https://doi.org/10.2352/EI.2023.35.9.IPAS-299

Blind denoising of dental X-Ray images

Mykola Ponomarenko[1], Oleksandr Miroshnichenko[2], Vladimir Lukin[2], Sergii Kryvenko[2] and Karen Egiazarian[1]
[1]Tampere University, Tampere, Finland
[2]National Aerospace University, Kharkiv, Ukraine

Abstract

The present study addresses the issue of automatic analysis and noise reduction in dental X-ray images obtained through the Morita system. These images are characterized by spatially correlated noise with an unknown spectrum and varying standard deviation across different regions of the image. To address this issue, we propose the utilization of two deep convolutional neural networks. The first network estimates the spectrum and level of noise for each pixel of a noisy image, predicting maps of noise standard deviation for three different image scales. The second network utilizes these maps as inputs to suppress noise in the image. Results obtained using both modeled and real-life images demonstrate that the proposed networks achieve a peak signal-to-noise ratio (PSNR) for dental X-ray images that is 2.7 dB better than the state-of-the-art denoising methods.

Introduction

Dental X-ray images are commonly used in the field of dentistry for the diagnosis and treatment of dental conditions. However, these images often contain noise that can negatively impact their diagnostic value. This noise on the images can be spatially correlated and may have an unknown spectrum and varying standard deviation across different regions of the image [1]. In this paper, we propose a novel approach for the automatic analysis and noise suppression of dental X-ray images obtained through the Morita system [2].

Dental X-ray images obtained through the Morita system undergo image processing that includes interpolation and geometric distortions [2]. As a result, the standard deviation of noise in these images varies across different regions. Additionally, the noise is spatially correlated and may have different spectra across different images. Furthermore, these images may contain frames and other artifacts of artificial origin which complicates a blind analysis of noise characteristics [2].

The noise suppression in digital images has been a central task in image processing for several decades [1]. Recently, convolutional neural networks (CNNs) have emerged as a powerful tool for image denoising, surpassing the capabilities of the human visual system [3-5]. However, these CNNs are effective in suppressing noise only when they have been trained with specific characteristics.

For example, the DRUNet [3] model is trained to suppress additive white Gaussian noise (AWGN) and requires a two-dimensional map of noise standard deviation (STD) values as an input. This map has to be estimated [6] using a separate method, such as SDNet, or be a priori known. The use of this map allows the DRUNet to effectively suppress noise with varying STD across an image. However, the DRUNet is not capable of effectively suppressing spatially correlated noise.

Some other CNNs, such as VDNet [4], do not require a noise map as an input, but instead evaluate it automatically. However, this comes at the cost of being limited to the noise spectrum and distribution for which the network was trained. The authors of VDNet provide two pre-trained models for AWGN and "real-life" noise, as trained on the SIDD image set.

The DRUNET+NLNET+M2 method [7] is efficient for the suppression of spatially correlated noise with an unknown spectrum, but it assumes that the noise spectrum and level are constant across the entire image.

Thus, there are currently no completely blind methods that provide reliable results for removing noise with a variable level and spectrum that are a priori unknown.

An approach we are proposing in this paper utilizes CNN to estimate a spectrum and level of noise for each pixel of a noisy image, and to predict maps of noise standard deviation for three different image scales. These maps are later used as inputs for a second CNN, which suppresses the noise in the image. We evaluate the performance of our proposed method using both modeled and real-life dental X-ray images.

We evaluate the performance of our proposed method using both modeled and real-life dental X-ray images and show that it outperforms the state-of-the-art denoising methods. This research has the potential to improve the diagnostic value of dental X-ray images and, ultimately, benefit patient care.

Proposed method of noise analysis and suppression

In practice, not only the standard deviation of noise but also its spectrum is often unknown and non-stationary. This presents a challenge for denoising methods, as they typically rely on specific assumptions about the noise characteristics.

