UDC 004.932
DOI: 10.15587/1729-4061.2023.275984

APPLYING MACHINE LEARNING TO IMPROVE A TEXTURE TYPE IMAGE

Jamalbek Tussupov
Doctor of Physical and Mathematical Sciences, Associate Professor*
Kairat Kozhabai
Corresponding author
Master*
E-mail: kairat_2747@mail.ru
Aigulim Bayegizova
Candidate of Physical and Mathematical Sciences, Associate Professor
Department of Radio Engineering, Electronics and Telecommunications**
Leila Kassenova
Candidate of Pedagogical Sciences, Associate Professor***
Zhanat Manbetova
Doctor filosofii PhD
Department of Radio Engineering, Electronics and Telecommunications
S. Seifullin Kazakh Agrotechnical University
Zhenis avenue str., 62, Nur-Sultan, Republic of Kazakhstan, 010011
Natalya Glazyrina
PhD, Assistant Professor
Department of Computer and Software Engineering**
Mukhamedi Bersugir
Candidate of Physical and Mathematical Sciences, Head of Department***
Miras Yeginbayev
Department of Computer Science
Toraighyrov University
Lomov str., 64, Pavlodar, Republic of Kazakhstan, 140008
*Department of Information Systems**
**L. N. Gumilyov Eurasian National University
Satpayev str., 2, Nur-Sultan, Republic of Kazakhstan, 010008
***Department of Information Systems and Technologies
Esil University
Zhubanova str., 7, Astana, Republic of Kazakhstan, 010000

The paper is devoted to machine learning methods that focus on texture-type image enhancements, namely the improvement of objects in images. The aim of the study is to develop algorithms for improving images and to determine the accuracy of the considered models for improving a given type of images. Although currently used digital imaging systems usually provide high-quality images, external factors or even system limitations can cause images in many areas of science to be of low quality and resolution. Therefore, threshold values for image processing in a certain field of science are considered.

The first step in image processing is image enhancement. The issues of signal image processing remain in the focus of attention of various specialists. Currently, along with the development of information technology, the automatic improvement of images used in any field of science is one of the urgent problems. Images were analyzed as objects: state license plates of cars, faces, sections of the field on satellite images.

In this work, we propose to use the models of Super-Resolution Generative Adversarial Network (SRGAN), Extended Super-Resolution Generative Adversarial Networks (ERSGAN). For this, an experiment was conducted, the purpose of which was to retrain the trained ESRGAN model with three different architectures of the convolutional neural network, i. e. VGG19, MobileNet2V, ResNet152V2 to add perceptual loss (by pixels), also add more sharpness to the prediction of the test image, and compare the performance of each retrained model. As a result of the study, the use of convolutional neural networks to improve the image showed high accuracy, that is, on average ESRGAN+MobileNETV2 – 91 %, ESRGAN+VGG19 – 86 %, ESRGAN+ +ResNet152V2 – 96 %

Keywords: image processing, SRGAN, ERSGAN, VGG19, MobileNet2V, ResNet152V2, machine learning, Super-Resolution

Received date 18.01.2023
Accepted date 28.03.2023
Published date 28.04.2023

***How to Cite**: Tussupov, J., Kozhabai, K., Bayegizova, A., Kassenova, L., Manbetova, Z., Glazyrina, N., Bersugir, M., Yeginbayev, M. (2023). Applying machine learning to improve a texture type image. Eastern-European Journal of Enterprise Technologies, 2 (2 (122)), 13–18 doi: https://doi.org/10.15587/1729-4061.2023.275984*

1. Introduction

At the same time, in the process of pattern recognition in images when making decisions, medical workers, agronomists, farmers, as well as specialists in important areas of life, face a number of problems: incomplete and inaccurate initial information; large variability in attributes and small sample sizes; limited decision time for conclusions. These factors often lead to errors in decision-making. In order to improve the efficiency and quality of experimental information processing, it is necessary to improve and modify methods for analyzing visual data, both to improve the quality of images and to improve the accuracy of object recognition.

