

# Image Classification Accuracy Analysis for Three-channel Remote Sensing Data

Fangfang Li<sup>a</sup>, Vladimir Lukin<sup>a</sup>, Galina Proskura<sup>a</sup>, Irina Vasilyeva<sup>a</sup> and Galina Chernova<sup>a</sup>

<sup>a</sup> National Aerospace University, 17 Chkalova Street, Kharkiv, 61070, Ukraine

## Abstract

Classification accuracy of remote sensing data depends on many factors including level of distortions if lossy compression is applied to original data. However, it is difficult to predict what compression ratio or characteristics of distortions have to be provided in order to ensure classification accuracy reduction due to lossy compression is appropriate for a given image. We show that, under certain conditions, the distortions introduced by many modern techniques of lossy compression can be described as additive white Gaussian noise (AWGN). Then, the task can be divided into two simpler subtasks. First, one needs to understand what is the maximal allowed level (mean square error, peak signal-to-noise ratio) of AWGN to ensure appropriate classification accuracy (or its deterioration). Second, how to provide such level of distortions for a given coder and a given image. The second task has been solved for many existing coders. Hence, the first task is addressed in this paper. We analyze four three-channel images composed of three bands of multispectral Sentinel-2 images. Then, AWGN with different variance values is added. Two methods of image pixel-wise classification are studied. The obtained results show that the classification accuracy depends on a classifier used and image complexity whilst dependence on noise variance is also obvious. Whilst PSNR of compressed images of the order 38 dB has practically no negative impact on classification accuracy, larger distortions might have a considerable negative impact. Several ways to attain a desired PSNR are mentioned.

## Keywords

Image complexity, performance, image lossy compression, noise, classification accuracy

## 1. Introduction

Remote sensing (RS) has become a popular tool for solving many important tasks of environment monitoring, estimation of sensed terrain parameters, control of agricultural plant state [1, 2], etc. One important stage of image processing is image classification [3, 4] that produces maps which can be the final product of remote sensing or a pre-final result if certain characteristics are then determined for objects (fragments) that belong to particular classes.

A general modern tendency in remote sensing is data size increasing. This is due to better spatial resolution of sensors, more frequent observation of terrains, larger number of channels (bands, polarizations) for which images are acquired. This leads to necessity to compress data especially at stages of their downlink transfer from satellite and aerial platforms. As known, there are lossless and lossy image compression techniques [5, 6]. Lossless compression often does not provide high enough compression ratio (CR). Meanwhile, lossy compression allows providing a larger and variable CR but by the expense of introduced distortions. There are many factors that determine how these distortions effect classification. At the first glance, it might be surprising but such situations are possible that lossy compression provides better classification than classification of uncompressed (original, compressed in

---

IntellTISIS'2022: 3rd International Workshop on Intelligent Information Technologies and Systems of Information Security, March 23–25, 2022, Khmelnytskyi, Ukraine

EMAIL: liff\_niat@yahoo.com (Fangfang Li); lukin@ai.kharkov.com (Vladimir Lukin); g.proskura@khai.edu (Galina Proskura); i.vasileva@khai.edu (Irina Vasilyeva); k505@ai.kharkov.com (Galina Chernova)

ORCID: 0000-0002-7392-586X (Fangfang Li); 0000-0002-1443-9685 (Vladimir Lukin); 0000-0001-8960-0421 (Galina Proskura); 0000-0002-1378-1104 (Irina Vasilyeva)



© 2022 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)

a lossless manner) data [7, 8]. Meanwhile, in most cases, lossy compression results in worse classification than for the corresponding uncompressed images. Thus, one needs to find a reasonable compromise between parameters of lossy compression (CR, metrics that characterize compressed image quality, resources and time spent on attaining a desired quality) and acceptable degradation of classification characteristics (accuracy).

Note that known results of studies carried out so far [5-10] allow expecting that classification accuracy depends upon a great number of factors: 1) properties of original images (number of channels, spectral range of sensing, total PSNR and particular PSNRs in component images); 2) properties of a sensed scene including the number of classes, mean size of objects, etc.; 3) classifier properties as type, used features, methodology of its learning or design, etc. This means that a complex research has to be carried out to answer the question on how to provide an appropriate quality of compressed images.

To get some preliminary answers, we perform a research for three-channel images which are the simplest variant of multichannel RS data. Pixel-wise classification is applied. Maximal likelihood (ML) and neural network (NN) based classifiers are considered. The novelty consists in methodology of simulating distortions introduced by lossy compression. We show that for several methods of image compression the distortions have distribution close to Gaussian and spatially they are almost uncorrelated. This allows simulating them as additive white Gaussian noise without carrying out numerous compression/decompression for obtaining different levels of distortions for the analyzed set of compression techniques. In turn, this allows saving time and obtaining generalized conclusions that, later, can be discussed more in detail for particular image compression techniques.

## **2. Brief analysis of distortions due to lossy compression and related studies**

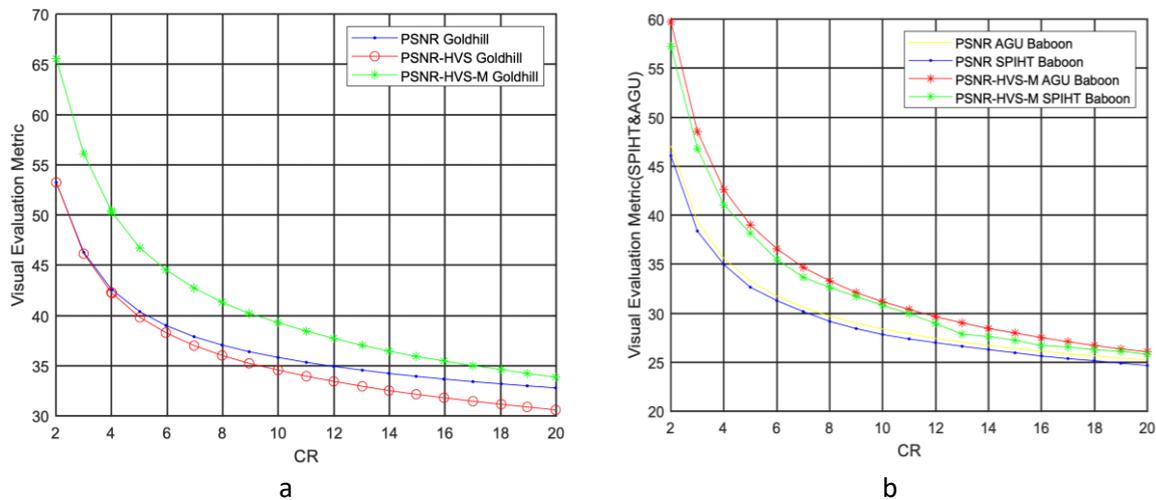
An inherent property of lossy compression is that it introduces distortions irrespectively of an image compression technique used. Meanwhile, statistical and spectral characteristics of such distortions have not been intensively studied yet.

