

Approach to the problem of the parameters optimization of the shooting system

L A Demidova^{1,2,*}, V A Sablina², and Y S Sokolova²

¹Moscow Technological Institute, Moscow, Russia

²State Radio Engineering University, Ryazan, Russia

*E-mail: demidova.liliya@gmail.com

Abstract. The problem of the objects identification on the base of their hyperspectral features has been considered. It is offered to use the SVM classifiers' ensembles, adapted to specifics of the problem of the objects identification on the base of their hyperspectral features. The results of the objects identification on the base of their hyperspectral features with using of the SVM classifiers have been presented.

1. Introduction

The problem of analysis of the hyperspectral information formed on the base of the hyperspectral images of the Earth's surface is one of the actual problems solved by the remote sensing systems [1, 2]. Also, the problem of the service development for search, indexing, cataloging and distribution of aerospace images should be solved simultaneously with the problem mentioned above [3, 4]. The russian „Resource-P” spacecrafts No. 1 – 3 with the hyperspectral equipment on the board give out the hyperspectral image (HSI) in the form of snapshots in 130 narrow adjoining ranges of visible area of spectrum. These snapshots are subjected to radiometric and atmospheric correction procedures, then the images are segmented. During the HSI processing much attention is paid to the questions of the objects identification on the base of their hyperspectral features. The datasets of the hyperspectral features (HSF) is formed on the basis of the segmented images.

Other paragraphs are indented (BodytextIndented style).

Every dataset is the text file containing information on the wavelength of the recorded radiation and the spectral reflection coefficient (SRC) (or the spectral brightness coefficient (SBC)) of the object. However, at the solution of the problem of the objects identification on the base of their HSF the use of the SRC dependence on wavelength is more preferable because the brightness feature doesn't depend on the shooting conditions in such degree as the SBC. Besides, unlike the SBC, for obtaining the SRC values the standard reflecting surface in sight of the analyzed object isn't necessary.

The HSF of the earth's surface object can be represented graphically in the form of a spectral curve reflecting the relationship between the wavelength and the SRC values of the analyzed object [1].

Figure 1 shows the main stages of the HSF obtaining of the analyzed object based on the HSI. Also, figure 1 shows the fragment of the text file with the object's HSF.

The Support Vector Machine (SVM) algorithm is the supervised machine learning algorithm. This algorithm is one of the boundary classification algorithms [5, 6]. Nowadays, it is used to solve different classification problems in various applications with great success.



The SVM classifier is used for training, testing, and classification. Satisfactory quality of training and testing allows using the SVM classifier to classify new objects. Choosing optimal parameters' values for the SVM classifier is a very significant problem. It is necessary to find the kernel function type, values of the kernel function parameters and value of the regularization parameter [5, 6]. The found optimal parameters' values participate in the formation of the classifying function $F(x)$, which compares the object to the class with the label from the set $\{-1; +1\}$. It is difficult to provide implementing of high-accuracy data classification with the use of the SVM classifier without adequate solution to this problem.