In accordance with the objectives of the manual, the emphasis is on visual recognition, and automatic recognition methods will be mentioned only as necessary. The most important principle of visual recognition is the analysis of improved objects in development and in the context of the environment, as well as the movement from the general to the particular, by sequentially studying the image as an information model of the subject area under consideration, presented as a whole. The main factor determining the quality of visual recognition is the skill level of the expert, which is composed of professional erudition and intuition – it is even argued that this type of activity borders on art. However, here, too, there is a certain methodology aimed at increasing reliability and reducing the number of errors. But with the development of artificial intelligence, automatic recognition of objects in images is an urgent task in all spheres of life.

As the image zoom ratio increases, the image quality deteriorates. Enlargement of objects and improvement of its quality in various applied tasks, for example, in medicine, space images of medium resolution, traffic, is one of the urgent problems. To solve this problem, machine learning methods were applied to automate the scaling and improvement of objects in images. This was the main goal of the study. The difference of this work from others is that other researchers in this field used classical methods [1–3]. In this work, machine learning methods based on artificial neural networks were used.

Upgrading the spatial resolution of medium-resolution images to the spatial resolution of high- and very high-resolution images is useful in a variety of applications. When the shooting height is increased, another effect is also observed – the radiation of large areas of objects suppresses the radiation of small ones, which gives a natural integration of the image, including smoothing the contours and merging groups of small objects. This is commonly referred to as a manifestation of iconic generalization. Visual generalization (visual and iconic) is natural, due to objective reasons and is perceived both by a human observer and by recording devices. However, many of the effects resulting from generalization remain difficult to explain. These include, for example, the "subsurface imaging" effect and other large-scale imaging effects. As noted above, multispectral images allow the use of various color synthesis options for visual and automated interpretation of objects on the earth's surface. There are many free medium-resolution images available (e.g. the multispectral instrument aboard Sentinel-2 and the operational ground-based thermal imager aboard Landsat-8). At present, the most significant volumes of medium-resolution space imagery data are provided by the LANDSAT-5/7/8 [3] and SENTINEL-2 [4] satellites.

Theoretical and applied research on the improvement of algorithms is related to the use of neural networks within a single computing technology for solving problems of assessing the state of an object (detection and classification of pathologies and neoplasms) on medical images (CXR-scanning, CT and MRI images), on space images to highlight homogeneous areas that are identified with negative factors, etc. At present, most high-resolution images and all very high-resolution images acquired by orbital sensors must be purchased at a high price. The results of the work can be applied in automated image analysis systems used in scientific research and in industry. Application in industry makes it possible to reduce the cost of analyzing product quality, and in some cases improve quality. The introduction of the results obtained into the systems for collecting, processing and transmitting information will contribute to the development of the scientific and technological complex.

During the study of X-ray images obtained from open-source databases in the field of medicine, the problem of improving their quality arises. Accurate visualization and identification of pathology directly depend on the good visibility of the images. Therefore, this issue remains relevant at any time. Due to the new technologies of today, they can be processed using neural network methods.

2. Literature review and problem statement

In order to improve the visual quality of images, [5] presented the results of the three key components of SRGAN – network architecture, collision loss and perception loss, and improved each of them to obtain Enhanced SRGAN (ESRGAN). In particular, the researchers introduce Residual-in-Residual Dense Block (RRDB) without batch normalization as the basic network construction unit. In order for the discriminator to be able to predict relative reality instead of absolute value, the authors borrowed the idea from the relativistic GAN. They improved the loss of perception by using pre-activation features. With these improvements, the proposed ESRGAN delivers consistently better visual quality with more realistic and natural-looking textures than SRGAN. However, with the addition of a convolutional neural network, the accuracy result was higher.

In [6], the authors extended the powerful ESRGAN into a practical recovery application that is trained on purely synthetic data. In particular, a high-order degradation modeling process is introduced to better model complex real-world degradations. The authors also considered common ringing and overshoot artifacts during the synthesis process. In this work, the researchers use a U-Net discriminator with spectral normalization to increase the capabilities of the discriminator and stabilize learning dynamics. Although the authors used the U-Net discriminator, they did not consider the scaling of objects.