Certainly, there is a lot of studies dealing with analysis of rate distortion curves and comparison of coder performance (see, e.g., [9-12] and references therein). The following main tendencies are demonstrated there: 1) image quality (characterized in terms of standard metrics as mean square error (MSE) and peak signal-to-noise ratio (PSNR) or visual quality metrics as SSIM, PSNR-HVS-M, FSIM and so on [13]) decreases if CR increases; 2) rate/distortion curves behave individually, the general tendencies are similar but particular curves can be quite different depending on image properties (complexity) [11]; 3) it might happen that for a given image one coder performs in the best way whilst for another image another coder is the best; 4) it is possible to provide a desired quality of an image subject to lossy compression according to a given quality metric in two iterations (steps) quite accurately [11] and with more iterations very accurately but by the expense of larger time and computational resources; 5) there is a certain correlation between conventional metrics that describe lossy compression and classification accuracy; however, this correlation is not strict and it is not perfectly established yet; it is mainly shown that lossy compression with a rather small compression ratio (CR) has a little impact on classification while a larger CR leads to considerable degradation [6].

Thus, let us start from considering available information on statistical and spectral characteristics of introduced distortions. For this purpose, it can be useful to analyze behavior of three rate/distortion curves for the coder AGU (<http://ponomarenko.info/agu.htm>). This is the DCT based coder that has relatively good performance (better than for JPEG and JPEG2000) due to three peculiarities: use of 32x32 pixel blocks, better coding of quantized DCT coefficients, and embedded deblocking of decompressed images. Metrics that we will jointly use are PSNR, PSNR-HVS (<http://ponomarenko.info/psnrhvs.htm>) and PSNR-HVS-M (<http://ponomarenko.info/psnrhvs-m.htm>) where HVS stands for human vision system and M relates to masking. The visual quality metrics PSNR-HVS and PSNR-HVS-M are organized in such a manner that they are equal to PSNR for the case of AWGN if masking effect is absent. If masking effect is present (it deals with noise masking by textures), then PSNR-HVS-M is larger than PSNR. In turn, PSNR-HVS and PSNR-HVS-M values are smaller than PSNR if noise or distortions are spatially correlated (and, for PSNR-HVS-M, if masking effect is absent or its influence is negligible).

Below we present the curves for the test image Goldhill (Figure 1, a). As one can see, for small CR the values of PSNR and PSNR-HVS are practically the same whilst PSNR-HVS-M is sufficiently larger. This is due to masking effect as well as thanks to the absence of spatial correlation of distortions. If CR is quite large (e.g., about 20) PSNR and PSNR-HVS-M are almost the same (masking effect becomes smaller and introduced errors become spatially correlated), PSNR-HVS is smaller than PSNR (due to spatial correlation of introduced errors).

Figure 1, b shows other examples. They are for two coders, AGU and SPIHT [14], the test image is Baboon which is more complex (textural) than Goldhill. The more complex structure of Baboon image leads to smaller PSNR values than for Goldhill for the same CR (e.g., compare the data for CR=10 in Figures 1, a and 1, b). Again, PSNR and PSNR-HVS-M values become practically equal for large CR. One more observation is that for entire range of CR values the coder AGU outperforms SPIHT according to both metrics (Figure 1, a) but the difference is not large.



**Figure 1:** Rate distortion curves for AGU coder for the test image Goldhill (a) and rate/distortion curves for the coders SPIHT and AGU for the test image Baboon (b)

Statistical and spatial correlation properties of introduced losses can be analyzed in another manner as well. One way is to obtain difference images as  $\Delta(i, j) = I(i, j) - I_c(i, j)$  where  $i = 1, \dots, I_{Im}$ ,  $j = 1, \dots, J_{Im}$ ,  $I_{Im}$  and  $J_{Im}$  denote the image size,  $I(i, j)$  and  $I_c(i, j)$  are original and compressed image values in the  $ij$ -th pixel.

Then, it is possible to carry out different operations of statistical and spectral-correlation analysis of the data array  $\{\Delta(i, j), i = 1, \dots, I_{Im}, j = 1, \dots, J_{Im}\}$ . In particular, it is possible to calculate mean and variance of integer variables  $\{\Delta(i, j), i = 1, \dots, I_{Im}, j = 1, \dots, J_{Im}\}$ , check its Gaussianity or determine skewness and kurtosis. We have carried out such preliminary analysis for several test images and for three values of quantization step (QS) for the AGU coder.

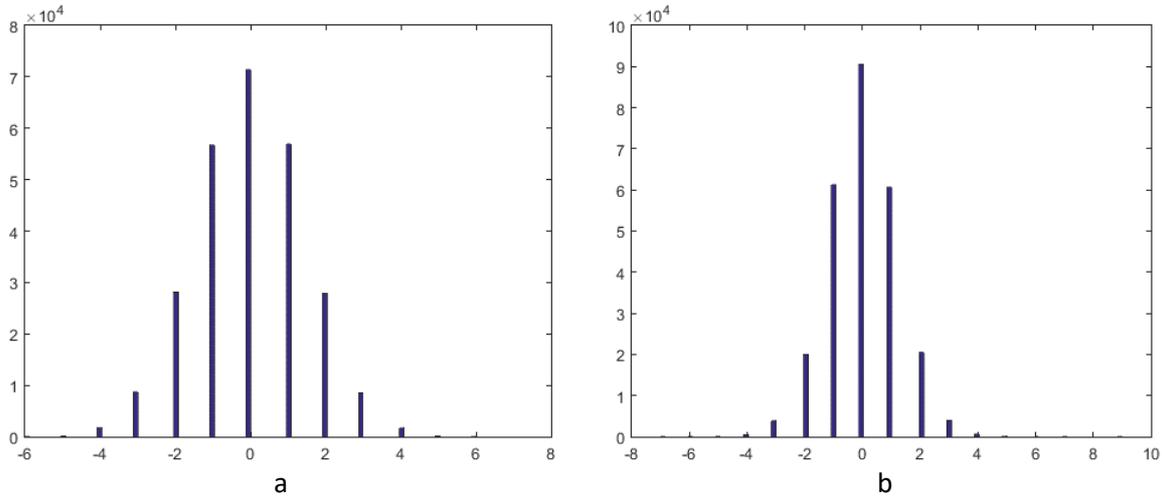
Data are presented in Table 1 for six grayscale test images where the images Baboon and Grass are the most textural and the images Fly and Pole are the simple structure ones.

**Table 1**  
Statistical characteristics of  $\{\Delta(i, j), i = 1, \dots, I_{Im}, j = 1, \dots, J_{Im}\}$ , QS=5

Test image	Statistical characteristics				CR
	Variance	Kurtosis	Skewness	St. deviation	
Baboon	2.12	2.98	-0.0077	1.45	2.22
Boat	1.95	3.00	-0.0062	1.39	3.21
Fly	1.43	3.33	0.0173	1.19	5.46
Grass	2.14	3.02	0.0051	1.46	1.65
Lenna	1.86	3.03	0.0013	1.36	4.06
Pole	1.26	3.58	0.0084	1.12	5.14

Analysis of the obtained data shows the following. First, for complex structure images, the variance of introduced losses is approximately equal to  $QS^2/12$  whilst variance is smaller for simpler structure images. Second, there is a sufficient difference (by almost 3 times) in CR. Third, skewness is almost equal to zero, mean (not shown in Table 1) is practically equal to zero as well. Thus, the distribution is symmetric and it is not surprising. Finally, kurtosis is about 3 for all images except the test image Pole and (in less degree) the test image Fly. Thus, distributions are close to Gaussian except of distributions for simple structure images for which a heavier tail takes the place.

