



**HAL**  
open science

## Lossy compression of multichannel remote sensing images with quality control

V. Lukin, I. Vasilyeva, S. Krivenko, F. Li, S. Abramov, O. Rubel, B. Vozel, K. Chehdi, K. Egiazarian

► **To cite this version:**

V. Lukin, I. Vasilyeva, S. Krivenko, F. Li, S. Abramov, et al.. Lossy compression of multichannel remote sensing images with quality control. *Remote Sensing*, 2020, 12 (22), pp.1-35. 10.3390/rs12223840 . hal-03040367

**HAL Id: hal-03040367**

**<https://hal.science/hal-03040367>**

Submitted on 28 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

# Lossy Compression of Multichannel Remote Sensing Images with Quality Control

Vladimir Lukin <sup>1</sup>, Irina Vasilyeva <sup>1</sup>, Sergey Krivenko <sup>1</sup> , Fangfang Li <sup>1</sup>, Sergey Abramov <sup>1</sup> , Oleksii Rubel <sup>1</sup> , Benoit Vozel <sup>2,\*</sup> , Kacem Chehdi <sup>2</sup> and Karen Eguiazarian <sup>3</sup>

<sup>1</sup> Department of Information and Communication Technologies, National Aerospace University, 61070 Kharkov, Ukraine; lukin@ai.kharkov.com (V.L.); i.vasilieva@khai.edu (I.V.); krivenkos@ieee.org (S.K.); liff@nchu.edu.cn (F.L.); s.abramov@khai.edu (S.A.); o.rubel@khai.edu (O.R.)

<sup>2</sup> Institut d'Électronique et des Technologies du numéRique, University of Rennes 1, UMR CNRS 6164, 22300 Lannion, France; kacem.chehdi@univ-rennes1.fr

<sup>3</sup> Computational Imaging Group, Tampere University, 33720 Tampere, Finland; karen.eguiazarian@tuni.fi

\* Correspondence: benoit.vozel@univ-rennes1.fr

Received: 25 October 2020; Accepted: 19 November 2020; Published: 23 November 2020



**Abstract:** Lossy compression is widely used to decrease the size of multichannel remote sensing data. Alongside this positive effect, lossy compression may lead to a negative outcome as making worse image classification. Thus, if possible, lossy compression should be carried out carefully, controlling the quality of compressed images. In this paper, a dependence between classification accuracy of maximum likelihood and neural network classifiers applied to three-channel test and real-life images and quality of compressed images characterized by standard and visual quality metrics is studied. The following is demonstrated. First, a classification accuracy starts to decrease faster when image quality due to compression ratio increasing reaches a distortion visibility threshold. Second, the classes with a wider distribution of features start to “take pixels” from classes with narrower distributions of features. Third, a classification accuracy might depend essentially on the training methodology, i.e., whether features are determined from original data or compressed images. Finally, the drawbacks of pixel-wise classification are shown and some recommendations on how to improve classification accuracy are given.

**Keywords:** remote sensing; lossy compression; image quality; image classification; visual quality metrics

## 1. Introduction

Nowadays, remote sensing (RS) is used in numerous applications [1–3] due to the following main reasons. Different types of useful information can be potentially retrieved from RS images, especially high resolution and multichannel data (i.e., a set of co-registered component images of the same territory acquired for different wavelengths, polarizations, even by different sensors [3–6]). The modern RS sensors often offer a possibility of fast and frequent data collection—good examples are multichannel sensors Sentinel-1 and Sentinel-2 which have been launched recently and started to produce a great amount of valuable RS data [7,8].

Therefore, the volume of RS data greatly increases due to the aforementioned factors: better spatial resolution, a larger number of channels, more frequent observations. This causes challenges in RS data processing that relate to all basic stages of their processing: co-registration, calibration, pre- or post-filtering, compression, segmentation, and classification [9,10]. One of the most serious challenges is the compression of multichannel RS images [11–13]. Compression is applied to diminish data size before their downlink transferring from spaceborne (more rarely, airborne, UAV) sensors,

to store acquired images in on-land centers of RS data collecting or special depositories, to pass images to potential customers. Lossless compression is often unable to meet requirements to compression ratio (CR) that should be provided, since, even in the most favorable situations of high inter-band correlation of component images [14], CR attained by the best existing lossless compression techniques reaches 4 ... 5 [12].

Lossy compression can provide considerably larger CR values [12,13,15–17] but at the expense of introduced distortions. It is always a problem to reach an appropriate compromise between compressed image quality (characterized in many different ways) and CR [11–13,18–20]. There are several reasons behind this. First, compression providing a fixed CR often used in practice [21,22] leads to compressed images whose quality can vary in wide limits [23]. For a given CR and simpler structure images, introduced distortions are smaller and compressed image quality is higher, whilst for images, having a more complex structure (containing many small-sized details and textures), losses are larger and, thus, image quality can be inappropriate since some important information can be inevitably lost, which is undesired. Certainly, some improvements can be reached due to the employment of a better coder, adaptation to image content [24], use of inter-channel correlation by three-dimensional (3D) compression [12,13,21,22], or some other means. However, the positive effect can be limited, and/or there can be some restrictions, e.g., the necessity to apply image compression standard [12,19,21].

Thus, we are more interested in a different approach that presumes lossy compression of multichannel RS images with a certain control of introduced distortions combined with simultaneous desire to provide a larger CR. Let us explain the advantages of such an approach, when it can be reasonable, and what the conditions are for its realization. Recall here that RS data compressed in a lossy manner can be useful if: (a) introduced losses do not lead to sufficient reduction of solving the final tasks of RS image processing such as segmentation, object detection, spectral unmixing, classification, parameter estimation [17,25–34]; (b) distortions due to lossy compression do not appear themselves as artifacts leading to undesired negative effects (e.g., appearing of ghost artifacts) in solving the aforementioned final task.

Below, we consider the impact of lossy compression on multichannel RS image classification [27,28,33–35]. Specifically, we focus on the case of a limited number of channels, e.g., color, multi-polarization, or multispectral images, due to the following reasons: (a) it is simpler to demonstrate the effects that take place in images due to lossy compression and how these effects influence classification just for the case of a small number of image components; (b) an accurate classification of multichannel images with a small number of components is usually a more difficult task than the classification of hyperspectral data (images with a relatively large number of components) because of a limited number of available features and the necessity to reliably discriminate classes in feature space.

Returning to the advantages of lossy compression of multichannel RS data, we can state the following. First, images compressed in a lossy manner might be classified better than original images [27,34]. Such an effect usually takes place if original images are noisy and coder parameters are adjusted so that the noise removal effect due to lossy compression [13,36–38] is maximal. However, even if the noise in images is absent (or, more exactly, peak signal-to-noise ratio (PSNR) values in original images are high and noise is not seen in visualized component images), the accuracy of image classification may remain practically the same for a relatively wide range of CR variation [27,34,38].

This range width can be different [39]. Its width depends on the following main factors:

- Used classifier, how efficient it is, how it is trained;
- Applied compression technique (coder);
- How well classes are discriminated, and which features are used;
- How many classes are present in each multichannel image;
- How “complex” is a considered image (is it highly textural, does it contain many small-sized or prolonged objects, what is the mean size of fragments belonging to each class).

Therefore, researchers have a wide field of studies intended on understanding how to set a CR or a coder's parameter that controls compression (PCC) for a given multichannel image, used lossy compression technique, and considered classifier. Here we would like to recall some already known aspects and dependencies that will be exploited below. First, for the same CR, considerably larger distortions can be introduced into images with higher complexity [40]. Moreover, larger distortions can be introduced into more complex and/or noisier images for the same PCC, for example, quantization step (QS) of coders based on orthogonal transforms (discrete cosine transform (DCT) or wavelets [40]). Second, the desired quality of a compressed image (a value of a metric that characterizes image quality) can be nowadays predicted and provided with an appropriate accuracy [40,41], at least, for some coders based on DCT as, e.g., the coder AGU [42]. Third, there exist certain dependencies between values of conventional (e.g., PSNR) and/or visual quality metrics and accuracy of correct classification of an entire image or classes [43]. Classification accuracy for classes mainly represented by large size homogeneous objects (water surfaces, meadows) better correlates with the conventional metrics whilst probability of correct classification for classes mainly represented by textures, small-sized and prolonged objects (forests, urban areas, roads, narrow rivers) correlates with visual quality metrics. Besides, classification accuracies for classes can depend on PCC or CR differently [38]. For the latter type of classes, the reduction of probability of correct classification with an increase of CR is usually faster. Fourth, different classifiers produce classification results that can significantly differ both in the sense of obtained classification maps and quantitative characteristics as the total probability of correct classification, probabilities for classes, and confusion matrices [43,44]. Fifth, there are quite successful attempts to predict the total probability of correct classification (TPCC) of compressed RS images based on the fractal models [28].

Aggregating all these, we assume the following:

- There are strict dependencies between distortions due to lossy compression (that can be characterized by certain metrics) and TPCC for multichannel RS images;
- Having such dependences, it is possible to recommend what metrics' values must be provided to ensure practically the same TPCC for compressed data as for original data;
- It is possible to provide the recommended metrics' values, at least, for some modern compression techniques based on DCT;
- Metric's thresholds might depend upon a classifier used and properties of a multichannel image subject to lossy compression;
- Classifier performance might also depend on how it is trained (for example, using an original or compressed multichannel image).

The goal of this paper is to carry out a preliminary analysis are these assumptions valid. The main contributions are the following. First, we show that it is worth applying lossy compression with distortions around their invisibility threshold to provide either some improvement of TPCC (achieved for simple structure images) or acceptable reduction of TPCC (for complex structure images) compared to TPCC for the corresponding original image. Moreover, we show how such compression can be carried out for a particular coder. Second, we demonstrate that classifier training for compressed data usually provides slightly better results than training for original (uncompressed) images.

The paper structure is the following. Section 2 deals with a preliminary analysis of the performance criteria used in lossy compression of RS images. In Section 3, two popular methods of RS data classification based on the maximum likelihood method (MLM) and neural network (NN) are described. Section 4 contains the results of classifier testing for the synthetic three-channel test image. Section 5 provides the results of experiments carried out for real-life three-channel image. Discussion and analysis of some practical peculiarities are presented in Section 6; the conclusions are shown in Section 7.

## 2. Performance Criteria of Lossy Compression and Their Preliminary Analysis

Consider an image  $I_{ij}^t$  where  $i, j$  denote pixel indices and  $I_{Im}, J_{Im}$  define the processed image size. If one deals with a multichannel image, index  $q$  can be used. Let's start with the consideration of the case of a single-component image. To characterize compression efficiency, we have used the following three metrics. The first one, traditional PSNR, is defined by

$$\text{PSNR}_{\text{inp}} = 10 \log_{10} \left( \text{DR}^2 / \sum_{i=1}^{I_{Im}} \sum_{j=1}^{J_{Im}} (I_{ij}^c - I_{ij}^t)^2 / (I_{Im} \times J_{Im}) \right) = 10 \log_{10} (\text{DR}^2 / \text{MSE}) \quad (1)$$

where DR denotes the range of image representation,  $I_{ij}^c$  is the compressed image, MSE is the mean square error of distortions introduced by lossy compression. In general, DR in remote sensing applications might sufficiently differ from 255, which is common for processing RGB color images. Below, we will mainly consider single and multichannel RS images having DR = 255 but will also analyze what should be done if DR is not equal to 255.

PSNR is the conventional metric widely used in many image processing applications. Meanwhile, the drawbacks of PSNR are well known. The main drawback of PSNR is that it is not adequate in characterizing image visual quality [45,46]. Currently, there are numerous other, so-called HVS (human vision system) metrics, that are considered to be more adequate [45–47]. Some of the existing HVS-metrics can be applied only to color images. As we need some metrics that should be good enough and applicable to single component images, we apply the metrics PSNR-HVS and PSNR-HVS-M. The former HVS-metric takes into account less sensitivity of human vision to distortions in high-frequency spectral components, the latter one also takes into consideration a masking effect of image texture and other heterogeneities [47]. These metrics are defined by

$$\text{PHVSM} = 10 \log_{10} (\text{DR}^2 / \text{MSE}_{\text{HVS}}) \quad (2)$$

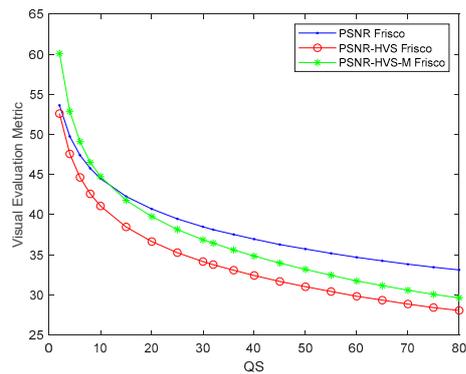
$$\text{PHVSM}_{\text{out}} = 10 \log_{10} (\text{DR}^2 / \text{MSE}_{\text{HVS-M}}) \quad (3)$$

where  $\text{MSE}_{\text{HVS}}$  and  $\text{MSE}_{\text{HVS-M}}$  are MSEs determined considering the aforementioned effects. If PSNR-HVS is approximately equal to PSNR, then the distortions due to lossy compression have properties similar to an additive white Gaussian noise. If PSNR-HVS is smaller than PSNR, this evidences that distortions are more similar to either non-Gaussian (heavy-tailed), or spatially correlated noise, or both [47] (see also [48] for details). If PSNR-HVS-M is considerably larger than PSNR, then a masking effect of image content is sufficient (most probably, an image is highly textural). In turn, if PSNR-HVS-M is approximately equal or smaller than PSNR, then a masking effect is absent, and/or distortions are like spatially correlated noise.

Figure 1a,c,e present three grayscale test RS images of different complexity—the image Frisco has the simplest structure whilst the image Diego is the most textural. Figure 1b,d,f represent dependences of the considered metrics in Equations (1)–(3) on quantization step which serves as PCC in the coder AGU [42] (this compression method employs 2D DCT in  $32 \times 32$  pixel blocks, the advanced algorithm of coding uniformly quantized DCT coefficients, and embedded deblocking after decompression). As one may expect, all metrics become smaller (worse) if QS increases. Meanwhile, for the same QS, compressed image quality is not the same. As an example, for the case of QS = 20, PSNR and PSNR-HVS-M for the test image Frisco are practically the same (about 40 dB). For the test images Airfield and Diego that have more complex structures, PSNR-HVS-M values are also about 40 dB whilst PSNR values are about 34 dB. This means the following: (1) there are masking effects, i.e., textures and heterogeneities still sufficiently mask introduced distortions; (2) the introduced distortions can be noticed by visual inspection (comparison of original and compressed images). Recall that distortion visibility thresholds are about 35 ... 38 dB according to the metric PSNR and about 40 ... 42 dB according to the metric PSNR-HVS-M [45].



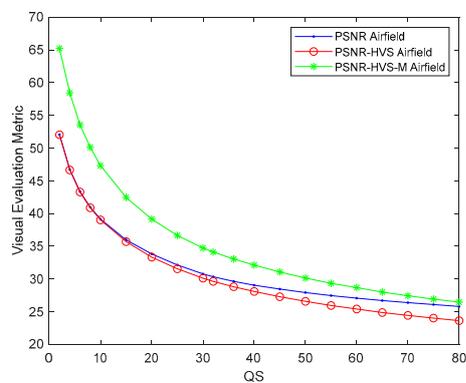
(a)



(b)



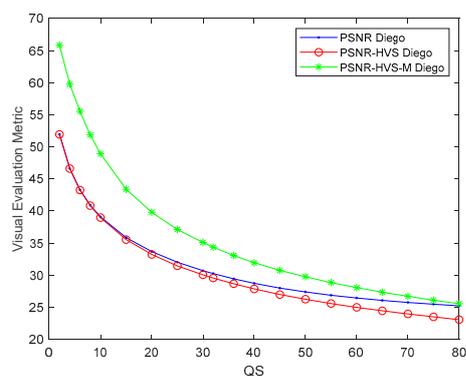
(c)



(d)



(e)



(f)

**Figure 1.** Test grayscale images Frisco (a), Airfield (c), and Diego (e), all of size  $512 \times 512$  pixels, and dependences of the considered metrics (PSNR, PSNR-HVS, PSNR-HVS-M, all in dB) on the quantization step (QS) for the coder AGU for these test images ((b), (d), and (f), respectively).

