

# The regression model for the procedure of correction of photos damaged by backlighting

A V Goncharova<sup>1</sup>, I V Safonov<sup>1</sup> and I A Romanov<sup>1</sup>

<sup>1</sup>National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe Shosse, 31, Moscow, Russia, 115409

e-mail: alen.gon4arowa@gmail.com

**Abstract.** In the paper, we propose an approach for selection a correction parameter for images damaged by backlighting. We consider the photos containing underexposed areas due to backlit conditions. Such areas are dark and have poorly discernible details. The correction parameter controls the level of amplification of local contrast in shadow tones. Besides, the correction parameter can be considered as a quality estimation factor for such photos. For an automatic selection of the correction parameter, we apply regression by supervised machine learning. We propose new features calculated from the co-occurrence matrix for the training of the regression model. We compare the performance of the following techniques: the least square method, support vector machine, random forest, CART, random forest, two shallow neural networks as well as blending and staking of several models. We apply two-stage approach for the collection of a big dataset for training: initial model is trained on a manually labeled dataset containing about two hundred of photos, after that we use the initial model for searching for photos damaged by backlit in social networks having public API. Such approach allowed to collect about 1000 photos in conjunction with their preliminary quality assessments that were corrected by experts if it was necessary. In addition, we investigate an application of several well-known blind quality metrics for the estimation of photos affected by backlit.

## 1. Introduction

A lot of photos are affected by various defects and need to be enhanced in an automatic manner to be more pleasant for observers. The most noticeable defects are the following: various issues with brightness and contrast, color misbalance, blurring and shaking, compression artifacts, high noise level, red eyes and other artifacts due to flash, color fringing, and geometrical distortions [1]. There are numerous methods for noise suppression (e.g. [2]), red-eye correction (e.g. [3]), and image sharpening (e.g. [4]), but there are just a few publications devoted to correction photos damaged by backlit. Photos taken in backlighting conditions has high global contrast, but local contrast in areas of shadow tones is quite low. Figure 1 demonstrates the photo affected by backlighting. One can see poorly distinguishable details in shadows. It is important to develop a method for enhancement of such images.

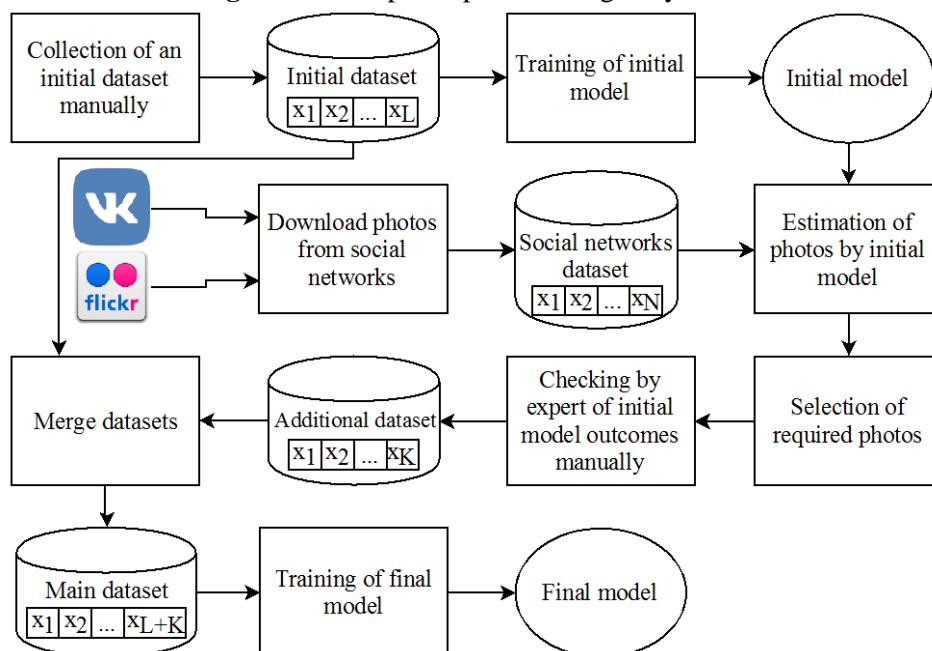
Paper [5] describes the technique for the correction of photos damaged by backlit. That method is based on a contrast stretching and alpha-blending of brightness of the initial image and an estimated reflectance. In the majority cases, the technique provides good visual outcomes, nevertheless, it has