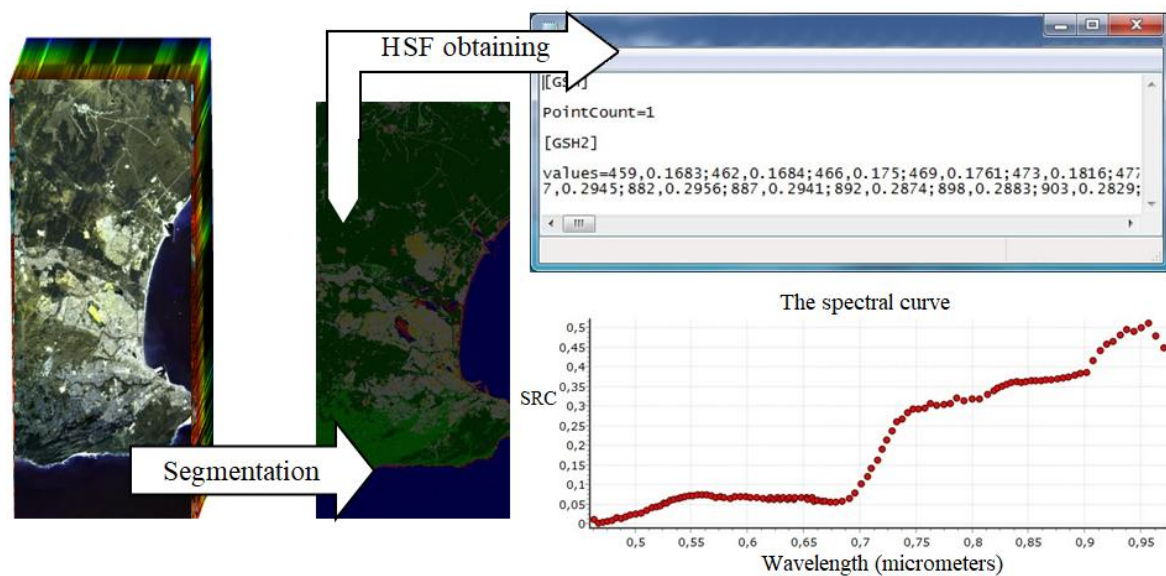


Figure 1. The main stages of the HSF obtaining for the analyzed object

Nowadays, the interest in the issue of increasing the accuracy of classifiers based on the machine learning algorithms by combining the capabilities of several classifiers and creating the classifiers' ensembles has increased significantly [7 – 9]. The training of the classifiers' ensemble is understood as the procedure of training of the final set of the private classifiers, whose private solutions are then combined to form the resulting classification decision on the basis of the aggregated classifier.

We propose to use the SVM classifiers' ensembles, formed on the basis of various strategies of the private classifiers' integration, to solve the identification problems of the objects on the basis of their HSF according to the HSI in order to improve the classification accuracy.

2. SVM algorithm

Let the experimental dataset be a set in the form of $\{(z_1, y_1), \dots, (z_s, y_s)\}$, in which each object $z_i \in Z$ ($i = \overline{1, s}$; s is the number of objects) is assigned to a number $y_i \in Y = \{+1; -1\}$ having a value of $+1$ or -1 depending on the class of the object z_i . It is assumed that every object z_i is mapped to q -dimensional vector of numerical values of features i (typically normalized by values from the interval $[0, 1]$) where z_i^l is the numeric value of the l -th feature for the i -th object ($i = \overline{1, s}$; $l = \overline{1, q}$) [5, 7]. It is necessary to use the kernel function $\kappa(z_i, z_\tau)$ to build the classifier $F: Z \rightarrow Y$, which compares the class to the number from the set $Y = \{+1; -1\}$ or some object from the set Z . In training of the SVM classifier it is necessary to determine the kernel function type $\kappa(z_i, z_\tau)$, values of the kernel parameters and value of the regularization parameter C , which allows finding a compromise between

maximizing of the gap separating the classes and minimizing of the total error [5 – 9]. A herewith typically one of the following functions is used as the kernel function $\kappa(z_i, z_\tau)$: linear function; polynomial function; radial basis function; sigmoid function [5, 6].

To build “the best” SVM classifier it is necessary to implement the numerous repeated training (for the training data set with S elements) and testing (for the test data set $s - S$ elements, $S < s$) on the different randomly generated training and test sets with following determination of the best SVM classifier in terms of the highest possible classification quality provision. The SVM classifier with satisfactory training and testing results can be used to classify new objects [5 – 9].

As a result of the training, the classification function is determined in the following form [5 – 7]:

$$f(z) = \sum_{i=1}^S \lambda_i \cdot y_i \cdot \kappa(z_i, z) + b. \quad (1)$$

The classification decision, associating the object z to the class -1 or $+1$, is adopted in accordance with the rule [5 – 7]:

$$F(z) = \text{sign}(f(z)) = \text{sign}\left(\sum_{i=1}^S \lambda_i \cdot y_i \cdot \kappa(z_i, z) + b\right). \quad (2)$$

3. The SVM ensembles

In most cases SVM classifier provides high quality of data classification. In exceptional cases the SVM ensembles can be used to increase the classification accuracy. The using of the SVM ensemble allows fulfilling the high-precision data classification, especially Big Data classification, with the acceptable time expenditures.