To partly confirm these conclusions, let us present histograms of distributions. Figure 2 shows the histograms for the test images Baboon (a) and Fly (b). They both look close to normal.



**Figure 2:** Histograms of  $\{\Delta(i,j), i=1,\dots,I_{Im}, j=1,\dots,J_{Im}\}$ ,  $QS=5$ , for the images Baboon (a) and Fly (b)

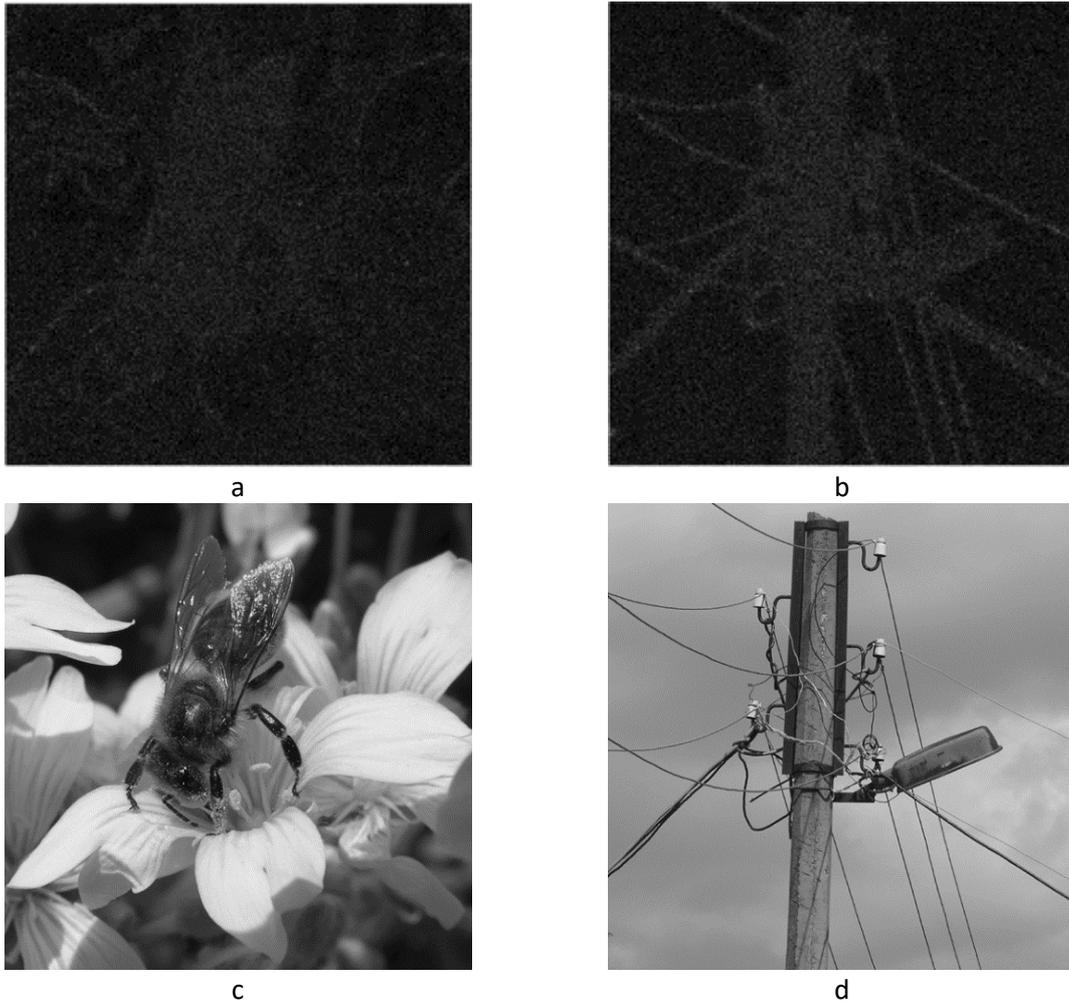
Magnified absolute values of differences is represented as image in Figure 3, a. Below (Figure 3, c) the test image Fly is given. Joint analysis shows that larger variations of differences are observed for locally active areas (pixels that correspond to edges, detail, textures) of the image Fly. The latter property appears itself even more obviously for the difference map for the test image Pole (Figure 3, b) where the more intensive distortions take place for edge and detail neighborhoods (compare the images in Figure 3, b and d, jointly). Analysis of central cross-sections of 2D autocorrelation functions and spectra of differences has shown that differences are practically spatially independent for  $QS=5$ .

Data for  $QS=15$  are collected in Table 2. Their analysis shows the following. First, for complex structure images variance of introduced losses is less than  $QS^2/12$  and variance is considerably smaller than  $QS^2/12$  for simpler structure images. The difference in CR becomes larger (the maximal and minimal values differ by almost 4 times). Second, skewness is close to zero, mean is close to zero, too. Hence, the distribution is close to symmetric. Third, kurtosis is about 3 for the most complex structure images. Meanwhile, for the test image Pole and Fly, the kurtosis is considerably larger than 3. Therefore, distributions can be close to Gaussian and they might also have a heavy tail.

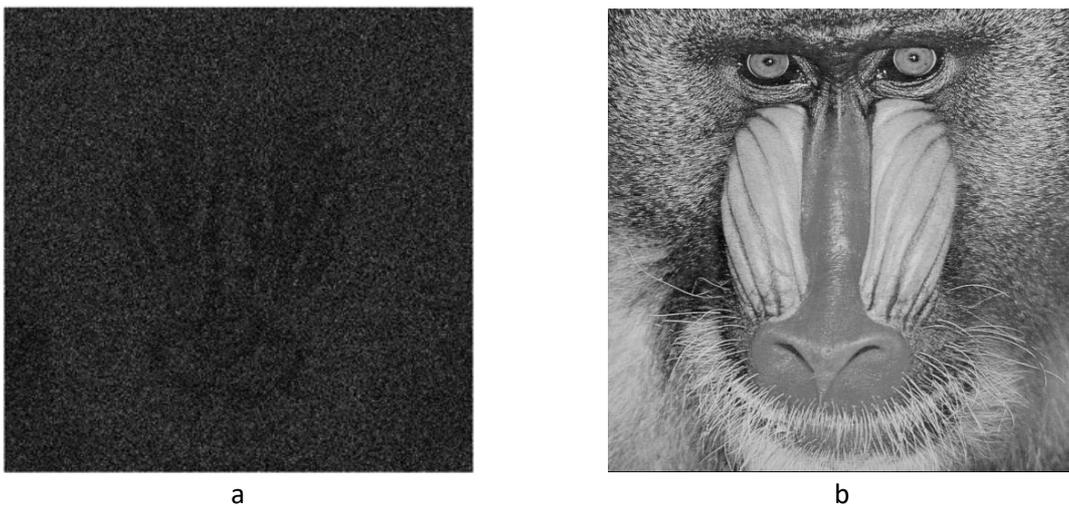
**Table 2**

Statistical characteristics of  $\{\Delta(i,j), i = 1, \dots, I_{Im}, j = 1, \dots, J_{Im}\}$ ,  $QS=15$