shortcomings. The most important parameter of the method is factor  $k_s$ , which controls the amplification of local contrast in shadows. This factor is calculated by means of decision tree based on features that originated from brightness histogram. The decision tree allows to obtain for  $k_s$  just five discrete values. Sometimes it leads to significant changes in correction power due to insignificant alterations in an image. Such effect is undesirable. Also, that decision tree was created based on several heuristic assumptions and was not verified on the big number of sample photos. Is that solution general for plenty of photos affected by backlit?

The aim of our paper is overcoming of disadvantages of the method from [5] by the development of a regression model for the estimation of  $k_s$  based on machine learning techniques. It is worth to note, also factor  $k_s$  can be treated as blind metrics for assessment of the visual quality of photos damaged by backlighting. In the paper, we discuss three subjects: an approach for collection of a representative dataset; a selection of method for creation of regression model; algorithms for calculation and selection of informative features for the model.



**Figure 1.** Example of photo damaged by backlit.



**Figure 2.** The scheme for dataset collection.

## 2. Collection of a dataset

In the best of our knowledge, there is no publicly available dataset containing photos deteriorated by backlighting. Moreover, a collection of a big dataset is a tiresome task because a relatively small number of such photos are in the Internet and social networks, so, manual search requires many efforts and a long time. We employ a concept of semi-supervised learning and self-training [6] for the

collection of the representative dataset. An initial dataset containing about two hundred of photos was collected and labeled manually. We trained the initial regression model using the initial dataset and random forest technique [7]. Nowadays many social networks provide an application interface for downloading of photos. We made a software tool for download of photos from flickr.com and vk.com [8]. The initial regression model estimates the downloaded photos and collects photos having required  $k_s$  to provide uniform distribution of images by the correction factor. Evidently, the initial model is not ideal, so, an expert checks outcome and makes re-labeling if it is necessary. Further, we add validated photos to the main dataset. Photos from the initial dataset are added to the main dataset as well. In this way, we collected about 1000 labeled images. The final model was trained on the main dataset. Figure 2 illustrates our approach for the dataset collection.

### 3. Analysis of methods for creation of a regression model

We have analyzed the application of the following methods for creation of regression model:

- linear least squares (LLS);
- support vector machine for regression (SVR) [9];
- k-nearest neighborhoods for regression (k-NN) [10];
- classification and regression tree (CART) [11];
- random forest [7];
- feedforward neural network [12];
- single-layer neural network [13];
- averaging of outcomes of enumerated above methods;
- blending [14]
- averaging of several outcomes of blending;
- stacking [15].

There are several measures to evaluate the performance of the regression models. We calculated the following five measures that evaluate conformance of outcomes of the regression model with experts' judgments.

1. Mean absolute error (*MAE*):

$$MAE(y, f) = \frac{1}{N} \sum_{i=1}^N |y_i - f_i|; \quad (1)$$

where  $N$  is the number of elements,  $f_i$  is the predicted value, and  $y_i$  is the true value.

2. Mean squared error (*MSE*):

$$MSE(y, f) = \frac{1}{N} \sum_{i=1}^N (y_i - f_i)^2. \quad (2)$$

3. Median absolute error (*MedAE*):

$$MedAE(y, f) = median(|y_1 - f_1|, \dots, |y_N - f_N|). \quad (3)$$

4. Pearson correlation coefficient ( $r$ ):

$$r(y, f) = 1 - \frac{\sum_{i=1}^N (y_i - \bar{y})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^N (f_i - \bar{f})^2}}. \quad (4)$$

where  $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$ ,  $\bar{f} = \frac{1}{N} \sum_{i=1}^N f_i$ .