After training, each classifier generates its own (private) classification decisions, same or different from the actual results of classification. Accordingly, the different private SVM classifiers correspond to the different classification accuracy. The quality of the received classification decisions can be improved on the base of ensembles of the SVM classifiers [7 – 9]. In this case, the finite set of privately trained classifiers must be learned. Then the classification decisions of these classifiers are combined. The resulting solution is based on the aggregated classifier. The majority vote method and the vote method based on the degree of reliability can be used as the rules (strategies) of the definition of the aggregated solutions.

The majority vote method is one of the most common and frequently used methods for combining decisions in the ensemble of classifiers. But this method does not fully use the information about the reliability of each private SVM classifier. For example, suppose that the SVM classifier ensemble aggregates the results of five private SVM classifiers, where values of the function $f(z)$ of the object z (1) obtained from the three private SVM classifiers, are negative (class -1), but very close to the neutral position, and values of the function $f(z)$ of the other two SVM classifiers are strongly positive (class $+1$), i.e. very far away from the neutral position. Then the result of the aggregated decision of the ensemble on the basis of “one classifier – one vote” is following: the object z belongs to the negative class (majority vote), although it is obvious, that the best and more appropriate choice for the object z is a positive class. Despite the good potential of the majority vote method for combining of the group of decisions, it is recommended to use other methods to increase the accuracy of classification.

Vote method based on the degree of reliability uses value of the function $f(z)$ for the object z obtained by each private SVM classifier. The greater the positive value of $f(z)$ in (1) returned by the SVM classifier, the more precisely the object z is determined in class $+1$, and the less negative value $f(z)$, the more precisely the object z is defined in class -1 . Values “ -1 ” and “ $+1$ ” for $f(z)$ indicate that the object z is situated on the boundary of the negative and positive classes, respectively.

When using an ensemble of classifiers for solving classification problems special attention should be paid to the methods of forming a set of private classifiers, which can later be used in the development of the final SVM classifier. It is experimentally confirmed [7 – 9], that the ensemble of

classifiers shows better accuracy than any of its private members, if private classifiers are accurate and varied. Therefore, the formation of the set of the private SVM classifiers is required: 1) to use the various kernel functions; 2) to build classifiers in the different ranges of change of the kernel parameters and regularization parameter; 3) to use various sets of training and test data. To select the appropriate members of the ensemble in the set of the trained SVM classifiers it is recommended to use the principle of maximum decorrelation. In this case the correlation between the selected classifications should be as small as possible. After training, each private j -th classifier from the k trained classifier will correspond to a certain array of errors: $e_{ij} = |y_{ij} - \tilde{y}_{ij}|$, where e_{ij} is the error of j -th classifier at i -th row of the experimental data set ($i = \overline{1, s}; j = \overline{1, k}$); y_{ij} is the classification decision (-1 or $+1$) of j -th classifier at i -th row of the experimental data set; y_{ij} is the real meaning of a class (-1 or $+1$), for which the i -th object is belong to.

The SVM classifiers not permitting an error on the experimental data set should be excluded from further consideration and from the remaining quantity of the SVM classifiers. It is necessary to select an appropriate number of private SVM classifiers with maximal variety. To solve this problem decorrelation maximization algorithm can be used. This algorithm provides a variety of private SVM classifiers, being used in the construction of the ensemble [7]. If the correlation between the selected classifiers is small, then the decorrelation is maximal.

Let there be an error matrix E of set of private SVM classifiers with size $s \times k$:

$$E = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1k} \\ e_{21} & e_{22} & \dots & e_{2k} \\ \dots & \dots & \dots & \dots \\ e_{s1} & e_{s2} & \dots & e_{sk} \end{bmatrix}, \quad (3)$$

where e_{ij} is the error of the j -th classifier at the i -th row of the experimental data set ($i = \overline{1, s}; j = \overline{1, k}$).