In [7], the authors show that ESRGAN produces higher quality images when trained on thematically classified images than when trained on a wide variety of examples. Other things being equal, the researchers showed that the algorithm performed better on some topics than on others. The authors paid attention to generative adversarial networks to improve the image, but the experiment was not carried out with the discriminator.

The work [8] is based on Super-Resolution Generative Adversarial Networks (SRGAN), where the authors modify the loss function and structure of the SRGAN network and propose an improved SRGAN (ISRGAN) that makes model training more stable and enhances the ability to generalize across locations and sensors. In the experiment, training and testing data were collected from two sensors (Landsat 8 OLI and Chinese GF 1) from different locations (Guangdong and Xinjiang in China). The proposed method was compared with the neighbor embedding (NE), sparse representation (SCSR) and SRGAN methods. The results show that the accuracy of the land cover classification after super-resolution has improved significantly, in particular the impenetrable surface class. For the study of satellite images, the scaling of the area under consideration is the basis of remote sensing of the earth, and this is the main object of study, which the authors did not take into account.

In [9], the authors proposed a new method for the automatic detection of COVID-19 using tomographic images (CT) and radiographic images (chest X-ray). To improve the performance of the detection system for this outbreak, the authors used two deep learning models: VGG and ResNet. The results of the experiments show that the models proposed by the authors achieved the best accuracy of 99.35 and 96.77 %, respectively, for VGG19 and ResNet50 with all chest radiographs. But the authors did not use the scaling of objects.

In [10], the authors sought to improve the images obtained from optical coherence tomography (OCT). However, instead of applying noise reduction techniques directly, the authors use several state-of-the-art ultra-high resolution techniques. The authors experimented with various SR imaging architectures, but no discriminator experiments were performed.

In [11], the authors propose a texture transformation network to simultaneously reduce image noise and improve spatial resolution in CT images. This network, called the Super-Resolution Texture Transformer (TTSR), is a deep learning super-resolution image reference method built on a Generative Adversarial Network (GAN). In this work, the authors compared the results of generative adversarial networks, but the experiment was not carried out with the addition of a discriminator.

The peculiarity of this work is that when studying image improvement, the unsupervised machine learning algorithm with convolutional neural networks showed higher accuracy than the unsupervised machine learning algorithm.

3. The aim and objectives of the study

The aim of the study is to improve images in various fields of science based on machine learning methods.

To achieve this aim, the following objectives are accomplished:

– to evaluate the effectiveness of using machine learning methods with convolutional neural networks for processing texture-type images;

– to implement improvements in selected areas of the image using machine learning methods, which allows you to determine the percentage of accuracy of image improvement in various fields of science.

4. Materials and methods of research

Single-Image Super-Resolution (SISR) is a computer vision task that reconstructs a high-resolution (HR) image from a low-resolution (LR) image. It can be used in various applications such as medical imaging, security, and surveillance imaging. Improving the quality of images used to solve important problems in various fields of science, which are the object of research in this work, such as electron microscope images for the chemical industry, satellite images for agricultural applications, X-ray and computed tomography images in the field of medicine and many other fields requires the use of new technologies.

The Enhanced Super Resolution GAN architecture (hereinafter ESRGAN [12]) is based on the GAN framework. During the study, convolutional neural networks were added to the model for image enhancement, such as VGG19 [13], MobileNet [14], and ResNet152V2 [15]. The architecture of this model with the addition of convolutional neural networks is shown in Fig. 1.

In this work, an additional descriptor is added to the GAN architecture model. This handle is executed before the activation function. A separate implementation scheme for a generative adversarial network (GAN) is shown in the figure below (Fig. 2).

