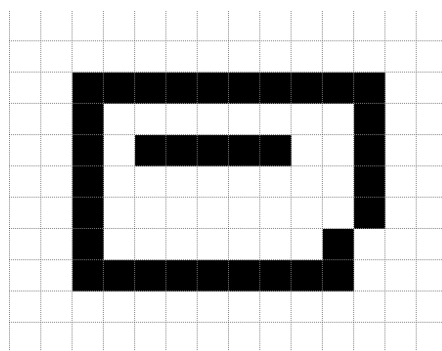


Figure 8: An example of forming an image of a figure from arrays of numbers of two perpendicular projections

If we neglect the difference between the two extreme horizontal segments, then this figure can be identified as a rectangle with a horizontal segment inside. For each geometric shape, you must assign an identifier, as well as assign it to a specific class. The form shown in Figure 8 can be attributed to a rectangle with a horizontal segment inside. If we consider real objects in the image, then, for example, a bus can be represented as a rectangle in a top view, a car can be represented by a quadrilateral with an internal quadrilateral that represents the roof, as well as additional internal quadrangles that represent the glass windows in the car.

The selection of objects by color arrangement does not always give the correct result, since a large number of individual objects in the image can be selected, each of which must be analyzed separately, which requires additional processing time, and it is also necessary to determine the geometric shape of the object. It is also difficult to determine the height of an object located on the ground from the top view, which does not make it possible to accurately identify the 3D geometric shape. However, it can be clearly defined that it belongs to one of the classes of geometric shapes of objects.

5. Selection of a group of objects based on the analysis of projections of contour objects

After selecting the edge pixels on the image, two projections are formed and analyzed. Previously, analysis was considered for a single selected object. However, it is possible to single out a group of objects at once by threshold analysis of the projections themselves. For selection, the analysis of quantities, the values of which were selected experimentally, is used. Let us introduce the symbols and formulas for determining the values of the case of horizontal projections:

$d(i)$ - number of non-zero pixels in row i ; μ - density test size;

$\delta(i) = (d(i-\mu) + d(i-\mu+1) + \dots + d(i-1) + d(i) + d(i+1) + \dots + d(i+\mu-1) + d(i+\mu)) / \mu$ -

density of the i -th element of the row (Figure 9, the density is shown in red);

θ - threshold for deleting a projection area (Figure 9, the threshold is shown in yellow, the supposed areas are in red frames in the image);

The allocation of informative areas is carried out according to the following rule

$$H_i^* = \begin{cases} 1, & \text{if } \delta(i) > 0 \\ 0, & \text{if } \delta(i) < 0 \end{cases}$$

For vertical projections, all quantities are determined similarly.

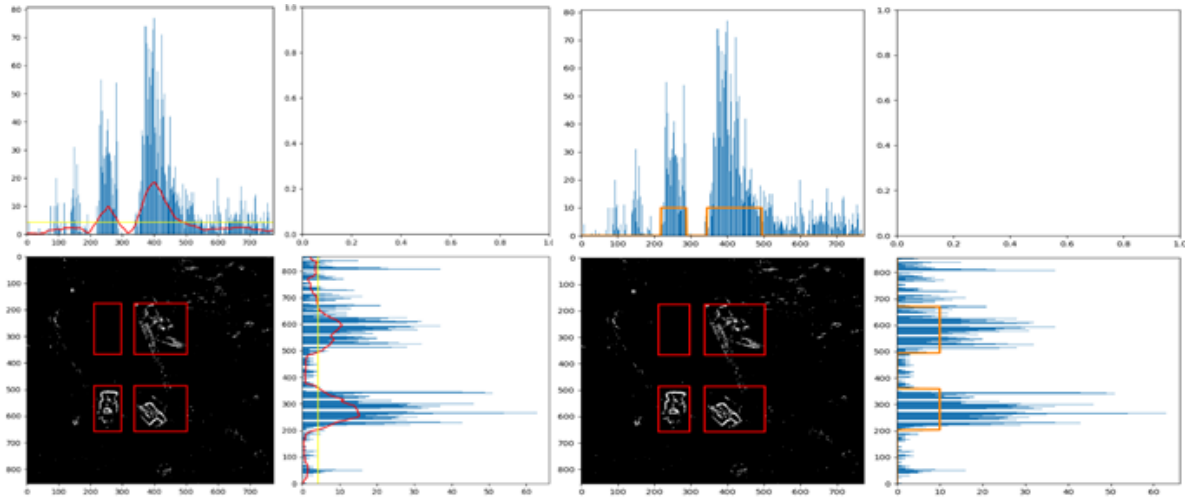


Figure 9: Analysis of the distribution density of single pixels in an image

It was empirically determined that the optimal threshold value Θ for deleting an informative area is determined by the following formula.

$$\Theta = \frac{(h+w)}{4,64},$$

were Θ - informative area deletion threshold: if the estimated area contains non-zero pixels less than Θ , it is removed (Figure 10, the area in the upper left corner has been removed);

h - plot height; w – width.