For example, as shown in Fig.1(a), the noise in an image obtained through phase imaging [8] is spatially correlated in the center of the image but is almost white near image corners. The standard deviation of noise is low in the center and increases towards the corners, and the distribution is near Gaussian in the center but becomes non-Gaussian (heavy-tailed) near the corners. Additionally, the noise at the corners can be considered as impulsive. In such cases, neither CNNs trained for noise with a specific distribution and spectrum, nor methods for estimating a noise standard deviation map, are effective.

a b

Fig. 1. Examples of images with unknown noise characteristics

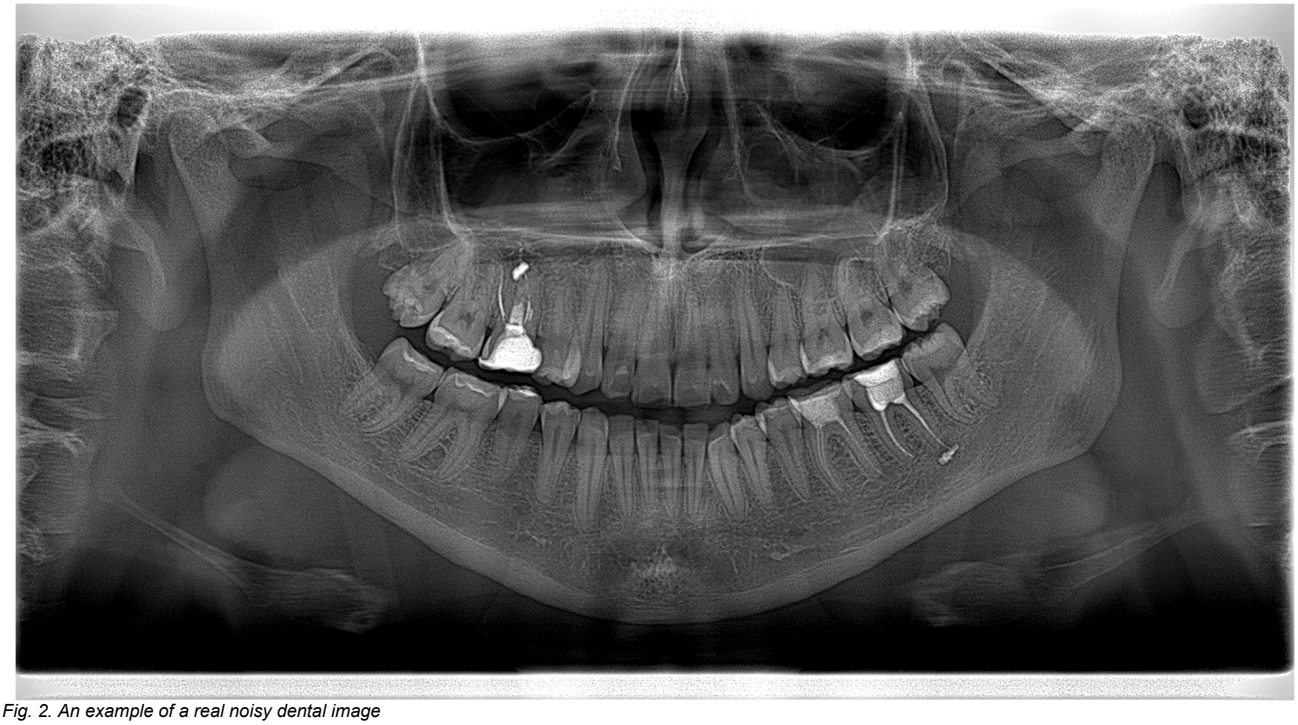


Fig. 2. An example of a real noisy dental image

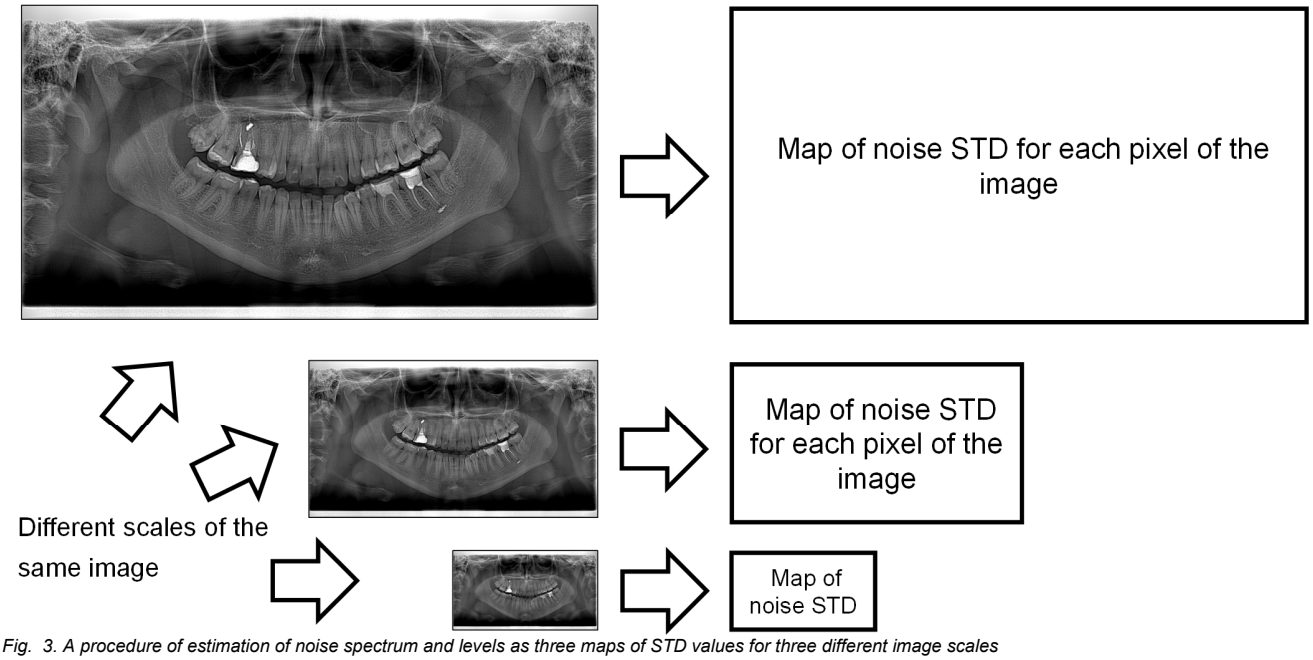


Fig. 3. A procedure of estimation of noise spectrum and levels as three maps of STD values for three different image scales