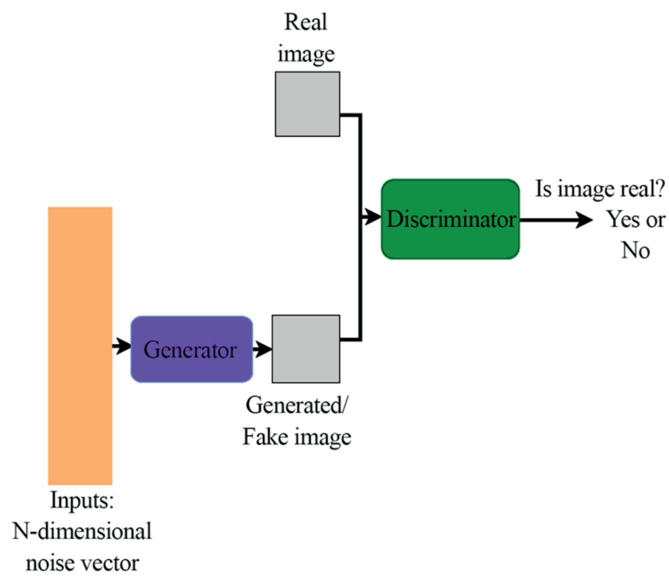


Fig. 1. Generative Adversarial Network architecture with the addition of convolutional neural networks

The ESRGAN model was used as a generative adversarial network. As input to this model, a low resolution (LR) image was trained, that is, an image with a size of 100×100. After the original image was trained using the ESRGAN model, it was retrained by adding the CNN model (Fig. 3).

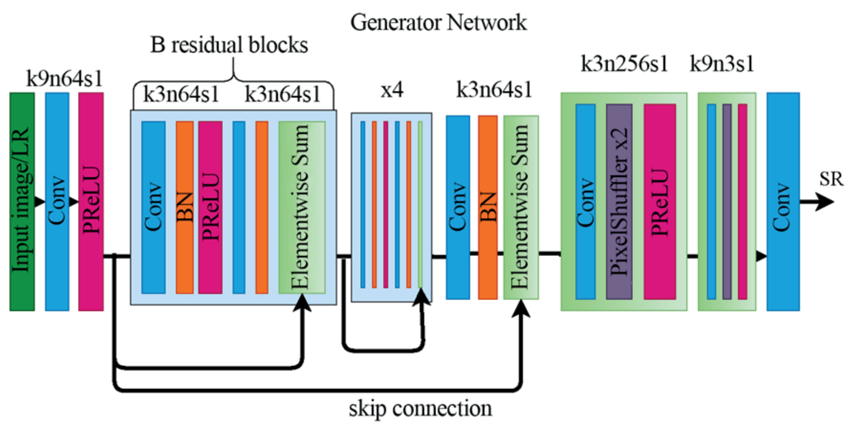


Fig. 2. Generator network architecture

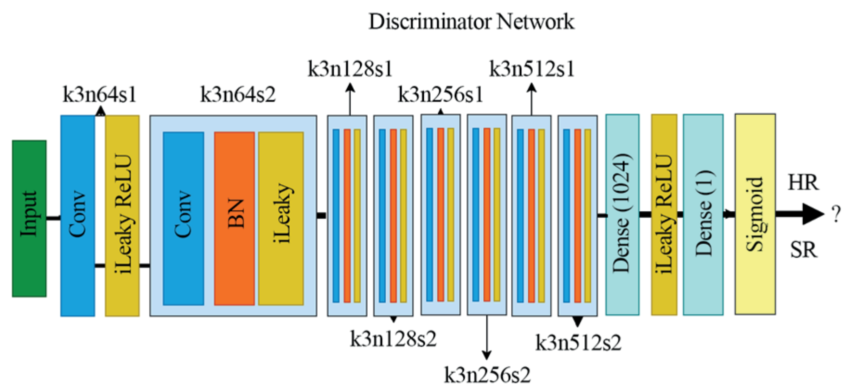


Fig. 3. The architecture of the discriminator networks with the number of feature maps (n) and stride (s) is specified for each convolutional layer

In this paper, an existing machine learning model, ESRGAN, was used to determine the loss of perception (by pixels) to add more sharpness to our predicted fake image. The combined use of the ESRGAN model with convolutional neural networks has shown that deeper network architectures can be difficult to train, but can potentially improve network accuracy significantly, as they allow very complex mappings to be modeled. To effectively train these deeper network architectures, batch normalization [16, 17] is often used to counteract internal covariant shift. Deeper network architectures have also been shown to improve performance for SISR.