Test image	Statistical characteristics				CR
	Variance	Kurtosis	Skewness	St. deviation	
Baboon	16.10	3.08	-0.0022	4.01	3.83
Boat	11.56	3.29	0.0126	3.40	7.37
Fly	5.90	7.79	-0.0991	2.43	11.93
Grass	18.14	3.03	-0.0002	4.25	2.44
Lenna	9.19	4.28	-0.1306	3.03	12.50
Pole	5.35	8.52	0.0169	2.31	10.02



**Figure 3:** Magnified absolute values of differences for the test images Fly (a) and Pole (b) and the test images Fly (c) and Pole (d)

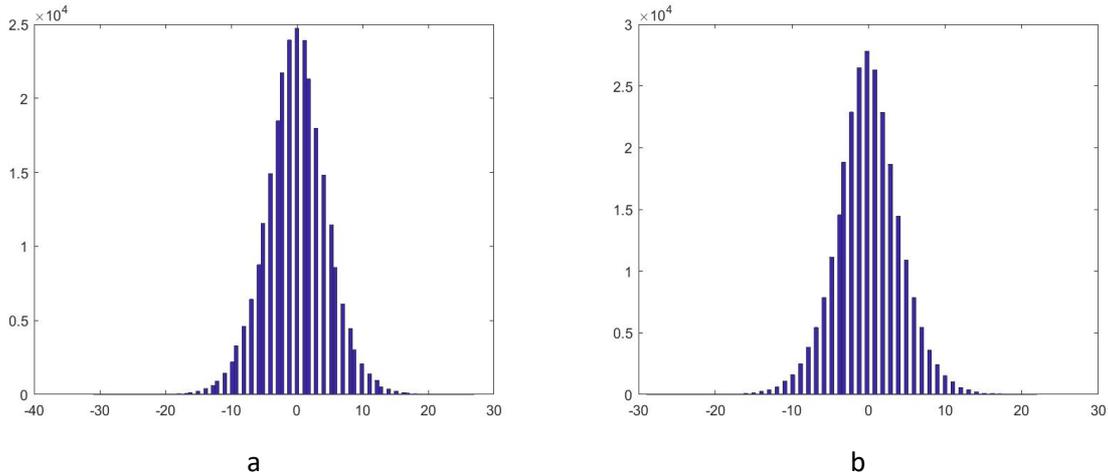


**Figure 4:** Magnified absolute values of differences for the test image Baboon,  $QS=15$  (a) and the test image Baboon (b)

Although the distribution of differences is close to Gaussian for the test image Baboon, there is a certain heterogeneity in local intensity of distortion fluctuations. The smallest intensity is observed in the central part (baboon's nose) where image is quasi-homogeneous (less textural, Figure 4).

Note that Gaussianity of the introduced distortions is observed not only for the coder AGU. We have done the tests for the coders ADCT (<http://ponomarenko.info/adct.htm>), SPIHT, and 3D version of AGU (see the histograms in Figure 5) that confirmed Gaussianity for CR that are not too large. Under this condition, AWGN can serve as a good model of introduced distortions. The model is approximately valid for rather small noise variance when PSNR exceeds 36 dB and it is visually unseen.

One interesting observation is that distortions are more intensive in locally active areas of images. This fact can be taken into account in later studied after more thorough analysis of dependence of distortion statistics on true image local activity. Meanwhile, it is possible to predict that the aforementioned phenomenon leads to a larger degradation of classification accuracy for classes represented by locally active image areas as textures, i.e. the class “Urban” or, in less degree, “Forest”.



**Figure 5:** Histograms of differences for blue (a) and green (b) components of color image Goldhill compressed by the 3D AGU coder

### 3. Used images and classification approach

We have already mentioned that classification results depend on many factors, including the used classifiers. Because of this, here we analyze the results obtained for two known classifiers with supervised learning.

The task of image classification is usually treated as creation of an optimal (or quasi-optimal) classifier that maps a set of available observations of class attributes (features) into a set of classes which is represented by unique names or numbers (indices)  $d(x): X \rightarrow A$ . The optimality is often treated as follows: if elements  $x$  from the observation space  $X$  are presented in the classification process, correct decisions have to be undertaken as often as possible.

Supervised classification procedures that we deal here presume the presence of training samples which are collections of pixels representing a recognizable pattern or a potential class. These are some areas in the image(s) which are well-defined and/or identified based on the ground truth data taken from the Earth's surface maps.

For the maximum likelihood (ML) approach [15, 16], the classification is based on obtaining feature distributions for each class and their analytical description. Features obviously have non-Gaussian distributions [16]. Because of this, densities have been approximated by Johnson's SB-distribution that have four unknown (variable) parameters that can be adjusted in iterative way.

Another considered approach to classification is based on a trained neural network [17]. Convolutional NN has been employed [17]. The used model includes a sequence of four hidden layers which are densely connected and have 64, 32, 16, and 8 neurons, respectively. The hidden layers are activated by the ReLU function. The output layer has linear activation function. The RMSProp optimizer for the MLP training has been employed.

To analyze the effect of lossy compression (using AWGM model) on classification results, four three-channel images of the size 512×512 pixels have been used. These images have been obtained from multichannel data acquired by Sentinel-2 satellite sensor (see these fragments in Figure 6). Visible range components have been used.

After visual analysis and using the maps of the corresponding regions, it has been assumed for all four fragments that in each image there are four classes of objects: 1 – Urban, 2 – Water, 3 – Vegetation, 4 – Bare soil. The RS images represent the territory fragments for Kharkiv (images SS2 and SS4) and its environs (images SS1 and SS3), Ukraine. At training stage (carried for each fragment individually), we have chosen relatively homogeneous area of image fragments representing each particular class. Pixels that correspond to each class were marked with conditional colors: Urban – yellow, Water – blue, Vegetation – green, Bare Soil – black. The obtained sets of marked pixels have been divided into two subsets, training and control (verification) samples that did not overlap. The marked areas (sets of reference pixels) have been divided into two subsets, which were used for training and assessing the quality of the classifier. The amount of pixels in the training samples were of the order of  $(4... 20) \times 10^3$  whilst the volumes of the verification samples were larger – about  $(7... 50) \times 10^3$  pixels.



**Figure 6:** Three-channel image fragments used in our analysis: SS1 (a), SS2 (b), SS3 (c), and SS4 (d)

## 4. Experimental results

Let us present some classification results for the ML classifier.

In accordance with this criterion, it is considered that the control sample  $\vec{x}^*$  (that is, the values of the classification features  $\vec{x}$  in the current pixel  $s$  of the image  $I(i, j)$ ) belongs to the class  $a_v$ ,  $1 \leq v \leq K$ , for which the likelihood function is maximum:

$$(\vec{x}^*; \vec{\theta}|a_v) = \max_{1 \leq k \leq K} \{f(\vec{x}^*; \vec{\theta}|a_k)\} \Rightarrow s \in a_v, \quad (1)$$

where  $f(\vec{x}; \vec{\theta}|a_k)$  is a conditional distribution density of features for the  $k$ -th class (more precisely, its estimate obtained at the stage of training the classifier);  $\vec{\theta}$  is a vector of distribution parameters.