Note that distortions' invisibility happens if QS is smaller than 18 ... 20 for grayscale images represented as 8-bit data arrays [23] (more generally, if  $QS \leq DR / (12 \dots 13)$ ). One should also keep in mind that the use of the same QS leads to sufficiently different CR. For example, for  $QS = 20$ , CR values are equal to 26.3, 4.6, and 4.5 for the considered test images Frisco, Airfield, and Diego,

respectively. These facts and dependences are given in Figure 1 confirm the basic statements presented in the Introduction.

Alongside differences in metrics' values for the same QS considered above for  $QS \approx 20$  (similar differences take place for  $QS > 20$ , compare the corresponding plots in Figure 1b,d,f), there are interesting observations for  $QS \leq 10$  (more generally,  $QS \leq DR / 25$ ). In this case,  $MSE \approx QS^2 / 12$  [20], PSNR exceeds 39 dB and PSNR-HVS-M is not smaller than PSNR and exceeds 45 dB. Thus, the introduced distortions are not visible and the desired  $PSNR_{des}$  (or, respectively,  $MSE_{des}$ ) can be easily provided by a proper setting of QS as  $QS \approx (12MSE_{des})^{1/2}$ .

Certainly, there are numerous quality metrics proposed so far. Below we carry our study based on PSNR-HVS-M because of the following reasons. First, it is one of the best component-wise quality metrics [49] that can be calculated efficiently. Second, the cross-correlation between the best existing quality metrics is high (see data in Supplementary Materials Section). Therefore, other quality metrics can be used in our approach under the condition of carrying out a corresponding preliminary analysis of their properties. Note that we have earlier widely used PSNR-HVS-M in image compression and denoising applications [23] utilizing well its main properties.

### 3. Considered Approaches to Multichannel Image Classification

As has been mentioned above, there are numerous approaches to the classification of multichannel images. Below we consider two of them. The first one is based on the maximum likelihood method (MLM) [50,51] and the second one relies on a neural network (NN) training [50]. In both cases, the pixel-wise classification is studied. There are many reasons behind using the pixel-wise approach and just these classifiers: (a) to simplify the classification task and to use only Q features ( $q = 1, \dots, Q$ ), i.e., the values of a given multichannel image in each pixel; (b) to show the problems of pixel-wise classification; (c) MLM and NN based classifiers are considered to be among the best ones [50].

The note that classifiers of RS data classification can be trained in different ways. One option is to train a classifier in advance using earlier acquired images, e.g., uncompressed ones stored for training or other purposes. Another option is to use a part of the obtained compressed image for classifier training and its use for entire image classification [39]. One can expect that classification results would be different where both options have both advantages and drawbacks.

#### 3.1. Maximum Likelihood Classifier

The classification (recognition) function for MLM is based on the calculation of some metric in the feature space

$$\Phi = \begin{cases} 1, & \rho(\vec{x}^*, \vec{x}_s) \geq \text{tr}; \\ 0, & \rho(\vec{x}^*, \vec{x}_s) < \text{tr}, \end{cases} \quad (4)$$

where  $\vec{x}^*$ ,  $\vec{x}_s$  are feature (attribute) vectors for current and sample objects (image pixel in our case);  $\rho(\bullet, \bullet)$  is a used metric of vector similarity; tr denotes the decision undertaking threshold.

As components of the feature vector, in general, one can consider both original ones (pixel (voxel) values of a multichannel image) and "derivative" features calculated on their basis (e.g., ratios of these values).

If  $\Phi = 1$ , then a current object (pixel) determined by the feature vector  $\vec{x}^*$ , is related to a class  $a_s$ .

Taking into account the stochastic nature of the features  $\vec{x}, \vec{x}^* \in R^c$ , information about each object class is contained in parameters  $\vec{\theta}$  of joint probability density functions (PDFs)  $f_c(\vec{x}; \vec{\theta})$ .

This information is used for creating statistical decision rules

$$L = \frac{f_c(\vec{x}^*; \vec{\theta} | a_2)}{f_c(\vec{x}^*; \vec{\theta} | a_1)} \geq \text{tr} \quad (5)$$

where  $L$  is the likelihood ratio. The threshold  $tr$  is determined by a statistical criterion used; for example, the Bayesian criterion provides getting the optimal classifier when the following information is available: PDFs for all sets of patterns, probabilities of each class occurrence, and losses connected with probabilities of misclassifications.

In most practical cases, for object description, statistical sample information is used, i.e., statistical estimates of PDFs obtained at the training stage are employed in likelihood ratio  $L$ . Then, sample size, reliability of information about observed objects, and efficiency of using this information in decision rule mainly determine the quality of undertaken decisions.

Training samples are formed using pixels representing a given class; usually, it is some area (or areas, a set of pixels) in an image identified based on some true data for a sensed terrain. The main requirement for the data of training samples is their representativeness—the pixels of the sample must correspond to one class on the ground; such a class should occupy a territory that is fairly well represented by pixels in the image with a given resolution. In other words, the number of pixels in the selected area of the image should ensure the adoption of statistically significant decisions.

When constructing a multi-alternative decision rule ( $K > 2$ ), the maximum likelihood criterion is used; the threshold  $tr$  in the Equation (5) is assumed to be equal to 1. The maximum likelihood criterion makes it possible to eliminate the uncertainty of the solution (when none of the classes can be considered preferable to the others), does not require knowledge of the a priori probabilities of the classes and the loss function, allows evaluating the reliability of the solutions, and can be easily generalized to the case of many classes. In accordance with this criterion, it is believed that the control sample (measured values of features in the current image pixel) belongs to the class,  $1 \leq u \leq K$ , for which the likelihood function (or rather, its estimate obtained at the training stage) is maximum:

$$f_c(\vec{x}^*; \vec{\theta} | a_u) = \max_{1 \leq k \leq K} \left\{ f_c(\vec{x}^*; \vec{\theta} | a_k) \right\} \Rightarrow \vec{x}^* \in a_u. \quad (6)$$

If the assumptions about the normal law of distribution of the feature vector are true, the maximum likelihood method Equation (6) provides optimal recognition. The assumption of normality of features, on the one hand, simplifies the estimation of the parameters of distribution laws (DL) and makes it possible to take into account the presence of mutual correlation relationships between data in different channels in models of classes. On the other hand, it is known that the distributions of real multichannel data are often non-Gaussian. Even with the assumption of the normalization of DL observations due to the central limit theorem, the nonlinearity introduced during the reception and primary processing of information, which manifests itself in a change in the shape and parameters of the observed distribution, cannot be ignored.

To increase the accuracy of approximating empirical non-Gaussian distributions of features and to take into consideration the possible presence of strong correlation relationships, it is advisable to apply multiparameter distribution laws based on non-linear transformations of normal random variables. These transformations include Johnson's transformations.

For describing a class of spectral features, we propose to apply Johnson  $S_B$ -distribution [52,53]:

$$f(x) = \frac{\eta\lambda}{\sqrt{2\pi}(x-\varepsilon)(\varepsilon+\lambda-x)} e^{-\frac{1}{2}(\gamma+\eta \ln \frac{x-\varepsilon}{\varepsilon+\lambda-x})^2} \quad (7)$$

where  $\eta$  and  $\gamma$  are two shape parameters connected with skewness and kurtosis ( $\eta > 0$ ,  $\gamma \in (-\infty, +\infty)$ );  $\varepsilon$  and  $\lambda$  are two scale parameters of a random variable  $x$  ( $\varepsilon \in (-\infty, +\infty)$ ,  $\varepsilon \leq x \leq \varepsilon + \lambda$ ).

Due to a large number of parameters, the model in Equation (4) is quite universal and able to approximate almost any unimodal distribution as well as a wide spectrum of bimodal distributions. Besides, since Johnson distributions are based on nonlinear transformations of normal random variables, their use in describing empirical distributions allows staying within correlation theory framework, which is important for the considered features characterized by non-Gaussian distributions (see examples below) and the presence of strong inter-channel correlation.

The estimates of the parameters for PDFs  $\vec{\theta}|a_k = \{\varepsilon_k, \lambda_k, \eta_k, \gamma_k\}$  can be found by a numerical optimization of the loss function minimizing integral MSE of empirical distribution (histogram) representation by the theoretical model in Equation (7).

The multidimensional variant of Equation (7) is the following [52]:

$$f_c(\vec{x}|a) = (2\pi)^{-c/2} \|\Xi\|^{-1/2} \prod_{v=1}^c \frac{\eta_v \lambda_v}{(x_v - \varepsilon_v)(\varepsilon_v + \lambda_v - x_v)} \times \exp\left[-\frac{1}{2} \sum_{v,k=1}^c \Xi_{v,k}^{-1} \left(\gamma_v + \eta_v \ln \frac{x_v - \varepsilon_v}{\varepsilon_v + \lambda_v - x_v}\right) \left(\gamma_k + \eta_k \ln \frac{x_k - \varepsilon_k}{\varepsilon_k + \lambda_k - x_k}\right)\right] \quad (8)$$

where  $c$  is the size of the feature vector  $\vec{x}$ ;  $\Xi$  is the sample correlation matrix.

Thus, to construct a multidimensional statistical model of correlated data, processing consists of several stages. At the first stage, a traditional statistical analysis is performed by estimating the moments of distributions and histograms of the components of the random vector. Sample estimates of the cross-correlation coefficients of the components  $\vec{x}$  are also found. At the second stage, one-dimensional statistical models of the form Equation (7) are constructed for each of the components of multidimensional data. The initial estimation of the distribution density (the histogram of the feature  $x_v$ ) is used as a target for the desired model  $f(x_v)$  described by the Johnson DL parameter vector  $\vec{\theta}$ . The multidimensional statistical model is built based on multidimensional histograms  $\hat{f}_c(\vec{x})$ . To construct a multidimensional model as in Equation (8) for a class  $a_k$ , one needs to evaluate the matrix of distribution parameters  $\Theta_k = [\varepsilon_{ck}, \lambda_{ck}, \eta_{ck}, \gamma_{ck}] : 4 \times c$  and the sample correlation matrix  $\Xi$ .

After the process of creation and estimation of the training samples, image pixels are sorted (related) to the classes based on the decision rule (#3). If the pixel-wise classification is applied, data for each pixel with coordinates  $(i, j)$  are processed separately (independently). Value vector for a current pixel  $\vec{x}_{ij}^*$  is put into mathematical models of the class target in Equation (5). The obtained results  $\hat{f}_c(\vec{x}_{ij}^*)|a_k, k = 1 \dots K$  are compared between each other and a maximal estimate of the likelihood function is chosen; its number is the number of the class the current pixel  $(i, j)$  is referred to.

### 3.2. NN Classifier Description

As the classifier is based on a neural network for image processing, a feedforward neural network is used. Nowadays there exist many NNs and approaches to their training and optimization [54]. We have chosen a simple but efficient NN for our application that can be easily implemented or placed on different platforms and devices. The employed classifier is an all-to-all connected perceptron combined with a self-organizing map to treat obtained weights or probabilities of the pixel belonging to one of the classes. The architecture details are the following: input layer for incoming data of three-color components, one hidden neuron layer with 30 fully connected neurons (this number has shown to perform equally well with larger possible numbers of hidden layer neurons). For the hidden layer, the tanh activation function is exploited. Scaled conjugate gradient backpropagation is used as a global training function. As has been mentioned above, the output layer of the used NN is a self-organized map fully connected with neurons of the hidden layer. It provides the mapping of output probabilities of the decision to the list of classes as vectors.

Training and validation processes have been performed in the following way. For this purpose, a certain image is divided into a set of pixel color value vectors. Hereby, the obtained set of vectors is permuted to introduce non-ordinary allocation for each training. For NN training, we have taken 70% of the produced set, the other 30% of data is used for validation. The self-dataset validation at this first stage has to be performed to prove the chosen NN architecture and NN parameters, and other training process peculiarities. We have considered different configurations of NN-based classifiers, especially varying the number of hidden layers, neurons, and connections among them. It has been established that complex architectures, like multi-layer perceptron, do not provide better efficiency of

the classification and even cause over-fitting using the same training dataset. As the result, the chosen architecture was NN with one hidden layer with a fully optimal number of training epochs equal to 50 for the given NN. The proposed classifier is easy to use, and it is fast (which is the crucial point for the application of classification to remotely sensed data). The overall training process was repeated 100 times with full permutation of the dataset and the obtained classifier has been applied to test images.

#### 4. Analysis of Synthetic Three-Channel Image

##### 4.1. MLM Classifier Results

Creating a test multichannel image, we have taken into account the following typical properties of classes that might take place in real-life RS images:

**Property 1.** *Features rarely have distributions close to Gaussian; since in our case, color component values (see the test color image and its components in Figure 2a–d, respectively) are the used features, then this property relates to them;*

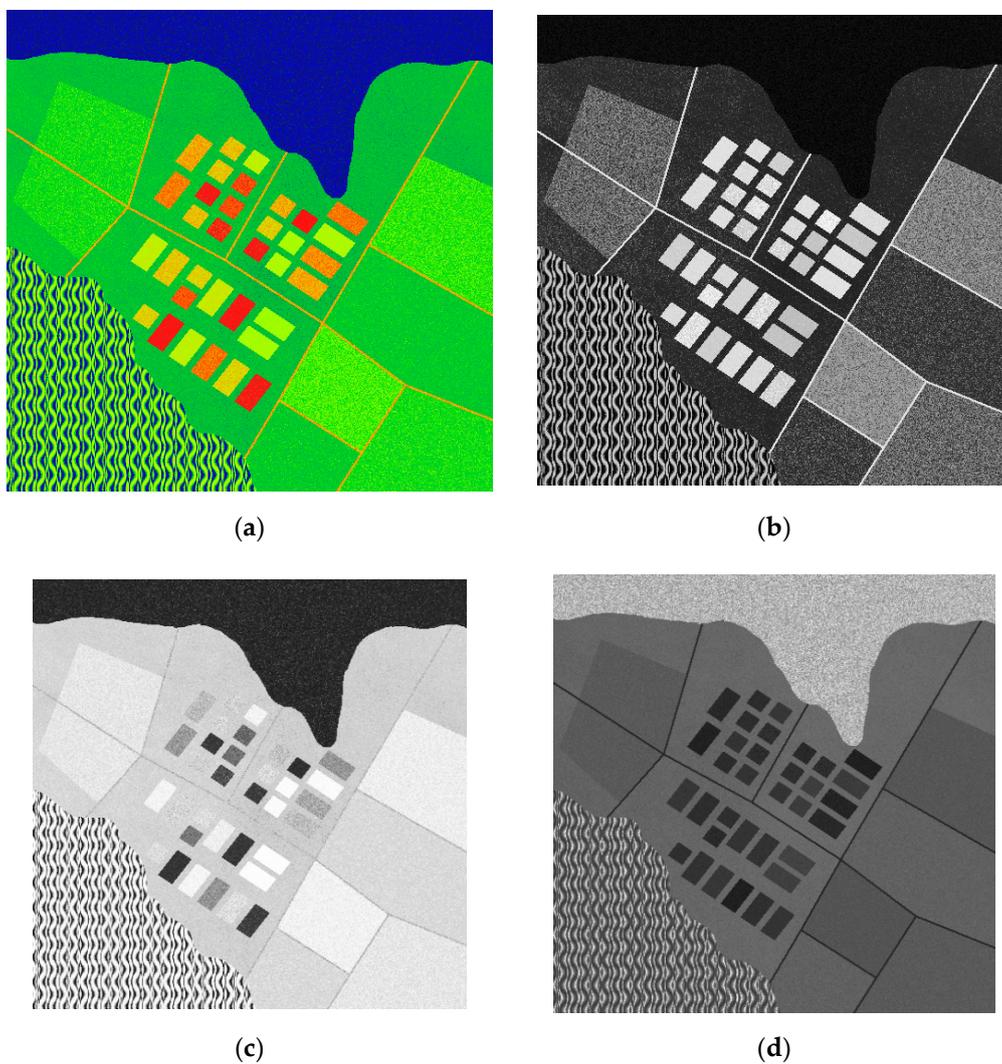
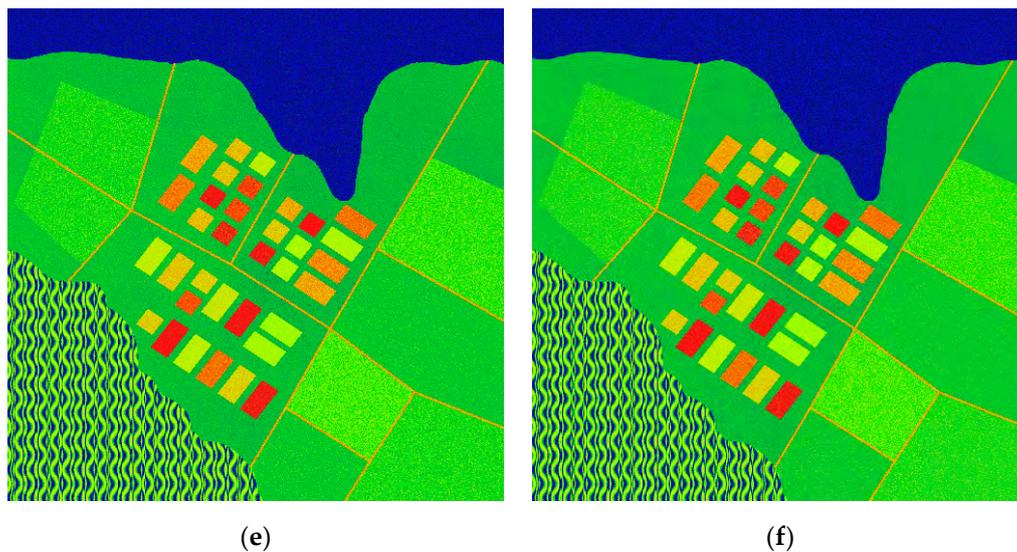