Another example is shown in Fig.1(b), where an image obtained through an image processing chain in the Canon EOS 250D digital camera. This image is automatically combined from several noisy images with a block matching procedure, resulting in visible residual spatially correlated noise with an unknown spectrum. The standard deviation of noise differs across different regions of the image, and the noise is partially smoothed due to a lossy JPEG compression.

In this study, we specifically consider the task of denoising of dental X-ray images of Morita system [2], such as the one shown in Fig. 2. The noise in these images is spatially correlated, and the standard deviation of noise varies across different regions.

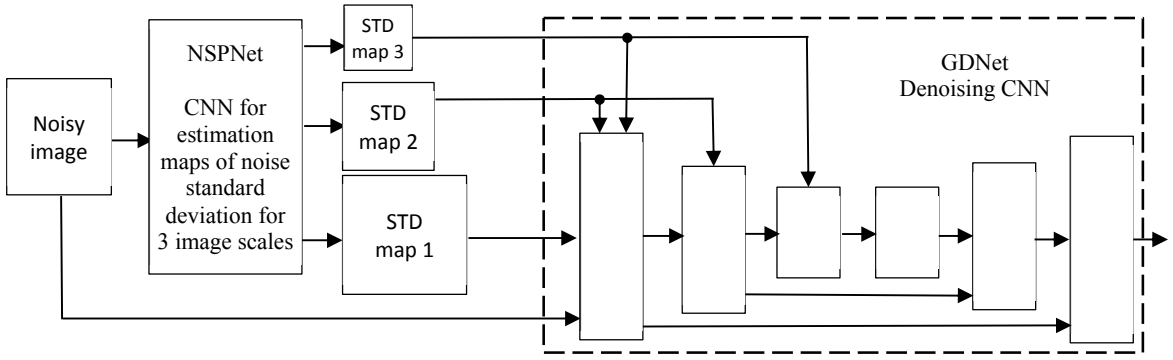


Fig.4. Structural scheme of the proposed blind denoising

Fig.5. Example of image with spatially correlated noise and noise spectrum (STD maps for three scales) estimated by NSPNet. True spectrum: σ_1=10, σ_2=8.5, σ_3=5.7. Estimated spectrum: σ_1=9.7, σ_2=8.3, σ_3=5.7

The proposed approach is based on the idea of obtaining maps of standard deviation levels for three different image scales (Fig. 3).