5. Normalized area under regression error characteristic curve (*AUC REC*) [16].

We used 5-fold cross-validation with stratification. All models were trained using the features from [5]. Table 1 contains the performance measures of different regression models. For *MAE*, *MSE* and *MedAE* smaller is better. For  $r$  and *AUC REC* larger is better. Random forest and SVR have the highest performance measures. However, the distribution of residuals of the regression model via random forest looks normal in comparison with residuals by SVR. In addition, random forest has a relatively small number of parameters for model adjustment. Thus, we have selected random forest for the creation of regression model for the estimation of amplification factor  $k_s$ .

**Table 1.** Performance measures of different regression models.

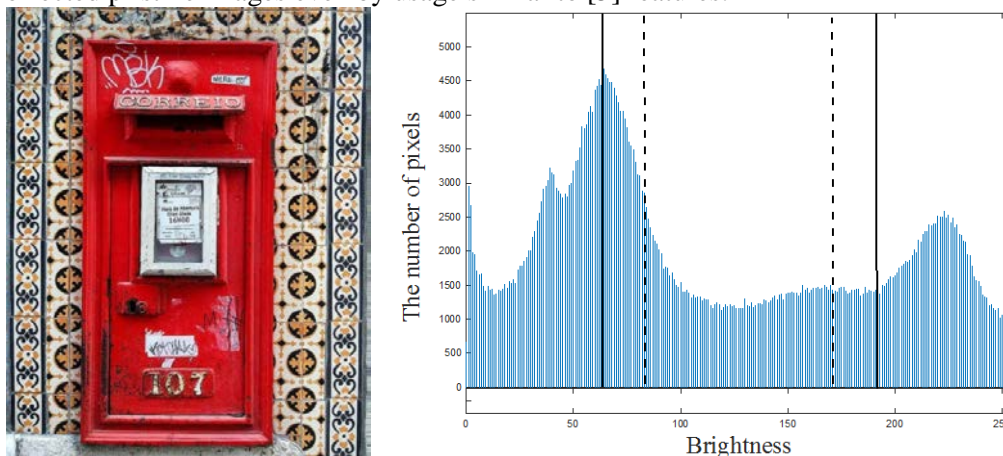
	<i>MAE</i>	<i>MSE</i>	<i>MedAE</i>	<i>r</i>	<i>normAUC</i>
LLS	15.2	350	13.3	0.59	0.694
SVR	15.1	377	<b>12.1</b>	0.54	<b>0.727</b>
k-NN	16.7	453	14.0	0.38	0.709
CART	18.5	559	15.3	0.41	0.723
Random forest	<b>14.8</b>	<b>334</b>	13.1	<b>0.61</b>	0.711
Feedforward neural network	15.0	350	12.9	0.59	0.706
Single-layer neural network	15.0	344	12.9	0.59	0.692
Averaging of methods	<b>14.8</b>	340	12.7	0.59	0.708
Blending	15.0	350	13.2	0.58	0.721
Averaging of several blendings	14.9	340	13.2	0.60	0.718
Stacking	14.9	339	13.4	0.60	0.714

#### 4. Features for the description of photos damaged by backlit

In the paper [5] feature set for estimation of amplification factor  $k_s$  is described. Those features are calculated from brightness histogram  $H$ . Typical histogram of a photo damaged by backlighting has relatively high peaks in shadow and/or highlights, but a gap in middle tones. Besides, the histogram in dark tones is asymmetric. To characterize such shape of histogram the following features were proposed: parts of the tones in the shadow and middle tones; parts of tones in the first and second halves of dark tones; the ratios of the histogram maxima in shadows, middle, and highlight tones per the global histogram maximum; locations of the histogram maxima in shadows and highlights. The entire dynamic range was divided uniformly on shadow, middle tones, and highlights.

