Combined Processing of Satellite and UAV Data to Increase the Classification Reliability

Irina Vasilyeva[a], Vladimir Lukin[a], Vyacheslav Kharchenko[a], and Andriy Nereta[b]

[a] *National Aerospace University, 17 Chkalova Street, Kharkiv, 61070, Ukraine*
[b] *Department of Environmental Protection and Nature Management of the Kharkiv Regional State Administration, 5 Svobody Square, Kharkiv, 61022, Ukraine*

Abstract

The reliability of remote sensing data classification is significantly affected by the quality of images and the efficiency of algorithms and methods used at all stages of their processing. At the same time, the complementarity of satellite data and the results of aerial photography from UAVs opens up wide opportunities for their joint use in order to improve the accuracy of decision-making about the observed class of objects. We discuss one of the options for the joint use of UAVs and satellite data for the tasks of integrated remote aerospace monitoring of natural and anthropogenic objects. Initially, regions of interest are identified on satellite images, where search objects are most likely to be located. To identify objects with known stochastic properties under conditions of significant a priori uncertainty, it is recommended to use a probabilistic filtering method that allows segmenting an original image into two classes: the "object of interest" and a generalized class that includes all other classes that differ from the searched object by characteristic features. Further, UAV data are involved in the analysis, which can provide reliable information about the presence (or absence) of a given object in the detected search area. This information is used to correct the results of satellite image recognition. The results of segmentation of satellite images (Sentinel-2) and classification of test aerial photographs by the maximum likelihood method are presented. It is shown that UAV data can be useful for analysing or calibrating satellite data, which is especially important for classification procedures based on nonparametric supervised learning methods.

Keywords 1

Remote sensing data, image classification, probabilistic filter, Johnson distribution, classification reliability

1. Introduction

At present, the level of development of technologies for remote sensing (RS) of the Earth, such as space optical and radar imaging, aerial photography from unmanned aerial vehicles (UAVs), allows creating complex multi-level models for monitoring objects and processes in various sectors of the economy. Aerospace images are used to solve many applied and research problems, in particular, geoecological monitoring of the state of natural resources, studying the dynamics of regional development, assessing the consequences of emergencies and so on. One of the important areas of remote sensing is forestry where the main tasks include, in particular, analysis of the state of forests, remote assessment of biomass, detection and recognition of deforestation, identification of zones of

IntelITSIS'2023: 4th International Workshop on Intelligent Information Technologies and Systems of Information Security, March 22–24, 2023, Khmelnytskyi, Ukraine
EMAIL: i.vasileva@khai.edu (Irina Vasilyeva); lukin@ai.kharkov.com (Vladimir Lukin); v.kharchenko@csn.khai.edu (Vyacheslav Kharchenko); a252825@gmail.com (Andriy Nereta)
ORCID: 0000-0002-1378-1104 (Irina Vasilyeva); 0000-0002-1443-9685 (Vladimir Lukin); 0000-0001-5352-077X (Vyacheslav Kharchenko); 0009-0006-4020-371X (Andriy Nereta)

erosion and soil degradation and other negative consequences of logging, as well as identification and monitoring of forest fires.

The solution of such problems is associated with the analysis of objects in images with their subsequent recognition and identification. The reliability of object recognition in images is mainly due to their spectral reflectivity (in the ranges of the electromagnetic spectrum used), the clarity of their boundaries, the degree of variability, and the presence of stable relationships with other objects. Significant factors in the reliability of recognition results are also the quality of images and the efficiency of the algorithms and methods used at all stages of their processing.

The complementarity of satellite and airborne data opens up wide opportunities for their joint use [1–3]. For example, the review article [3] classifies strategies for the so-called synergy between UAVs and satellite data, discusses emerging trends in detail and suggests ways to overcome existing limitations and how to fully exploit the potential of this synergy.

In this paper, it is proposed to discuss one of the options for the joint use of data obtained from UAVs and satellites to solve the problems of integrated remote aerospace monitoring of natural and anthropogenic objects of interest. At the same time, the accompanying tasks are the choice (synthesis) of effective data analysis procedures and the study of ways to combine the obtained analysis results in order to improve the accuracy of decisions undertaken at regional level.

The variant of the strategy for the joint use of UAV and satellite data proposed for discussion is based on the results of the study of three-channel images, which are the simplest version of multi-channel remote sensing data.

The applied technique is based on per-pixel classification and considering of statistical classifiers [4] using probabilistic filter and maximum likelihood (ML). The novelty lies in the methodology for combining the results of segmentation of satellite images and the classification of aerial photographs from UAVs. It is shown that by increasing the resolution and detailing of the analyzed images, it is possible to increase the reliability of recognition of objects under observation.

2. Advantages and disadvantages of satellite and UAV data

As it is known, the processing of satellite images makes it possible to reveal many important characteristics of the Earth surface: the properties of mineral composition of open rocks, the areas of water and wind erosion of soils, the vegetation cover and the degree of its degradation, the state of water bodies (pollution, floods), the location of forest fires, burnt areas, illegal logging, etc.

The main advantages of satellite imagery are the global coverage of the Earth surface and the relative availability of remote sensing data, since a large amount of information is publicly offered on the Internet. In addition, coordinate referencing and geometric correction of information coming from satellites is carried out in an automated mode using orbital data.

The best characteristics of spatial information about the Earth surface, which is in open and free access, are currently provided by Landsat 8 and Sentinel-2 satellites [5]. The Landsat 8 satellite provides the collection and storage of optical multispectral images of medium resolution (30 meters per pixel) for, at least, 5 years with the preservation of geometry, calibration, coverage, spectral characteristics, images quality and data availability at a level similar to previous satellites of the Landsat program. Landsat 8 forms images using a wide range of spectral channels (9 bands of visible light and near infrared (IR) and 2 bands of far (thermal) IR). Satellites of the Sentinel-2 series are equipped with an optical-electronic multispectral sensor for surveying with a resolution of 10 to 60 m in the visible, near-IR and short-wave IR zones of the spectrum, including 13 spectral channels. The width of the observation strip is 290 km, the survey frequency is 10 days, but the series of images "overlap", therefore, in fact, the same area is recorded every 2–3 days.

The main disadvantages of open information provided by satellites are quite low imaging frequency (several days or more), dependence on weather conditions got optical imaging, and relatively low spatial resolution of images.

Along with satellite imagery, aerial photography with the help of UAVs is becoming increasingly popular [6]. Ultralight aircraft UAVs are mainly used to create orthophotomaps of territories and digital terrain models, and to monitor extended objects. The main advantages are high cruising speed, significant flight range and autonomy. Helicopter-type UAVs (especially quadcopters and hexacopters) are mainly used for monitoring small areas or inspecting complex structures, laser scanning and thermal imaging. Their main advantages are their small size, the possibility of launching them from any sites and their hovering over the object of examination, as well as an increased payload.

