MSC 68T07 DOI: 10.14529/jcem240305

SUPPRESSING OF IMAGE DIGITAL NOISE USING A NEURAL NETWORK BASED ON U-NET

A. A. Kuznetsov, South Ural State University, Chelyabinsk, Russian Federation, kuznetsovaa@susu.ru

A digital image is a computer representation of an optical image. The process of obtaining a digital image using digital cameras is always accompanied by noise. Removing noise from an image is an important stage in digital image processing, since noise at large values degrades the quality of the image and complicates subsequent analysis of the data on it. Noise in the image can occur due to environmental factors, ISO sensitivity, camera sensor and so on. The purpose of this research is to create a method for improving the visual quality of images by reducing the noise presented in them. This method, based on neural networks, will work with RAW images, converting them into RGB images. The resulting RGB image will be noise-free. The performance of the presented technique is evaluated in terms of noise reduction and image detail preservation. Experimental results demonstrate the effectiveness of the proposed denoising method in achieving significant noise reduction while maintaining image details.

Keywords: RAW Images; U-Net; image denoise.

Introduction

This article presents a method for denoising RAW images. This method uses a neural network to create an RGB image without digital noise from the RAW image. Unlike standard image formats such as JPEG or PNG, RAW images contain all the data captured from the camera's sensor, without any modification or compression. Photosensitive matrices use sensors, each of which can distinguish the brightness of only one color. Accordingly, since there is one sensor per pixel of the matrix, one pixel represents one color in a RAW image. Moreover, the color depth of RAW images is most often 14 bits per color, that is, 14 bits are allocated for color encoding. While standard RGB images have a color depth of 8 bits. The arrangement of colors in such an image can be different, but the most popular scheme is the Bayer pattern [1]. The classic version uses red, blue and green. In order to get a familiar RGB image, you need to use the demosaicing algorithm [2], for example one of the most popular is AMaZE. The disadvantage of these algorithms is that their use violates the initially Gaussian noise distribution [3], which degrades the quality of the RGB image. This paper proposes a neural network-based method to suppress image noise without this drawback. The proposed neural network architecture is based on U-Net [4]. The novelty of this noise reduction method lies in the fact that the initial data for the neural network are RAW images. In this case, noise is not introduced into the real image through mathematical modeling, as can be found in other studies [5], but images with real noise obtained from a digital matrix are used.

There are several advantages for neural network processing of raw data. Neural networks excel at processing raw, unstructured data by automatically extracting relevant features and learning complex, nonlinear patterns without the need for manual feature engineering. This end-to-end learning approach allows neural networks to be highly scalable and adaptable, making them well-suited for a wide range of applications, from computer vision and natural language processing to speech recognition and predictive analytics.

Overall, working with raw data allows neural networks to leverage their capabilities for automatic feature learning, flexibility, generalization, End-to-End Learning, adaptability, and interpretability. These benefits improve the performance, scalability, and reliability of neural network models in a wide range of applications.

1. Methods Used for Noise Suppression

Let's look at some of the most popular methods of noise suppressing.

Median filtering is a widely used technique in image processing for noise reduction, especially in the presence of salt-and-pepper noise. Here are the advantages and disadvantages of using median filtering:

Advantages:

- Effective noise reduction: Median filtering is particularly effective at removing salt-and-pepper noise, which appears as randomly occurring white and black pixels in an image.
- Preservation of edges: Unlike some other noise reduction techniques, median filtering tends to preserve edges and fine details in the image, making it suitable for retaining important features.
- Robustness to outliers: Median filtering is robust to outliers because it uses the median value of the pixel values in the neighborhood, making it less affected by extreme noise values.
- Simplicity and efficiency: The algorithm for median filtering is relatively simple and computationally efficient, making it suitable for real-time or near-real-time applications.