However, there are high-quality images that have values of those features close to values for photos affected by backlit. Sometimes normal images have a histogram that has peaks near the boundaries of dynamic range and a valley between them. Figure 3 shows an example of such pristine photo and its histogram of brightness. The method from [5] makes dark areas in the photo a lighter. It is necessary to prevent the modification of dark tones in normal images. To overcome the undesirable correction of high-quality photos, we propose to use another set of features. We modify histogram-based features from [5] and introduce new features originated from the co-occurrence matrix.

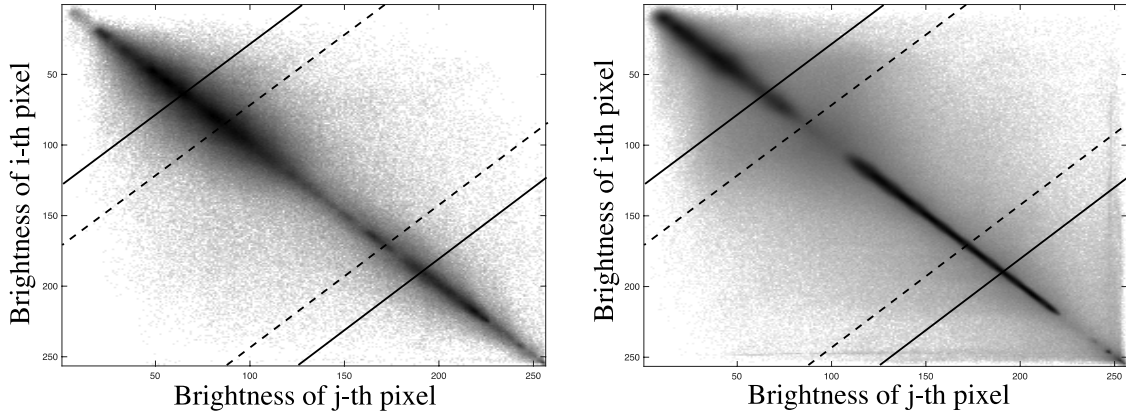
In figure 3 one can see that peaks in the left and right parts of the histogram are shifted towards to middle tones. We propose to treat as shadow tones a quarter of the leftmost part of the dynamic range instead of one third. Highlight tones are a quarter of the rightmost part of the dynamic range. Middle tones occupy half of the range. Solid lines in histogram in figure 3 demonstrate this division for dark, middle and light tones. Dashed lines show a similar differentiation in [5]. This simple alteration of sub-ranges for shadow, highlights, and middle tones allows to decrease considerably the number of falsely corrected pristine images even by usage similar to [5] features.



**Figure 3.** The high-quality photos and corresponding brightness histogram which looks like the histogram of a photo damaged by backlit.

One more opportunity is a division of range on more than three sub-ranges and the calculation of features for each range. We propose to employ nine evenly distributed sub-ranges.

We argue that extraction of features from the co-occurrence matrix is a prospective approach for the description of photos affected by backlit. Neighboring pixels of natural images have close values, and non-zero elements locate along the main diagonal of the co-occurrence matrix, as a rule. For photos damaged by backlighting, the co-occurrence matrix has the highest concentrations of elements in left-top or right-bottom corners as well as gaps in the central part of the main diagonal. Figure 4 shows the co-occurrence matrices for undistorted and damages by backlighting photos.