Since the projections of objects can overlap each other in one of the directions, the method does not immediately unambiguously perform recognition and determine clear boundaries. However, a repeated launch in certain areas refines the found boundaries and selects objects without losing information or, conversely, storing extra elements that do not belong to the object (Figure 11).

When processing image projections, simple threshold removal showed low efficiency, since areas that were inside the object were removed or, conversely, single pixels remained, perceived by the model as informative areas (Figure 12). Determining the density helps to reduce the “estimation” of single pixels and, conversely, smooth empty pixels inside the informative area. And thresholding leaves only the necessary objects on the image. This behavior is shown in previous figures.

6. Conclusion

A system for recognizing objects on satellite images has been successfully implemented in the work. The results of the experiments confirmed the effectiveness and advantages of the developed method in comparison with other neural network approaches. One of the key advantages of the implemented system is its speed. Compared to other approaches that require significant computational resources for training and direct work, the developed method offers an effective solution that allows you to increase the processing speed. This is important given the volume and complexity of such images. As a result of the experiment, the threshold for deleting an informative area is determined if the proposed area contains non-zero pixels less than a given value, which allows you to select several objects at the same time in one image. The developed system for recognizing objects on satellite images has demonstrated its efficiency and speed advantages. Given the limited computing resources, the developed method is an attractive option for the implementation of systems for processing and recognizing satellite images.

In future studies, the authors plan to study and analyze images of the earth's surface to highlight and recognize objects fixed at an angle of less than 90° , as well as increase the number of projections, which will significantly improve recognition accuracy.

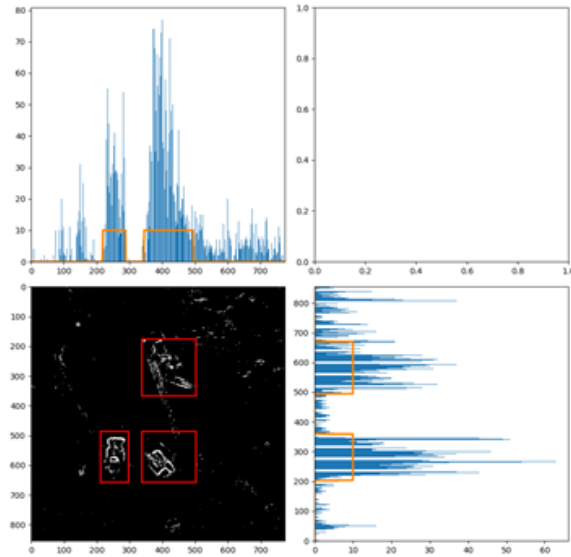


Figure 10: Thresholding example to remove redundant areas

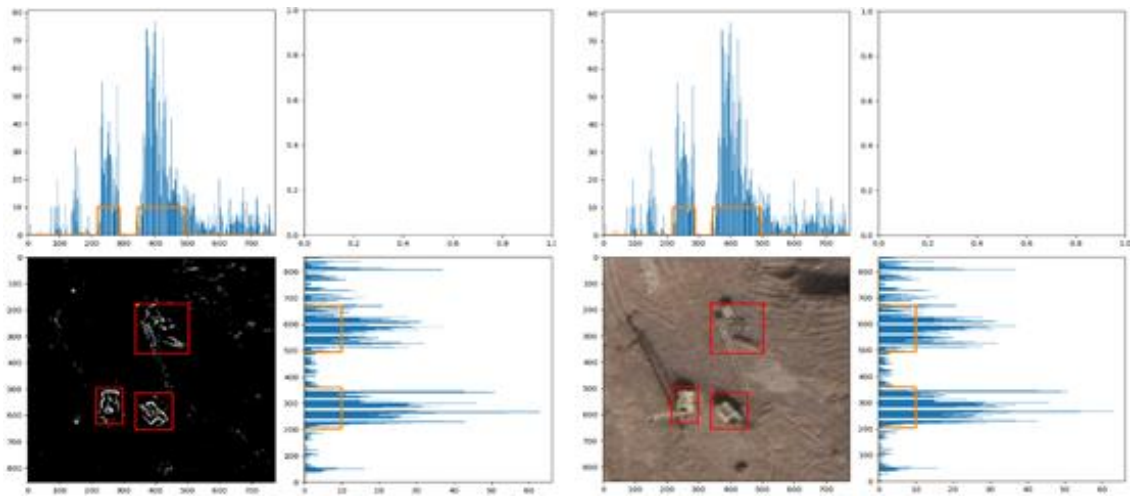


Figure 11: An example of refining the distribution of object boundaries

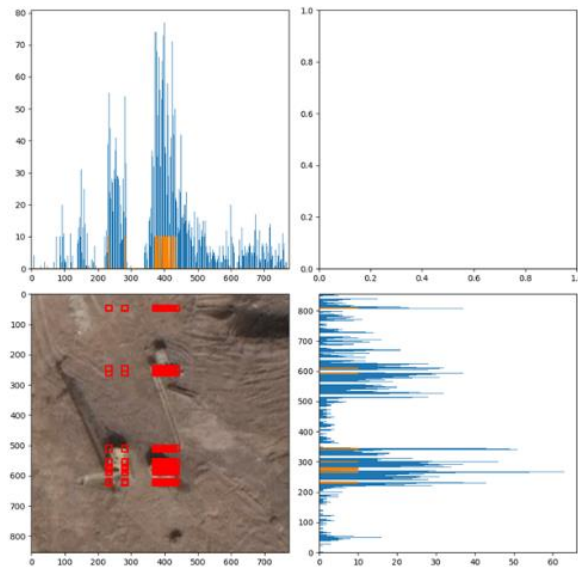


Figure 12: An example of the selection of informative objects based on the analysis of the density of the distribution of pixels on the projections