Disadvantages:

- Blurring of fine details: In areas of the image with high levels of noise, median filtering can lead to a loss of fine details and textures, resulting in a slightly blurred appearance.
- Inability to handle uniform noise: Median filtering is not as effective for reducing Gaussian noise or uniform noise, which may require other denoising methods to complement its effectiveness.
- Blocky artifacts: In the presence of salt-and-pepper noise at higher noise densities, median filtering can introduce blocky artifacts around the noisy pixels, affecting the overall visual quality of the image.
- Parameter sensitivity: The effectiveness of median filtering can be sensitive to the choice of filter window size, and selecting an inappropriate window size may lead to suboptimal noise reduction.

Gaussian filtering is a method used in image processing and computer vision to reduce noise and blur an image. It involves applying a Gaussian blur to the image, which means convolving the image with a Gaussian function to achieve the desired smoothing effect. This is done by replacing each pixel's value with a weighted average of its neighboring pixels, with the weights determined by the Gaussian function. The result is a softened, less noisy image that can be easier to work with for certain algorithms and applications. Gaussian filtering is a fundamental technique in image processing and is commonly used for problems such as edge detection, feature extraction, and image enhancement. Here are the advantages and disadvantages of using Gaussian filtering:

Advantages:

- Smoothing of Noise: Gaussian filtering effectively smoothens out noise in digital images, making it particularly useful for reducing Gaussian noise, which is common in digital imaging.

- Preservation of Image Details: When carefully tuned, Gaussian filtering can reduce noise while preserving important details and edges in the image, resulting in a visually pleasing output.

- Adjustable Control: The standard deviation (sigma) parameter in the Gaussian filter allows for control over the amount of smoothing, offering flexibility in balancing noise reduction and preservation of image features.

- Simple Implementation: Gaussian filtering is a relatively simple and computationally efficient technique, making it widely used in image processing applications and accessible in software libraries.

Disadvantages:

- Blurring of Details: In some cases, aggressive application of Gaussian filtering can lead to excessive blurring of image details, especially when using a smaller standard deviation (sigma) value.

- Non-Adaptive: Gaussian filtering treats all pixels within the kernel window equally, which means it does not adapt to the local variations in image content. This can result in over-smoothing of important details, particularly in regions with high frequency or fine texture.

- Limited Effectiveness for Certain Noise Types: While Gaussian filtering is effective for reducing Gaussian noise, it may not be as effective for other types of noise such as salt-and-pepper noise or periodic noise. Different noise reduction techniques may be required for these cases.

- Trade-off between Noise Reduction and Detail Preservation: Adjusting the standard deviation (sigma) to reduce noise may also lead to a trade-off in terms of preserving fine details, and finding the right balance can be challenging.

- Oversmoothing in Homogeneous Regions: In regions of the image with relatively uniform pixel values, Gaussian filtering can result in oversmoothing, potentially leading to loss of subtle variations in texture and contrast.

The proposed method, while maintaining the advantages of "median filtering", will have no parameter sensitivity and will not blur fine details of the image. This is because the method will not have parameters, and the neural network will handle the fine details preservation. Method also will be free from blurring and oversmoothing. Since this method has no parameters, the trade-off between noise reduction and detail preservation will be eliminated. The novelty of the algorithm will also lie in the fact that it is optimized for low-power single-board computers without a GPU.

2. The Proposed Method for Suppressing Digital Noise in an RAW Image

The proposed method uses convolutional neural network, based on U-Net, that is commonly used for image segmentation problems. The U-Net architecture consists of two parts (Fig. 1). The first part compresses the image and increases the number of channels. The second part does the opposite, enlarging the image and reducing the number of image channels. In general, U-Net is known for its ability to efficiently segment images with limited training data and has become a popular choice for many segmentation problems.

Thanks to its architecture, the neural network trains well on small samples. This is important because obtaining several images for training a neural network, which would be absolutely similar to each other, but have different noise levels, is possible only in a static scene with a rigid camera mount, which is problematic. The ability to use contextual information allows local details to be preserved, which in turn helps to remove noise and images. Also, the relatively low number of parameters due to the use of 1x1 convolutions has a good effect on performance.