On the basis of the error matrix E (3) the following assessments can be calculated [7]:

- mean:

$$\bar{e}_j = \frac{1}{s} \sum_{i=1}^s e_{ij} \quad (j = \overline{1, k}); \quad (4)$$

- variance:

$$V_{jj} = \frac{1}{s} \sum_{i=1}^s (e_{ij} - \bar{e}_j)^2 \quad (j = \overline{1, k}); \quad (5)$$

- covariance:

$$V_{jt} = \frac{1}{s} \sum_{i=1}^s (e_{ij} - \bar{e}_j) \cdot (e_{it} - \bar{e}_t) \quad (j = \overline{1, k}; t = \overline{1, k}); \quad (6)$$

Then the elements r_{ij} of the correlation matrix with size $k \times k$ are calculated as:

$$r_{ij} = V_{ij} / \sqrt{V_{ii} \cdot V_{jj}}; \quad (7)$$

where r_{ij} is the correlation coefficient, representing the degree of correlation of t -th and j -th classifiers ($j = \overline{1, k}; t = \overline{1, k}$); $r_{jj} = 1$ ($j = \overline{1, k}$).

Using the correlation matrix R it is possible for each private j -th classifier to calculate the plural-correlation coefficient ρ_j , which characterizes the degree of correlation of the j -th and all other $(k-1)$ classifiers with numbers t ($t = \overline{1, k}; t \neq j$) [7]:

$$\rho_j = \sqrt{1 - |R|/R_{jj}} \quad (j = \overline{1, k}), \quad (8)$$

where $|R|$ is the determinant of the correlation matrix R ; R_{jj} is the cofactor of the element r_{jj} of the correlation matrix R .

A quantity ρ_j^2 called the coefficient of determination. It shows the proportion of the variation of the analyzed variable, which is explained by variation of the other variables. The coefficient of determination ρ_j^2 can take values from 0 to 1. The closer the coefficient to 1, the stronger the relationship between the analyzed variables (in this case, between private classifiers) [7]. It is believed that there is a dependency, if the coefficient of determination is not less than 0.5. If the coefficient of determination greater than 0.8, it is assumed that high dependence exists.

For selection of private SVM classifiers for integration into the ensemble it is necessary to determine the threshold θ . Thus, the j -th private classifier must be removed from the list of classifiers if the coefficient of determination ρ_j^2 satisfies to condition $\rho_j^2 > \theta$ ($j = \overline{1, k}$). If it is necessary to identify the most various classifiers, generating decisions with the most different arrays of errors on the experimental data set, thresholds θ , satisfying to condition $\theta < 0.7$ should be selected. The additional considerations can be also taken into account to avoid the exclusion of insufficient or excessive number of private SVM classifiers.

The decorrelation maximization algorithm can be summarized into the following steps [7].

Step 1. To calculate the matrix V and the correlation matrix R with formulas (5), (6) and (7) respectively.

Step 2. To calculate the multiple correlation coefficients ρ_j ($j = \overline{1, k}$) with (18) for all classifiers.

Step 3. To remove classifiers, for which $\rho_j^2 > \theta$ ($j = \overline{1, k}$), from the list of classifiers.

Step 4. To repeat iteratively steps 1 – 3 for the remaining classifiers in the list until for all classifiers the condition $\rho_j^2 \leq \theta$ ($j = \overline{1, k}$) will not satisfied.

As a result, the list of classifiers used to form the ensemble will consist of m ($m \leq k$) private classifiers.

For classifiers selected in the ensemble, it is necessary to carry out:

- the rationing of degrees of the reliability;
- the strategy search for the integration of members of the ensemble;
- the calculation of the aggregated decision of the ensemble.