7. References

- [1] Fan Q, Zhuo W, Tang C-K, Tai Y-W, Few-shot object detection with attention-RPN and multi-relation detector, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) IEEE, (2020), pp. 4013-4022.
- [2] Ye Y, Chen H, Zhang C, Hao X, Zhang Z, Sarpnet: Shape attention regional proposal network for lidar-based 3d object detection: *Neurocomputing*, No. 379, 2020, pp. 53-63
- [3] Hongchun Yuan, Hui Zhou, Zhenyu Cai, Shuo Zhang, Ruoyou Wu, Dynamic Pyramid Attention Networks for multi-orientation object detection. *Journal of Internet Technology* Vol. 23 No. 1, January 2022, pp. 81-90.
- [4] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 39, No. 6, pp. 1137-1149, June, 2017.
- [5] Tarshin V. A., Sydorenko R., Sotnikov O. M., Pevtsov G. V., Megelbey G. V., Lupandin V. A., Kovalchuk V. A. (2015). The method of selection of informative sections of images based on the theory of fractal analysis. Patent of Ukraine for utility model No. 100560, BI No. 14. from 27.07.2015 p.
- [6] Fractal analysis of processes, structures and signals. Collective monograph / Under 5 ed. R.E. Pashchenko // Kharkiv: HOOO "NEO "Ecoperspektiva", 2006. - 348 p.
- [7] Wang Z., Bovik A. C., Lu L. Why is image quality assessment so difficult? // *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2002. V. 4. P. 3313-3316.
- [8] Orda M. V., Abramov S. V., Poltorak M. F. (2017). A method of automated detection of planar objects and informative areas on halftone raster images of the earth's surface. Patent of Ukraine for a utility model No. 121816, Byul. No. 23 dated 11.12.2017
- [9] Pat. WO 2009143651 (A1), IPC7 G06T 5/00. Fast image segmentation using region merging 30 with a k-nearest neighbor graph / Mantao X. [CN], Qiyong G. [CN], Hongzhi L. [CN], Jiwu Z. [CN].
- [10] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, (2016) You only look once: Unified, real-time object detection, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788.
- [11] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, Ssd: Single shot multibox detector, 2016 European Conference on Computer Vision (ECCV), Amsterdam, Netherlands, 2016, pp. 21-37.
- [12] T-Y Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 42, No. 2, pp. 318-327, 2020.
- [13] G. Zhang, S. Lu, W. Zhang, CAD-Net: A context-aware detection network for objects in remote sensing imagery, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 57, No. 12, pp. 10015-10024, December, 2019
- [14] L. Zhou, H. Wei, H. Li, W. Zhao, Y. Zhang, Y. Zhang, Arbitrary-Oriented Object detection in remote sensing images based on polar coordinates, *IEEE Access*, Vol. 8, pp. 223373-223384, November, 2020
- [15] X. Yang, J. Yang, J. Yan, Y. Zhang, T. Zhang, Z. Guo, X. Sun, K. Fu, SCRDet: Towards More Robust Detection for Small, Cluttered and Rotated Objects, 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 2019, pp. 8231-8240
- [16] R. L. Motornyuk, S. Bilan. The Moving Object Detection and Research Effects of Noise on Images Based on Cellular Automata With a Hexagonal Coating Form and Radon Transform. *Handbook of Research on Intelligent Data Processing and Information Security Systems. Handbook of Research on Intelligent Data Processing and Information Security Systems*. Edited by Bilan, S. M., & Al-Zoubi, S. I. Hershey, USA: IGI Global(2019): 330 – 359.
- [17] Sobel, I.E. Camera Models and Machine Perception, PhD Thesis, Stanford Univ, (1970): 60.
- [18] Prewitt, J.M.S. "Object Enhancement and Extraction". *Picture processing and Psychopictorics*. Academic Press. (1970): 75-149.
- [19] Stepan Bilan. (2022) Operators for Edge Detection in an Image Based on Technologies of Cellular Automata, International Conference "Information Technology and Interactions" (IT&I-2022). Workshops Proceedings Kyiv, Ukraine, 2022. Vol. 3384: 142-150
- [20] Motornyuk R.L., & S. Bilan. (2019). Methods for Extracting the Skeleton of an Image Based on Cellular Automata With a Hexagonal Coating Form and Radon Transform. In *Handbook of Research on Intelligent Data Processing and Information Security Systems*, ed. by S. M. Bilan, & S. I. Al-Zoubi, 289 – 329, USA: IGI Global