The U-net architecture with 32x32 resolution at its lowest resolution is presented in figure 1. The encoder transforms the input data into a latent representation containing information about the input data. The decoder uses this hidden representation to generate output. In a modified version of U-Net, the encoder takes an 8x8 piece of the RAW image. The decoder generates a 8x8 RGB image.

Batch normalization layers have been added to the U-net based neural network. These layers improve the learning process and improve the quality of the neural network. These layers normalize the input data of each layer by subtracting the mean and dividing by the standard deviation.

Fully connected layers were also added to the model. Advantages of these layers that have improved the quality of the neural network are flexibility, object extraction, model capacity, generalization, nonlinear transformations, universal approximation theorem [6].

In general, fully connected layers significantly increase the efficiency of the resulting neural network.

The last change to the original architecture is the addition of a dropout [7] layer. It is a regularization technique used in neural networks to prevent overfitting. The regularization property of this method helped prevent model overfitting. The generalization property forced the neural network to study features that were more reliable for its work. The effects of reducing coadaptation and ensemble improved the quality of the neural network.

The result of all the changes is a neural network adapted for the problem of converting RAW images to RGB with noise suppression. The architecture of the modernized network is shown in Fig. 2.

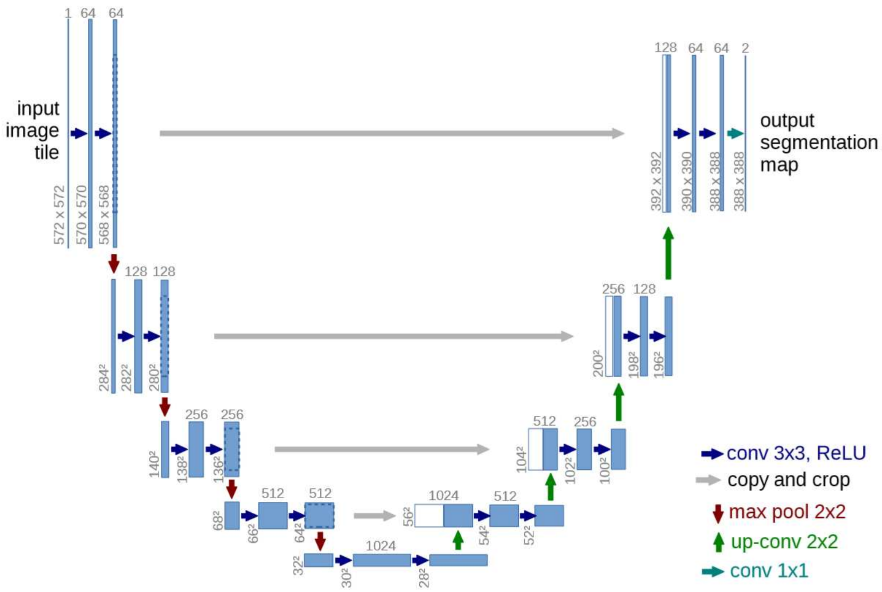


Fig. 1. U-Net architecture

As a result of the work done, a digital noise reduction method was built based on the convolutional neural network, which removes noise from images. To evaluate the performance of the neural network, the DHT algorithm (an algorithm using discrete Hartley transform) was used. This algorithm converts RAW image to RGB image.