5. Results of applying machine learning methods to improve texture-type images

5. 1. Evaluation of the effectiveness of using machine learning methods with convolutional neural networks for processing texture-type images

The model was trained using a prepared database of 20,936 images containing 12,396 medical images from an open database, 1,715 satellite images, and 6825 images from various classes (state license plates, human faces, buildings). The architecture was trained under a variation of the ESRGAN model with CNN augmented learning. At the same time, the best performance achieved 97.33 % success in identification with super enhanced images (SR). The chosen architecture of the model makes it possible to process and improve the image up to four times the best quality in various fields of science. Table 1 shows the average percentage improvement in the images of the considered models.

Table 1

Average accuracy for ESRGAN and ESRGAN+CNN training

Initial images / Algorithms	ESRGAN	ESRGAN+ +MobileNETV2	ESRGAN+ +VGG19	ESRGAN+ +ResNet152V2
Medical images	83 %	92 %	86 %	97 %
Satellite imagery	82 %	91 %	85 %	96 %
Other images (state numbers of cars, human faces, buildings, roads)	84 %	93 %	87 %	98 %

As shown in Table 1, the ESRGAN algorithm with augmented learning by convolutional neural networks gave the best image processing result. So, we can say that when processing and improving images in various fields of science, the proposed algorithm is effective.

5. 2. Implementation of improvement in selected areas of the image using machine learning methods

In recent years, there has been an active improvement of neural network models, which demonstrate their versatility and high efficiency, solving a number of problems, including in the field of image enhancement. Thus, an algorithm for improving image quality based on neural network approaches was implemented. During the experiment, 20,936 images were considered. Images of different subject areas were trained to see a comparative improvement in image quality. The following 100×100 (Fig. 4) images are embedded as source images.

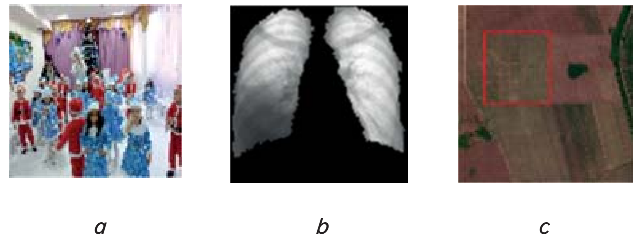


Fig. 4. Original image with a resolution of 100×100: *a* – source image for face recognition; *b* – X-ray image; *c* – space image

Human images (Fig. 4, *a*), X-ray images (Fig. 4, *b*) and space images (Fig. 4, *c*) were trained as the initial image of the ESRGAN model (Fig. 4). The images obtained as a result of training using the ESRGAN model are presented in the figure below (Fig. 5).

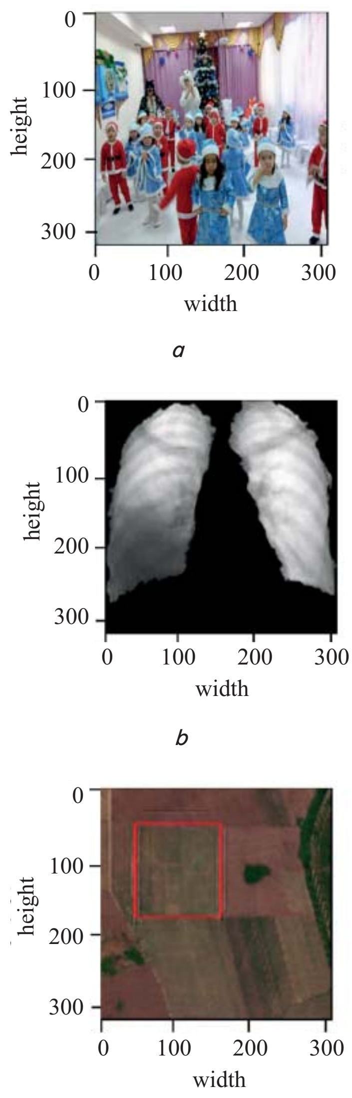
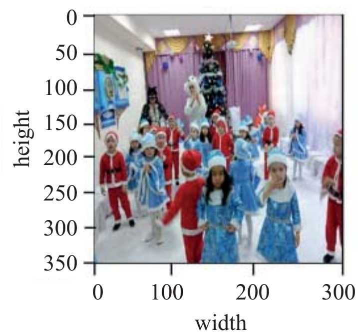
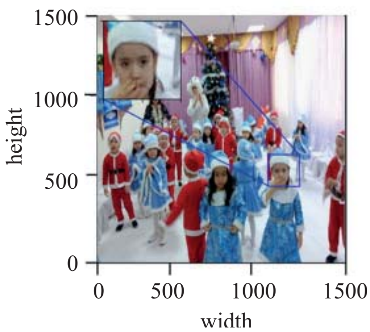