Figure 2. Cont.



**Figure 2.** Three-channel test image: original (a), its three components (b–d), compressed with providing PSNR-HVS-M = 36 dB (e), compressed with providing PSNR-HVS-M = 30 dB (f).

**Property 2.** Objects that relate to a certain class are rarely absolutely homogeneous (even water surface usually cannot be considered as absolutely homogeneous); many factors lead to diversity (variations) of pixel values as variations of physical and chemical characteristics of reflecting or irradiating surface, noise, etc.

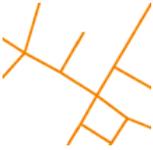
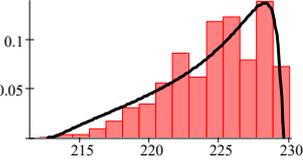
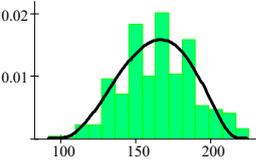
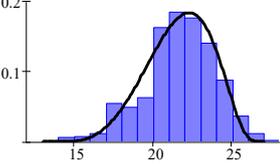
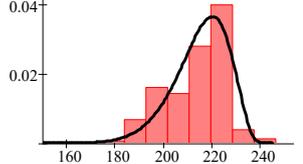
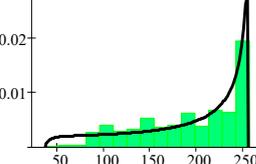
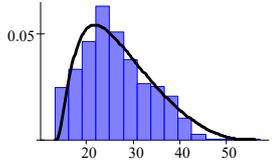
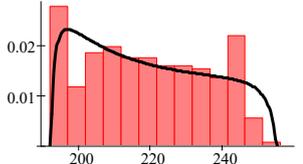
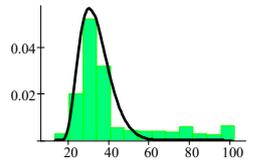
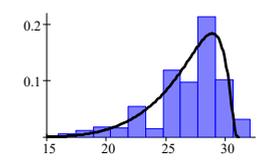
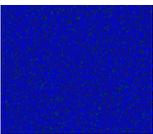
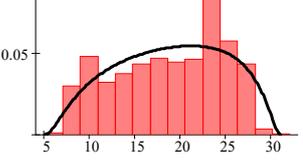
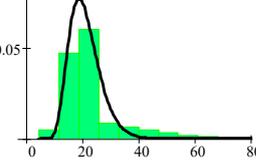
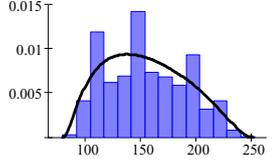
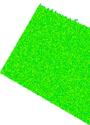
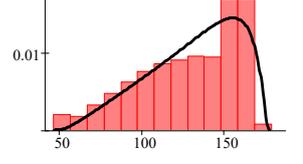
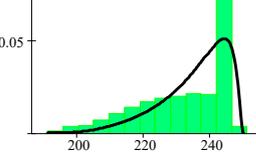
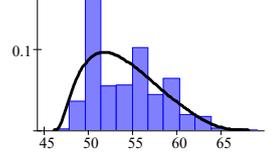
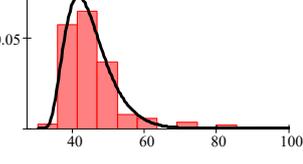
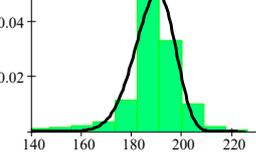
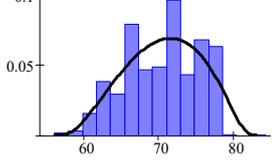
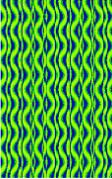
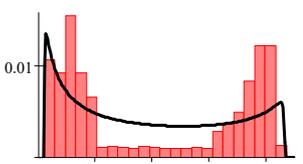
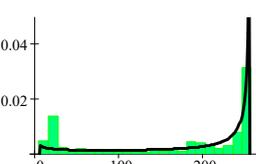
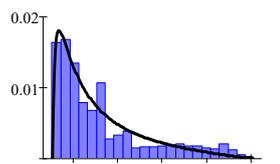
**Property 3.** Features overlap (intersect) in the feature space; just this property usually leads to misclassifications especially if such overlapping takes place for many features;

**Property 4.** The numbers of pixels belonging to different classes in each RS image can differ a lot, it might be so that a percentage of pixels belonging to one class is less than 1% whilst for another class, there can be tens of percent of the corresponding pixels.

**Property 5.** There can be objects of two or even more classes in one pixel [54,55] that requires the application of unmixing methods but we do not consider such, more complex, situations in our further studies (usually more than three components of multichannel RS data are needed to carry out efficient unmixing).

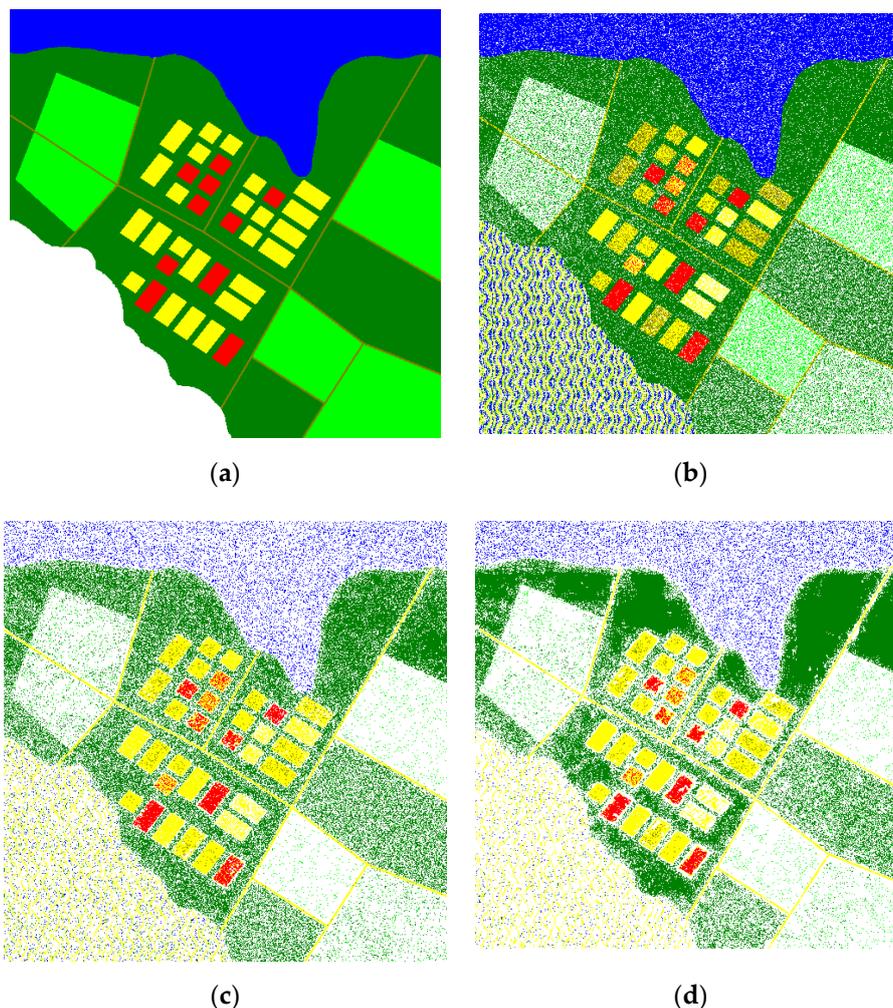
Let us show that we have, in a more or less extent, incorporated these properties in our test image. Table 1 contains data about seven classes, histograms for them, and approximations using Johnson  $S_B$ -distribution. Class 1 can be conditionally treated as “Roads”, Class 2—as agricultural fields (Field-Y), Class 3—as agricultural fields of another color (Field-R), Class 4—as “Water”, Class 5—as “Grass”, Class 6—as “Trees”, Class 7—as some very textural or heterogeneous class like urban area or vineyard (“Texture”).

**Table 1.** Information about classes.

#	Class Map (Object Sample)	Red (Histogram and Approximation)	Green (Histogram and Approximation)	Blue (Histogram and Approximation)
1				
2				
3				
4				
5				
6				
7				

As one can see, Property 1 (non-Gaussian distribution) holds for practically all particular distributions. Property 2 (internal heterogeneity) takes place for all classes, especially for Classes 2 and 7. Property 3 (intersection of features) takes place for many classes—all three features for Classes 1 and 2, all three features for Classes 5 and 6; features of Class 7 overlap practically with all other classes (intersection of features can be also seen from a comparison of component images where some classes can be hardly distinguished, e.g., Classes 2 and 4 in Green component, Figure 2c). Property 4 is also observed—there is a small percentage of pixels belonging to Class 1 whilst quite many pixels belong to Classes 4, 5, and 7.

Figure 3a presents the true map of classes where Class 1 is shown by light brown, Class 2—by yellow; Class 3—by red, Class 4—by blue; Class 5—by light green, Class 6—by dark green, Class 7—by white.



**Figure 3.** Classification results: true map (a); classification map for original image (b); classification map for image compressed with PSNR-HVS-M = 36 dB (c); classification map for image compressed with PSNR-HVS-M = 30 dB (d).

MLM classifier applied to original image pixel-wise (Figure 3a) produces a classification map given in Figure 3b. As one can see, there are quite many misclassifications. Classes 4 and 6 are classified in the best way (probabilities of correct classification  $P_{44}$  and  $P_{66}$  are equal to 0.85 and 0.82, respectively) whilst Class 5 is recognized in the worst manner ( $P_{55} = 0.19$ ). Classes 1 and 7 are not recognized well ( $P_{11}$  and  $P_{77}$  are equal to 0.48 and 0.42, respectively). Classes 2 and 3 are characterized with  $P_{22} = 0.67$  and  $P_{33} = 0.77$ , respectively.

Quite low probabilities of classification can be explained by several factors. Some of them have been already mentioned above—intersections in feature space are sufficient. One more possible reason is that the MLM classifier is based on distribution approximations and these approximations can be not perfect (see histograms and their approximations in Table 1).

Quite many pixels are erroneously related to Class 7 (“Texture”) for which distributions are very wide and they intersect with distributions for other classes.

Figure 2e,f presents images compressed providing PSNR-HVS-M  $\approx 36$  dB and PSNR-HVS-M  $\approx 30$  dB. In these cases, CR values are about 4.5 and 8.9, respectively (for component-wise lossy compression). For the case of PSNR-HVS-M  $\approx 36$  dB, introduced distortions are visible [45]. They mostly appear themselves as smoothing of variations in quasi-homogeneous regions (consider the fragments for Class 6 (dark blue) for images in Figure 2a,e). The effects of such suppression of noise or high-frequency variations due to lossy compression are known for lossy compression [36,37] and they might have even a positive effect for classification (under certain conditions). For the case of PSNR-HVS-M  $\approx 30$  dB, the distortions due to lossy compression (see the compressed image in Figure 2f) appear themselves in variations’ smoothing and edge/object smearing (this can be seen well for Class 1). Clearly, such effects might have an impact on classification.

Probabilities of correct classification for classes are given in Table 2. Note that the MLM classifier has been “trained” using distribution approximations obtained for the original (uncompressed) three-channel image. As one can see, probabilities for particular classes depend on compressed image quality and compression ratio in a different manner.  $P_{11}$  steadily decreases and becomes close to zero. This happens because of two reasons. First, features for Class 1 sufficiently intersect with features for Class 2. Second, due to lossy compression, feature distribution after compression differs from feature distribution for the original image (for which the MLM classifier has been trained) where this difference increases with the reduction of PSNR-HVS-M and CR increase. This is illustrated by distributions presented in Table 3 for Class 1 and, partly, in Table 4 for Class 2.

**Table 2.** Probabilities of correct classifications for particular classes depending on image quality.

PSNR-HVS-M, dB	$P_{11}$	$P_{22}$	$P_{33}$	$P_{44}$	$P_{55}$	$P_{66}$	$P_{77}$
$\infty$ (original image)	0.475	0.667	0.773	0.854	0.193	0.822	0.415
48	0.190	0.781	0.686	0.384	0.093	0.809	0.579
45	0.147	0.795	0.652	0.308	0.077	0.777	0.623
42	0.101	0.804	0.615	0.253	0.058	0.712	0.665
36	0.031	0.802	0.568	0.163	0.040	0.545	0.777
30	0.001	0.753	0.478	0.177	0.042	0.556	0.843

**Table 3.** Illustration of distribution changes due to lossy compression for Class 1.

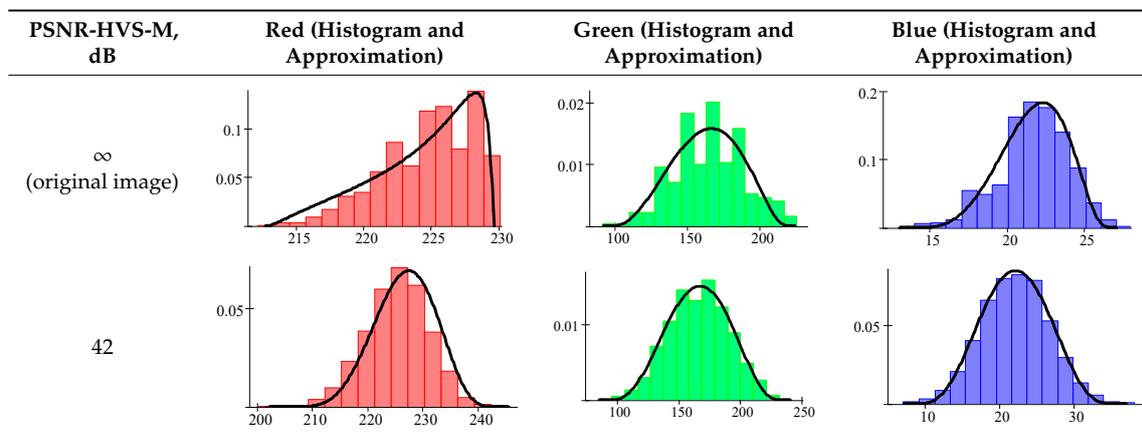


Table 3. Cont.