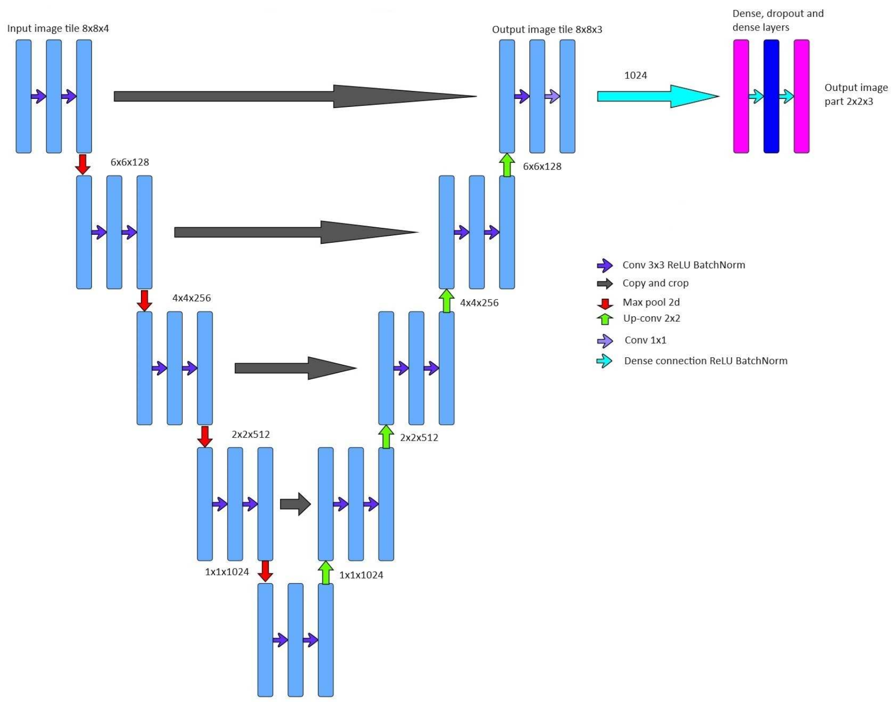


Fig. 2. The resulting neural network architecture

Before being fed into the neural network, the image undergoes preparation. The raw image is an array of Bayer color filters. The resulting array is then divided into overlapping 8x8 squares. Each square is then fed to the input of the neural network. The resulting 2x2 squares are assembled into RGB images. There are two reasons why a neural network's input is an 8x8 square. Firstly, due to the architecture, the minimum image piece can be 8x8. Also, the sides of the image part must be multipliers of 8. In the process of testing the neural network on image sections of different sizes, it turned out that the 8x8 size is the most effective. It is large enough to capture all the necessary features to calculate the four internal pixels, and contains a minimum of noise that degrades the performance of the neural network.

The U-Net like many other neural networks can be run on NVIDIA video adapters for fast and efficient performance. When running U-Net on NVIDIA GPUs, the network can take advantage of the parallel processing capabilities of the GPUs to accelerate training and performance. In this paper neural network was trained and tested on NVIDIA 1080ti video adapter.

By using TensorFlow library with CUDA support on NVIDIA GPU, the proposed model was trained and executed with high performance. The parallel processing power of NVIDIA GPUs enables the model to process large volumes of image data in parallel, leading to faster training times and quicker inference speeds. The dataset contains 124 images for training and 16 for validating. All images have the same 6024x4020 resolution. Training time was 1 hour 27 minutes. It takes 27 seconds to get RGB image without noise from RAW image.

3. Experiments

To train the neural network, images from the website www.dpreview.com were used, which presents test images for various digital cameras. The advantage of images obtained from this source is that they contain images of the same scene taken with unchanged lighting parameters and camera settings, with the exception of exposure time and ISO sensitivity parameter, which is responsible for the amount of noise in the image.

Images taken from a Canon 80D camera were used. An image in JPG format obtained at ISO 100, the least noisy and high-quality image, was chosen as the standard (Fig. 3).