Fig. 5. Results of training with Enhanced Super-Resolution Generative Adversarial Networks: *a* – face recognition image; *b* – X-ray image; *c* – space image

The quality of images obtained when training the ESRGAN model did not improve significantly. Therefore, the result of retraining by adding ESRGAN+ResNet152V2 neural network methods to this model is shown in the figure below (Fig. 6).



a



b

Fig. 6. Comparison of results: *a* — enhanced super-resolution generative adversarial networks; *b* — enhanced super-resolution generative adversarial networks with convolutional neural network added

The image result obtained from the ESRGAN model is shown in Fig. 6, *a* and the retraining result by adding convolutional neural network methods is shown in Fig. 6, *b*. MobileNetV2, VGG119, ResNet152V2 methods were used as convolutional neural network methods. On average, CNNs showed VGG119=147.95, MobileNetV2=86.09, ResNet152V2=48.09 – pixel loss during overfitting. As a result of the experiment, we notice that the quality of the image obtained by the ResNet152V2 method improves significantly. Solving the problem of improving image quality using neural networks, the choice was made in favor of convolutional neural networks, since they are better than others in coping with image improvement tasks.

6. Discussion of the results of using machine learning algorithms and methods for image processing

Given the fact that neural networks are widely covered in the community of researchers and developers of image visualization algorithms, this paper presents an algorithm for computational technologies based on neural networks: in the field of medicine, agriculture and object recognition. So in the future, researchers in these areas will be able to highlight features, classify, recognize objects, segment images, as well as detect boundaries and analyze visual content. To raise awareness of how neural networks can be applied in these areas, image pre-processing has been carried out for further research and practical applications in related areas.

The original acquired images are low-frequency LR (100, 100, 3) (Fig. 4) images. Using the existing ESRGAN model, we increased the original image size by a factor of 4 (scaled) while maintaining the quality, i.e. we obtained high HR frequencies (400, 400, 3) (Fig. 5). By adding a CNN network to the trained model, the model was retrained and the resulting HR increased the frequency by a factor of 4, achieving an improved super SR (1600, 1600, 3) frequency image (Fig. 6, *b*) resolution.

With the help of efficient pre-processing algorithms and computer vision, the basic ESRGAN model can be extended to provide image enhancements and show sufficient accuracy for the use of the proposed algorithms in different areas of life [9].

An important limitation of the proposed algorithms is the computational resource. That is, when using the NVIDIA GEFORCE GTX 1650 Ti video card, 30 minutes were spent to train the ESRGAN model, while 6–7 hours were spent for the proposed algorithm. As a result of the experiments, it was found that one of the proposed models is highly accurate, that is, of the trained models, the ESRGAN algorithm with the ResNet152V2 convolutional neural network showed the highest accuracy.

The proposed algorithm has reduced the level of errors when improving the image and can be used to create a database in various subject areas of science.

The advantage is the ease of implementation of the algorithm, i.e. the ability to add convolutional neural networks before activating the function. This architecture uses model parameters more efficiently than others considered, is faster to assemble, and outperforms a deep neural network with a large sequence of parameters.

7. Conclusions

1. Data was pre-trained and trained with Enhanced Super-Resolution Generative Adversarial Networks models and in combination with Convolutional Neural Networks (MobileNETV2, VGG19, ResNet152V2). For the experiment, medical, satellite, and other images were taken as the original image. A study with the combined use of convolutional neural networks showed high accuracy scores. That is, the combined use of the ESRGAN algorithm with convolutional neural networks showed, for example, on average ESRGAN+MobileNETV2 – 91 %, ESRGAN+VGG19 – 86 %, ESRGAN+ResNet152V2 – 96 % accuracy.