Value of the reliability $f_j(z)$, which is defined for the object z by the j -th classifier, falls into the interval $(-\infty, +\infty)$. The main drawback of such values is that in the ensemble the private classifiers with large absolute value are often dominated in the final decision of the ensemble. To overcome this drawback, the rationing is carried out: the transformation of values of degrees of reliability in the interval $[0; 1]$ is fulfilled. In the case of binary classification in the rationalization for the object z the values of the reliability of its membership to positive class (labeled +1) $g_j^+(z)$ and to negative class $g_j^-(z)$ are determined. These values can be determined by the formulas [7]:

$$g_j^+(z) = \frac{1}{1 + e^{-f_j(z)}}, \quad (9)$$

$$g_j^-(z) = 1 - g_j^+(z). \quad (10)$$

The selected private classifiers are combined into the ensemble using $g_j^+(z)$ and $g_j^-(z)$ ($j = \overline{1, m}$) in accordance with one of the following five strategies [7].

1. Maximum strategy:

$$A(z) = \begin{cases} 1, & \text{if } \max_{j=1,m} g_j^+(z) \geq \max_{j=1,m} g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \quad (11)$$

2. Minimum strategy:

$$A(z) = \begin{cases} 1, & \text{if } \min_{j=1,m} g_j^+(z) \geq \min_{j=1,m} g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \quad (12)$$

3. Median strategy:

$$A(z) = \begin{cases} 1, & \text{if } \frac{1}{m} \sum_{j=1}^m g_j^+(z) \geq \frac{1}{m} \sum_{j=1}^m g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \quad (13)$$

4. Mean strategy:

$$A(z) = \begin{cases} 1, & \text{if } \sum_{j=1}^m g_j^+(z) \geq \sum_{j=1}^m g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \quad (14)$$

5. Product strategy:

$$A(z) = \begin{cases} 1, & \text{if } \prod_{j=1}^m g_j^+(z) \geq \prod_{j=1}^m g_j^-(z), \\ -1, & \text{otherwise.} \end{cases} \quad (15)$$

The value $A(z)$ is an aggregated measure of the reliability's value of the SVM classifier ensemble. It can be used to integrate the members of the ensemble [7].

The learning algorithm of the ensemble of the SVM classifiers can be summarized into the following steps.

Step 1. To divide the experimental data set into k training data sets: TR_1, \dots, TR_k .

Step 2. To learn k private SVM classifiers with the different training data sets TR_1, \dots, TR_k and to obtain k private SVM classifiers (ensemble members).

Step 3. To select m ($m \leq k$) SVM classifiers from k classifiers using the decorrelation maximization algorithm.

Step 4. To determine values of m classification functions for each selected private SVM classifier: $f_1(z), \dots, f_m(z)$.

Step 5. To transform values of degrees of reliability, using (19) and (20), for the positive class $g_1^+(z), \dots, g_m^+(z)$ and for the negative class $g_1^-(z), \dots, g_m^-(z)$.

Step 6. To determine the aggregated value $A(z)$ of the reliability of the SVM classifier ensemble using (11) – (15).

This algorithm, used for the weak SVM classifiers, will provide a better quality of the classification accuracy than accuracy of any single private classifier used for aggregation.

The problem of choosing of the threshold θ is very important. Value θ for which all five rules of classification (11) – (15) show stable improvement of the classification quality must be chosen as the threshold value θ^* ($\theta^* < 0.7$). Thus the use of each of the five rules leads to improvement of the classification quality resulting in the reduction of the number of erroneous decisions, when the smaller number of private classifiers, corresponding to the threshold value θ^* , is applied. Such stable $\theta' > \theta^*$ improvement of the classification quality isn't observed for all examined values θ' (for which).

It should be noted, that the majority vote rule may be used for decisions, obtained using the classification rules (11) – (15), to determine the required threshold value θ^* .

4. Experimental studies

The proposed algorithms were combined into the united software named as “Intellectual Classification”, which allows fulfilling the intellectual classification of objects using the SVM algorithm and providing the ability of combining of the several private classifiers into an ensemble to improve the classification quality.

In the course of research, the identification of objects according to the hyperspectral shooting data was made. In particular, the detection problem of water objects and anthropogenic ones was solved.