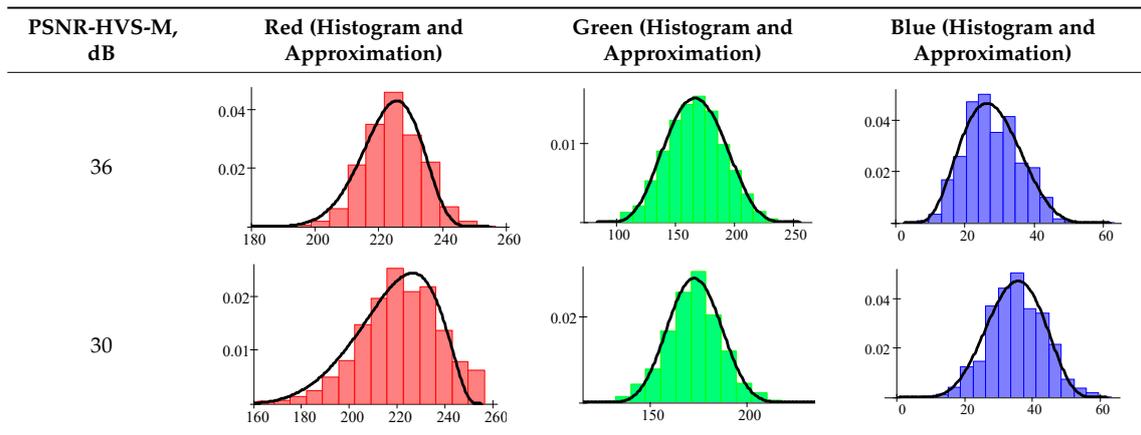
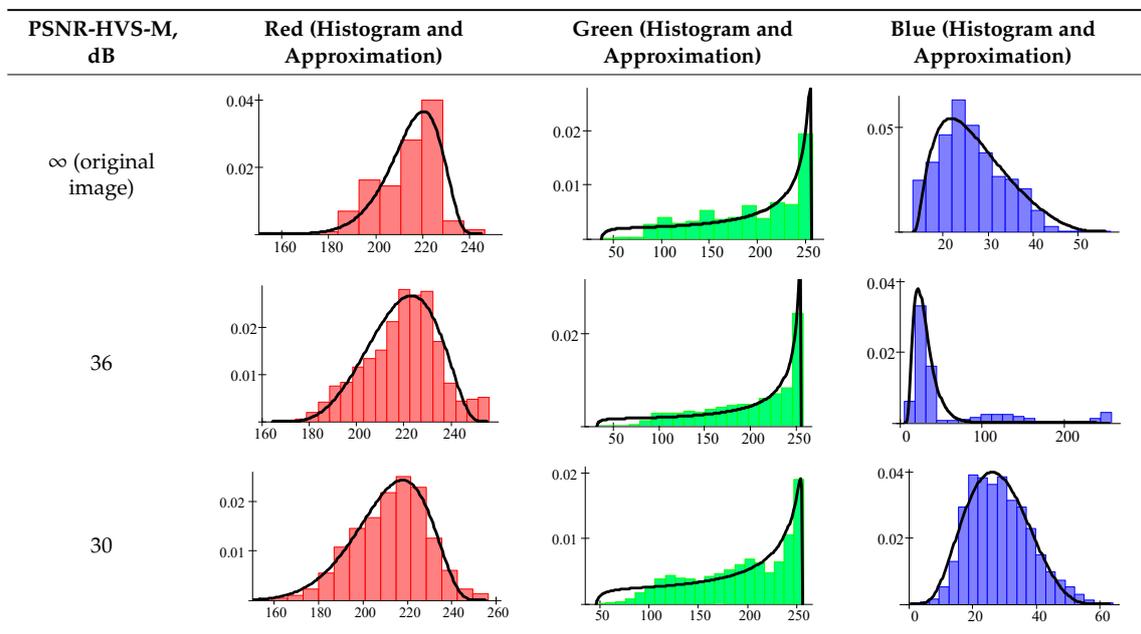


Table 4. Illustration of distribution changes due to lossy compression for Class 2.



Analysis of data in Table 2 also shows the following. There are some classes (Classes 3–6) for which a larger CR and smaller PSNR-HVS-M (that correspond to more sufficient lossy compression) results in a reduction of probabilities of correct classification (analyze data in columns for  $P_{33}$ ,  $P_{44}$ ,  $P_{55}$ , and  $P_{66}$ ). Meanwhile, there are classes, for which probabilities of correct classification increase—see data for  $P_{77}$ . This is because of the decrease in misclassifications in the texture area (see the maps in Figure 3c,d). There is also one class (Class 2) for which  $P_{22}$  is the largest for PSNR-HVS-M about 40 dB (due to noise filtering effect) but it reduces for smaller PSNR-HVS-M (larger CR).

One probable reason why the MLM classifier trained for original images and applied to compressed ones loses efficiency is that feature distributions observed for compressed data differ from those ones observed for original RS data and used for classifier training.

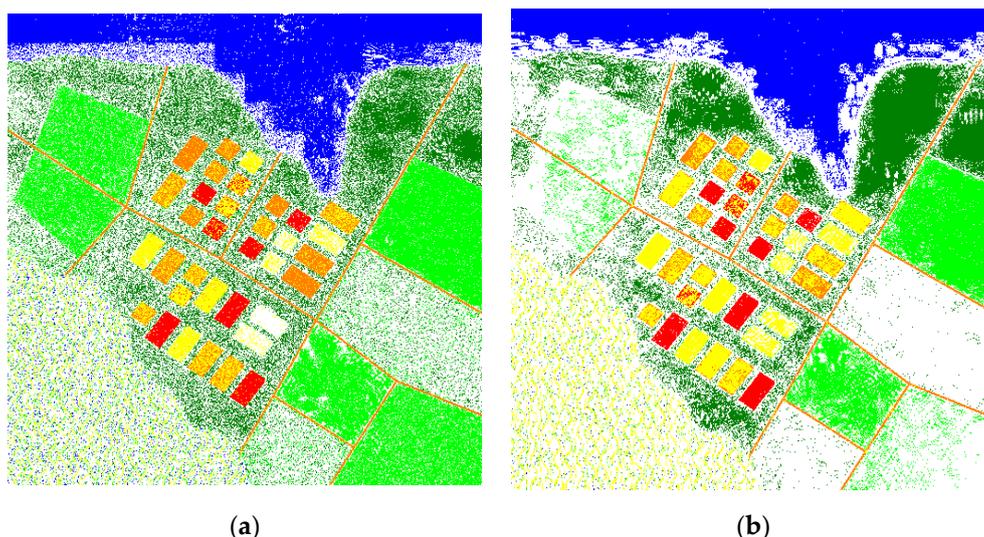
Then, one can expect that the MLM classifier has to be trained for compressed images. We have carried out the corresponding study and the results are presented in Table 5. The conclusion that follows from the analysis of data in Table 5 might seem trivial—it is needed to train the MLM classifier for compressed images with the same characteristics of compression. Following this rule leads to a considerable improvement in classification accuracy. For example, if original data are used in training and then the MLM classifier is applied to data compressed providing PSNR-HVS-M = 42 dB,

then  $P_{cc} = 0.501$ . At the same time, if the image compressed providing PSNR-HVS-M = 42 dB was used in training,  $P_{cc}$  radically increases and becomes equal to 0.579. Even more surprising results are observed for cases of images compressed producing PSNR-HVS-M equal to 36 and 30 dB.  $P_{cc}$  for the corresponding training reaches 0.597 and 0.527, respectively. Note that  $P_{cc} = 0.627$  for original image classified using training for the original image. Then, it occurs that for images compressed with PSNR-HVS-M about 40 dB almost the same  $P_{cc}$  is observed if they are trained for the corresponding compressed images.

**Table 5.** Total probabilities of correct classification  $P_{cc}$  for different images depending upon “training” data.

Image Used for Training		Total Probabilities of Correct Classification for Compressed Images		
		PSNR-HVS-M = 42 dB	PSNR-HVS-M = 36 dB	PSNR-HVS-M = 30 dB
0	Original (uncompressed)	0.501	0.434	0.442
1	Compressed with providing PSNR-HVS-M = 42 dB	0.579	0.566	0.563
2	Compressed with providing PSNR-HVS-M = 36 dB	0.519	0.597	0.599
3	Compressed with providing PSNR-HVS-M = 30 dB	0.408	0.446	0.527

Classification maps obtained for this methodology of training are presented in Figure 4a,b. They can be compared to the maps in Figure 3c,d, respectively. This comparison clearly demonstrates that training for the corresponding compressed data is preferable. Many classes are recognized sufficiently better. One interesting point is that many classifications might appear in the neighborhoods of sharp edges in multichannel images (near edges of Class 4). This is because such edges are smeared by lossy compression and pixels with “intermediate values” that can be referred to as “wrong classes” appear.



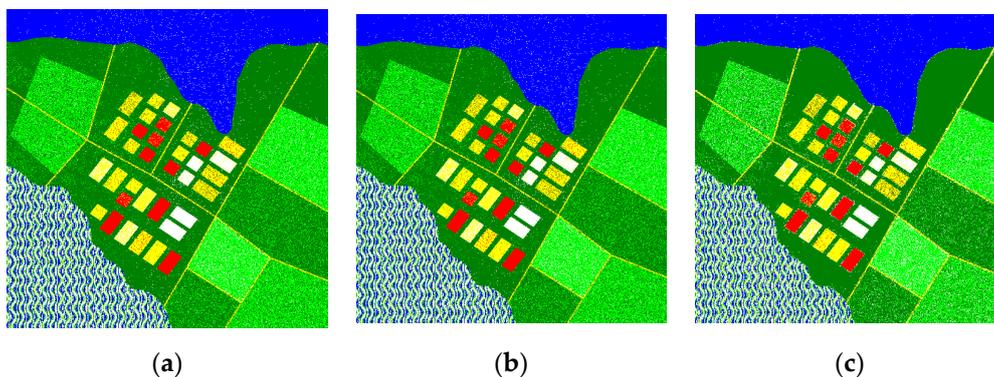
**Figure 4.** Classification maps for the compressed image with providing PSNR-HVS-M = 36 dB using maximum likelihood method (MLM) classifier trained for the same compressed image (a); compressed image with providing PSNR-HVS-M = 30 dB using MLM classifier trained for the same compressed image (b).

#### 4.2. NN-Based Classification

Let us start by analyzing data for classifying the original three-channel image using NN training for this image. The obtained confusion matrix is presented in Table 6. The corresponding map is given in Figure 5a. As one can see, several classes are recognized well: Field-R (Class 3), Water (class 4), Trees (Class 6). There are many misclassifications for Classes “Roads” (with Class Field-Y that has practically the same colors) and “Texture” (its pixels can be erroneously related to classes “Water”, “Trees”, “Grass”, and, more rarely, Field-Y). The results are relatively bad for Class 2 and Class 5. Thus, even for this almost ideal case (classified image is the same as that one used for training, the NN-based classifier able to work with non-Gaussian distributions of features is applied), the classification results are far from being perfect.

**Table 6.** Confusion matrix for the original image classified by neural networks (NN) trained for this image.

Class	Probability of Decision						
	Road	Field-Y	Field-R	Water	Grass	Trees	Texture
Road	<b>0.356</b>	0.645	0	0	0	0	0
Field-Y	0.109	<b>0.608</b>	0.002	0	0	0	0.281
Field-R	0	0.004	<b>0.996</b>	0	0	0	0
Water	0	0	0	<b>0.986</b>	0	0	0.014
Grass	0	0	0	0	<b>0.744</b>	0.245	0.012
Tree	0	0	0	0	0.072	<b>0.923</b>	0.005
Texture	0	0.03	0	0.26	0.116	0.153	<b>0.441</b>



**Figure 5.** Classification results: for compressed image (PSNR-HVS-M = 42 dB) (a), for compressed image (PSNR-HVS-M = 36 dB) (b); for compressed (PSNR-HVS-M = 30 dB) (c).

Taking into account the results for the MLM classifier, we also considered the case when the NN classifier has been trained for the original image and then applied to compressed images. The data obtained for the cases of PSNR-HVS-M equal to 42 dB, 36 dB, and 30 dB are given in Tables 7–9, respectively. Analysis of data in Table 7 shows the following. As was expected, all probabilities have changed. Let us consider the diagonal elements marked by Bold that correspond to probabilities of correct classification of particular classes. In this sense, there are interesting observations (compare the corresponding data in Tables 6 and 7): some probabilities have decreased sufficiently (see  $P_{11}$ ), some decreased slightly ( $P_{33}$ ,  $P_{44}$ ,  $P_{66}$ ,  $P_{77}$ ). Meanwhile, some probabilities have slightly increased ( $P_{22}$  and  $P_{55}$ ).

**Table 7.** Confusion matrix for compressed image classified by NN trained for the original image and applied for the compressed image (PSNR-HVS-M = 42 dB).

Class	Probability of Decision						
	Road	Field-Y	Field-R	Water	Grass	Trees	Texture
Road	<b>0.299</b>	0.699	0	0	0	0	0
Field-Y	0.089	<b>0.627</b>	0	0	0	0	0.276
Field-R	0	0.018	<b>0.982</b>	0	0	0	0
Water	0	0	0	<b>0.980</b>	0	0	0.020
Grass	0	0	0	0	<b>0.753</b>	0.219	0.028
Tree	0	0	0	0	0.084	<b>0.911</b>	0.006
Texture	0	0.03	0	0.26	0.116	0.153	<b>0.441</b>

**Table 8.** Confusion matrix for compressed image classified by NN trained for original image and applied for the compressed image (PSNR-HVS-M = 36 dB).

Class	Probability of Decision						
	Road	Field-Y	Field-R	Water	Grass	Trees	Texture
Road	<b>0.271</b>	0.728	0.001	0	0	0	0
Field-Y	0.078	<b>0.631</b>	0.012	0	0.001	0	0.279
Field-R	0	0.027	<b>0.973</b>	0	0	0	0
Water	0	0	0	<b>0.979</b>	0	0	0.021
Grass	0	0	0	0	<b>0.721</b>	0.22	0.059
Tree	0	0	0	0	0.078	<b>0.915</b>	0.007
Texture	0	0.049	0	0.253	0.123	0.148	<b>0.426</b>

**Table 9.** Confusion matrix for compressed image classified by NN trained for the original image and applied for the compressed image (PSNR-HVS-M = 30 dB).

Class	Probability of Decision						
	Road	Field-Y	Field-R	Water	Grass	Trees	Texture
Road	<b>0.217</b>	0.766	0	0	0	0	0.017
Field-Y	0.067	<b>0.646</b>	0.007	0	0.005	0	0.275
Field-R	0	0.05	<b>0.946</b>	0.004	0	0	0
Water	0	0	0	<b>0.992</b>	0	0	0.008
Grass	0	0	0	0	<b>0.658</b>	0.21	0.131
Tree	0	0	0	0	0.051	<b>0.925</b>	0.025
Texture	0	0.067	0	0.243	0.141	0.131	<b>0.418</b>

Analysis of data in Tables 8 and 9 shows that lossy compression negatively influences  $P_{11}$ , i.e., Probability of correct classification for the Class “Road”. Recall that the same effect has been observed for the MLM. With a larger CR (smaller provided PSNR-HVS-M), probabilities  $P_{33}$  and  $P_{77}$  continue to decrease.  $P_{22}$  slowly increases.  $P_{44}$  remains practically the same and very high.  $P_{66}$  also remains practically the same. Finally,  $P_{55}$  decreases. Thus, we can state that being trained for the original image, the NN classifier continues working well for PSNR-HVS-M about 40 dB and then its performance starts making worse faster (with further increase of CR).