This will provide a three-values estimate of the noise spectrum for each pixel, which, although rough, may be sufficient for an effective noise suppression.

The proposed method is based on two CNNs (see Fig. 4), which are trained and applied successively. However, the CNNs can be applied separately for usage in other tasks of image processing. Architectures of both CNN are like those in the DRUNet [3].

The first CNN (NSPNet) in the pair predicts three two-dimensional STD maps of noise: for a full-size image ("STD map1"), for the image downscaled two times ("STD map2"), and for the image downscaled four times ("STD map3"). Together, these three noise maps provide an estimate of noise spectrum and noise STD for each pixel of the image.

The second CNN (GDNet) suppresses the noise, taking as inputs noisy image and estimated STD map1, STD map2 and STD map3. The maps "STD map2" and "STD map3" are also connected to inner layers of the network, which process the corresponding image scales.

For training of GDNet we used STD maps estimated by already pre-trained NSPNet. It provides an additional robustness to the trained GDNet.

Matlab scripts with detailed structure of the designed NSPNet and GDNet are available in https://ponomarenko.info/ddxi.html.

Training of the proposed networks

For training of NSPNet, we used noise-free images with artificially added noise. In this way, ground truth maps "STD map1", "STD map2", and "STD map3" were known.

To train the NSPNet, we selected 4238 images from various datasets as described below. We utilized 1000 images from the Tampere21 image database [9], which comprises of 1000 near noise-free color images. The second set of 1000 images has been selected from four datasets: KonIQ10k [10], FLIVE [11], NRTID [12], and SPAQ [13]. The merged mean opinion scores (MOS) [14] of these four databases were used to select high-quality images with the best MOS values. The remaining images were collected from various sources: 217 images from the Flickr2K database [15], 123 images from the Waterloo Exploration Database [16], 103 images from the DIV2k database [17], and 1795 images from various photo hostings. The KonCept512 metric [30], pre-trained on six databases (KonIQ10k [10], Live-in-the-Wild [18], FLIVE [11], NRTID [12], HTID [19], and SPAQ [13]), was used to collect high-quality images from the above-mentioned image sources.

For the simplicity of noise modeling in the training sets, spatially correlated noise was modeled as $n \otimes k$, where n represents the noise and k represents a rotationally symmetric Gaussian low-pass filter with a standard deviation σ_k. According to this model, the noise level is characterized by the σ value, while the degree of correlation between noise values in neighboring image pixels, referred to as the "noise grain", is characterized by the σ_k value.

In the training data, 20% of the noise was AWGN, 60% was spatially correlated noise with varying levels of correlation, and 20% was high-frequency noise.

7% of the noise was additive noise with a constant STD, 7% was multiplicative noise with a constant STD, and 7% was Poisson noise with a constant STD. The remaining 79% was additive noise with a STD map of a randomly selected shape.

Mean variance for a given patch was selected as $(40\,\mathrm{abs}(\mathrm{randn}))^2$, where $\mathrm{abs}(x)$ represents the absolute value of x, randn is a Matlab function that generates a random value with a Gaussian distribution and standard deviation of 1.

This training provided the NSPNet with the ability to effectively estimate the noise spectrum and level in a blind manner.

The GDNet was trained on the same patches with input STD maps estimated by the pre-trained NSPNet. This training scheme provided the GDNet with partial robustness to estimation errors made by the NSPNet.

The training was carried out in the Matlab 2022 environment using a custom training loop, resulting in 100,000 iterations with a minibatch size of 32. Mean squared error (MSE) was used as the loss function to minimize the number of outliers in the STD maps estimates. Pre-trained NSPNet and GDNet are available in https://ponomarenko.info/ddxi.html.

Analysis of effectiveness of the proposed NSPNet

To evaluate the quality of the NSPNet training, we applied the network to estimate the standard deviation of AWGN with a constant STD. Table 1 presents the results of a comparison of the NSPNet with the best methods for this task (using the same image set as in the SDNet paper [9]). The results show that the NSPNet demonstrates estimation precision that is very close to the state-of-the-art SDNet network, which was specifically designed for AWGN.