**Figure 4.** Co-occurrence matrices for undistorted (left) and damages (right) photos.

We propose to divide the co-occurrence matrix into three parts orthogonally to the main diagonal similar to shadow, middle tones and highlights sub-ranges in a histogram. It is possible to divide the main diagonal in three equal parts (see dashed lines in figure 4), but it is preferable to set range of middle tones twice larger than for dark and bright tones (see solid lines in figure 4). The following features are calculated from the co-occurrence matrix  $G$ , where pixels situated in distance 3 pixels in column is considered. Fractions of  $G$  in dark, light, and middle tones:

$$S_1 = \sum_{i=0}^{127} \sum_{j=0}^{127-i} G(i,j)/(M \times N) \quad (5)$$

$$S_2 = 1 - (S_1 + S_3), \quad (6)$$

$$S_3 = \sum_{i=128}^{255} \sum_{j=383-i}^{255} G(i,j)/(M \times N) \quad (7)$$

where  $M$  is the number of rows and  $N$  is the number of columns of the image.

Fractions of  $G$  in sub-regions of dark tones:

$$S_{11} = \sum_{i=0}^{63} \sum_{j=0}^{63-i} G(i,j)/(M \times N) \quad (8)$$

$$S_{12} = S_1 - S_{11}. \quad (9)$$

Ratios of maxima in dark tones and maxima in bright tones to global maxima:

$$M_1 = \frac{\max_{\substack{i=0...127, \\ j=0...127-i}} (G(i,j))}{\max_{\substack{i=0...255, \\ j=0...255}} (G(i,j))}, \quad (10)$$

$$M_3 = \frac{\max_{\substack{i=128...255, \\ j=383-i...255}} (G(i,j))}{\max_{\substack{i=0...255, \\ j=0...255}} (G(i,j))}. \quad (11)$$

Locations of the matrix  $G$  maxima in the left and in the right parts:

$$P_1 = \max_{\substack{i=0...127, \\ j=0...127-i}} (G(i,j)), \quad (12)$$

$$P_3 = \max_{\substack{i=128...255, \\ j=383-i...255}} (G(i,j)). \quad (13)$$

The following features are the number of elements of co-occurrence matrix  $G$ , which are the greater than the threshold which equals to the average value of  $G$ :

$$thres(y, G_0) = \begin{cases} 1 & y \geq G_0 \\ 0 & y < G_0 \end{cases} \quad (14)$$

$$G_0 = M \times (N - 3)/256^2, \quad (15)$$

$$A_1 = \sum_{i=0}^{127} \sum_{j=0}^{127-i} thres(G(i, j), G_0), \quad (16)$$

$$A_2 = \sum_{i=0}^{127} \sum_{j=127-i}^{255} thres(G(i, j), G_0) + \sum_{i=128}^{255} \sum_{j=0}^{383-i} thres(G(i, j), G_0) \quad (17)$$

$$A_3 = \sum_{i=128}^{255} \sum_{j=383-i}^{255} thres(G(i, j), G_0). \quad (18)$$

where  $M$  is the number of rows and  $N$  is the number of columns.

Also, we calculate the ratios of the last three features:

$$R_1 = A_2/A_1, \quad (19)$$

$$R_2 = A_2/A_3, \quad (20)$$

$$R_3 = A_1/A_3. \quad (21)$$

Additionally, the same features (5-21) can be extracted from big fragments of image. We divide a photo into 2 equal parts vertically and 3 parts horizontally. Totally we calculate about two hundred of features from the histogram and the co-occurrence matrix for the entire image and its fragments.

## 5. Enrichment of feature set

To enrich our feature set we applied a combination of features with the following feature selection. The idea of combining features is based on the assumption that some of the combinations, for example, sum, product, ratio of different features and their squaring, can provide an improvement to the regression model. Therefore, it is advisable to try as many such combinations as possible.

In an ideal case, features should be informative that is they should have a high correlation with amplification factor  $k_s$  from the ground truth. Obviously, informativeness of various features is very different. A use of non-informative features can lead to the worsening of the regression model. It is necessary to select "good" features and drop "bad" ones. For feature selection, we employed a greedy addition of features to random forest regression model. If the addition of a feature to model leads to decreasing of MAD, then we remain the feature, otherwise, we drop it.