Figure 5 shows classification results for the test image compressed providing three different values of PSNR-HVS-M. The results are quite similar. However, some differences can be noticed:

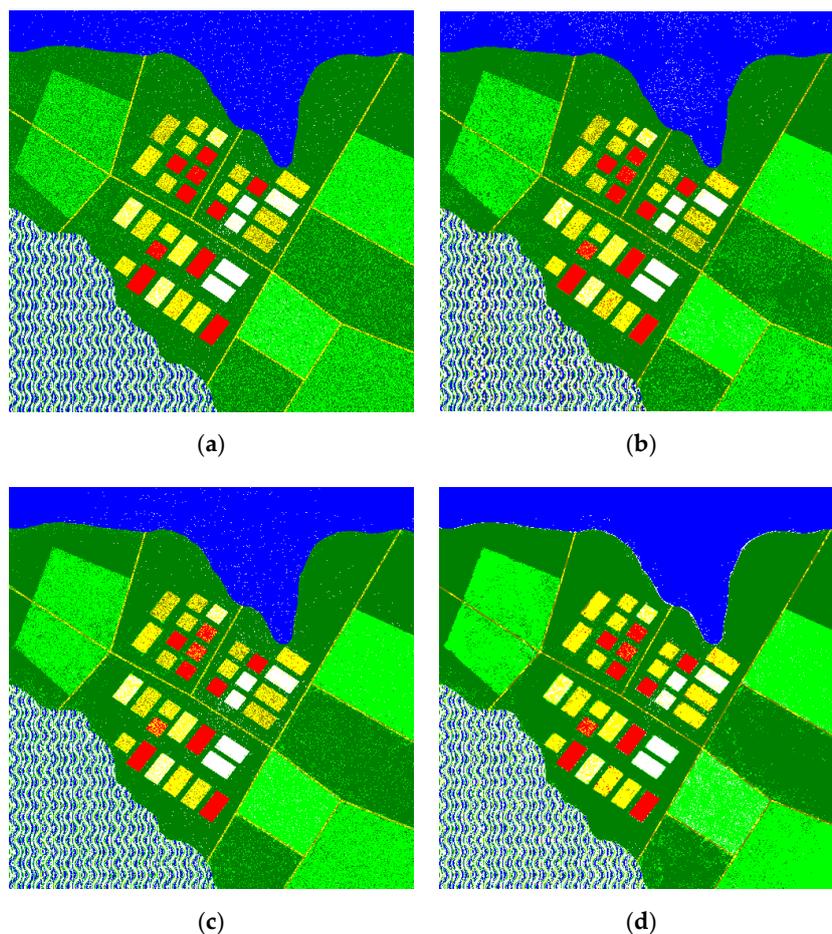
1. Less misclassifications are observed for PSNR-HVS-M = 30 dB for the class “Water”.
2. There are quite many misclassifications around the edge of Class “Water”.
3. More misclassifications appear for the class “Grass” when CR increases.
4. Many pixels and even some objects that belong to Class 2 (Field-Y) are classified as “Texture” (white color). This example one more time shows the problems in image classification if an image

to be classified contains a class like “Texture”. The problems occur for the class “Texture” itself and the classes for which there is an intersection of features with features of the class “Texture”.

Let us now check what happens if compressed images are used for training. Table 10 presents the confusion matrix obtained for the case of NN training using image compressed providing PSNR-HVS-M = 42 dB. These data can be compared to the corresponding data in Table 7. If compressed image is employed for training, probabilities  $P_{11}$ ,  $P_{55}$ , and  $P_{77}$  are sufficiently better, probability  $P_{22}$  is sufficiently worse, other probabilities are slightly worse. Thus, the quality of classification, in general, slightly improves. The classification map is presented in Figure 6b.

**Table 10.** Confusion matrix for compressed image classified by NN trained for this image (PSNR-HVS-M = 42 dB).

Class	Probability of Decision						
	Road	Field-Y	Field-R	Water	Grass	Trees	Texture
Road	<b>0.455</b>	0.545	0.001	0	0	0	0
Field-Y	0.137	<b>0.561</b>	0.014	0	0	0	0.289
Field-R	0	0.026	<b>0.974</b>	0	0	0	0
Water	0	0	0	<b>0.961</b>	0	0	0.039
Grass	0	0	0	0	<b>0.820</b>	0.165	0.015
Tree	0	0	0	0	0.068	<b>0.929</b>	0.004
Texture	0	0.029	0	0.234	0.124	0.153	<b>0.461</b>



**Figure 6.** Results of image classification by NN trained for the same image: original (a), compressed with PSNR-HVS-M = 42 dB (b), compressed with PSNR-HVS-M = 36 dB (c), compressed with PSNR-HVS-M = 30 dB (d).

The next considered case is the classification of the image compressed producing PSNR-HVS-M = 36 dB. The obtained probabilities are given in Table 11, the classification map is presented in Figure 6c. The probabilities of correct classification in Table 11 are almost the same as the corresponding values in Table 10. Comparing the data in Tables 8 and 11, it is possible to state that it is worth using a compressed image for training (most probabilities of correct classification for classes are better in this case). The classification map is given in Figure 6c and it is quite similar to those in Figure 6a,b. This means that there is no considerable degradation of classification accuracy compared to the previous two cases.

**Table 11.** Confusion matrix for compressed image classified by NN trained for this image (PSNR-HVS-M = 36 dB).

Class	Probability of Decision						
	Road	Field-Y	Field-R	Water	Grass	Trees	Texture
Road	<b>0.507</b>	0.492	0.002	0	0	0	0
Field-Y	0.127	<b>0.562</b>	0.005	0	0.001	0	0.305
Field-R	0	0.069	<b>0.931</b>	0	0	0	0
Water	0	0	0	<b>0.989</b>	0	0	0.01
Grass	0	0	0	0	<b>0.901</b>	0.087	0.013
Tree	0	0	0	0	0.039	<b>0.956</b>	0.005
Texture	0	0.029	0	0.236	0.142	0.141	<b>0.452</b>

Finally, the last case is the classification of the image compressed providing PSNR-HVS-M = 30 dB, i.e., with considerable distortions. The confusion matrix is given in Table 12, the classification map is represented in Figure 6d. Compared to the corresponding data in Table 9, a sufficient increase of  $P_{11}$ ,  $P_{55}$ ,  $P_{66}$ , and  $P_{77}$  is observed,  $P_{22}$ ,  $P_{33}$ , and  $P_{44}$  have increased a little too. This means that it is worth using the compressed image for training, especially if compression is carried out with a large CR.

**Table 12.** Confusion matrix for compressed image classified by NN trained for this image (PSNR-HVS-M = 30 dB).

Class	Probability of Decision						
	Road	Field-Y	Field-R	Water	Grass	Trees	Texture
Road	<b>0.692</b>	0.304	0	0	0.002	0.001	0.01
Field-Y	0.055	<b>0.654</b>	0.009	0	0.019	0	0.263
Field-R	0	0.033	<b>0.967</b>	0.004	0	0	0
Water	0	0	0	<b>0.994</b>	0	0	0.006
Grass	0	0	0	0	<b>0.939</b>	0.035	0.026
Tree	0	0	0	0	0.023	<b>0.972</b>	0.005
Texture	0	0.049	0	0.221	0.113	0.131	<b>0.486</b>

Attentive analysis of the classification map in Figure 6d allows noticing quite many misclassifications near the edge between classes “Water” and “Trees” where these misclassified pixels are related to “Texture”. We associate this effect with edge smearing and other effects that happen near high contrast edges if the compression ratio is quite high.

#### 4.3. Preliminary Conclusions

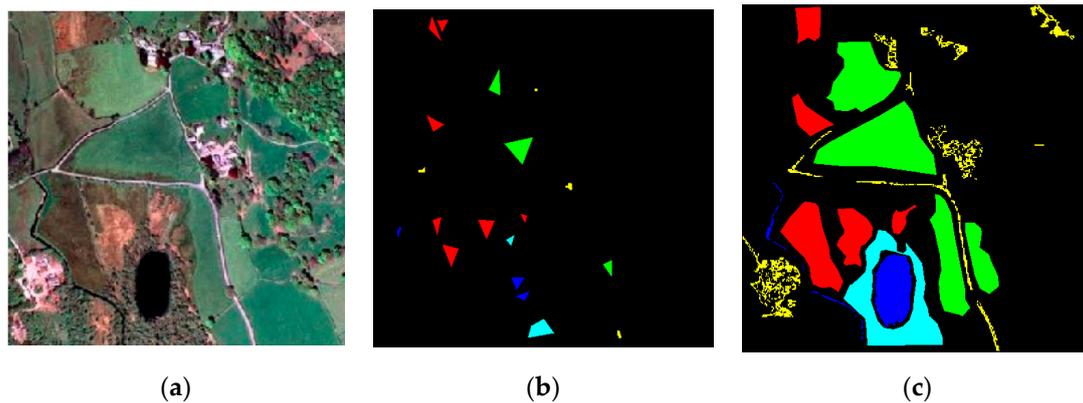
Summarizing the obtained results and conclusions, we can state the following:

1. Classification by any classifier is not perfect if features intersect in feature space and/or if not all possible features are used; here we mean that pixel-wise classification is unable to exploit spatial information that can be useful for improving classification accuracy;
2. If a classifier is trained for original (uncompressed) data and then applied to compressed images, then more compression (a larger CR, a smaller PSNR or PSNR-HVS-M) leads to classification worsening (in general); meanwhile, the difference in Pcc and probabilities of correct classification for particular classes can be negligible; classification accuracy for some particular classes can even improve due to noise/variation suppression (if noise or high frequency variations are present);
3. It is worth training a classifier (at least, MLM and NN-based classifiers) for “conditions” they will be applied; these conditions can be described by quality metrics’ values that are planned to be provided at compression stage or QS value that will be used;
4. Different classifiers can produce sufficiently different classification maps for the same RS data; quantitative parameters that characterize classification can be considerably different as well; a good example are data for Class 6—the MLM classifier does not recognize it well enough whilst the NN classifier performs sufficiently better, the situation is the opposite for Class 7;
5. Lossy compression negatively influences classes that are mainly associated with small-sized and prolonged objects such as roads, narrow rivers, and cricks, small houses; the reason is that lossy compression smears such kinds of objects and changes feature distributions; in this sense, the use of visual quality metrics and setting desired values for them at compression stage can be useful to avoid sufficient degradation of classification accuracy for such classes;
6. For pixel-wise classification, the misclassifications usually appear as isolated pixels; if distortions due to compression are larger, the misclassifications might appear as groups of pixels, especially in the neighborhoods of high contrast edges in RS images; this is undesired and it is worth avoiding such situations that mostly happen if distortions due to lossy compression are visible (in our experiments, if PSNR-HVS-M = 30 dB and, more rarely, if PSNR-HVS-M = 36 dB);
7. This runs us into the idea that, possibly, it is worth carrying out lossy compression either without visually noticeable distortions or with distortions relating to partial suppression of noise in quasi-homogeneous image regions; this can be controlled by either setting the desired PSNR-HVS-M ( $\approx 40$  dB) or the corresponding PCC for the used coder (e.g., QS for AGU).

Certainly, these conclusions are based on analysis of only one test image, and more studies are needed. One opportunity is to check them for real-life RS data.

### 5. Analysis of Real-Life Three-Channel Image

Our experiments with real-life images have been done using a three-channel Landsat TM image earlier used in our studies [38,39]. This test image (shown in pseudo-color representation in Figure 7a) contains three component images acquired in optical bands with central wavelengths equal to 0.66  $\mu\text{m}$ , 0.56  $\mu\text{m}$ , and 0.49  $\mu\text{m}$ , respectively. They have been associated with R, G, and B components of the color image that has a size of 512  $\times$  512 pixels. The image has five recognizable classes, namely, “Soil” (Class 1), “Grass” (Class 2), “Water” (Class 3), “Urban (Roads and Buildings)” (Class 4), and “Bushes” (Class 5). The image fragments used in classifier training are shown in Figure 7b whilst the pixel used for verification of classifiers are marked by the same colors in Figure 7c. Details concerning the numbers of pixels are given in [39]. Class 1 is marked by red color, Class 2—by green, Class 3—by dark blue; Class 4—by yellow, Class 5—by azure.



**Figure 7.** The three-channel image in pseudo-color representation (a); pixel groups used for training (b), pixel groups used for verification (c).

As one can see, the total number of classes is less than for the test image in the previous section, but some classes are the same or similar. Class 1 is similar to Field-R, Class 5—to Class “Texture”. As in test RS data, color features for many classes intersect (consider Classes “Grass” and “Soil”, Classes “Soil” and Bushes”). This means that classification is not a simple task.

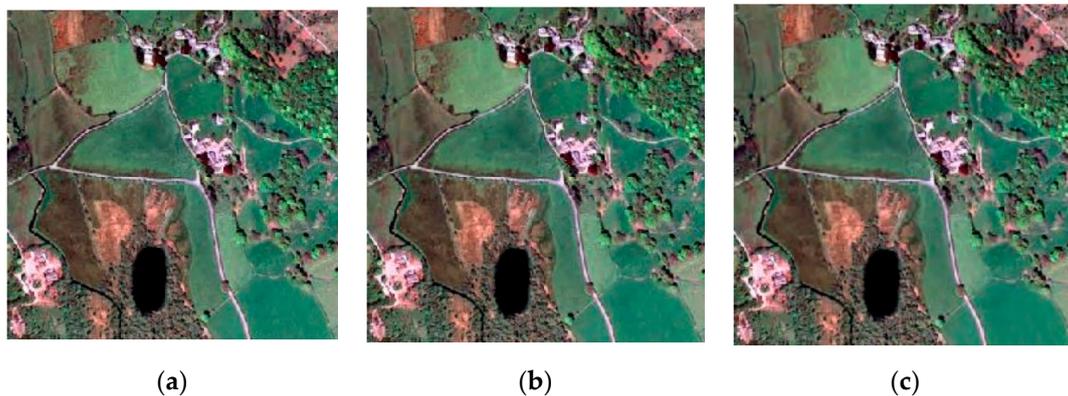
A part of data concerning lossy compression of this image component-wise are presented in Table 13 (more details can be found in [39]). As one can see, to provide  $\text{PSNR-HVS-M}_{\text{des}} = 42$  dB, one has to use  $\text{QS} \approx 17.3$ . To provide  $\text{PSNR-HVS-M}_{\text{des}} = 30$  dB,  $\text{QS} \approx 29.3$ . This is in good agreement with the data in Figure 1 for other test images. Thus, it is possible to give recommendations on what QS to use for providing  $\text{PSNR-HVS-M}_{\text{des}}$  (see details in [23]).

**Table 13.** Compression parameters.

Component	PSNR-HVS-M Desired	QS	PSNR-HVS-M Provided	CR
Pseudo-red	42 dB	17.372	41.941	6.008
	36 dB	29.825	36.059	8.952
	30 dB	55.558	30.007	16.503
Pseudo-green	42 dB	17.285	41.936	5.994
	36 dB	29.856	35.992	8.968
	30 dB	55.134	29.996	16.445
Pseudo-blue	42 dB	17.227	41.847	6.016
	36 dB	29.729	35.998	8.993
	30 dB	55.014	30.0008	16.577

Note that the provided CR can be quite large. It is about 6 for  $\text{PSNR-HVS-M}_{\text{des}} = 42$  dB and reaches 16.5 for  $\text{PSNR-HVS-M}_{\text{des}} = 30$  dB.