Table 1. Relative error of STD estimation of AWGN with a constant STD

σ	IEDD [20]	PCA [21]	WTP [22]	VDNet [4]	SDNet [9]	NSPNet
3	0.407	0.573	0.507	1.782	**0.150**	0.192
5	0.204	0.274	0.269	0.905	0.094	**0.091**
7	0.114	0.166	0.174	0.559	0.062	**0.050**
10	0.073	0.097	0.122	0.323	0.034	**0.028**
15	0.071	0.053	0.075	0.168	**0.013**	0.021
20	0.077	0.038	0.057	0.105	**0.016**	0.020
30	0.114	0.036	0.055	0.066	**0.011**	0.012
50	0.157	0.069	0.103	0.089	**0.009**	**0.009**
75	0.207	0.172	0.176	0.150	**0.007**	0.009

Figure 4 illustrates an example of noise spectrum and level estimation by the NSPNet for an image with a spatially correlated noise with a fixed STD. The median values of the STD for the ground truth STD maps and the estimated STD maps (σ_1, σ_2, and σ_3) can be compared. For this image, the NSPNet estimates the noise spectrum with a high degree of precision. A visual analysis of the estimated maps confirms that image fine details and textures introduce minimal errors in the estimated STD maps.

Analysis of effectiveness of noise suppression by the proposed approach

We modeled the noise typical of X-ray images obtained through the dental Morita system. The noise consists of a mixture of spatially correlated additive Gaussian noise with σ=2 and spatially correlated Poisson noise, as well as geometric distortions. This model was used to generate 20 sets of 3 noisy images (with 3 different noise spectra) for 20 manually selected ground truth noise-free images.

Fig.6 shows example of simulated noise free and noisy images with a noise like in dental X-Ray images with σ_k=0.9.

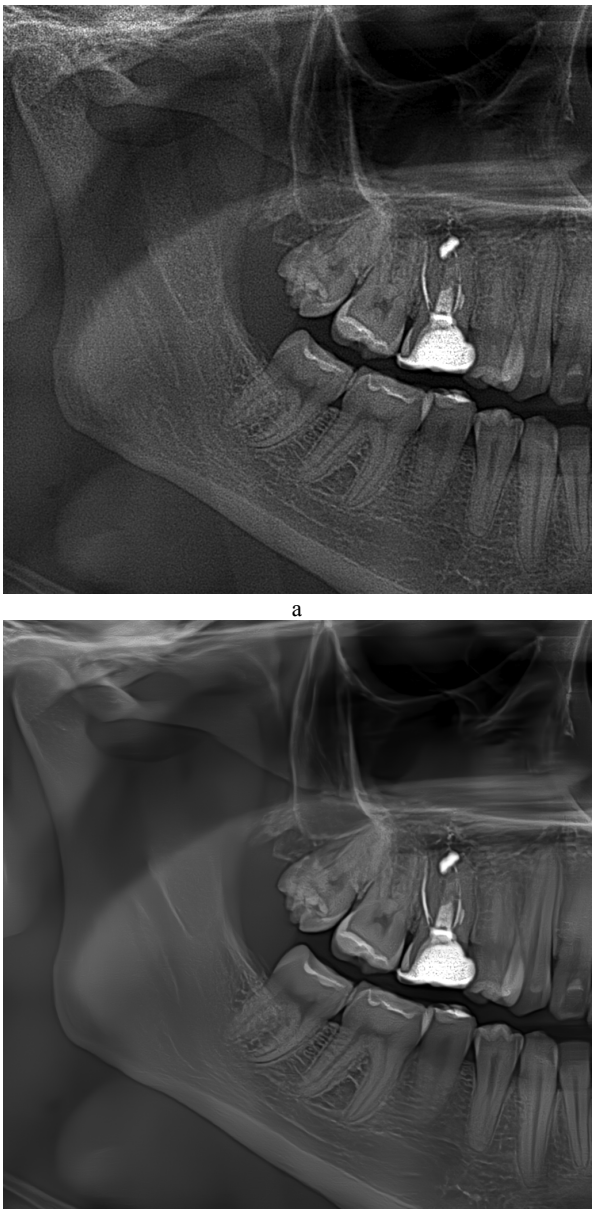
Fig. 6. Simulated noise free image and simulated noisy image with a noise from dental X-Ray image

The test set generated was used to compare the proposed NSPNet+GDNet with several state-of-the-art denoising methods. The results are presented in Table 2. The parameter σ_k in this context is related to the degree of correlation or "grain" of the spatially correlated noise.