The reliability of classification is characterized by the probabilities of correctly recognized pixels and errors. The statistical estimates of these probabilities are the percentage of correctly and falsely recognized image pixels. Reference images are used for evaluation.

The percentage of correctly recognized pixels of the  $k$ -th class is calculated as follows:

$$Q_k = \frac{1}{N_k} \sum_{i,j:s(i,j) \in a_k} C_k(\vec{x}^*), \quad (2)$$

where  $N_k$  is the number of pixels belonging to the  $k$ -th class on the reference image;  $C_k(\vec{x}^*) = 1$  if the classifier decision is correct, otherwise  $C_k(\vec{x}^*) = 0$ .

The estimate of the total probability of correct recognition based on the classification results of a test image containing images of  $K$  classes is found by the formula

$$\hat{P}_t = \frac{1}{K} \sum_{k=1}^K Q_k. \quad (3)$$

This criterion is considered the most general and suitable for many recognition problems with a limited number of classes.

For the image SS1, the verification areas are shown in Figure 7, a.

The classification map for the original image (without added AWGN) is presented in Figure 7, b. As one can see, even for noise-free image, the classification is not perfect. There are quite many misclassifications, especially for the classes Urban and Bare Soil. There are also quite many misclassifications for the classes Water and Vegetation.

Classification errors are due to several reasons. First, the spectral composition of each pixel is a set of spectral characteristics of the objects that form this pixel. Significant distortions in the obtained spectra are also introduced by the fact that different parts of the surface are in different conditions during the shooting. Secondly, the classes in the feature space are not uniquely separable (in other words, the decision-making areas in favor of one or another class intersect). Thirdly, the quality of training samples also affects - these samples may be not representative enough, that is, they may not adequately describe the stochastic properties of classes. Finally, the accepted class standard models may not accurately approximate empirical distributions.

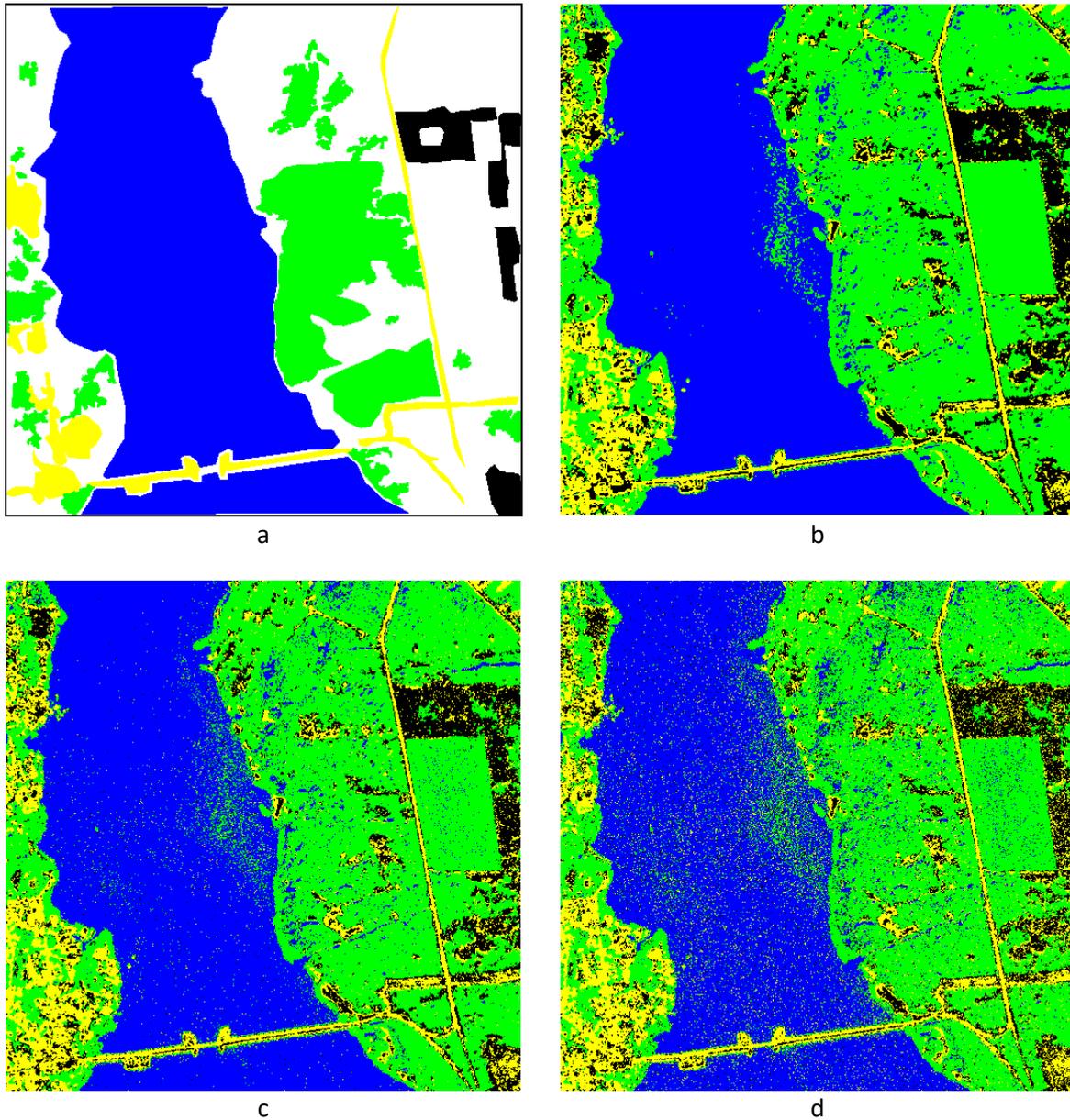
The confusion matrix of the original image SS1 is presented in Table 3.

**Table 3**

Confusion matrix [18] for the ML classifier applied to original SS1 image fragment

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0.695	0.001	0.049	0.255
Water	0.001	0.979	0.019	0.001
Vegetation	0.026	0.029	0.9	0.044
Bare soil	0.124	0	0.025	0.85

The classification results for images to which noise has been artificially added are given in Figure 7, c and d. Their comparison between each other and to the classification map in Figure 7, b clearly demonstrate that the percentage of misclassifications increases when AWGN variance becomes larger. This fact can be also confirmed quantitatively.



**Figure 7:** Classification results for SS1 fragment: verification areas (a); classification maps for original (b) and noisy ( $\sigma^2=9$ , c;  $\sigma^2=25$ , d) images, ML classifier

The confusion matrix for AWGN variance equal to 49 is presented in Table 4. Consider its diagonal elements. Probability of correct classification for the first class  $P_{11}$  has slightly increased, but the probabilities  $P_{22}$ ,  $P_{33}$ , and  $P_{44}$  for the corresponding classes have considerably reduced. Percentage of misclassifications has greatly increased.