In addition, the UAV can be quickly delivered to the place of image registration. The use of light UAVs does not require specially prepared sites for take-off and landing, and for medium and heavy UAVs, the object of observation can be located at a distance of several hundred kilometers. When registering from low altitudes, even relatively simple cameras make it possible to obtain a spatial resolution of several centimetres. With the exception of extreme weather conditions (fog, heavy rain), UAVs can be considered almost all-weather, and when equipped with infrared optics, even around the clock.

Thus, the advantages of using UAVs are economy, efficiency and the possibility of obtaining high-resolution images. The main disadvantages of UAVs include a smaller viewing radius and their increased accident rate – this is due to the lack of effective on-board obstacle recognition and collision avoidance systems. In addition, many UAV models are equipped with imperfect autopilots (to reduce the cost and weight of onboard equipment). Small UAVs do not allow for stable geometric survey parameters. As a result, UAV images can have high detail, brightness and contrast, but low photogrammetric quality. This complicates the tasks of coordinate referencing and geometric transformation of aerial photographs by reference points.

Additional problems are created by gaps in the legislation on UAVs in many countries – issues of certification, insurance, registration, as well as the confidentiality of data that can be accessed using UAV images, are not fully regulated [7].

3. Used approaches to image classification

The task of identifying an object of observation on a satellite image is directly related to the problem of choosing informative classification features and effective recognition algorithms [8]. At the same time, classification features that express the quantitative information of interest are the basis for distinguishing classes of objects.

In cases where the external characteristics of the shape of objects are of interest, the description of the boundaries of the regions corresponding to these objects in the images is used to represent classes. In particular, the boundaries of deforestation areas are rectilinear; their shape is often close to a rectangle or polygon, which is not typical for natural objects. Burnt areas are characterized by sharply intermittent borders, which is determined mainly by the terrain, since on hills fire penetrates into the forests to a great depth.

If the recognition procedure is based on the properties of the area itself, such as its color and texture, then it is represented by internal characteristics (the set of image elements that make up this area). Sometimes, in practice, to increase the reliability of recognition, both methods of representation are used simultaneously [9].

When classifying satellite images, the most common approach is based on spectral (brightness) features. It is based on the fact that the brightness of chromatic objects in different spectral zones is not the same and is characterized by a coefficient of spectral brightness. Thus, the value of each pixel of the bitmap corresponds to the spectral brightness of the object for a certain region of the electromagnetic spectrum. Although different objects have specific spectral characteristics, their properties can change over time, depending on the height of the Sun, the transparency of the atmosphere, the angle of view of the imaging system, landscape relief, etc.

Thus, the images of terrain objects are formed depending on the spatial distribution of their spectral brightness and changes in the optical-geometric conditions of observation. In addition, the RS transmission channel can introduce additional spatial-frequency distortions, which leads to a decrease in spatial resolution. This also leads to blurring the structure of images of objects and radiometric distortions due to light scattering in the atmosphere. As a result, tone scales and color rendition are distorted; image contrast is reduced; at that, the degree of distortion is different in different spectral zones. The variability of the spectral brightness of objects, as well as the ambiguity of the transfer characteristics of imaging systems, leads to an increase in stochasticity and, accordingly, to a decrease in the reliability of image classification.

In addition to direct measurement of spectral features and interpretation of their values, their combinations and various functional transformations can be used to identify objects. Transformations of the initial spectral features of objects are performed in order to emphasize and display the differences in classes; it is assumed that each pixel of the object's image retains the spectral features of the entire object, and their purposeful combinations in different spectral zones contribute to its identification. Thus, for satellite monitoring of agricultural crops, the technology for monitoring changes in vegetation indices has become widespread [10]. The calculation of most vegetation indices is based on a certain transformation of the values obtained in the two most stable sections of the curve of the spectral reflectance of plants.

Accordingly, the first step in solving the problem of object recognition is the choice of classification features, the information content (separating properties) that largely determines the reliability of the decisions made.

The next step is the construction of a decision rule in the space of informative features. At the same time, the choice of the classification method depends both on the dimension of the feature space and the method of describing the classification features as well as on the completeness and reliability of a priori information about object classes.

When classifying remote sensing data, methods of both supervised classification (classification with a teacher) and clustering methods (classification without a teacher) are used. In supervised classification, to assign a feature vector to one of the mutually exclusive classes (the so-called "hard" classification), the classifier is trained to distinguish between these classes based on training samples representing each of the classes. If, with varying degrees of confidence, the current feature vector is assigned to several classes, then we are talking about a fuzzy classification. Fuzzy classification methods are used in the case of mixed pixels (i.e., if the pixels are the weighted average of the spectral characteristics of the area of the earth's surface, within which there are different classes of objects). Like traditional ("hard") classification, fuzzy classification also uses a learning procedure, but the difference is that it is necessary to introduce more information about the classes that make up the pixel into the set of reference data. Fuzzy classification algorithms use the logic of building membership functions or fuzzy convolution operations that allow assignment of classes within a sliding window [11].

The learning process is actually the process of building decision functions (decision rules) for classes of objects. In the theory of statistical decisions, all types of decision rules are based on the formation of the likelihood ratio L and its comparison with a certain threshold c, the value of which is determined by the accepted quality criterion. The statistical approach assumes that at one point in the feature space with a non-zero a priori probability, implementations belonging to different classes may appear. This is due to the inevitable random errors in the registration and transmission of observational data, as well as the fact that features are fundamentally random variables. The latter is typical for remote sensing problems.

The problem of synthesizing statistical decision rules and the problem of analyzing their quality are closely related to the role of a priori information. Thus, among the methods of supervised classification, the techniques that take into account the probability of the presence in the image of objects belonging to a certain class are widely used. To develop the corresponding algorithms, experimental data on the relationship between the spectral brightness of objects and their properties are used. For example, the spectral brightness of soils clearly correlates with their particle size distribution and moisture content.

The most detailed and complete specification of a priori information is that the total set of classes A, conditional (by class) distributions of features $f(\vec{x}|a_k)$, a priori probabilities of occurrence of objects from different classes in the observation sample $P(a_k)$, as well as the loss function, which quantitatively expresses the sum of generalized losses due to erroneous decisions, are considered to be known. This level of a priori information corresponds to the Bayesian direction in statistical decision theory [4]. Real recognition problems are characterized by the presence of uncertainty due to the stochastic nature of the features, noise and interference during the registration and transmission of observational data, changes in the angle and position of the object, etc. There may be no reliable information about the total set of observed classes, about a priori probabilities of classes; also, as a rule, the type and parameters of real (but not supposed theoretically) distributions of random features of classes are previously unknown.

The most problematic task is to construct a decision rule under conditions of complete a posteriori uncertainty, that is, in the absence of complete and reliable information about the number, type, and spectral characteristics of the observed classes. Therefore, the first step in the unsupervised learning process is to divide the data into subgroups (clusters); in this case, data with similar values of features are combined into one group. To assess the degree of similarity (or difference), various distance metrics in the feature space are employed; it is considered that the distance between indications in one group (one class) will be significantly less than the distance between indications from different groups. A typical example of a classical clustering algorithm is the K-means algorithm (and its modifications), which is widely used to partition large amounts of multivariate data [12].