The base of spectral standards containing 220 HSFs of natural and artificial objects was formed on the basis of the real hyperspectral shooting data from the „Resource-P” spacecraft No. 1. Two datasets such as AqwaObj and AntroObj have been created on the basis of these spectral standards.

The text file with the HSF (figure 1) contains information represented as the two-element tuples separated by a semicolon. In each tuple, the first number is the wavelength (nm) in the form (it is integer number), the second number is the value of the SRC corresponding to the wavelength, represented as the real number from the range [0; 1] with 4 bits of the fractional part, separated from the entire part by a dot. The data in the tuple is separated by a comma. The tuples are ordered by increasing of the wavelength values.

Firstly, all standard HSFs were combined into the common dataset of 220 objects with 127 wavelengths (the SRCs of all objects for these wavelengths were different from zero). The wavelengths were distributed in the range from 401 to 971 nm with non-uniform step. The AqwaObj dataset was created from this dataset by specifying the membership class for the uploaded HSF. The membership class of the real water object was designated as the “class 1”, the class of any other object was designated as the “Class 2”. Similarly, the AntroObj dataset was created.

The AqwaObj dataset is the sample of 220 HSFs, 53 of which belong to water objects (class 1), the remaining 167 HSFs belong to objects of different nature (class 2).

The AntroObj dataset is the sample of 220 HSFs, 80 of which belongs to anthropogenic objects (class 1), the remaining 140 HSFs belong to objects of different nature (class 2).

To develop the SVM classifier, it is necessary: to determine the value of the regularization parameter, which allows finding a compromise between the maximization of the strip separating the classes and minimizing of the overall error of the SVM classifier; to specify the kernel function; to determine the parameters values of the kernel function; to form the training and test sets.

The classification quality can be estimated using the following indicators: overall accuracy $Accur$; sensitivity Se ; specificity Sp ; F_1 -measure; AUC , calculated at the test set (AUC_{test}); a number of errors of the I-st and II-nd type (Er_I and Er_{II}), number of errors at the training and test sets (Er_{train} and Er_{test}).

The kernels with polynomial, radial basis, sigmoid and linear kernel functions with the parameters values set by default were included in the search. The test set was randomly generated and contained 20% of objects from the initial dataset. The remaining 80% of objects from the initial dataset were included in the training set.

Table 1 shows the results of the SVM classifiers development. For each dataset the values of $Accur$ (overall accuracy), F_1 -measure, the overall number of errors of the classifier that showed the best classification results (the highest values of $Accur$ and F_1 -measure, the smallest number of errors at the training and test sets) are selected.

Table 1 shows that even at the kernel parameters values set by default the SVM algorithm allows to construct the classifier with the overall classification accuracy exceeding 90% for all datasets that demonstrates its high generalizing ability. Moreover, for the AqwaObj dataset the overall classification accuracy is equal to 98.64%, therefore, there is no need for increase in accuracy of classification and search of optimum parameters values for this dataset.

For the AntroObj dataset we developed 18 private SVM classifiers with the radial basis function rbf and the different parameters values of C and σ . The different training and test sets randomly

generated from the original data set were used. The different training and test sets randomly generated from the original data set were used. The values of parameters and quality indicators of 18 private classifiers have been shown in the table 2. At the training for each private SVM classifier the training set was formed in a random way on the basis of the initial experimental dataset. The number of objects in the test set was equal to 10% – 30% of the initial number of objects in the initial experimental dataset.