RAW images were downloaded at various ISOs: 100, 200, 400, 800, 1600, 3200, 6400, 12800, 25600. These were the original data. It is worth noting that as ISO increases, the amount of noise increases proportionally. At ISO 12800 and 25600, not only does noise increase, but colors also become distorted. Therefore, the neural network learns not only to suppress noise, but also to restore colors in the image. Examples of noisy images are shown in Fig. 4:

Noise level can be measured as signal-to-noise ratio (PSNR) [8]:

$$PSNR = 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right),$$

where

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}|I_{i,j} - K_{i,j}|^2.$$

Fig. 3. Standart image

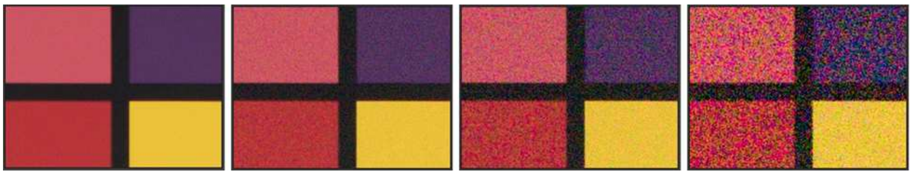


Fig. 4. Sections of images with ISO 100, 1600, 6400, 25600

Testing was carried out on a noisy RAW image with ISO 25600. The standard was an image in JPG format obtained from RAW with ISO 100 through in-camera conversion in a digital camera. The image obtained as a result of applying the neural network and the areas with an eightfold scale are shown in Fig. 5.

When looking at the images in more detail, you can see the difference in the images. Fig. 6 shows some areas of the image obtained using different methods. For comparison, sections have been added from a camera image obtained by in-camera conversion of a RAW image at ISO 25600 to JPG format.

The amount of noise in the original RAW image with ISO 25600 is very high, as a result of which traditional image processing methods, along with noise reduction, also blur the image. This is characterized by blurring of contrast boundaries and loss of fine details in the image. In-camera noise reduction is optimized for the specific camera model on which it operates. Despite this fact, the result of such noise reduction cannot be called acceptable; the image retains details and contrast boundaries, but the noise level still remains high. Using a neural network to convert a high-noise RAW image at high ISO produces an image that is close to the reference. This is achieved due to the fact that the RAW image contains a large amount of information, which is available to the neural network. Unlike a neural network, traditional noise reduction methods are adapted for images in JPG format. To use them, one must first convert the RAW image using some kind of demosaicing method and only then apply it to the resulting image. For comparison