We have carried out the analysis for six values of PSNR-HVS-M (45, 42, 39, 36, 33, and 30 dB). Figure 8 presents the considered three-channel image (in pseudo-color representation) compressed providing PSNR-HVS-M equal to 42 dB, 36 dB, and 30 dB. They are all very similar between each other and similar to the original image. A more attentive analysis allows finding differences for images in Figure 8b,c that mainly appear themselves in the neighborhoods of small-sized objects and high contrast edges.



**Figure 8.** Three-channel image in pseudo-color representation: compressed with providing PSNR-HVS-M = 42 dB (a), compressed with providing PSNR-HVS-M = 36 dB (b); compressed with providing PSNR-HVS-M = 30 dB (c).

### 5.1. MLM Classifier Results

Let us start by considering the results of applying the MLM classifier trained for the original image. Tables 14–17 present the obtained data in the form of confusion matrices. As it is seen, probabilities of correct classification for particular classes vary from 0.75 to 0.99 and the smallest probabilities take place for rather heterogeneous classes “Soil” and “Bushes”.

**Table 14.** Classification probabilities for the MLM method trained for the original image and applied to it.

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.747</b>	0.027	0	0.037	0.189
Grass	0.173	<b>0.812</b>	0	0.008	0.007
Water	0	0	<b>0.967</b>	0	0.033
Urban	0.004	0	0	<b>0.989</b>	0.007
Bushes	0.091	0.068	0.013	0.016	<b>0.812</b>

**Table 15.** Classification probabilities (confusion matrix) for the MLM method trained for the original image and applied to the compressed image (PSNR-HVS-M = 42 dB).

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.732</b>	0.025	0	0.043	0.199
Grass	0.177	<b>0.808</b>	0	0.008	0.006
Water	0	0	<b>0.963</b>	0	0.037
Urban	0.005	0	0	<b>0.987</b>	0.008
Bushes	0.118	0.07	0.014	0.018	<b>0.779</b>

If the same classifier is applied to a compressed image (PSNR-HVS-M=42 dB), the results are slightly worse (see data in Table 15). Reduction of probabilities of correct classification by 0.002 ... 0.033 takes place. The largest reduction is observed for the most heterogeneous class “Bushes”.

**Table 16.** Classification probabilities (confusion matrix) for the MLM trained for the original image and applied to the compressed image (PSNR-HVS-M = 36 dB).

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.744</b>	0.028	0	0.043	0.185
Grass	0.185	<b>0.799</b>	0	0.009	0.007
Water	0	0	<b>0.959</b>	0	0.041
Urban	0.005	0	0	<b>0.985</b>	0.010
Bushes	0.133	0.074	0.016	0.018	<b>0.758</b>

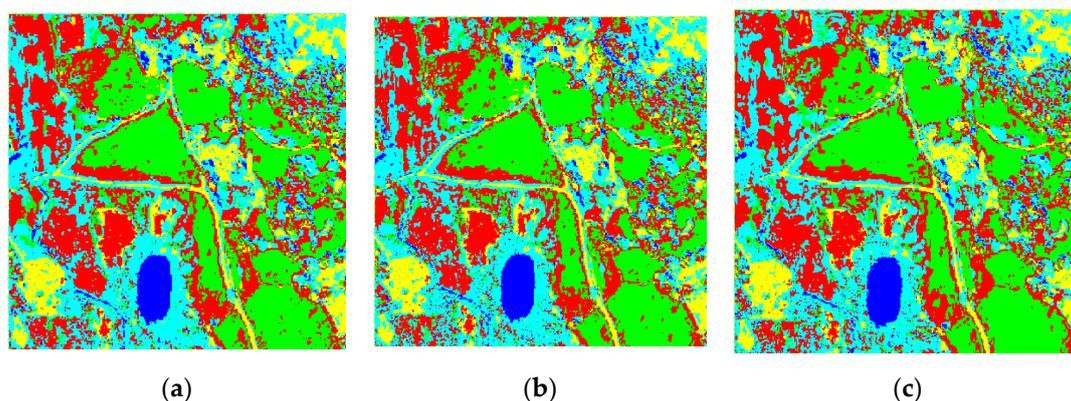
**Table 17.** Classification probabilities (confusion matrix) for the MLM trained for the original image and applied to the compressed image (PSNR-HVS-M = 30 dB).

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.724</b>	0.026	-	0.044	0.206
Grass	0.178	<b>0.809</b>	0	0.006	0.007
Water	-	0	<b>0.942</b>	0	0.058
Urban	0.007	-	0	<b>0.982</b>	0.011
Bushes	0.145	0.079	0.016	0.019	<b>0.741</b>

Consider now the data obtained if the MLM classifier is applied to the image compressed with visible distortions (PSNR-HVS-M = 36 dB). The confusion matrix is given in Table 16. Most probabilities of correct classification for particular classes continue to decrease although this reduction is not essential—up to 0.021 compared to the previous case (Table 15).

Finally, let us analyze data for a compressed image (PSNR-HVS-M = 30 dB) to which the MLM classifier has been applied. The data are presented in Table 17. The tendency is the same—most probabilities of correct classification for particular classes continue to decrease but, again, the reduction is not large.

Classification maps for three images are presented in Figure 9. They are quite similar although some differences can be found. Compression does not lead to radical degradation of classification accuracy.

**Figure 9.** Classification results: for original image (a), for compressed image (PSNR-HVS-M = 36 dB) (b), for compressed (PSNR-HVS-M = 30 dB) (c).

One can be interested in the behavior of  $P_{cc}$  depending on compression. For original image  $P_{cc} = 0.865$ ; for compressed images it equals to 0.853, 0.854, 0.853, 0.849, 0.842, and 0.839 for PSNR-HVS-M equal to 45, 42, 39, 36, 33, and 30 dB, respectively. So, it is possible to state that lossy compression providing PSNR-HVS-M about 40 dB does not lead to sufficient reduction of  $P_{cc}$  for the considered case. Moreover, if training is done for the image compressed with the same conditions as

an image subject to classification, classification results can improve. For example, if training has been done for the compressed image (PSNR-HVS-M = 36 dB) and then applied to this image (to verification set of pixels),  $P_{cc}$  increases to 0.855.

We have analyzed distributions of features before and after compression. One reason why  $P_{cc}$  does not radically reduce with CR increase is that the corresponding distributions do not differ a lot. The largest differences are observed for the classes “Soil” and “Bushes”. It might be slightly surprising that the class “Urban” is recognized so well. The reason is that, for the considered image, features for this class do not sufficiently overlap with features for other classes.

## 5.2. NN Classifier Results

Consider now the results of the NN-based classification of the real-life image. The ideal case data (NN trained for original image is applied to the original image) are given in Table 18. The results can be compared to the data in Table 14. The NN classifier better recognizes the classes “Soil” and “Grass”, the results for the class “Water” are approximately the same.

**Table 18.** Classification probabilities for the NN-based method trained for the original image and applied to it.

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.950</b>	0	0	0	0.05
Grass	0	<b>1</b>	0	0	0
Water	0	0	<b>0.952</b>	0	0.048
Urban	0.098	0	0	<b>0.767</b>	0.135
Bushes	0.187	0.009	0.02	0.009	<b>0.774</b>

Suppose now that this classifier (NN trained for the original image) is applied to compressed images. The results for different qualities of compressed data are presented in Tables 19–21. If PSNR-HVS-M is equal to 42 dB or 36 dB, the results keep practically the same. Only the probability of correct classification for “Bushes” steadily decreases. Reduction of classification accuracy occurs to be larger for the image compressed with PSNR-HVS-M equal to 30 dB. Mainly, reduction takes place for the classes “Water” and “Urban”.

**Table 19.** Classification probabilities (confusion matrix) for the NN-based method trained for the original image and applied to the compressed image (PSNR-HVS-M = 42 dB).

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.942</b>	0	0	0	0.058
Grass	0	<b>0.999</b>	0	0	0.001
Water	0	0	<b>0.942</b>	0	0.058
Urban	0.01	0	0	<b>0.765</b>	0.136
Bushes	0.207	0.016	0.018	0.009	<b>0.751</b>

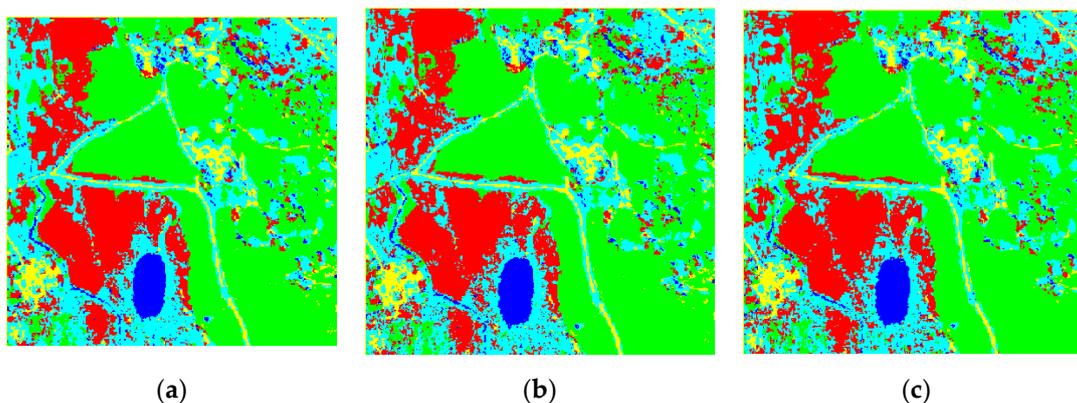
**Table 20.** Classification probabilities (confusion matrix) for the NN-based method trained for original image and applied to the compressed image (PSNR-HVS-M = 36 dB).

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.944</b>	0	0	0	0.057
Grass	0	<b>1</b>	0	0	0
Water	0	0	<b>0.941</b>	0	0.059
Urban	0.108	0	0	<b>0.761</b>	0.131
Bushes	0.236	0.022	0.017	0.007	<b>0.718</b>

**Table 21.** Classification probabilities (confusion matrix) for the NN-based method trained for the original image and applied to the compressed image (PSNR-HVS-M = 30 dB).

Class	Probability of Decision				
	Soil	Grass	Water	Urban	Bushes
Soil	<b>0.941</b>	0	0	0	0.059
Grass	0	<b>0.999</b>	0	0	0.001
Water	0.001	0	<b>0.919</b>	0	0.08
Urban	0.099	0	0	<b>0.747</b>	0.153
Bushes	0.247	0.028	0.014	0.004	<b>0.707</b>

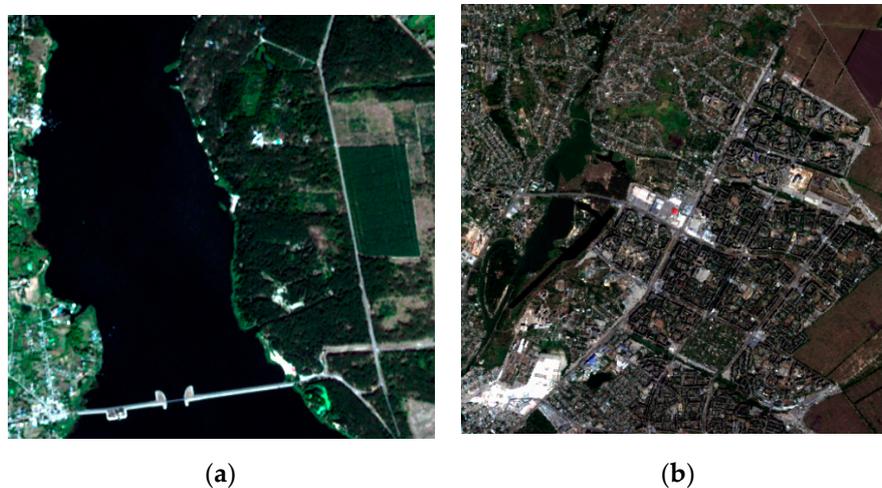
Some classification maps are presented in Figure 10. They do not differ a lot from each other. Some pixels that belong to the class “Water” for the narrow river (see the left low corner in Figure 10c) “disappear” (become misclassified). This is because of the effects of smearing prolonged objects due to lossy compression. The MLM classifier training for compressed images was performed using training samples for fragments shown in Figure 7b. Due to the usage of the same data for both training and validation, noticeable classification improvement can be observed (at least, for several classes).

**Figure 10.** Classification results provided by NN-based method: for original image (a), for compressed image (PSNR-HVS-M = 36 dB) (b); for compressed (PSNR-HVS-M = 30 dB) (c).

### 5.3. Brief Analysis for Sentinel-2 Three-Channel Images

One can be interested whether the observed dependences hold for other images or their fragments, other imagers, and other compression techniques. Sentinel-2 offers wide possibilities to check this since it provides a huge amount of data that can be exploited for different purposes. For our study, we have taken three-channel images of the Kharkiv region (Ukraine) in visible range acquired on 30 August 2019 when there were practically no clouds (images are available at [56]). The analyzed  $512 \times 512$  pixel fragments are for the neighborhood of Staryi Saltiv (45 km north-east from Kharkiv, Ukraine, set 1) and north part of Kharkiv (set 2)—see Figure 11. The main reason for choosing these fragments is

the availability of ground truth data that allow easy marking of four typical classes: Urban, Water, Vegetation, and Bare Soil. One more reason is that the image in Figure 11a is considerably less complex (textural) than the image in Figure 11b.



**Figure 11.** Fragments of Sentinel-2 images of Staryi Saltiv (a) and north part of Kharkiv (b).

First, these fragments have been compressed component-wise with providing a set of PSNR-HVS-M values using three coders: AGU used in the previous part of this paper, SPIHT [57] which can be considered as an analog of JPEG2000 standard, and Advanced DCT coder (ADCTC) [58] that uses partition scheme optimization for better adaptation to image content. Basic data on compression performance are given in Table 22. The observed dependences are predictable. CR for all coders increases if desired PSNR-HVS-M reduces. CR values for different components for the same desired PSNR-HVS-M and set differ but not considerably. CR for AGU is usually slightly larger than for SPIHT (for the same conditions), ADCTC outperforms both coders. CR values for set 2 images are several times smaller than for the corresponding set 1 images due to the higher complexity of set 2 RS data.

**Table 22.** CR comparison for real-life data for three coders, both sets.

Color	Desired PSNR-HVS-M	CR for AGU	CR for SPIHT	CR for ADCTC
Pseudo-Red, set 1	45.000	10.338	9.467	11.306
	39.000	18.383	16.631	20.362
	33.000	32.929	29.303	36.465
Pseudo-Red, set 2	45.000	3.317	3.283	3.498
	39.000	4.520	4.762	4.886
	33.000	7.121	7.360	7.986
Pseudo-Green, set 1	45.000	7.179	5.957	7.513
	39.000	12.666	10.914	13.513
	33.000	23.907	20.833	26.030
Pseudo-Green, set 2	45.000	3.214	3.150	3.363
	39.000	4.356	4.564	4.684
	33.000	6.784	7.105	7.573
Pseudo-Blue, set 1	45.000	6.665	5.727	7.216
	39.000	11.942	10.336	12.740
	33.000	24.914	22.346	27.191
Pseudo-Blue, set 2	45.000	3.212	3.252	3.418
	39.000	4.314	4.698	4.785
	33.000	6.882	7.380	7.817

Second, classifier training has been performed and probabilities of correct classification for all four classes as well as the total probability of correct classification have been obtained. For the NN classifier and AGU coder, the data are presented in Tables 23 and 24. Here and below abbreviation SS1 means that classification has been done for an original (uncompressed) image 1 whilst, e.g., SS1\_45 means that classification has been done for the image compressed with providing PSNR-HVS-M = 45 dB. Training has been done for the original image. As one can see, compression results for the Set 1 image practically do not depend on compressed image quality. Even for PSNR-HVS-M = 30 dB probabilities of correct classification are practically the same as for original data. Meanwhile, for the Set 2 image, the situation is another. There is a tendency for classification accuracy degradation if compressed image quality becomes worse. This is mainly due to the reduction of correct classification probabilities for more heterogeneous classes, e.g., vegetation. For this image, it is possible to recommend compression with providing PSNR-HVS-M about 42 dB to avoid considerable reduction of classification accuracy.