Table 1. The results of the SVM classifier development with the values of the kernel function parameters set by default

Dataset ($s \times q$) (Train/Test)	The kernel Function type	The quality assessment of the SVM classifier										
		Accur, %	F_1 , %	Se , %	Sp , %	AUC_{test} , %	SV	The number of errors				
								Er_{train}	Er_{test}	Er_I	Er_{II}	Overall sum
AqwaObj (220×127) (176/44)	<i>lin</i>	95.91	91.59	92.45	97.01	0.9339	38	8	1	4	5	9
	<i>poly</i>	98.64	97.20	98.11	98.80	0.9181	31	0	3	1	2	3
	<i>rbf</i>	98.18	96.23	96.23	98.80	0.9256	148	0	4	2	2	4
	<i>mlp</i>	6.06	71.82	3.77	93.41	0.3281	31	48	14	51	11	62
AntroObj (220×127) (176/44)	<i>lin</i>	71.82	68.37	83.75	65.00	0.8047	143	53	9	13	49	62
	<i>rbf</i>	90.45	87.72	93.75	88.57	0.7294	154	10	11	5	16	21
	<i>mlp</i>	64.09	42.34	36.25	80.00	0.3945	80	62	17	51	28	79

At the testing it was found, that the private classifiers indicate the classification accuracy ranged from 84.09% to 92.73%, and the initial values of the determination coefficient, calculated for all 18 private classifiers, are in the range from 0.615 to 0.925. As a result, the threshold values θ were examined from the range [0.60; 0.95] with step 0.05. Values of the classification parameters corresponding to the different threshold values θ for the AntroObj dataset are given in the table 3.

Table 2. The values of parameters and quality indicators of the private classifiers (AntroObj dataset)

Classifier number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
C	1	1	1	1	0.5	0.5	0.5	1.5	1.5	1.5	1.5	1.5	2	2	2	2	2	0.5
σ	1	1.5	2	2.5	1	1.5	2	0.5	1	1.5	2	2.5	0.5	1	1.5	2	2.5	0.5
Accur	89.09	87.73	87.73	84.09	87.27	85.91	85.00	91.82	92.27	89.09	87.27	86.36	90.45	92.73	90.00	89.09	88.18	90.45
SV	141	141	140	137	163	147	158	151	146	127	135	136	157	159	140	131	129	184
Train	154	176	187	198	176	165	198	154	165	165	187	198	165	187	176	187	198	187
Er_{train}	8	15	19	27	19	19	26	0	3	13	20	24	0	6	12	15	21	4
Test	66	44	33	22	44	55	22	66	55	55	33	22	55	33	44	33	22	33
Er_{test}	16	12	8	8	9	12	7	18	14	11	8	6	21	10	10	9	5	17
Se	86.25	86.25	90.00	91.25	97.50	88.75	93.75	83.75	90.00	91.25	91.25	92.50	76.25	92.50	93.75	88.75	91.25	98.75
Sp	90.71	88.57	86.43	80.00	81.43	84.29	80.00	96.43	93.57	87.86	85.00	82.86	98.57	92.86	87.86	89.29	86.43	85.71
Er_I	13	16	19	28	26	22	28	5	9	17	21	24	2	10	17	15	19	20
Er_{II}	11	11	8	7	2	9	5	13	8	7	7	6	19	6	5	9	7	1
AUC_{test}	0.804	0.779	0.819	0.719	0.883	0.853	0.641	0.785	0.775	0.789	0.698	0.752	0.766	0.708	0.870	0.808	0.925	0.615

θ_0 0.451 0.606 0.765 0.787 0.608 0.689 0.833 0.211 0.381 0.682 0.679 0.707 0.175 0.399 0.585 0.605 0.751 0.160

Table 3. The values of the classification quality indicators for the different thresholds of the determination coefficient (AntroObj dataset)