## 6. Results

We tested and compared to each other all feature sets described above. In the paper, we do not show intermediate outcomes, because they have a huge size. The best result demonstrates model which uses features  $S_3, S_{11}, M_1, M_3, P_1, P_3, A_2, A_3, R_1$  and ratio  $S_{12}/M_1$ , where the features are calculated for whole image as well as its fragments.

As was mentioned, the amplification factor for dark tones  $k_s$  can be treated as blind quality factor for images damaged by backlit. At present, a sufficient number of blind image quality metrics have been proposed. Such metrics make possible to evaluate the quality based on the image only without the reference. Part of the metrics is developed to assess the single factor affecting the quality, for example, the blurriness level [17]. Other ones claim to be universal. We compared the correlation coefficient  $r$  between experts' judgments about  $k_s$  and several measures for quality assessment.

The following existing algorithms for non-reference quality assessment were analyzed. Blind Image Quality Index (BIQI) [18] implements a two-step approach to assess the quality of photographs. This method is based on usage of features originated from natural scene statistics (NSS) in the wavelet domain and assumptions that photos of natural scenes have determined statistic characteristics, these characteristics are changed due to distortions, and type and strength of distortion can be predicted. The first stage of BIQI is a classification type of defect. The second one is a numerical quality assessment

by means of the regression model. Support vector machine (SVM) is applied for training classification and regression models. Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [19] uses features from NSS in the spatial domain. One regression model for all distortions is trained by SVM. Oriented Gradients Image Quality Assessment (OG-IQA) method [22] analyzes the correlation of oriented gradients in the spatial domain. It was speculated, orientations of local gradients change predictably for distorted images of natural scenes. One regression model for all distortions is trained via Adaptive Boosting for decision trees. Natural Image Quality Evaluator (NIQE) [20] does not use distorted images for training. In this method, multivariate Gaussian (MVG) model based on NSS features in the spatial domain is calculated for pristine photos only. The quality of the estimated image is estimated as the distance between its MVG and pre-calculated MVG of several undistorted images. Integrated Local Natural Image Quality Evaluator (IL-NIQE) [21] algorithm exploits the same idea as NIQE, but IL-NIQE operates with color channels of photo in salient local patches.

**Table 2.** Correlation coefficient between experts' judgments and measures for quality assessment.

	Decision tree, [5]	BIQI	BRISQUE	NIQE	ILNIQE	OG-IQA	Proposed
Correlation coefficient, $r$	0.47	-0.22	0.05	0.13	0.23	0.13	<b>0.69</b>

Table 2 contains the correlation coefficient between ground truth and measures for quality evaluation. Amplification factor  $k_s$  serves as quality metrics for algorithm from [5] and for proposed regression model. Our method provides the best conformance to experts' judgments. Well-known universal image quality measures have low correlation with ground truth and cannot be used for quality characterizing of photos damaged by backlighting.

Table 3 contains outcomes of comparison of three methods for the estimation of amplification factor  $k_s$ . We analyzed decision tree from [5] as a baseline, regression model by means of random forest trained in features from [5], and random forest model trained on proposed features. According to all performance measures our technique outperforms considered alternatives.

**Table 3.** Comparison of methods for estimation of  $k_s$ .

	<i>MAE</i>	<i>MSE</i>	<i>MedAE</i>	<i>r</i>	<i>normAUC</i>
Decision tree from [5]	21.5	800	17.0	0.47	0.700
Random forest with features from [5]	14.8	334	13.1	0.61	0.711
Random forest with proposed features	<b>13.5</b>	<b>282</b>	<b>11.7</b>	<b>0.69</b>	<b>0.724</b>

## 7. Conclusion

Photos taken by backlighting conditions need to be enhanced to improve their pleasantness. There is a technique for the correction of dark tones. The automatic adjustment of the parameter for amplification of dark tones is needed for that method. For this purpose, we propose a feature set, extracted from the co-occurrence matrix, and regression model via random forest. For the collection of training dataset, we employed semi-supervised paradigm for obtaining of required photos from social networks. The regression model developed outperforms the baseline method. In addition, our model produces smooth continuous values of amplification factor rather than step-wise discrete values in the existing method. The proposed algorithm is intended for photo enhancement software.