**Table 23.** Probabilities of correct classification depending on compressed image quality for Set 1, AGU coder, NN classifier.

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0.774	0.777	0.775	0.777	0.778	0.773	0.786
Water	0.998	0.998	0.998	0.999	0.999	0.999	0.998
Vegetation	0.915	0.913	0.914	0.915	0.918	0.921	0.927
Bare soil	0.809	0.811	0.809	0.809	0.812	0.798	0.801
P <sub>total</sub>	0.874	0.875	0.874	0.875	0.877	0.873	0.878

**Table 24.** Probabilities of correct classification depending on compressed image quality for Set 2, AGU coder, NN classifier.

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0.877	0.872	0.871	0.865	0.861	0.861	0.864
Water	0.654	0.659	0.662	0.661	0.654	0.644	0.647
Vegetation	0.857	0.825	0.815	0.801	0.788	0.797	0.788
Bare soil	0.890	0.867	0.865	0.869	0.866	0.860	0.870
P <sub>total</sub>	0.820	0.810	0.803	0.799	0.792	0.791	0.792

We have also analyzed the possibility of using compressed images for training. The probability of correct classification has improved by about 0.01 for Set 1 image and remained practically the same for Set 2 image.

Another part of our study relates to classification by MLM (the same maps have been used for training the NN and MLM). MLM has been applied to data compressed by three aforementioned coders. For the AGU coder, the obtained data are presented in Tables 25 and 26.

**Table 25.** Probabilities of correct classification depending on compressed image quality for Set 1, AGU coder, ML classifier.

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0.695	0.727	0.727	0.735	0.745	0.743	0.771
Water	0.979	0.99	0.992	0.992	0.993	0.992	0.988
Vegetation	0.9	0.896	0.896	0.895	0.899	0.901	0.902
Bare soil	0.85	0.852	0.851	0.849	0.852	0.831	0.826
P <sub>total</sub>	0.856	0.866	0.866	0.868	0.872	0.867	0.872

**Table 26.** Probabilities of correct classification depending on compressed image quality for Set 2, AGU coder, ML classifier.

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0.917	0.914	0.913	0.913	0.91	0.913	0.918
Water	0.923	0.902	0.893	0.885	0.873	0.873	0.857
Vegetation	0.691	0.667	0.647	0.645	0.634	0.631	0.619
Bare soil	0.856	0.819	0.815	0.817	0.812	0.802	0.812
P <sub>total</sub>	0.847	0.826	0.817	0.815	0.807	0.805	0.801

It is seen that lossy compression leads to a positive effect for the Set 1 image. For all classes except Bare Soil, the probabilities of correct classification improve or remain the same. Total probability also improves and remains at approximately the same level for all considered qualities of the compressed image (Table 25). Meanwhile, for the Set 2 image, larger CR leads to a steady reduction of total probability and decreasing of probabilities for most classes Table 26).

We have also checked whether or not it is worth carrying out training for compressed images instead of uncompressed ones. The answer, as earlier, is yes. In particular, for images compressed with providing PSNR-HVS-M = 42 dB. The total probabilities equal to 0.893 for the Set 1 data and 0.858 for the Set 2 image. Thus, it is worth using those data in training that have been obtained with the same compression conditions as images subject to classification.

Data obtained for the SPIHT coder are presented in Tables 27 and 28. Their analysis shows the following. Again, lossy compression has a small impact on the classification of the Set 1 data. An optimum is observed for PSBR-HVS-M about 40 dB. The impact of compression for the Set 2 data is also small. For some classes, probabilities of correct classification improve, for others become slightly worse. In aggregate, the total probability remains almost the same.

**Table 27.** Probabilities of correct classification depending on compressed image quality for Set 1, SPIHT coder, ML classifier.

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0.695	0.721	0.724	0.728	0.735	0.742	0.759
Water	0.979	0.991	0.991	0.993	0.993	0.992	0.985
Vegetation	0.9	0.899	0.901	0.9	0.904	0.908	0.906
Bare soil	0.85	0.856	0.864	0.858	0.858	0.848	0.816
P <sub>total</sub>	0.856	0.866	0.87	0.87	0.872	0.873	0.867

**Table 28.** Probabilities of correct classification depending on compressed image quality for Set 2, SPIHT coder, ML classifier.

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0.917	0.915	0.915	0.917	0.916	0.92	0.924
Water	0.923	0.918	0.913	0.903	0.892	0.884	0.875
Vegetation	0.691	0.686	0.688	0.692	0.697	0.699	0.707
Bare soil	0.856	0.847	0.858	0.862	0.853	0.861	0.859
P <sub>total</sub>	0.847	0.841	0.843	0.843	0.839	0.841	0.841

Consider now the data obtained for ADCTC. They are presented in Tables 29 and 30. Their analysis shows the same tendencies as those observed for the SPIHT coder.

**Table 29.** Probabilities of correct classification depending on compressed image quality for Set 1, ADCTC, ML classifier.

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0.695	0.723	0.725	0.734	0.736	0.751	0.76
Water	0.979	0.992	0.994	0.995	0.995	0.995	0.995
Vegetation	0.9	0.898	0.9	0.901	0.901	0.91	0.913
Bare soil	0.85	0.863	0.855	0.85	0.864	0.84	0.827
P <sub>total</sub>	0.856	0.869	0.869	0.87	0.874	0.874	0.874

**Table 30.** Probabilities of correct classification depending on compressed image quality for Set 2, ADCTC coder, ML classifier.

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0.917	0.914	0.911	0.91	0.911	0.912	0.915
Water	0.923	0.918	0.912	0.906	0.909	0.901	0.898
Vegetation	0.691	0.674	0.688	0.682	0.664	0.67	0.693
Bare soil	0.856	0.848	0.848	0.847	0.843	0.859	0.868
P <sub>total</sub>	0.847	0.838	0.84	0.836	0.832	0.836	0.844

## 6. Discussion

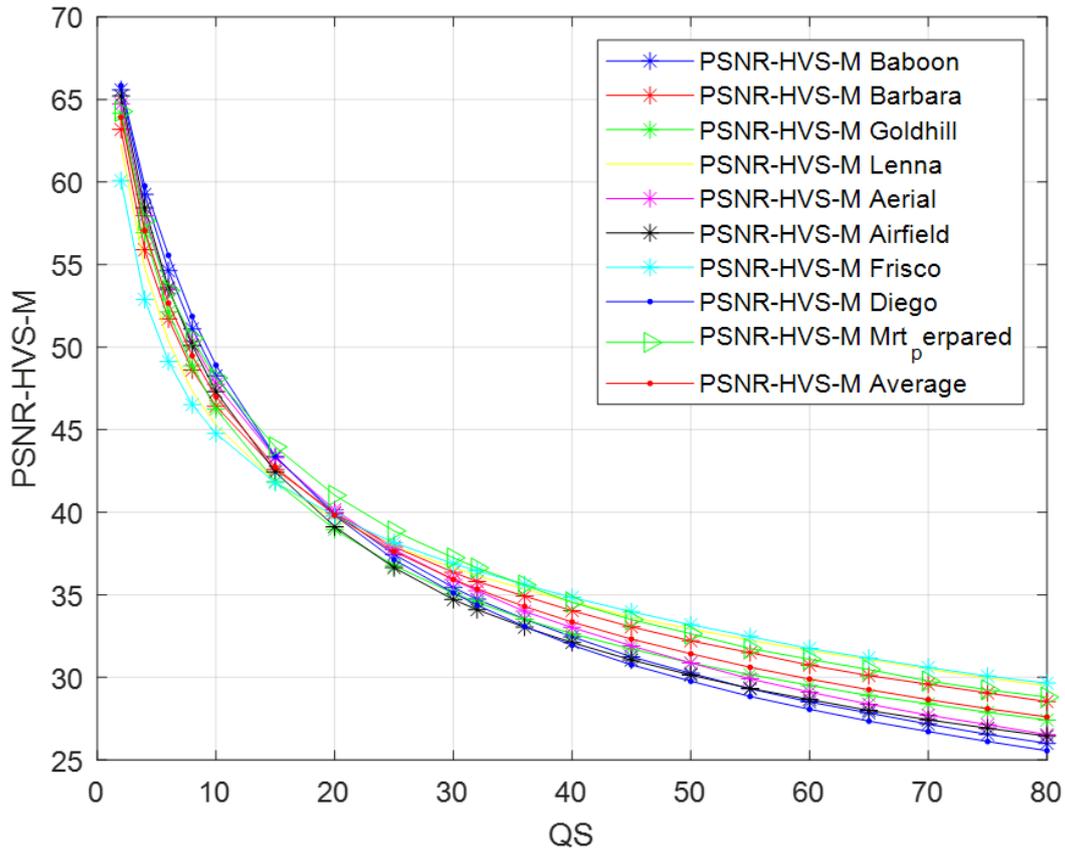
Above, we have considered tendencies for one test and three real-life three-channel images presented as 8-bit 2D data for each component. In practice, images can be presented differently, for example, by 16-bit data [13] or by 10-bit data after certain normalization [59]. Then, a question arises how to provide the desired PSNR-HVS-M (e.g., 40 dB) for a given multichannel image to be compressed. To answer this question, let us consider some data. First, Figure 11 taken from [60] presents the dependences of PSNR-HVS-M on QS for nine 8-bit grayscale test images of different complexity for AGU. The average curve obtained for these nine test images is also given. It is seen that this average curve allows the approximate setting of QS to provide the desired PSNR-HVS-M. For example, to provide PSNR-HVS-M  $\approx$  43 dB, one has to set  $QS_{rec} \approx 15$ . To provide PSNR-HVS-M  $\approx$  40 dB, it is possible to set  $QS_{rec} \approx 20$ . If D is not equal to 255, then the recommended QS is

$$QS_{recD} = QS_{rec}D/255, \quad (9)$$

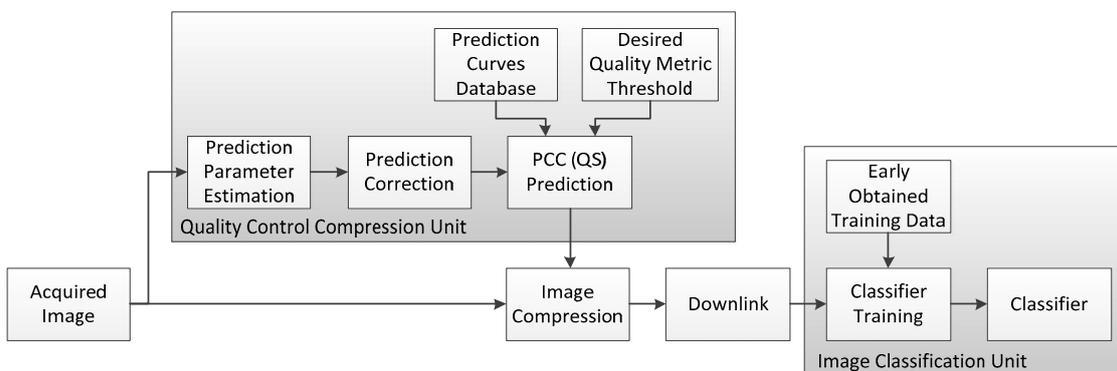
where  $QS_{rec}$  is determined from the average curve in Figure 11. As it follows from the analysis of data in Figure 11, the use of  $QS_{rec}$  or  $QS_{recD}$  provides PSNR-HVS-M approximately, errors can be up to 1...2 dB depending upon the complexity of an image to be compressed. Such accuracy can be treated as acceptable, since, as it is shown in the previous Section, change of PSNR-HVS-M by even 2 dB does not lead to radical changes of  $P_{cc}$  and probabilities of correct classification for particular classes. Second, if errors in providing the desired PSNR-HVS-M are inappropriate, accuracy can be improved by applying a two-step procedure proposed in [60]. As it has been shown in Table 13, QS that should be used in component-wise compression is practically the same for all components. This means that it is enough to determine  $QS_{rec}$  or  $QS_{recD}$  for one component image and then apply it for compressing other components (this can save time and resources at the data compression stage). Moreover,  $QS_{rec}$  or  $QS_{recD}$  determined according to recommendations given above can be used in joint compression of all components of multichannel images or groups of components [59] by the 3D version of AGU. In this case, the positive effect is twofold. First, a larger CR is provided compared to the component-wise compression. Second, a slightly larger quality of the compressed image can be ensured.

Figure 12 demonstrates the dependences PSNR-HVS-M on QS for different test images. Figure 13 shows the RS data processing flowchart. Acquired images (e.g., on-board) are subject to “careful” lossy compression in Quality Control Compression Unit where PCC (e.g., QS) is determined using

the desired threshold for a chosen quality metric (e.g., 40 dB for PSNR-HVS-M) and rate/distortion curves obtained in advance (like those in Figure 12). PCC corrections can be done if, e.g., images are normalized before compression. In the Image Classification Unit, RS data are subject to classification where the classifier can be trained using either earlier processed images or data that have been just received. If time limitations are not strict, the second option seems preferable (according to results obtained in our analysis).



**Figure 12.** Dependences PSNR-HVS-M on QS for nine test images (see the list in the upper part of the plot) and the average curve for AGU.



**Figure 13.** Flowchart of the proposed processing approach.

Another question that has arisen several times is the following—is it possible to improve classification accuracy? One answer that easily follows from our analysis is that the pixel-wise classification (at least, its simple version used by us) does not allow exploiting information from neighboring pixels. Such information can be of different types and it can be used differently. For example, a pixel can be preliminarily classified as belonging to a homogeneous, textural, or locally active area (by locally active, we mean pixel belonging to a small-sized object or edge or their neighborhood). Then, different features can be determined and used for such pre-classified pixels. However, this approach is out of the scope of this paper.

There are also two other approaches possible. They have been proposed in [39]. First, several elementary classifiers (for example, MLM, NN, and SVM ones) can be applied in parallel, and then their outputs can be combined. Second, post-processing of classification results with preliminary detection of edges can be performed. Note that these approaches can lead to a sufficient (up to 0.1 . . . 0.2) increase of  $P_{cc}$ , especially if it is low for originally classified data. Meanwhile, more experiments for verifying the methods developed in [39] are needed.

Finally, the last question is how to improve CR without losing classification performance for multichannel RS data? In our opinion, 3D compression should be applied. However, in this case, additional studies are also desired.

## 7. Conclusions

We have considered the task of lossy compression of multichannel images by introducing different levels of distortions to analyze their influence on classification accuracy. Two classifiers, namely MLM and NN, have been studied. The DCT-based coder AGU has been used. In addition, two other transform-based coders have been employed with a brief analysis of their performance. One artificial and three real-life three-channel images have been thoroughly analyzed. Component-wise compression has been addressed with quality controlled (characterized) by visual quality metric PSNR-HVS-M.

The following has been shown for the case of classifier training for original (uncompressed) data;

- there is no sufficient decrease in classification accuracy if images are compressed without visible distortions (PSNR-HVS-M exceeds 40 . . . 42 dB); in some cases, TPCC can even increase compared to TPCC for original (uncompressed) data;
- lossy compression might have no considerable negative impact on classification accuracy even if distortions are visible (PSNR-HVS-M < 40 dB) but this happens if a compressed image has a simple structure;
- distortions of the aforementioned level result in different CR depending upon image complexity; this means that images with complex structure have to be compressed with smaller CR to avoid considerable reduction of  $P_{cc}$ ;
- dependences of probabilities of correct classification of particular classes on compression parameters are different; for classes mostly represented by large-size objects, lossy compression is not so crucial; it might even happen that lossy compression can improve classification accuracy if noise or high-frequency variations are partly suppressed due to it;
- for classes represented by small-sizes and prolonged objects as well as some textures, lossy compression might lead to sufficient reduction of classification accuracy. This is due to the image smearing effect that transforms features.