Table 2. Denoising of simulated noise for dental X-Ray images for 3 different level of noise correlation, PSNR, dB

Denoiser	σ_k=0.7	σ_k=0.8	σ_k=0.9
Noisy image	29.2	29.2	29.2
NSPNet+GDNet	**36.5**	**36.5**	**36.7**
DRUNet with a fixed STD map [3]	35.2	34.6	34.0
Nonlocal Mean with a fixed threshold [23]	33.9	33.4	32.9
Edge-preserving Gaussian bilateral filter (Matlab)	33.5	33.4	33.2
DCT Filter with a fixed threshold [24]	33.3	32.8	32.3
Wavelet Filter with a fixed threshold (Matlab)	31.8	31.5	31.2
DRUNET+NLNET+M2 [7]	30.2	29.9	29.7
Restormer, sigma 25 [5]	30.2	29.9	29.8
Restormer, blind [5]	29.2	29.2	29.2

It is noteworthy that both blind and fixed-sigma versions of the state-of-the-art Restormer denoiser [5] do not perform well in this case, as they were trained for AWGN. We have not included many contemporary noise reduction techniques in this table because they leave the noise almost untouched.



a

b

Fig. 7. Example of denoising of real-life dental image of Morita system

In contrast, the combination of NSPNet + GDNet is effective and provides a peak signal-to-noise ratio (PSNR) that is better than that of the DRUNet by up to 2.7 dB for larger noise correlations.

Fig. 7 illustrates the result of denoising a fragment of a real-life dental X-ray image obtained through the Morita system. There is no visible residual noise in the denoised image.

Conclusions

In this paper, we proposed a novel approach for blind denoising of dental X-ray images that are distorted by spatially correlated noise with unknown spectrum and standard deviation that varies for different image regions. We proposed two deep convolutional neural networks, NSPNet and GDNet, that allow us to provide an efficient solution for this problem.

The NSPNet estimates spectrum and level of noise for each pixel of a noisy image, predicting maps of noise STD for three image scales. The GDNet uses these maps as inputs to suppress noise in the image.

We tested our approach on a set of simulated dental X-ray images and compared it to the state-of-the-art denoising methods. The results showed that our proposed approach provides PSNR that is better than other modern denoising methods by over 2.7 dB.

The proposed approach is highly effective and can be used to denoise any kind of spatially correlated noise with unknown spectrum and standard deviation that varies for different image regions. Our method is general enough to be adapted to other imaging modalities and is expected to have a wide range of applications in various fields such as medicine, biology, and materials science.

References

[1] J. Astola, P. Kuosmanen, Fundamentals of Nonlinear Digital Filtering, CRC press, 288 p., 1997.

[2] V. Abramova, S. Krivenko, V. Lukin, O. Krylova, "Analysis of Noise Properties in Dental Images", Proceeding of IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO), pp. 511-515, 2020.

[3] K. Zhang, Y. Li, W. Zuo, L. Zhang, L.Van Gool, R. & Timofte, "Plug-and-play image restoration with deep denoiser prior", IEEE Transactions on Pattern Analysis and Machine Intelligence, 17 p, 2021.

[4] Z. Yue, H. Yong, Q. Zhao, L. Zhang, and D. Meng, "Variational denoising network: Toward blind noise modeling and removal," 2019, arXiv:1908.11314.

[5] S.W. Zamir, A. Arora, S. Khan, M. Hayat, F.S. Khan, & M.H. Yang, "Restormer: Efficient transformer for high-resolution image restoration", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5728-5739, 2022.

[6] V. Lukin, S. Abramov, N. Ponomarenko, M. Uss, M. Zriakhov, B. Vozel, K. Chehdi, J. Astola, Methods and automatic procedures for processing images based on blind evaluation of noise type and characteristics, Journal of applied remote sensing, 5(1):053502, 2011.