Another approach that allows splitting the feature space into arbitrary-shaped clusters is based on the assumption that the initial data is a sample from a multimodal distribution, the so-called mixture of basis functions. Each class is interpreted as a unimodal population, and classified observations as a sample from a mixture of such populations; in this case, the number of mixture components determines the number of classes K, and the specific weights of these components determine their a priori probabilities. Normal distributions (Gaussian Mixture Model, GMM) are usually chosen as basis functions [13].

Along with the statistical approach to the classification of RS data, methods based on machine learning, including methods based on deep neural networks. The use of neural networks is due to the absence of the need for a priori information about the input data, the possibility of forming nonlinear boundaries of decision-making areas in the feature space, as well as their resistance to errors when processing incomplete or partially incorrect input images [14].

A problem in the processing of satellite images using neural networks is the spatial size of objects. To train a neural network, a large training set is required. At the same time, even relatively large objects on satellite images can occupy only a few pixels (depending on their spatial resolution).

Despite the variety of methods and approaches to pattern recognition, the development of algorithms for the automated classification of RS images faces a number of specific problems. In particular, when using pattern-matching methods, the problem lies in the lack of a complete database of reference class descriptions, which is associated with the spatiotemporal instability of the characteristics of objects used for recognition. Thus, in the process of observation, the color and illumination of an object, its size and shape change (depending on the viewing angle and due to partial shading or masking by other objects). In addition, the heterogeneity and variability of the background make it difficult to apply contour detection algorithms and statistical segmentation methods. In object recognition methods, there is a number of algorithms based on singular points, invariant to affine transformations of images, resistant to changes in illumination, survey position and image noise. However, such algorithms are inefficient in recognizing patterns without a well-defined texture [15, 16]. Methods based on the correlation analysis of the temporal sequence of images do not allow one to reliably separate the background and the object of observation, because not only the object is moving, but also the background image. Correlation is rarely used in cases where an arbitrary rotation of the recognized object is possible, since if the rotation data is unknown, then in order to find the best match between the pattern and the object, it will be necessary to analyze all possible rotations of the pattern. In addition, size normalization requires spatial scaling, which requires large computing resources.

In statistical recognition, we face the lack of complete and reliable a priori information about the observed objects, which inevitably affects the effectiveness of decision rules. Thus, the values of a priori probabilities of classes are usually determined based on expert estimates or a preliminary analysis of the observed territory. In this case, an incorrect assessment of a priori probabilities can lead to a significant change in the value of the decision threshold, which, in turn, leads to an increase in the probability of recognition errors. In the absence of a priori information about the type and parameters of distributions of attributes, one has to accept a hypothesis about the type of distribution based on theoretical assumptions. So, in some software systems (for example, PlanetaMonitoring [17]), the vector features of objects are described by multidimensional normal distributions, and the decision about the class is made according to the maximum likelihood criterion. At the same time, the distributions of real multichannel data often have a non-Gaussian type. Even under the assumption that the distribution of observations is normalized according to the central limit theorem, one cannot exclude the nonlinearity introduced when receiving data, which leads to a change in the type and parameters of the observed distribution. In addition, the presence of sensor noise "blurs" the features of objects and distorts their distribution. Finally, it is far from always possible to determine all members of the entire set of classes of observed objects. Even if, based on the results of a preliminary analysis of the situation in the area being probed, many classes have been established, the possibility of a new (unrecorded) class of objects appearing on the scene of the survey cannot be completely ruled out. For methods based on neural networks, a significant problem is to find a well-prepared sufficiently large test set for training. In addition, the process of training a neural network usually takes more time than image classification by classical algorithms.

Thus, the reliability of recognition of objects in the image is largely determined by their spectral properties, the severity of the boundaries, the degree of variability, as well as the presence of stable relationships with other objects. The quality of the images and the efficiency of the algorithms and methods used at all stages of image processing are also significant factors in the reliability of the results. At the same time, it is impossible to guarantee the absolute reliability of the recognition results, it is only possible to minimize the probability of erroneous decisions, including through the joint use of satellite data and UAV images.

4. Proposed method for identifying objects on satellite images

To identify objects with known stochastic properties under conditions of significant a priori uncertainty about the surrounding environment, it is recommended to use the probabilistic filtering method described in [18]. In this case, it is assumed that there are only two classes of objects in the sample set: "the given object A" and the generalized class "not object A", which includes, in the general case, an infinite number of classes of objects that differ from the object A by their characteristic features. To recognize an object A, it is sufficient to know only the distribution parameters of the vector feature for the class, to which this object belongs. Statistical estimates of these parameters are found from the classified sample at the training stage.

To describe the features characterized by non-Gaussian distributions and the presence of correlation relationships, it is advisable to use multi-parameter distributions based on nonlinear transformations of normal random variables. Due to the large number of parameters, such models make it possible to adequately approximate a wide range of unimodal and even bimodal distributions. Among the universal probabilistic models is the Johnson distribution system $\{S_L, S_B, S_U\}$ [19] based on non-linear transformations of the form:

$$z = \gamma + \eta\tau(x, \varepsilon, \lambda); -\infty < \gamma < +\infty, \eta > 0, -\infty < \varepsilon < +\infty, \lambda > 0,$$

which allows obtaining a normalized normal value of z from a random variable x.

The Johnson transform has four parameters: ε is the displacement parameter, λ is the scale parameter, η and γ are the shape parameters; the function $\tau(\blacksquare)$ is determined by the type of transformation. To describe real observation data, the values of which are always limited to a certain range, it is advisable to choose the Johnson S_B-transform:

$$\tau(x, \varepsilon, \lambda) = ln[(x - \varepsilon)/(\lambda + \varepsilon - x)].$$

In this case, the distribution density of the random variable x is described by the formula

$$f(x) = \frac{\eta\lambda}{\sqrt{2\pi}(x-\varepsilon)(\lambda+\varepsilon-x)} exp\left[-\frac{1}{2}\left(\gamma + \eta ln\left(\frac{x-\varepsilon}{\lambda+\varepsilon-x}\right)\right)^2\right].$$

The disadvantage of the Johnson distribution is the absence of a direct connection between the estimates of empirical moments (mean, variance, skewness, and kurtosis) with distribution parameters. Methods for estimating the parameters of the Johnson distribution are iterative and are reduced to solving an optimization problem, the purpose of which is to minimize the mean square error when approximating the empirical distribution by the accepted model.