Classification quality indicator	Strategy	The threshold value of the determination coefficient					
		0.95-0.85	0.8	0.75	0.7	0.65	0.6
<i>Accur</i> , %	Majority vote	91.82	91.82	91.82	91.82	91.82	91.82
	Maximum and minimum	96.36	96.36	96.36	96.82	97.73	98.64
	Median and sum	92.27	92.27	94.09	95.45	96.82	98.64
	Product	92.27	92.27	94.09	95.45	96.82	98.64
<i>Se</i> , %	Majority vote	97.50	97.50	97.50	97.50	97.50	97.50
	Maximum and minimum	96.25	96.25	96.25	96.25	97.50	98.75
	Median and sum	95.00	95.00	97.50	97.50	98.75	100
	Product	95.00	95.00	97.50	97.50	98.75	100
<i>Sp</i> , %	Majority vote	88.57	88.57	88.57	88.57	88.57	88.57
	Maximum and minimum	96.43	96.43	96.43	97.14	97.86	98.57
	Median and sum	90.71	90.71	92.14	94.29	95.71	97.86
	Product	90.71	90.71	92.14	94.29	95.71	97.86
<i>Er_I</i>	Majority vote	16	16	16	16	16	16
	Maximum and minimum	5	5	5	4	3	2
	Median and sum	13	13	11	8	6	3
	Product	13	13	11	8	6	3
<i>Er_{II}</i>	Majority vote	2	2	2	2	2	2
	Maximum and minimum	3	3	3	3	2	1
	Median and sum	4	4	2	2	1	0
	Product	4	4	2	2	1	0
Number of classifiers in the ensemble		18	17	14	13	10	7

The optimal threshold value θ^* for the reviewed example equals to 0.6, since for the threshold value $\theta^* = 0.6$ all five classification rules (11) – (15) give the stable improvement of the classification quality when the number of classifiers reduces to the number corresponding to the threshold value $\theta^* = 0.6$. In this case, the finite number of classifiers in the SVM ensemble is equal to 7. A further decrease in the number of classifiers is not feasible (due to a further sharp decrease in their number and a substantial reduction of their variety).

The use of all five strategies with $\theta^* = 0.3$ allowed classifying correctly 98.64% of the objects of the initial dataset. At the same time, the maximum classification accuracy of one of the private SVM classifiers, used in the SVM ensemble, was equal to 92.73%, and the accuracy reached with the use of the majority vote rule was equal to 91.82%. Thus, the use of the SVM ensemble allowed increasing the classification accuracy almost by 6% compared to the maximum classification accuracy of one of the private classifiers in the SVM ensemble.

5. Conclusions

The SVM ensembles based on the decorrelation maximization algorithm for the different strategies of the decision-making on the data classification and the majority vote rule allow reducing the accident classification decision received by one classifier, and help to improve the classification accuracy. The shortcomings of some private classifiers are compensated by strengths of others private classifiers thanks to combination of their results. Classifiers counterbalance the results' accident of each other, finding the most plausible output classification decision. It allows finding the best classification result with minimum classification error.

The experimental results obtained on the basis of the hyperspectral shooting data confirm, that the SVM ensembles allow increasing the classification quality of objects by their GSFs. However, it is necessary to create the representative dataset of the HSFs and use it to learn the SVM classifiers.

References

- [1] Demidova L A and Truhanov S V 2015 *Contemporary Engineering Sciences* **8** (20) 885.
- [2] Grigorieva O, Brovkina O, Mochalov V, Akhtman Y, Zelentsov V, Potryasaev S, Kozyr I and Belova N 2016 *The 4th International Workshop on Simulation for Energy, Sustainable Development and Environment*.
- [3] Taganov A I, Kolesenkov A I, Babaev S I and Sablina V A (2017) *6th Mediterranean Conference on Embedded Computing (MECO) Proceedings*. Budva, Montenegro pp. 130–133.
- [4] Sablina V A, Novikov A I, Nikiforov M B and Loginov A A (2013) *2nd Mediterranean Conference on Embedded Computing (MECO) Proceedings*, Budva, Montenegro pp.117–120.
- [5] Chapelle O, Vapnik V, Bousquet O and Mukherjee S 2002 *Machine Learning* **46** 131.
- [6] Yu L, Wang S, Lai K K, and Zhou L 2008 *Springer* 244.
- [7] Demidova L, Nikulchev E and Sokolova Yu 2016 *International Journal of Advanced Computer Science and Applications* **7**(5) 294.
- [8] Graf H P, Cosatto E, Bottou L, Durdanovic I and Vapnik V 2005 *Advances in Neural Information Processing Systems* **17** 521.
- [9] Eastaff M S and Premalatha P 2015 *Proceedings of the UGC Sponsored National Conference on Advanced Networking and Applications* pp. 191–193.