**Table 4**

Confusion matrix for SS2 image fragment corrupted by AWGN with  $\sigma^2=49$

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0.745	0.0	0.05	0.205
Water	0.07	0.756	0.086	0.088
Vegetation	0.101	0.053	0.771	0.076
Bare soil	0.368	0	0.034	0.598

This means that even for  $\sigma^2=49$  (PSNR about 31 dB) the degradation of classification accuracy due to noise (that simulate distortions due to lossy compression) is sufficient. To carry out a more detailed study, we have calculated probabilities for particular classes as well as total probability  $P_t$  for several values of noise variance. They are presented in Table 5.

**Table 5**

Particular class and total probabilities of correct classification for image SS1 with different noise intensity, ML classifier

Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.70	0.72	0.74	0.75	0.76
Water	0.98	0.95	0.86	0.76	0.61
Vegetation	0.90	0.88	0.83	0.77	0.65
Bare soil	0.85	0.81	0.71	0.60	0.43
Total	0.856	0.839	0.784	0.718	0.615

Analysis of the presented data shows that added distortions might increase the probability of correct classification for a particular class. However, for most classes, there is reduction of probabilities of correct classification as well as total probability. While for  $\sigma=3$  the reduction is appropriate (0.02 – 0.04), it becomes too large for already  $\sigma=5$  and catastrophic for  $\sigma=10$ .

It is impossible to make conclusions based on only one image. So, consider other test data. Table 6 presents the probabilities for the image fragment SS2. Again, the distortions have practically no negative impact on correct classification for the class Urban. Meanwhile, it has considerable negative impact on classification results for all three other classes leading to noticeable degradation of correct classification for  $\sigma=3$  and catastrophic degradation for  $\sigma \geq 5$ .

**Table 6**

Particular class and total probabilities of correct classification for image SS2 with different noise intensity, ML classifier

Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.92	0.91	0.91	0.91	0.91
Water	0.92	0.87	0.78	0.67	0.56
Vegetation	0.69	0.65	0.60	0.55	0.46
Bare soil	0.86	0.78	0.65	0.52	0.38
Total	0.847	0.805	0.735	0.663	0.557

Table 7

contains data for the image fragment SS3. Only for the class Urban there is no negative impact of the noise on classification accuracy. For other classes, the negative impact is sufficient. The analysis shows that it is strongly undesired to have noise with  $\sigma > 3$ , i.e.  $PSNR \leq 38$  dB for compressed images.

**Table 7**

Particular class and total probabilities of correct classification for image SS3 with different noise intensity, ML classifier

Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.91	0.92	0.92	0.92	0.92
Water	0.99	0.87	0.74	0.63	0.52
Vegetation	0.90	0.90	0.85	0.76	0.63
Bare soil	0.71	0.60	0.47	0.34	0.21
Total	0.877	0.822	0.743	0.662	0.568

Finally, the data for the image fragment SS4 are collected in Table 8. There is  $P_{11}$  increase if noise variance increases. For other classes, there is a sufficient reduction of probabilities of correct

classification. We have carried out special study to understand why this happens. To our opinion, this is because distributions for the classes Water and Bare Soil (and, in less degree, for the class Vegetation) are quite narrow. Then, if noise is present, the distorted pixel values (features) can fall into Urban class feature area which is very wide.

**Table 8**

Particular class and total probabilities of correct classification for image SS4 with different noise intensity, ML classifier

Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.65	0.65	0.66	0.68	0.69
Water	0.86	0.79	0.71	0.63	0.52
Vegetation	0.94	0.89	0.84	0.73	0.53
Bare soil	0.80	0.74	0.64	0.55	0.42
Total	0.811	0.768	0.712	0.642	0.540

## 5. Discussion

Summarizing the results, we can state the following. First, the total probability of correct classification  $P_i$  has a steady tendency to reduction if noise variance increasing. The reduction can be acceptable for  $\sigma$  about 3 but it becomes too large for larger  $\sigma$ . Probably, reduction also depends on image complexity. However, to determine this dependence, we need more images and a parameter (or parameters) able to characterize image complexity adequately.

Second, misclassifications are “point-wise” and this is due to pixel-wise classification of RS images. Maybe, the use of spatial information (either at getting the features or at post-classification stage) can improve classification. This will make it possible to include in the decision rules not only the spectral values of individual pixels, but also the complex brightness-structural characteristics of the groups of pixels that form the segments.

This use of spatial information can be realized in different ways:

1. as use of spatial features at classification stage [19] (it is also worth mentioning here that only three input features (i.e., pixel color) are used in classification and this can be not enough);
2. by filtering of decompressed images before classification;
3. as post-classification processing [20].

Consider now the results for the NN-classifier.

The classification maps for the test image fragment SS1 are shown in Figure 8. It is seen that, in general, they look better than for the ML-classifier (compare the maps in Figure 7, c and Figure 8, b and in Figure 7, d and Figure 8, c). There are less misclassifications, at least, for the class Water.

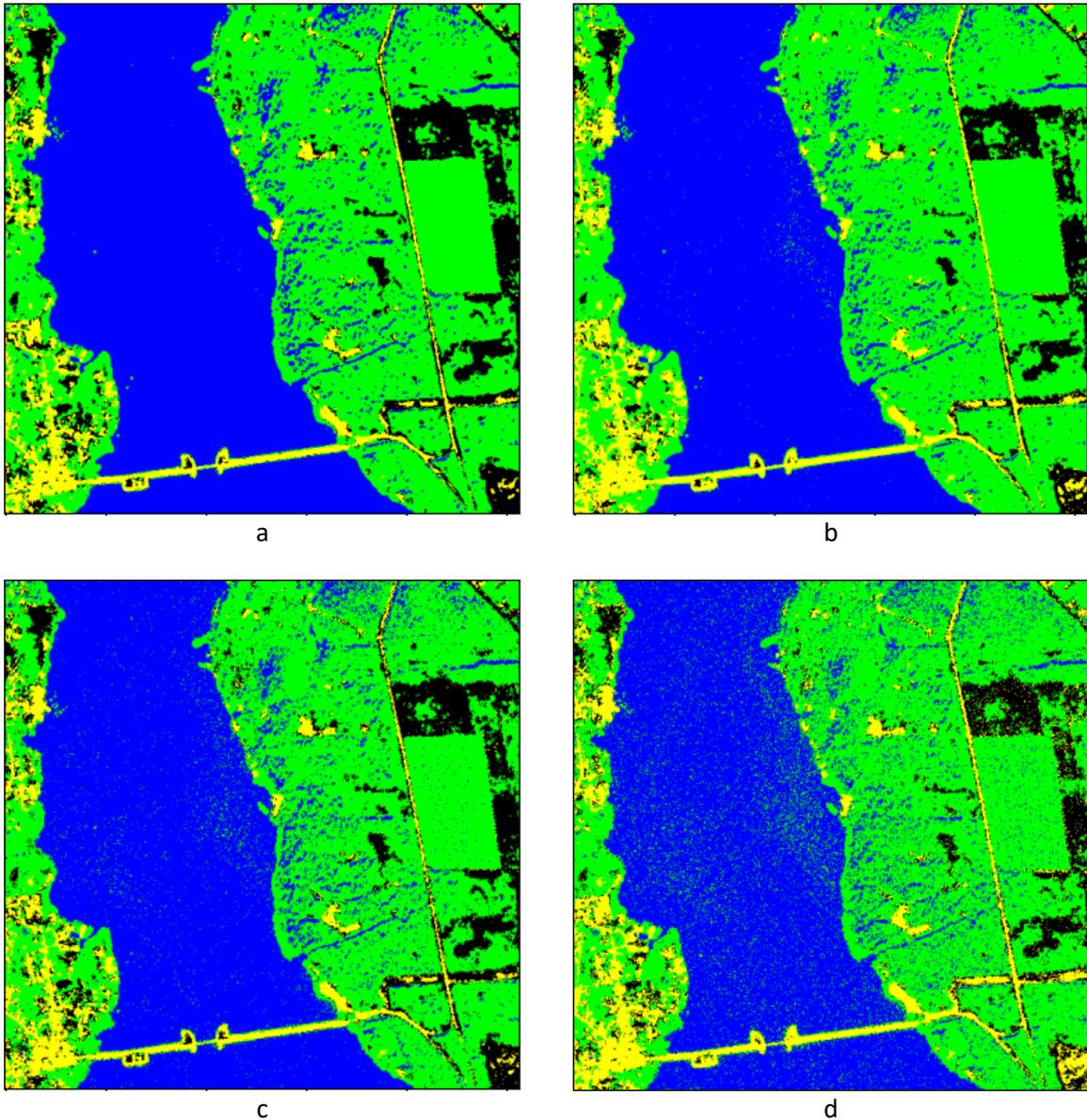
Quantitative data summarizing the results for the test image SS1 are given in Table 9.