Using the Johnson S_B-transform, you can get a non-Gaussian "probabilistic window" that allows choosing an object with a non-Gaussian distribution of its feature x:

$$W(x) = exp\left\{-\frac{1}{2}\left[\gamma_W + \eta_W ln\left(\frac{x-\varepsilon_W}{\lambda_W+\varepsilon_W-x}\right)\right]^2\right\}, \quad (1)$$

where ε_W , λ_W , η_W , γ_W are the filter parameters, the values of which are set equal to the corresponding parameters of the Johnson S_B-distribution for the desired object A.

Then, the signal at the filter (1) output

$$y = \begin{cases} 0; & x \le \varepsilon_W; \\ W(x|\varepsilon_W, \lambda_W, \eta_W, \gamma_W); & x \in \,]\varepsilon_W, \varepsilon_W + \lambda_W[; \\ 0; & x \ge \varepsilon_W + \lambda_W \end{cases}$$

has a distribution density that depends both on the statistical characteristics of the features of the object and the degree of consistency of its distribution parameters with the filter parameters; at that, the interval $x \in \,]\varepsilon_W, \varepsilon_W + \lambda_W[$ forms the passband of the probabilistic filter.

If multidimensional models of the shape are used to describe the features of objects

$$f_p(\vec{x}) = (2\pi)^{-p/2}|\mathbf{R}|^{-1/2} \prod_{\kappa=1}^{p} \frac{\eta_\kappa \lambda_\kappa}{(x_\kappa - \varepsilon_\kappa)(\lambda_\kappa + \varepsilon_\kappa - x_\kappa)} \times$$
$$\times\, exp\left[-\frac{1}{2}\sum_{\kappa,\upsilon=1}^{p} R_{\kappa\upsilon}^{-1} \left(\gamma_\kappa + \eta_\kappa ln\left(\frac{x_\kappa - \varepsilon_\kappa}{\lambda_\kappa + \varepsilon_\kappa - x_\kappa}\right)\right)\left(\gamma_\upsilon + \eta_\upsilon ln\left(\frac{x_\upsilon - \varepsilon_\upsilon}{\lambda_\upsilon + \varepsilon_\upsilon - x_\upsilon}\right)\right)\right], \quad (2)$$

where p is the dimension of the feature vector; $\mathbf{R}$ is the sample correlation matrix, then the non-Gaussian "probabilistic window", consistent in form and parameters with the distribution density of the selected object, has the form

$$W(\vec{x}) = exp\left\{-\frac{1}{2}\sum_{\kappa=1}^{p}\sum_{\upsilon=1}^{p}\left[\left(\gamma_\kappa + \eta_\kappa ln\left(\frac{x_\kappa - \varepsilon_\kappa}{\lambda_\kappa + \varepsilon_\kappa - x_\kappa}\right)\right)\right] \times\right.$$
$$\left.\times\, R_{\kappa\upsilon}^{-1}\left(\gamma_\upsilon + \eta_\upsilon ln\left(\frac{x_\upsilon - \varepsilon_\upsilon}{\lambda_\upsilon + \varepsilon_\upsilon - x_\upsilon}\right)\right)\right\}. \quad (3)$$

The output of the probabilistic filter can be interpreted as the probability P_{ij} that the current pixel with spatial coordinates (i, j) belongs to the object A. Then, after scaling the range [0, 1]→[0, 255], we get a grayscale image on which "the degree similarity" of image pixels to a given object A is displayed in gradations of brightness.

The decision rule $W(\vec{z}^*) \ge 1/\pi$ splits the interval of possible values of the output signal of the probabilistic filter (3) $P = W(\vec{z}) \in [0, 1]$ into two areas:

1) $\Gamma_1 \in [1/\pi, 1]$ is the decision-making area on the presence of such a class of objects, the distribution parameters of which coincide with the parameters of the probabilistic filter;

2) $\Gamma_2 \in [0, 1/\pi\,[$ is the decision-making area on the presence of an unknown class of objects.

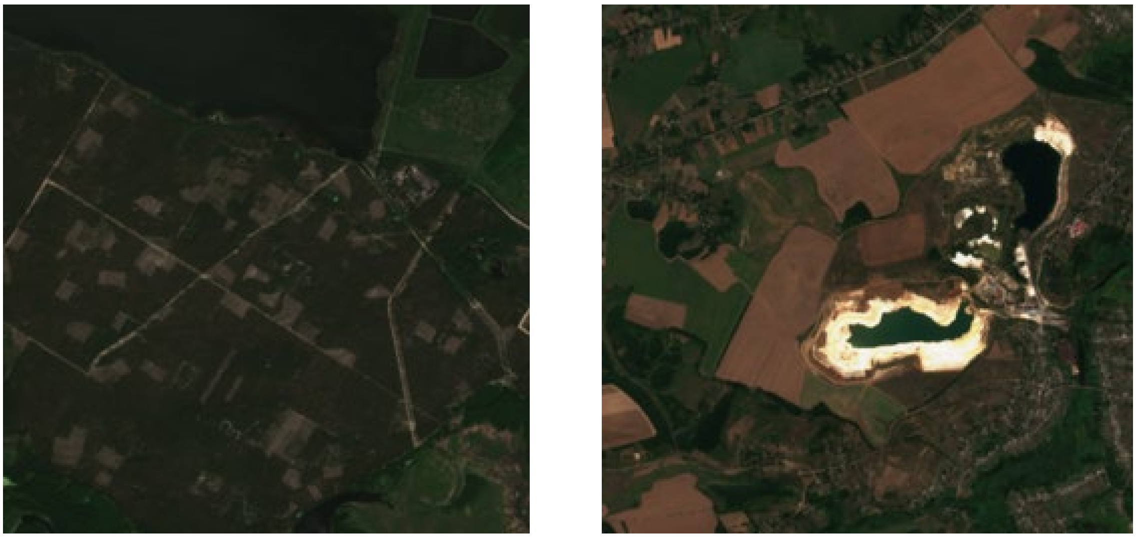
Thus, after thresholding, the results of probabilistic filtering are a binary image in which white pixels (equal to 1 for a binary image or 255 for a grayscale image representations) correspond to the desired object, and black pixels (equal to 0) to the background (generalized class "not object A").

5. Analysis of the results of probabilistic filtering of satellite images

The study of the effectiveness of the probabilistic filtering procedure was carried out on multichannel images obtained from the Sentinel-2 satellite. In each image, it was required to recognize a certain class of objects of natural and anthropogenic origin: forest clearings (Figure 1, a) and quartz quarries (Figure 1, b). In the problem being solved, the objects of analysis were image pixels with certain brightness values in the conditional spectral channels R, G, B.