[7] M. Ponomarenko, O. Miroshnichenko, V. Lukin, K. Egiazarian, "Blind Estimation and Suppression of Additive Spatially Correlated Gaussian Noise in Images", European Workshop on Visual Information Processing, 6 p., 2021.

[8] P. Kocsis, I. Shevkunov, V. Katkovnik, H. Rekola, K. Egiazarian, "Single-shot pixel super-resolution phase imaging by wavefront separation approach", Optics Express, 29(26):43662-43678, 2021.

[9] S.G. Bahnemiri, M. Ponomarenko, K. Egiazarian, "Learning-based Noise Component Map Estimation for Image Denoising", IEEE Signal Processing Letters, pp. 1407-1411, 2022.

[10] V. Hosu, H. Lin, T. Sziranyi, and D. Saupe, "Koniq-10k: An ecologically valid database for deep learning of blind image quality assessment," IEEE Trans. Image Process., vol. 29, pp. 4041–4056, 2020.

[11] Z. Ying, H. Niu, P. Gupta, D. Mahajan, D. Ghadiyaram, and A. Bovik, "From patches to pictures (paq-2-piq): Mapping the perceptual space of picture quality," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2020, pp. 3575–3585.

[12] N. Ponomarenko, O. Eremeev, V. Lukin, and K. Egiazarian, “Statistical evaluation of no-reference image visual quality metrics,” in Proc. IEEE 2nd Eur. Workshop Vis. Inf. Process., 2010, pp. 50–54.

[13] Y. Fang, H. Zhu, Y. Zeng, K. Ma, and Z. Wang, “Perceptual quality assessment of smartphone photography,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2020, pp. 3677–3686.

[14] A. Kaipio, M. Ponomarenko, and K. Egiazarian, “Merging of MOS of large image databases for no-reference image visual quality assessment,” in Proc. IEEE 22nd Int.Workshop Multimedia Signal Process., 2020, pp. 1–6.

[15] B. Lim, S. Son, H. Kim, S. Nah, and K.Mu Lee, “Enhanced deep residual networks for single image super-resolution,” in Proc. IEEE Conf. Comput. Vis. pattern Recognit. workshops, 2017, pp. 136–144.

[16] K. Ma et al., “Waterloo exploration database: New challenges for image quality assessment models,” IEEE Trans. Image Process., vol. 26, no. 2, pp. 1004–1016, Feb. 2017.

[17] E. Agustsson and R. Timofte, “Ntire 2017 challenge on single image superresolution: Dataset and study,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2017, pp. 126–135.

[18] D. Ghadiyaram and A. C. Bovik, “Massive online crowdsourced study of subjective and objective picture quality,” IEEE Trans. Image Process., vol. 25, no. 1, pp. 372–387, Jan. 2016.

[19] M. Ponomarenko et al., “Color image database htid for verification of no-referencemetrics: Peculiarities and preliminary results,” in Proc. IEEE 9th Eur. Workshop Vis. Inf. Process., 2021, pp. 1–6.

[20] M. Ponomarenko, N. Gapon, V. Voronin, K. Egiazarian, “Blind estimation of white Gaussian noise variance in highly textured images”, Proceedings of International Conference Image Processing: Algorithms and Systems, 5p., 2018.

[21] S. Pyatykh, J. Hesser, and L. Zheng, "Image Noise Level Estimation by Principal Component Analysis", IEEE Transactions on Image Processing, pp. 687-699, 2013.

[22] X. Liu, M. Tanaka and M. Okutomi, "Noise Level Estimation Using Weak Textured Patches of a Single Noisy Image", IEEE International Conference on Image Processing, 2012, pp. 665-668.

[23] A. Buades, B. Coll, & J.M. Morel, “Non-local means denoising”, Image Processing On Line, 1:208-212, 2011.

[24] R. Öktem, K. Egiazarian, V.V. Lukin, N.N. Ponomarenko, & O.V. Tsymbal, “Locally adaptive DCT filtering for signal-dependent noise removal”, EURASIP Journal on Advances in Signal Processing, 2007:1-10, Springer, 2007.