Most probabilities are sufficiently better than for the ML classifier (compare the data in Table 5 and Table 9). For  $\sigma \leq 5$ , the results can be considered acceptable. For all classes, a general tendency to degradation of classification results is observed if AWGN variance increases.

**Table 9**

Particular class and total probabilities of correct classification for image SS1 with different noise intensity, NN classifier

Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.80	0.80	0.75	0.79	0.73
Water	0.99	0.98	0.97	0.93	0.92
Vegetation	0.93	0.92	0.88	0.81	0.75
Bare soil	0.77	0.76	0.73	0.71	0.64
Total	0.873	0.865	0.833	0.805	0.760



**Figure 8:** Classification results for SS1 fragment maps for original (a) and noisy ( $\sigma^2=9$ , b;  $\sigma^2=25$ , c;  $\sigma^2=49$ , d) images, NN classifier

The classification results for the test image fragment SS2 are presented in Table 10. In general, they are slightly worse than for the ML classifier (compare data in Table 10 to the corresponding data in Table 6). Reduction of all particular probabilities as well as  $P_t$  takes the place if noise intensity increases and this reduction can be considered appropriate only for  $\sigma=3$ .

**Table 10**

Particular class and total probabilities of correct classification for image SS2 with different noise intensity, NN classifier

Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.92	0.90	0.87	0.84	0.81
Water	0.71	0.64	0.52	0.53	0.37
Vegetation	0.62	0.60	0.55	0.52	0.44
Bare soil	0.90	0.87	0.80	0.74	0.67
Total	0.788	0.753	0.660	0.633	0.598

Table 11 contain data of SS3 image fragment classification. The classification by NN is worse than for the ML classifier for the class Urban (compare data in Table 7

and Table 11). For other classes and, in aggregate, the NN classifier performs better. Classification accuracy reduction is acceptable if  $\sigma=3$  but for larger considered values of AWGN standard deviation the reduction is too large. Thus, again we come to necessity to provide PSNR of compressed data about 38 dB or larger, i.e. to ensure invisibility of introduced distortions (this usually happens if PSNR exceeds 36 dB).

**Table 11**

Particular class and total probabilities of correct classification for image SS3 with different noise intensity, NN classifier

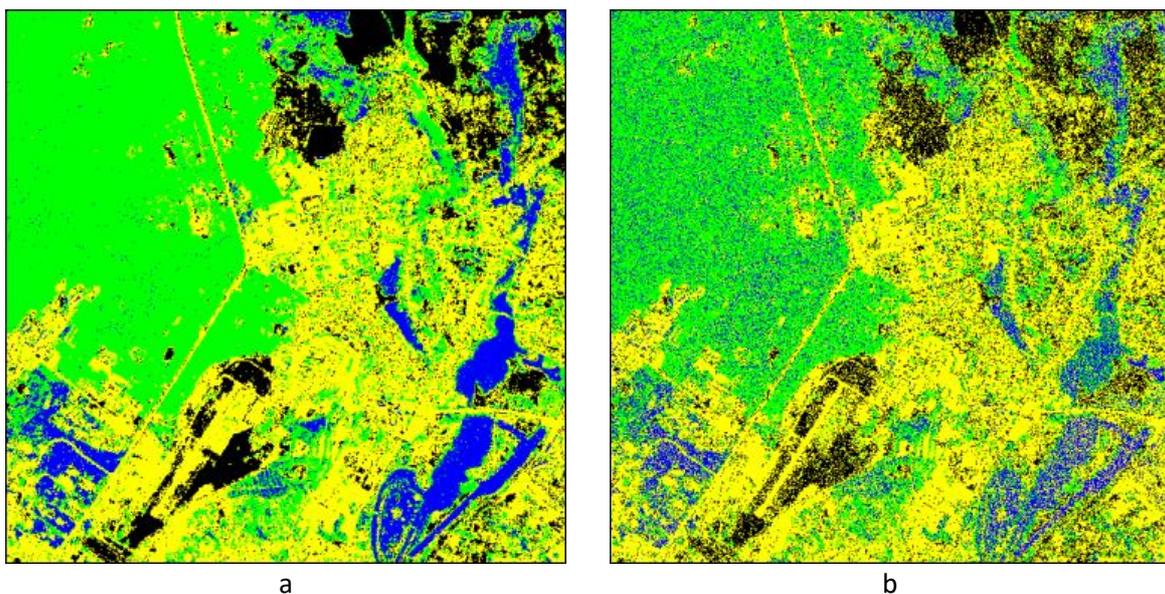
Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.86	0.86	0.83	0.83	0.76
Water	0.99	0.95	0.84	0.75	0.68
Vegetation	0.99	0.94	0.87	0.82	0.78
Bare soil	0.83	0.83	0.79	0.72	0.71
Total	0.918	0.895	0.833	0.780	0.733

Finally, Table 12 presents the classification results for the image fragment SS4. The results, in general, are comparable for the considered classifiers.

**Table 12**

Particular class and total probabilities of correct classification for image SS4 with different noise intensity, NN classifier

Class	Original	$\sigma=3$	$\sigma=5$	$\sigma=7$	$\sigma=10$
Urban	0.83	0.81	0.77	0.72	0.62
Water	0.87	0.64	0.49	0.41	0.29
Vegetation	0.95	0.92	0.89	0.82	0.69
Bare soil	0.82	0.74	0.67	0.67	0.54
Total	0.878	0.778	0.705	0.653	0.535



**Figure 9:** Classification maps for the original image (a) and image contaminated with AWGN with noise variance equal to 49

As for the SS4 image, the NN classifier recognizes the class water worse but the class Vegetation better compared to the ML classifier. There are a lot of misclassifications between the classes Water and Vegetation. One possible reason is that the images have been acquired at the end of August when water basins in the city were blossoming (see the classification maps in Figure 9. Even for  $\sigma=3$ , the reduction of  $P_t$  is too large.