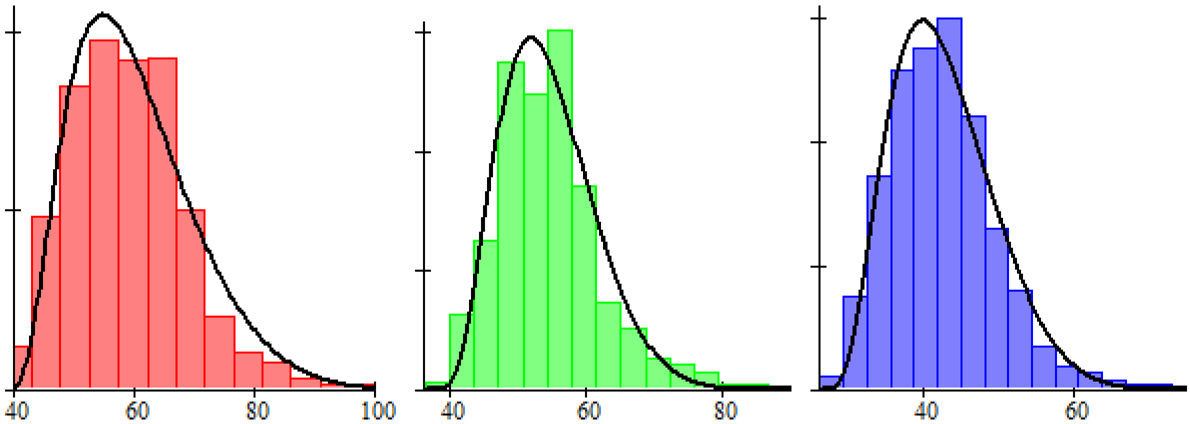
Supervised learning methods were used to adjust the parameters of the probabilistic filter. Based on the actual data on the Earth surface, test areas were identified (relatively homogeneous areas that represent objects that can be distinguished in the images). The sets of reference pixels were divided into two subsets, which were used to train and evaluate the quality of the classifier. It was assumed that these subsets might partially overlap. As the results of the statistical analysis of the training samples showed, the spectral features of the classes are characterized by non-Gaussian distributions and the presence of strong correlations, therefore, for their analytical description, the multivariate Johnson S_B-distribution (2) was used. For example,Figure 2 shows histograms of the components of the vector feature {R, G, B} for the class "forest clearing" for the test satellite image (Figure 1, a) and plots of the Johnson S_B-distributions approximating these histograms.

Estimates of the parameters of the distributions of attributes of classes were found by numerical optimization of the objective function, minimizing the total quadratic error of the representation of the histogram by the theoretical distribution. The obtained estimates were used as parameters of the probabilistic filter (3) tuned to select the corresponding object in the test image.



a b

Figure 1: Test satellite images containing the objects "forest clearing" (a) and "quartz quarry" (b)



a b c

Figure 2: Histograms and their approximations for the spectral features of the "forest clearings" class on the test satellite image (Figure 1, a): R (a), G (b), B (c)

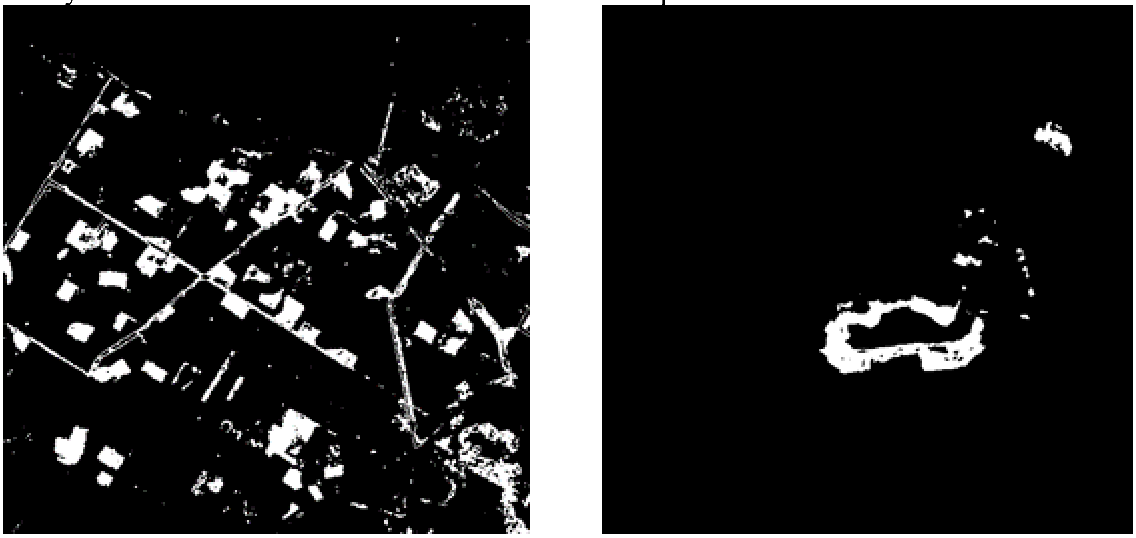
When implementing the pixel-by-pixel procedure of binary statistical image segmentation, the vector of values of the spectral features of the current pixel with spatial coordinates (i, j) was the input of the probabilistic filter (3); the output of the probabilistic filter (3) was compared with the decision threshold. However, the classification results have showed that the choice of a threshold value equal to $1/\pi$ leads to insufficiently efficient detection of search objects due to the high probability of an object omission error. Therefore, the threshold value was experimentally set at 0.1 (in this case, the estimate of the probability of detecting the desired object against the background of other classes was maximum and was 0.902 for forest clearings, and 0.896 for a quartz quarry).

Binary images representing the results of threshold probabilistic filtering of test images are shown in Figure 3. In these images, white pixels correspond to the desired object, and black pixels to the background (generalized class “not object A”).

The results of detecting the contours of objects on the original test images (Figure 1) using binary images (Figure 3) are shown in Figure 4.

Thus, it is possible to select the boundaries of the area corresponding to the search object in the image based on the results of image segmentation using a probabilistic filter (3), consistent with the estimates of the parameters of the Johnson S_B-distribution, approximating the empirical distributions of the spectral characteristics of this object.

An image representing the results of probabilistic filtering may contain pixels that are erroneously assigned to the search object, or vice versa, pixels of a given object may be erroneously recognized as a “background image”. This can lead to significant deformations of the contours of search objects in the image, their relative displacement, overlapping of some object details with others and changes in their relative sizes. To increase the reliability of decisions made on the basis of satellite images, it is necessary to use additional information that UAV data can provide.



a b

Figure 3: Results of detection of objects "forest clearings" (a) and "quartz quarry" (b)

a b

Figure 4: Results of detecting the contours of objects "forest clearings" (a) and "quartz quarry" (b)

6. Discussion on the use of UAV data to improve classification

As can be assumed from the results of segmentation of satellite images presented in Figure 3, a and Figure 4, a, among the selected objects that were assigned to the class "forest clearings", there are also recognition errors: roads, areas of natural degradation of the forest cover, burnt areas, windblows. To verify the reliability of the detection of forest clearings, it is necessary to involve UAV data, which may well replace the on-site observations. So, satellites allow you to determine the location of “zones of interest”, and UAVs provide additional information to clarify the characteristics of search objects.