## 8. References

- [1] Safonov I V, Kurilin I V, Rychagov M N and Tolstaya E V 2018 *Adaptive Image Processing Algorithms for Printing* (Singapore: Springer Nature)
- [2] Thang P C and Kopylov A V 2018 Tree-serial parametric dynamic programming with flexible prior model for image denoising *Computer Optics* **42(5)** 838-845 DOI: 10.18287/2412-6179-2018-42-5-838-845
- [3] Safonov I V, Rychagov M N, Kang K and Kim S H 2008 Automatic red eye correction and its quality metric *Proc. of SPIE* **6807** 68070W

- [4] Safonov I V, Rychagov M N, Kang K and Kim S H 2008 Adaptive sharpening of photos *Proc. of SPIE* **6807** 68070U
- [5] Safonov I V 2006 Automatic correction of amateur photos damaged by backlighting *Proc. of GraphiCon* 80-89
- [6] Triguero I, García S and Herrera F 2015 Self-labeled techniques for semi-supervised learning: taxonomy, software and empirical study *Knowledge and Information systems* **42(2)** 245-284
- [7] Breiman L 2001 Random forests *Machine learning* **45(1)** 5-32
- [8] Goncharova A V 2018 No-reference metrics for the quality of images damaged by backlighting *Proc. of DSPA* **2** 568-572
- [9] Drucker H, Burges C J, Kaufman L, Smola A J and Vapnik V 1997 Support vector regression machines *Proc. of NIPS* 155-161
- [10] Altman N S 1992 An introduction to kernel and nearest-neighbor nonparametric regression *The American Statistician* **46(3)** 175-185
- [11] Breiman L, Friedman J, Stone C J and Olshen R A 1984 *Classification and regression trees* (Boca Raton: CRC Press)
- [12] MATLAB feedforwardnet URL: <https://www.mathworks.com/help/deeplearning/ref/feedforwardnet.html> (01.04.2019)
- [13] MATLAB linearlayer URL: <https://www.mathworks.com/help/deeplearning/ref/linearlayer.html> (01.04.2019)
- [14] Jahrer M, Töscher A and Legenstein R 2010 Combining predictions for accurate recommender systems *Proc. of ACM SIGKDD* 693-702
- [15] Wolpert D H 1992 Stacked generalization *Neural networks* **5(2)** 241-259
- [16] Bi J and Bennett K P 2003 Regression error characteristic curves *Proc. of ICML* 43-50
- [17] Asatryan D G 2017 Image blur estimation using gradient field analysis *Computer Optics* **41(6)** 957-962 DOI: 10.18287/2412-6179-2017-41-6-957-962
- [18] Moorthy A K and Bovik A C 2010 A two-step framework for constructing blind image quality indices *IEEE Signal processing letters* **17(5)** 513-516
- [19] Mittal A, Moorthy A K and Bovik A C 2012 No-reference image quality assessment in the spatial domain *IEEE Transactions on Image Processing* **21(12)** 4695-4708
- [20] Mittal A, Soundararajan R and Bovik A C 2013 Making a “completely blind” image quality analyzer *IEEE Signal Processing Letters* **20(3)** 209-212
- [21] Zhang L, Zhang L and Bovik A C 2015 A feature-enriched completely blind image quality evaluator *IEEE Transactions on Image Processing* **24(8)** 2579-2591
- [22] Liu L, Hua Y, Zhao Q, Huang H and Bovik A C 2016 Blind image quality assessment by relative gradient statistics and adaboosting neural network *Signal Processing: Image Communication* **40** 1-15