Summarizing the obtained results, we can state the following.

First, there is a sufficient dependence of classification accuracy on image complexity. For complex structure images, sufficient reduction of accuracy can take the place even if PSNR is about 38 dB. For simpler structure images, images compressed with PSNR about 35 dB can produce acceptable classification accuracy.

Second, the NN classifier produced better classification accuracy than the ML classifier for simple structure images. Meanwhile, the accuracy of both used classifiers was approximately the same for complex structure images. It might be also so that classifiers produce essentially different probabilities for different classes.

Third, training has been done for the original (undistorted) images. Meanwhile, it has been shown in [20] that training for compressed images is able to provide certain benefits. Note that sometimes we have only compressed images subject to further classification. Thus, this aspect has to be additionally studied.

## 6. Conclusions

In this paper, analysis of classification accuracy has been performed for the ML and NN classifiers applied pixel-wise to three-channel images. The peculiarity of analysis is that a coder is not specified but the effects of lossy compression are simulated as AWGN. Such an opportunity stems from statistical and spectral analysis of introduced distortions carried out for several techniques of lossy image compression.

The performed research shows that the AWGN presence leads to reduction of classification accuracy in general and for most particular classes (at least, those ones that have rather narrow distribution of features). This reduction becomes sufficient if PSNR of distorted images is about 38 dB for complex structure images and for slightly smaller PSNR if the structure of an image is simple. This approximately corresponds to PSNR-HVS-M=45 dB.

In the future, we plan to analyze more test images, in particular, those one produced by other multispectral sensors. It is also expected that classification accuracy can be improved using spatial information in one or another way.

## 7. References

- [1] D. K. Pillai, P. Swarnalatha and P. Eds. "New Computational Models for Image Remote Sensing and Big Data" *Big Data Analytics for Satellite Image Processing and Remote Sensing*, IGI Global (2018): 1–21.
- [2] N. Kussul, M. Lavreniuk, A. Shelestov, S. Skakun, Crop inventory at regional scale in Ukraine: Developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery, *European Journal of Remote Sensing*, 51(1), (2018) 627-636. doi:10.1080/22797254.2018.1454265.
- [3] M. Li, S. Zang, B. Zhang, S. Li, C. Wu, A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information, *European Journal of Remote Sensing*, (2014) 389-411. doi: 10.5721/EuJRS20144723.
- [4] J. García-Sobrino, V. Laparra, J. Serra-Sagrìstà, X. Calbet, G. Camps-Valls, Improved Statistically Based Retrievals via Spatial-Spectral Data Compression for IASI Data, *IEEE Transactions on Geoscience and Remote Sensing*, 57 (8), (2019) 5651-5668. doi: 10.1109/TGRS.2019.2901396.
- [5] I. Blanes, E. Magli, J. Serra-Sagrìstà, A Tutorial on Image Compression for Optical Space Imaging Systems, *IEEE Geoscience and Remote Sensing Magazine* 2, (2014) 8-26. doi: 10.1109/MGRS.2014.2352465.

- [6] M. Radosavljevic, B. Brkljac, P. Lugonja, V. Crnojevic, Z. Trpovski, Z. Xiong, D. Vukobratovic, Lossy Compression of Multispectral Satellite Images with Application to Crop Thematic Mapping: A HEVC Comparative Study, *Remote Sensing*, 12, (2020) 1590. doi: 10.3390/rs12101590.
- [7] M. Conoscenti, R. Coppola, and E. Magli. "Constant SNR, rate control, and entropy coding for predictive lossy hyperspectral image compression." *IEEE Transactions on Geoscience and Remote Sensing* 54 (2016): 7431-7441. doi:10.1109/TGRS.2016.2603998.
- [8] N. Ozah, A. Kolokolova, Compression improves image classification accuracy, in *Artificial Intelligence. Canadian AI 2019. Lecture Notes in Computer Science*, (2019) 525–530. doi: 10.1007/978-3-030-18305-9\_55.
- [9] B. Yochai, M. Tomer, Rethinking Lossy Compression: The Rate-Distortion-Perception Tradeoff, in: *Proceedings of the 36th International Conference on Machine Learning, Long Beach, California, PMLR 97*, 2019.
- [10] C. Lerner, H. Cheng, Rate-Distortion Approach to Bit Allocation in Lossy Image Set Compression, 2014. URL: <https://www.cs.uleth.ca/~cheng/papers/iwssip2014.pdf>
- [11] F. Li, V. Lukin, K. Okarma, Y. Fu, J. Duan, Intelligent lossy compression method of providing a desired visual quality for images of different complexity, in: *Proceedings of 2021 International Conference on Applied Mathematics, Modeling and Computer Simulation, AMMCS 2021, Wuhan, China, 2021*, 500-505.
- [12] I. Siqueira, G. Correa, M. Grellert, Rate-Distortion and Complexity Comparison of HEVC and VVC Video Encoders, in: *2020 IEEE 11th Latin American Symposium on Circuits & Systems (LASCAS)*, 2020. doi:10.1109/LASCAS45839.2020.9069036.
- [13] J. Gazagnaire, Survey of image quality metrics from the perspective of detection and classification performance, in: *Proceedings of the SPIE XXII conference on Detection and Sensing of Mines, Explosive Objects, and Obscured Targets, SPIE Digital library, Proceedings Volume 10182, Anaheim, California, USA, 2017*. doi: 10.1117/12.2262665.
- [14] D. Napoleon, S. Sathya, M. Praneesh, and M. Siva Subramanian, Remote sensing image compression using 3D-SPIHT algorithm and 3D-OWT, *International Journal on Computer Science and Engineering* (2012) 899-908. doi: 10.13140/RG.2.1.2009.7045.
- [15] P. S. Sisodia, V. Tiwari, A. Kumar, Analysis of Supervised Maximum Likelihood Classification for remote sensing image, in: *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, 2014, pp. 1-4. doi: 10.1109/ICRAIE.2014.6909319.
- [16] V. Makarichev, I. Vasilyeva, V. Lukin, B. Vozel, A. Shelestov, N. Kussul, Discrete Atomic Transform-Based Lossy Compression of Three-Channel Remote Sensing Images with Quality Control, *Remote Sensing*, 14, (2022) 125. doi: 10.3390/rs14010125.
- [17] C.M. Bishop, *Pattern Recognition and Machine Learning*, New York, Springer, 2006.
- [18] R. Congalton, K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 3rd ed., CRC/Taylor & Francis, Boca Raton, FL, 2019.
- [19] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson. "A Survey on Spectral-Spatial Classification Techniques Based on Attribute Profiles." *IEEE Transactions on Geoscience and Remote Sensing* 53.5 (2015): 2335–2353. doi:10.1109/TGRS.2014.2358934.
- [20] 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.