A UAV for aerial photography must have an autopilot on board capable of maintaining the required shooting parameters (route, camera tilt angles, percentage of longitudinal and transverse overlap, altitude, etc.) in a wide range of weather conditions. One of the most important tasks of flight support is the task of UAV route planning [20, 21]. This task is to determine a set of points in space that would correspond to a given UAV flight trajectory taking into account the restrictions on the flight time and on the maximum number of reference points while ensuring the flight safety (avoiding possible obstacles).

The solution of the problem of spatial referencing of images involves not only the identification of various images of an object, but also the determination of the coordinates of a number of points of interest. As it is known, most numerical algorithms for processing coordinates, including standard photogrammetric procedures, use conjugate points as input data. As such points, one can choose the characteristic points of the contours, in which the maximum curvature is achieved. In addition, a separate area selected as a result of segmentation (probabilistic filtering) can be described by its centre of gravity, which is invariant to rotation, scaling, distortion and noise in the image. At the same time, segmentation accuracy can significantly affect the result of subsequent positioning. Thus, it is necessary to determine the geographical coordinates of the points of interest from the characteristic points of the contours of the selected areas on satellite images. Further, according to the obtained coordinates, the optimal UAV flight route should be compiled. The UAV in automatic mode is capable of flying over a given route with reference to GPS coordinates. Additionally, the UAV flight parameters can be monitored by a ground control station.

The data provided by the UAV includes images in four spectra: R (red), G (green), B (blue) and IR (infrared). Modern UAVs can be also equipped with multispectral cameras, thermal imagers and various detectors, which expands the possibilities for interpreting observational data and increases the reliability of decisions based on them. To monitor the terrain, turret-type video cameras are most preferred, mounted on gyro-stabilized platforms under the UAV fuselage and providing a circular view of the lower hemisphere. Cameras mounted fixedly in the wing or under the fuselage of the UAV are used to capture certain areas of the terrain.

The data obtained from the UAV is proposed to be used instead of on-site surveys to obtain reliable information about the presence (or absence) of the object of observation in the given spatial coordinates. This information can then be used to correct the results of satellite image recognition (image pixel relabelling).

To analyse the possibility of refining the results of segmentation of satellite images according to UAV data, we used two fragments of a three-channel aerial photograph containing images of deforestation in the Kharkiv region (https://mykharkov.info/news/lesnaya-mafiya-kto-stoit-za-vyrubkoj-lesa-na-milliard-griven-76727.html).

On the test images TU-1 and TU-2 (Figure 5), three classes of objects were selected: “Road”, “Forest” and “Forest clearing”.

The histograms of the spectral components {R, G, B} and their approximations for the “Forest clearing” class, specified on the TU-1 test image obtained from the UAV, are shown in Figure 6. If we compare these results with the results of training the classifier on a satellite image (Figure 2), then significant differences between them are obvious. These differences in the “spectral pattern” of the search object are due to distortions of the spectral features during image processing, as well as various conditions during the registration of the analysed images.

The size of the training samples was from 1500 to 5000 pixels, the size of the control samples was from 3000 to 25000.

The results of recognition of these classes by the maximum likelihood criterion [4, 22] are shown in Figure 7. In this case, pixel-by-pixel supervised statistical classification was performed; to obtain the reference descriptions of the spectral features for the classes, we used Johnson S_B-distributions (2).

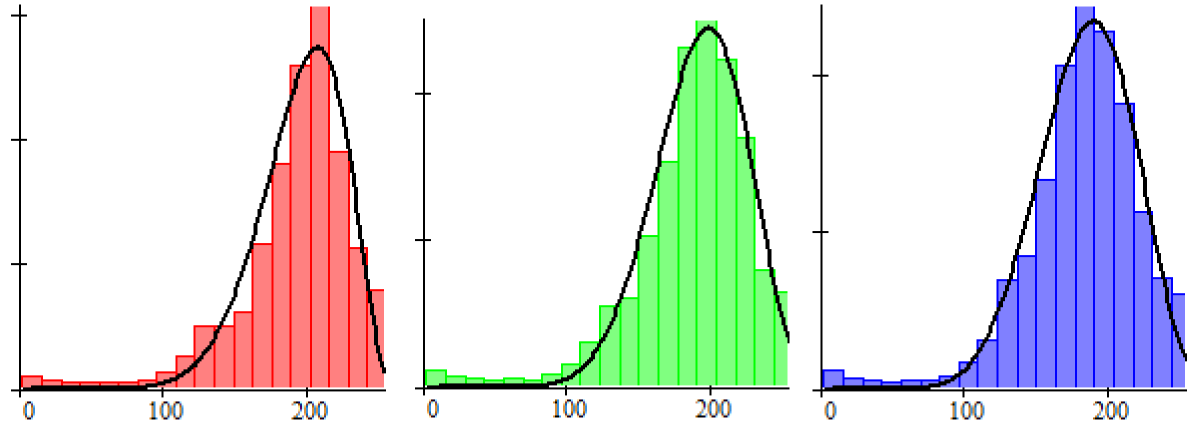
According to the decision rule, value vector for a current pixel $\vec{x}_{ij}^{*}= \{R_{ij}, G_{ij}, B_{ij}\}$ is put into mathematical models of class etalons (2) $f_p(\vec{x}_{ij}^{*}|a_k)$, where $\{a_k\}_{k=1}^{3}$ is the set of object classes. The obtained results are compared between each other and a maximal estimate of the likelihood function is chosen; its index determines the number of the class to which the current pixel belongs:

$$f(\vec{x}_{ij}^{*}|a_v) = \max_{1 \le k \le 3}\{f(\vec{x}_{ij}^{*}|a_k)\} \Rightarrow \vec{x}_{ij}^{*} \in a_v.$$

One of the simplest metrics for evaluating the accuracy of a multi-alternative classifier is the percentage of correctly recognized patterns. In the case of pixel-by-pixel image classification, the evaluation is performed for each pixel of the control sample. Correct recognition is considered to obtain a decision corresponding to a specified class.

a b

Figure 5: Test images of deforested areas obtained from the UAV: TU-1 (a) and TU-2 (b)



a b c

Figure 6: Histograms and their approximations for the spectral features of the “Forest clearing” class on the test UAV image TU-1: R (a), G (b), B (c)

Statistical estimates of class recognition probabilities on the test aerial images (Figure 5) are given in Table 1 and Table 2.

Pixel classification errors, as can be seen in Figure 7, visually manifest themselves in the violation of the topological properties of the reference sets (the appearance of disconnected components, isolated points, open contours, holes in homogeneous areas). One of the possible ways to improve the results of class recognition is to complicate the structure of the classifier. For example, decision rules can include not only the spectral values of individual image pixels, but also the brightness-geometric and structural characteristics of images for objects from various classes. Nevertheless, even a pixel-by-pixel statistical classification makes it possible to quite confidently distinguish between the “Road” and “Forest clearing” classes on aerial images. Thus, the probabilities of making an erroneous decision in favour of the “Forest clearing” class, provided that the objects (pixels) of the control sample belong to the “Road” class, are respectively 0.138 (for the TU-1 image) and 0.097 (for the TU-2 image). The probabilities of erroneously assigning objects (pixels) of the “Forest clearing” class to the “Road” class are respectively 0.054 (for the TU-1 image) and 0.16 (for the TU-2 image).