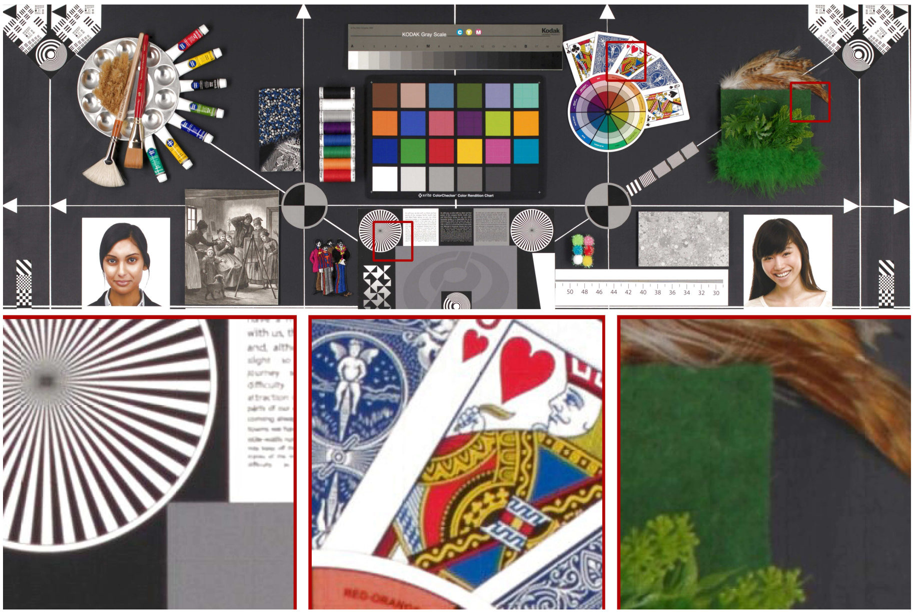


Fig. 5. The result of the neural network work

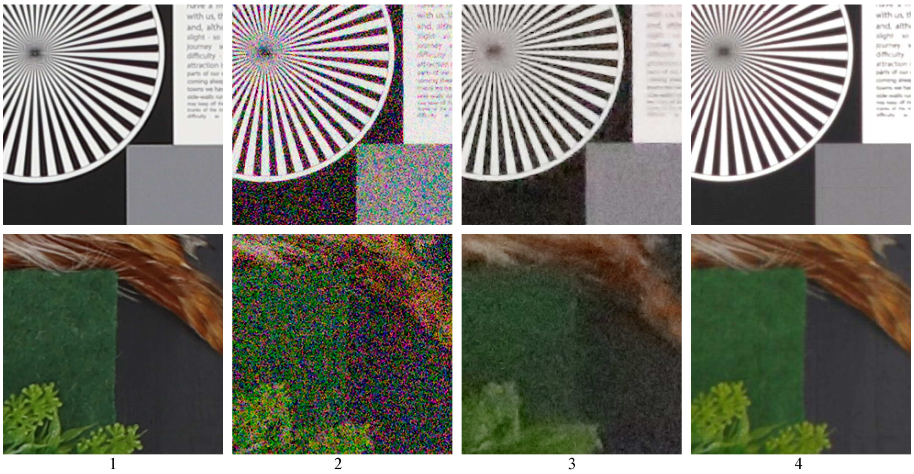
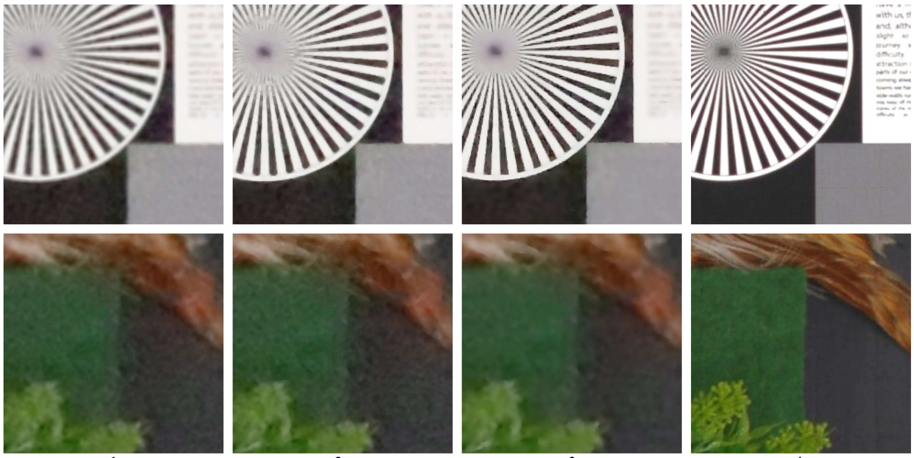


Fig. 6. Comparison of images: 1 – reference; 2 – noisy at ISO 25600; 3 – in-camera noise reduction at ISO 25600; 4 – noisy at ISO 25600 after neural network

with the neural network, several well-known denoising methods were used, implemented in the OpenCV computer vision library [9]: averaging filter, median filter, bilateral filter. Sections of images after filtering with the specified filters and neural network are presented in Fig. 7.



1 2 3 4

Fig. 7. Comparison of the filtering efficiency of an image shot at ISO 25600: 1 – averaging filter; 2 – median filter; 3 – bilateral filter; 4 – neural network

The noise reduction efficiency of the presented filters is higher than that of the noise reduction method built into a digital camera. However, the averaging, median and bilateral filters retained less detail in the image, unlike the neural network filter. It has a high level of noise reduction and preserves image details. An assessment of the noise suppression efficiency of the methods PSNR is presented in table which shows comparison of performance of neural network and three other popular ones on 4 different images. Neural network outperforms all of them in all cases.