The situation is slightly different if the training is done for compressed data. Then, it is worth using for training such compressed data that are characterized by the same quality (for example, approximately the same PSNR-HVS-M value) as data to which classification will be applied. Then, even for PSNR-HVS-M about 36 dB, the classification can be almost as good enough as for the original image classified by the classifier trained for original data.

It has been also shown that the main dependencies are observed for different coders (in particular, compression techniques based on DCT and wavelets) and for data provided by different multispectral

systems. We also believe that other than PSNR-HVS-M visual quality metrics can be applied within the proposed approach to lossy compression. Thus, the approach is quite general.

Finally, practical recommendations on how to set coder parameters to avoid sufficient reduction of classification accuracy are given. Possibilities of classification improvement are briefly mentioned and directions of future research are described.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-4292/12/22/3840/s1>, The Spearman Rank-Order Correlations Coefficient (SROCC) values for 49 quality metrics for TID 2013 database.

**Author Contributions:** I.V. created and tested MLM classifier, O.R. created and tested NN-based classifier, F.L. performed compression with desired characteristics; S.K. participated in model data design, S.A. was responsible for simulation results analysis, V.L. put forward the idea of coder parameter selection and prepared the paper draft, B.V. was responsible for real-life image result analysis, K.C. proposed methodology of investigation, K.E. carried out editing and supervision. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Mielke, C.; Boshce, N.K.; Rogass, C.; Segl, K.; Gauert, C.; Kaufmann, H. Potential Applications of the Sentinel-2 Multispectral Sensor and the ENMAP hyperspectral Sensor in Mineral Exploration. In Proceedings of the EARSeL eProceedings, Warsaw, Poland, 13 February 2014; Volume 13, pp. 93–102.
- Schowengerdt, R.A. *Remote Sensing: Models and Methods for Image Processing*, 3rd ed.; Academic Press: San Diego, CA, USA, 2007; p. 515.
- Joshi, N.; Baumann, M.; Ehammer, A.; Fensholt, R.; Grogan, K.; Hostert, P.; Jepsen, M.R.; Kuemmerle, T.; Meyfroidt, P.; Mitchard, E.T.A.; et al. A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Rem. Sens.* **2016**, *8*, 23. [[CrossRef](#)]
- Deledalle, C.; Denis, L.; Tabti, S.; Tupin, F. MuLoG, or How to Apply Gaussian Denoisers to Multi-Channel SAR Speckle Reduction? *IEEE Trans. Image Process.* **2017**, *26*, 4389–4403. [[CrossRef](#)] [[PubMed](#)]
- Ponmaryov, V.; Rosales, A.; Gallegos, F.; Loboda, I. Adaptive vector directional filters to process multichannel images. *Ieice Trans. Commun.* **2007**, *90*, 429–430. [[CrossRef](#)]
- Zhong, P.; Wang, R. Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2269–2275. [[CrossRef](#)]
- First Applications from Sentinel-2A. Available online: [http://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Sentinel-2/First\\_applications\\_from\\_Sentinel-2A](http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/First_applications_from_Sentinel-2A) (accessed on 10 January 2019).
- Kussul, N.; Lemoine, G.; Gallego, F.J.; Skakun, S.; Lavreniuk, M.; Shelestov, A. Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2500–2508. [[CrossRef](#)]
- Bioucas-Dias, J.M.; Plaza, A.; Camps-Valls, G.; Scheunders, P.; Nasrabadi, N.; Chanussot, J. Hyperspectral Remote Sensing Data Analysis and Future Challenges. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–36. [[CrossRef](#)]
- Chi, M.; Plaza, A.; Benediktsson, J.A.; Sun, Z.; Shen, J.; Zhu, Y. Big Data for Remote Sensing: Challenges and Opportunities. *Proc. IEEE* **2016**, *104*, 2207–2219. [[CrossRef](#)]
- Christophe, E. Hyperspectral Data Compression Tradeoff. *Opt. Remote Sens. Adv. Signal Process. Exploit. Tech.* **2011**, *7*, 9–29.
- Blanes, I.; Magli, E.; Serra-Sagrista, J. A Tutorial on Image Compression for Optical Space Imaging Systems. *IEEE Geosci. Remote Sens. Mag.* **2014**, *2*, 8–26. [[CrossRef](#)]
- Zemliachenko, A.N.; Kozhemiakin, R.A.; Uss, M.L.; Abramov, S.K.; Ponomarenko, N.N.; Lukin, V.V. Lossy compression of hyperspectral images based on noise parameters estimation and variance stabilizing transform. *J. Appl. Remote Sens.* **2014**, *8*, 25. [[CrossRef](#)]

14. Manolakis, D.; Lockwood, R.; Cooley, T. On the Spectral Correlation Structure of Hyperspectral Imaging Data. In Proceedings of the IGARSS 2008—2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008; pp. II-581–II-584.
15. Lam, K.W.; Lau, W.; Li, Z. The effects on image classification using image compression technique. *Amsterdam, Int. Arch. Photogramm. Remote Sens.* **2000**, *33*, 744–751.
16. Penna, B.; Tillo, T.; Magli, E.; Olmo, G. Transform Coding Techniques for Lossy Hyperspectral Data Compression. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1408–1421. [[CrossRef](#)]
17. Kaur, A.; Wasson, V. Compressing Land Images using Fractal Lossy Compression. *Int. J. Eng. Comput. Sci.* **2014**, *3*, 8806–8811.
18. Taubman, D.; Marcellin, M. *JPEG2000 Image Compression Fundamentals, Standards and Practice*; Kluwer Academic Publishers: Boston, MA, USA, 2002; p. 777.
19. Santos, L.; Lopez, S.; Callico, G.; Lopez, J.; Sarmiento, R. Performance evaluation of the H.264/AVC video coding standard for lossy hyperspectral image compression. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2012**, *5*, 451–461. [[CrossRef](#)]
20. Krivenko, S.; Krylova, O.; Bataeva, E.; Lukin, V. Smart Lossy Compression of Images Based on Distortion Prediction. *Telecommun. Radio Eng.* **2018**, *77*, 1535–1554. [[CrossRef](#)]
21. Khelifi, F.; Bouridane, A.; Kurugollu, F. Joined spectral trees for scalable SPIHT-based multispectral image compression. *IEEE Trans. Multimed.* **2008**, *10*, 316–329. [[CrossRef](#)]
22. Zemliachenko, A.; Kozhemiakin, R.; Abramov, S.; Lukin, V.; Vozel, B.; Chehdi, K.; Egiazarian, K. Prediction of compression ratio for DCT-based coders with application to remote sensing images. *J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 257–270. [[CrossRef](#)]
23. Zemliachenko, A.; Ponomarenko, N.; Lukin, V.; Egiazarian, K.; Astola, J. Still Image/Video Frame Lossy Compression Providing a Desired Visual Quality. *Multidimens. Syst. Signal Process.* **2015**, *27*, 697–718. [[CrossRef](#)]
24. Shi, C.; Wang, L.; Zhang, J.; Miao, F.; He, P. Remote Sensing Image Compression Based on Direction Lifting-Based Block Transform with Content-Driven Quadtree Coding Adaptively. *Rem. Sens.* **2018**, *10*, 999. [[CrossRef](#)]
25. Balasubramanian, R.; Ramakrishnan, S.S. Wavelet application in compression of a remote sensed image. In Proceedings of the 2013 the International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE 2013), Nanjing, China, 26–28 July 2013; pp. 659–662.
26. Jagadeesh, P.; Kumar, S.; Nedumaan, J. Accepting the Challenges in Devising Video Game Leeway and Contrivance. *Int. J. Inf. Secur.* **2015**, *14*, 582–590.
27. Ozah, N.; Kolokolova, A. Compression improves image classification accuracy. In Proceedings of the 32nd Canadian Conference on Artificial Intelligence, Canadian AI 2019, Kingston, ON, Canada, 28–31 May 2019; pp. 525–530.
28. Chen, Z.; Ye, H.; Yingxue, Z. Effects of Compression on Remote Sensing Image Classification Based on Fractal Analysis. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4577–4589. [[CrossRef](#)]
29. García-Sobrino, J.; Laparra, V.; Serra-Sagristà, J.; Calbet, X.; Camps-Valls, G. Improved Statistically Based Retrievals via Spatial-Spectral Data Compression for IASI Data. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5651–5668. [[CrossRef](#)]
30. Gimona, A.; Poggio, L.; Aalders, I.; Aitkenhead, M. The effect of image compression on synthetic PROBA-V images. *Int. J. Remote Sens.* **2014**, *35*, 2639–2653. [[CrossRef](#)]
31. Perra, C.; Atzori, L.; De Natale, F.G.B. Introducing supervised classification into spectral VQ for multi-channel image compression. In Proceedings of the IGARSS 2000, IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment (Cat. No.00CH37120), Honolulu, HI, USA, 24–28 July 2000; pp. 597–599. [[CrossRef](#)]
32. Izquierdo-Verdiguier, E.; Laparra, V.; Gómez-Chova, L.; Camps-Valls, G. Encoding Invariances in Remote Sensing Image Classification With SVM. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 981–985. [[CrossRef](#)]
33. Zabala, A.; Pons, X. Impact of lossy compression on mapping crop areas from remote sensing. *Int. J. Remote Sens.* **2013**, *34*, 2796–2813. [[CrossRef](#)]

34. Zabala, A.; Pons, X.; Diaz-Delgado, R.; Garcia, F.; Auli-Llinas, F.; Serra-Sagrista, J. Effects of JPEG and JPEG2000 Lossy Compression on Remote Sensing Image Classification for Mapping Crops and Forest Areas. In Proceedings of the 2006 IEEE International Symposium on Geoscience and Remote Sensing, Denver, CO, USA, 31 July–4 August 2006; pp. 790–793. [CrossRef]
35. Popov, M.A.; Stankevich, S.A.; Lischenko, L.P.; Lukin, V.V.; Ponomarenko, N.N. Processing of hyperspectral imagery for contamination detection in urban areas. In *Environmental Security and Ecoterrorism*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 25–26.
36. Ponomarenko, N.; Krivenko, S.; Lukin, V.; Egiazarian, K.; Astola, J. Lossy Compression of Noisy Images Based on Visual Quality: A Comprehensive Study. *Eurasip. Jasp.* **2010**, *2010*, 976436. [CrossRef]
37. Al-Shaykh, O.K.; Mersereau, R.M. Lossy compression of noisy images. *IEEE Trans. Image Process* **1998**, *7*, 1641–1652. [CrossRef]
38. Lukin, V. Processing of Multichannel RS data for Environment Monitoring. *Proc. Nato Adv. Res. Workshop Geogr. Inf. Process. Vis. Anal. Env. Secur.* **2009**, *1*, 129–138.
39. Proskura, G.; Vasilyeva, I.; Fangfang, L.; Lukin, V. Classification of Compressed Multichannel Images and Its Improvement. In Proceedings of the 2020 30th International Conference Radioelektronika (RADIOELEKTRONIKA), Bratislava, Slovakia, 15–16 April 2020; pp. 62–67.
40. Lukin, V.; Zemliachenko, A.; Krivenko, S.; Vozel, B.; Chehdi, K. *Lossy Compression of Remote Sensing Images with Controllable Distortions. Satellite Information Classification and Interpretation*; Rustam, B.R., Ed.; IntechOpen: Londone, UK, 2018. [CrossRef]
41. Krivenko, S.S.; Abramov, S.K.; Lukin, V.V.; Vozel, B.; Chehdi, K. Lossy DCT-based compression of remote sensing images with providing a desired visual quality. In *Image and Signal Processing for Remote Sensing XXV*; International Society for Optics and Photonics: Bellingham, WA, USA, 2019; p. 11155. [CrossRef]
42. Ponomarenko, N.N.; Lukin, V.V.; Egiazarian, K.; Astola, J. DCT Based High Quality Image Compression. In *Scandinavian Conference on Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 1177–1185.
43. Lukin, V.; Abramov, S.; Krivenko, S.; Kurekin, A.; Pogrebnyak, O. Analysis of classification accuracy for pre-filtered multichannel remote sensing data. *J. Expert Syst. Appl.* **2013**, *40*, 6400–6411. [CrossRef]
44. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC Press: Florida, FL, USA, 1999.
45. Lukin, V.; Ponomarenko, N.; Egiazarian, K.; Astola, J. Analysis of HVS-Metrics' Properties Using Color Image Database TID2013. *Proc. Acivs* **2015**, *9386*, 613–624.
46. Lin, W.; Jay Kuo, C.C. Perceptual visual quality metrics: A survey. *J. Vis. Commun. Image Represent.* **2011**, *22*, 297–312. [CrossRef]
47. Ponomarenko, N.; Silvestri, F.; Egiazarian, K.; Carli, M.; Astola, J.; Lukin, V. On between-coefficient contrast masking of DCT basis functions. In Proceedings of the Third International Workshop on Video Processing and Quality Metrics, Scottsdale, AZ, USA, 25–26 January 2007; Volume 7.
48. Available online: <http://ponomarenko.info/psnrhvs.htm> (accessed on 20 November 2020).
49. Ieremeiev, O.; Lukin, V.; Okarma, K.; Egiazarian, K. Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images. *Remote Sens.* **2020**, *12*, 2349. [CrossRef]
50. Sun, J.; Yang, J.; Zhang, C.; Yun, W.; Qu, J. Automatic remotely sensed image classification in a grid environment based on the maximum likelihood method. *Math. Comput. Model.* **2013**, *58*, 573–581. [CrossRef]
51. Sisodia, P.S.; Tiwari, V.; Kumar, A. Analysis of Supervised Maximum Likelihood Classification for remote sensing image. In *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*; IEEE: Piscataway, NY, USA, 2014; pp. 1–4. [CrossRef]
52. Parresol, B.R. *Recovering Parameters of Johnson's  $S_B$  Distribution*; US Department of Agriculture, Forest Service, Southern Research Station: Asheville, NC, USA, 2003; Volume 9.
53. Lemonte, A.J.; Bazán, J.L. New class of Johnson SB distributions and its associated regression model for rates and proportions. *Biom. J.* **2016**, *58*, 727–746. [CrossRef]
54. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
55. Bioucas-Dias, J.M. Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 354–379. [CrossRef]

56. Available online: [https://apps.sentinel-hub.com/eo-browser/?lat=46.45&lng=34.12&zoom=6&time=2019-11-03&preset=3\\_NDVI&datasource=Sentinel-2%20L1C](https://apps.sentinel-hub.com/eo-browser/?lat=46.45&lng=34.12&zoom=6&time=2019-11-03&preset=3_NDVI&datasource=Sentinel-2%20L1C) (accessed on 20 November 2020).
57. Taubman, D.; Marcellin, M.W. *JPEG2000 Image Compression: Fundamentals, Standards and Practice*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2002. [[CrossRef](#)]
58. Ponomarenko, N.N.; Egiazarian, K.O.; Lukin, V.V.; Astola, J.T. High-Quality DCT-Based Image Compression Using Partition Schemes. *IEEE Signal Process. Lett.* **2007**, *14*, 105–108. [[CrossRef](#)]
59. Zemliachenko, A.; Lukin, V.; Ieremeiev, O.; Vozel, B. Peculiarities of Hyperspectral Image Lossy Compression for Sub-band Groups. In Proceedings of the 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), Lviv, Ukraine, 2–6 July 2019; pp. 918–923.
60. Li, F.; Krivenko, S.; Lukin, V. A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression. In Proceedings of the 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, 25–29 February 2020; pp. 502–506. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).