Thus, the use of UAV data in combination with satellite images expands the possibilities for monitoring emergency situations, studying the anthropogenic impact on the environment, and, in particular, controlling deforestation.

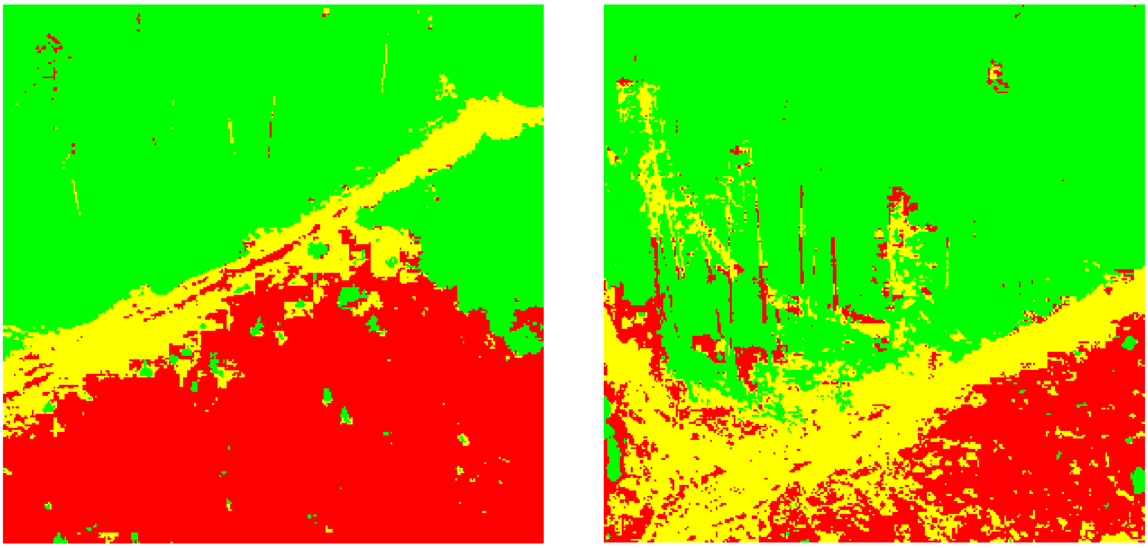


Figure 7: The results of pixel-by-pixel classification by the maximum likelihood method: class map for TU-1 (a) and class map for TU-2 (b). Each class is marked with a specific colour: "Forest" with green, "Road" with yellow, and "Forest clearing" with red

Table 1
Confusion matrix for the ML classifier applied to the test image TU-1 in Figure 4, a

Class	Probability of Decision		
	Road	Forest	Forest clearing
Road	**0.862**	0	0.138
Forest	0.004	**0.980**	0.016
Forest clearing	0.054	0.004	**0.942**

Table 2
Confusion matrix for the ML classifier applied to the test image TU-2 in Figure 4, b

Class	Probability of Decision		
	Road	Forest	Forest clearing
Road	**0.903**	0	0.097
Forest	0.002	**0.994**	0.005
Forest clearing	0.160	0.008	**0.832**

Additionally, it should be noted that the disadvantage of pixel-by-pixel classification algorithms is that the decision on the object class is made for each individual pixel without taking into account the features of the local image structure. These algorithms are quite suitable for detecting compact (point) objects in images localized in a small part of space, but, as a rule, they are not efficient enough for reliable selection of linear (spatially extended) and area objects, since errors in pixel-by-pixel classification lead to distortions of the geometric shape and sizes of objects, their connectivity, as well as changes in the spatial distribution of elements of the image structure.

One of the possible ways to improve the reliability of pixel-by-pixel classification of images is an integrated approach proposed in [23]. This approach consists in the combined use of the results of image classification using various decision-making algorithms. Each of the elementary classifiers generates its own result matrix at the output, which is considered as a “classification layer”. Then, for each pixel of the image, a vector of ambiguous decisions about the class is formed. To select a specific solution, it is necessary to determine the relationship between various combinations of outcomes of elementary classifiers and the set of classes. After that, the decision is made in favour of that class, for which the resulting combination of outcomes is the most possible. As shown in [23], this approach can significantly increase the probability of recognizing poorly distinguishable classes and improve the reliability of image classification. To reduce pixel-by-pixel classification errors, one can also use local spatial methods for post-classification processing of segmented images [24, 25].

This makes it possible to include in the decision rules not only the values of individual pixels, but also the brightness-geometric characteristics of groups of pixels (segments) that form connected objects in the image.

Verification of the effectiveness and feasibility of applying such post-processing procedures is the subject of further research.

7. Conclusions

The proposed strategy for combining data from satellites and UAVs involves the use of satellite images to identify a given class of observation objects (for example, forest clearings, sand pits, burnt areas, reservoirs, etc.). To check and correct the recognition results (taking into account possible classification errors due to the presence of mixed pixels, a small number of pixels representing the search object, the non-representativeness of training samples, insufficient separability of classes in the feature space, etc.), it is advisable to use the advantages of aerial photography from the UAV . The flexibility of the UAV allows choosing the conditions for collecting additional data about the objects you are looking for: the type of sensor, viewing angles, spatial resolution (from centimeter to millimeter), time and frequency of registration of observations. At the same time, since the task is to search for (select) objects of a given class, and not to form an overview image, the small width of the UAV swath (several square kilometers) is of no fundamental importance.

Thus, based on preliminary results of detection of given objects from satellite data, “suspicious” areas or “zones of interest” are formed, where search objects are most likely located. After detecting the contours of the “zones of interest” and their coordinate referencing, the problem is reduced to determining the optimal UAV flight route along the supposed boundaries of the objects being sought. In this case, UAV data are used to obtain reliable information about the presence (or absence) of a given object at the specified spatial coordinates. This information is further exploited to correct the results of recognition of satellite images. In addition, the analysis of the results of checking the quality of the classification of satellite images according to UAV data will make it possible to further correct the recognition procedure, for example, to supplement the algorithm of pixel-by-pixel statistical segmentation by spectral features with a subsequent processing stage that implements an object-oriented classification of a segmented image. This makes it possible to introduce additional information about the geometric shape and size of objects, their topological connectivity, and the spatial distribution of structure elements into the decision rules. Thus, it is supposed to increase the reliability of recognition of given objects.

Hence, UAVs are an affordable solution for obtaining reliable data for analysis or calibration of satellite data, which is especially important when using classification procedures based on non-parametric supervised learning methods. Integration of satellite and UAV based technologies for image classification and recognition can be added some stationary technologies, in particular, sensor networks, 5/6G and so on. In this case, task of structural and parametrical optimizing integrated system could be formulated and solved as a next step of research.