Table 1

Methods performance

Averaging	Median	Bilateral	Neural network
28.63	28.68	28.65	30.37
28.56	28.31	29.21	29.92
27.28	27.92	28.32	29.86
28.36	28.16	27.86	30.21

Despite all the advantages of a neural network, it still has disadvantages. The first and most significant drawback is the speed of the noise reduction algorithm based on a neural network. When using a GPU, an image of 6000x4000 pixels was processed by the neural network in 27 seconds. When using only the processor power, this time increases by two orders of magnitude, while traditional algorithms cope with the same image within a few seconds.

The second drawback is the periodically appearing "artifacts" in the image and color conversion errors (Fig. 8). They are associated with insufficient sample size and (possibly) shortcomings of the neural network architecture. In the future, it is planned to continue work and eliminate these shortcomings.

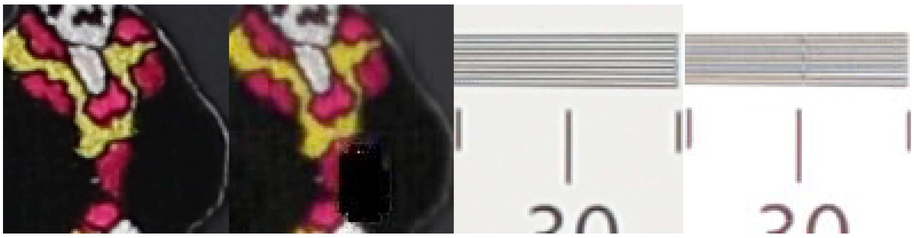


Fig. 8. An example of an “artifact”: 1 – reference image; 2 – image after the neural network. Example of color conversion error: 3 – reference image; 4 – image after neural network

4. Conclusion

There was developed a method using a neural network to convert a RAW image to JPG with suppression of the noise component. It can serve as a replacement for demosaicing algorithms, since the resulting images are obtained directly from RAW images.

The presented method has greater noise reduction efficiency and preserves image detail better than the traditional noise reduction algorithms presented for comparison. The method has prospects for further development and potential for practical application for processing RAW images.

Further development of the created method is aimed at: eliminating conversion artifacts, eliminating color conversion errors and reducing the consumption of computer computing resources.

References

1. De Haan G., Kwaaitaal-Spassova G., Larragy M., Ojo O., Schutten R. Television Noise Reduction IC. *IEEE Transactions on Consumer Electronics*, 1998, vol. 44, no. 1, pp. 143–154. DOI: 10.1109/30.663741.
2. Devies A., Fennessy P. *Digital Imaging for Photographers*. Focal Press, Waltham, 2001.
3. Surin V.A. About Processing Noisy Contrast Images. *Bulletin of South Ural State University. Series: Mathematics. Mechanics. Physics*, 2021, vol. 13, pp. 14–21. DOI: 10.14529/mmph210102 (in Russian)
4. U-Net: Convolutional Networks for Biomedical Image Segmentation. Available online: https://arxiv.org/abs/1505.04597 (accessed on 28 November 2023)
5. Gilyazetdinov E.V., Konovalov B.D. Removing Noise from Images Using Neural Networks. *E-Scio*, 2019, vol. 6, pp. 367–377. (in Russian)
6. Huang G., Liu Z., Maaten L., Weinberger K. Densely Connected Convolutional Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269. DOI: 10.1109/CVPR.2017.243
7. Srivastava N., Hinton, G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal Of Machine Learning Research*, 2014, vol. 15, pp. 1929–1958.
8. Boukhayma A., Caizzone A., Enz C. A CMOS Image Sensor Pixel Combining Deep Sub-Electron Noise with Wide Dynamic Range. *IEEE Electron Device Letters*, 2020, vol. 41, no. 6, pp. 880–883. DOI: 10.1109/LED.2020.2988378.

9. Garcia G.B., Suarez O.D., Aranda J.L.E., Tercero J.S., Gracia I.S., Enano N.V. *Learning Image Processing Using OpenCV*. Packt Publishing, Birmingham, Mumbai, 2015.