8. References

[1] J. Jiang, K. Johansen, Y.-H. Tu, M. F. McCabe, Multi-sensor and multi-platform consistency and interoperability between UAV, Planet CubeSat, Sentinel-2, and Landsat reflectance data, GIScience & Remote Sensing 59:1 (2022) 936–958. doi:10.1080/15481603.2022.2083791.

[2] K. Jain, A. Pandey, Calibration of Satellite Imagery with Multispectral UAV Imagery, J Indian Soc Remote Sens 49 (2021) 479–490. doi: 10.1007/s12524-020-01251-z.

[3] E. Alvarez-Vanhard, T. Corpetti, T. Houet, UAV & satellite synergies for optical remote sensing applications: A literature review, Science of Remote Sensing 3 (2021) 100019. doi:10.1016/j.srs.2021.100019.

[4] K. Fukunaga, Introduction to Statistical Pattern Recognition., New York, Academic Press, 1990.

[5] J. Sigurdsson, S.E. Armannsson, M.O. Ulfarsson, J. R. Sveinsson, Fusing Sentinel-2 and Landsat 8 Satellite Images Using a Model-Based Method, Remote Sensing 14(13) (2022) 3224. doi: 10.3390/rs14133224.

[6] H. Yao, R. Qin, X. Chen Unmanned Aerial Vehicle for Remote Sensing Applications – A Review, Remote Sensing 11(12) (2019) 1443. doi:10.3390/rs11121443.

[7] D. Lee, D.J. Hess, M.A. Heldeweg, Safety and privacy regulations for unmanned aerial vehicles: A multiple comparative analysis, Technology in Society 71 (2022) 102079. doi:10.1016/j.techsoc.2022.102079.

[8] M. Mehmood, A. Shahzad, B. Zafar, A. Shabbir, N. Ali, Remote Sensing Image Classification: A Comprehensive Review and Applications, Mathematical Problems in Engineering 2022 (2022) 5880959. doi:10.1155/2022/5880959.

[9] F. Bianconi, A. Fernández, R.E. Sánchez-Yáñez, Special Issue Texture and Color in Image Analysis, Applied Sciences 11(9) (2021) 3801. doi:10.3390/app11093801.

[10] V.S. Bhagat, A.K. Kadam, S. Kumar, Analysis of Remote Sensing based Vegetation Indices (VIs) for Unmanned Aerial System (UAS): A Review, Remote Sens. Land 3 (2020) 58–73. doi:10.21523/GCJ1.19030202.

[11] L. Li, T. Xu, Y. Chen, Fuzzy Classification of High Resolution Remote Sensing Scenes Using Visual Attention Features. Comput Intell Neurosci 2017 (2017) 9858531. doi: 10.1155/2017/9858531.

[12] W. Chen, C. He, C. Ji, M. Zhang, S. Chen, An improved K-means algorithm for underwater image background segmentation, Multimedia Tools and Applications 80 (2021) 21059–21083. doi:10.1007/s11042-021-10693-7.

[13] H. Bi, H. Tang, G. Yang, H. Shu, J.-L. Dillenseger, Accurate image segmentation using Gaussian mixture model with saliency map, Pattern Analysis and Applications 21 (2018) 869–878. doi:10.1007/s10044-017-0672-1.

[14] M. Alam, J. Wang, G. Cong, Y. Lv, Y. Chen, Convolutional Neural Network for the Semantic Segmentation of Remote Sensing Images, Mobile Networks and Applications 26 (2021) 200–215. doi:10.1007/s11036-020-01703-3.

[15] O.V. Tsymbal, V.V. Lukin, N.N. Ponomarenko, A.A. Zelensky, K.O. Egiazarian, J.T. Astola, Three-State Locally Adaptive Texture Preserving Filter for Radar and Optical Image Processing, EURASIP J. Adv. Signal Process 2005 (2005) 851609. doi:10.1155/ASP.2005.1185.

[16] C.C. Hung, E. Song, Y. Lan, Image Texture, Texture Features, and Image Texture Classification and Segmentation, in: Image Texture Analysis, Springer, Cham, 2019, pp. 3–14. doi:10.1007/978-3-030-13773-1_1.

[17] V.V. Asmus, A.A. Buchnev, V.A. Krovotyntsev, V.P. Pyatkin, G.I. Salov, Planetamonitoring Software Complex in Applied Remote Sensing Problems, in: Optoelectron. Instrument. Proc. 54, 2018, pp. 222–229. doi:10.3103/S8756699018030020.

[18] A. Popov, O. Pogrebnyak, Radar Target Recognition by Probabilistic Filtering, in: Proc. SPIE, Proceedings Volume 5542, 2004, pp. 459–467. doi:10.1117/12.558627.

[19] N. Johnson, Systems of Frequency Curves Generated by Methods of Translation, Biometrika 36(1/2), (1949) 149–176. doi: 10.2307/2332539.

[20] S.K.K. Hari, S. Rathinam, S. Darbha, K. Kalyanam, S.G. Manyam, D. Casbeer, Optimal UAV Route Planning for Persistent Monitoring Missions, in: IEEE Transactions on Robotics, volume 37, no. 2, pp. 550–566. doi:10.1109/TRO.2020.3032171.

[21] I. Kliushnikov, H. Fesenko, V. Kharchenko, O. Illiashenko, O. Morozova. UAV Fleet Based Accident Monitoring Systems with Automatic Battery Replacement Systems: Algorithms for Justifying Composition and Use Planning, International Journal of Safety and Security Engineering, volume 11, no. 4, 2021, pp. 319–328. doi:10.18280/ijsse.110404.

[22] V. Lukin, I. Vasilyeva, S. Krivenko, F. Li, S. Abramov, O. Rubel, B. Vozel, K. Chehdi, K. Egiazarian, Lossy Compression of Multichannel Remote Sensing Images with Quality Control, Remote Sensing 12 (2020) 3840. doi: 10.3390/rs12223840.

[23] V. Lukin, G. Proskura, I. Vasilyeva, Improvement of Multichannel Image Classification by Combining Elementary Classifiers, in: Proc. Internat. Scientific-Practical Conf. on Problems of Infocommunications. Science and Technology, Kiev, Ukraine, 2019, pp. 666-670. doi:10.1109/PICST47496.2019.9061380.

[24] X. Pan, J. Zhao, J. Xu, An End-to-End and Localized Post-Processing Method for Correcting High-Resolution Remote Sensing Classification Result Images, Remote Sensing 12(5) (2022) 852. doi:10.3390/rs12050852.

[25] X. Cheng, H. Liu, A Novel Post-Processing Method Based on a Weighted Composite Filter for Enhancing Semantic Segmentation Results. Sensors (Basel) 20(19) (2020) 5500. doi:10.3390/s20195500.