Anton A. Kuznetsov, post graduate student, Department of Applied Mathematics and Programming, South Ural State University (Chelyabinsk, Russian Federation), kuznetsovaa@susu.ru

Received December 25, 2023

УДК 004.932 **DOI: 10.14529/jcem240305**

АЛГОРИТМ ПОДАВЛЕНИЯ ЦИФРОВОГО ШУМА НА ИЗОБРАЖЕНИИ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ СЕТИ НА ОСНОВЕ U-NET

А. А. Кузнецов, Южно-Уральский государственный университет, г. Челябинск, Российская Федерация

Цифровое изображение – это компьютерное представление оптического изображения. Процесс получения цифрового изображения с помощью цифровых фотоаппаратов всегда сопровождается шумами в оцифрованном изображении. Удаление шума из изображения – важный этап цифровой обработки изображений, поскольку сильный шум ухудшает качество изображения и усложняет последующий анализ данных на нем. Шум на изображении может возникать из-за факторов окружающей среды, чувствительности ISO, сенсора камеры и т.д. Целью данного исследования является создание метода улучшения визуального качества изображений за счет уменьшения присутствующего в них шума. Этот метод, основанный на нейронных сетях, будет работать с изображениями RAW, преобразуя их в изображения RGB. В полученном RGB-изображении не будет шума. Эффективность представленной методики оценивается с точки зрения снижения шума и сохранения деталей изображения. Результаты экспериментов демонстрируют эффективность предложенного метода шумоподавления в достижении значительного снижения шума при сохранении деталей изображения.

Ключевые слова: изображения в формате RAW; U-Net; шумоподавление.

Литература

1. De Haan, G. Television Noise Reduction IC / G. De Haan, G. Kwaaitaal-Spassova, M. Larragy, O. Ojo, R. Schutten // IEEE Transactions on Consumer Electronics. – 1998. – V. 44, № 1. – P. 143–154.
2. Devies, A. Digital Imaging for Photographers / A. Devies, P. Fennessy. – Waltham: Focal Press, 2001.
3. Сурин, В.А. Об обработке зашумленных контрастных изображений / В.А. Сурин // Вестник Южно-Уральского государственного университета. Серия: Математика. Механика. Физика. – 2021. – Т. 13, № 1. – С. 14–21.

4. U-Net: Convolutional Networks for Biomedical Image Segmentation. Available online: https://arxiv.org/abs/1505.04597 (accessed on 28 November 2023)
5. Гилязетдинов, Э.В. Устранение шумов в изображениях с помощью нейронных сетей / Э.В. Гилязетдинов, Б.Д. Коновалов // E-Scio. – 2019. – Т. 6, С. 367–377.
6. Huang, G. Densely Connected Convolutional Networks / G. Huang, Z. Liu, L. Maaten, K. Weinberger // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – P. 2261–2269.
7. Srivastava, N. Dropout: A Simple Way to Prevent Neural Networks from Overfitting / N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov// Journal Of Machine Learning Research. – 2014. – V. 15, P. 1929–1958.
8. Boukhayma, A. A CMOS Image Sensor Pixel Combining Deep Sub-Electron Noise with Wide Dynamic Range / A. Boukhayma, A. Caizzone, C. Enz // IEEE Electron Device Letters. – 2020. – V. 41, № 6, P. 880–883.
9. Гарсия, Г. Обработка изображений с помощью OpenCV / Г. Гарсия, О. Суарес, Х. Аранда, Х. Терсеро, И. Грасиа, Н. Энано. – М.: ДМК–Пресс, 2016.

Кузнецов Антон Андреевич, аспирант, кафедра прикладной математики и программирования, Южно-Уральский государственный университет (г. Челябинск, Российская Федерация), kuznetsovaa@susu.ru

Поступила в редакцию 25 декабря 2023 г.