2. Many scientific studies use the ESRGAN model to improve the image. To improve or restore images that are researched in different fields of science, specialists in this field are needed as an expert. In some cases, noise in images can be important information. Therefore, while improving images, we may lose important objects. In our work, for selected images where no experts are required, convolutional neural networks were added to the existing ESRGAN model, and a better result was achieved than the model itself. When performing calculations with the addition of the ResNet152V2 convolutional neural network, with improvement on pre-prepared 20,936 images, the ESRGAN + ResNet152V2 algorithm gives an accuracy result of up to 96 %.

Conflict of interest

The authors declare that they have no conflict of interest regarding this research, whether financial, personal, authorship or otherwise that could affect the research and its results presented in this paper.

Financing

The study was conducted without financial support.

Data availability

The manuscript has no associated data.

Acknowledgments

The author expresses gratitude to the Scientific and Production Center of Grain Farming named after A. I. Barayev for providing data on agricultural crops of Northern Kazakhstan in the preparation of this paper.

References

1. Yessenova, M., Abdikerimova, G., Adilova, A., Yerzhanova, A., Kakabayev, N., Ayazbaev, T. et al. (2022). Identification of factors that negatively affect the growth of agricultural crops by methods of orthogonal transformations. Eastern-European Journal of Enterprise Technologies, 3 (2 (117)), 39–47. doi: https://doi.org/10.15587/1729-4061.2022.257431
2. Yessenova, M., Abdikerimova, G., Baitemirova, N., Mukhamedrakhimova, G., Mukhamedrakhimov, K., Sattybaeva, Z. et al. (2022). The applicability of informative textural features for the detection of factors negatively influencing the growth of wheat on aerial images. Eastern-European Journal of Enterprise Technologies, 4 (2 (118)), 51–58. doi: https://doi.org/10.15587/1729-4061.2022.263433
3. Yessenova, M., Abdikerimova, G., Ayazbaev, T., Murzabekova, G., Ismailova, A., Beldeubayeva, Z. et al. (2023). The effectiveness of methods and algorithms for detecting and isolating factors that negatively affect the growth of crops. International Journal of Electrical and Computer Engineering (IJECE), 13 (2), 1669. doi: https://doi.org/10.11591/ijece.v13i2.pp1669-1679
4. Yerzhanova, A., Kassymova, A., Abdikerimova, G., Abdimomynova, M., Tashenova, Z., Nurlybaeva, E. (2021). Analysis of the spectral properties of wheat growth in different vegetation periods. Eastern-European Journal of Enterprise Technologies, 6 (2 (114)), 96–102. doi: https://doi.org/10.15587/1729-4061.2021.249278
5. Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V., Murayama, Y., Ranagalage, M. (2020). Sentinel-2 Data for Land Cover/Use Mapping: A Review. Remote Sensing, 12 (14), 2291. doi: https://doi.org/10.3390/rs12142291
6. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C. et al. (2019). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. Computer Vision – ECCV 2018 Workshops, 63–79. doi: https://doi.org/10.1007/978-3-030-11021-5_5
7. Wang, X., Xie, L., Dong, C., Shan, Y. (2021). Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). doi: https://doi.org/10.1109/iccvw54120.2021.00217
8. Clabaut, ., Lemelin, M., Germain, M., Bouroubi, Y., St-Pierre, T. (2021). Model Specialization for the Use of ESRGAN on Satellite and Airborne Imagery. Remote Sensing, 13 (20), 4044. doi: https://doi.org/10.3390/rs13204044
9. Zouch, W., Sagga, D., Echtioui, A., Khemakhem, R., Ghorbel, M., Mhiri, C., Hamida, A. B. (2022). Detection of COVID-19 from CT and Chest X-ray Images Using Deep Learning Models. Annals of Biomedical Engineering, 50 (7), 825–835. doi: https://doi.org/10.1007/s10439-022-02958-5
10. Yamashita, K., Markov, K. (2020). Medical Image Enhancement Using Super Resolution Methods. Computational Science – ICCS 2020, 496–508. doi: https://doi.org/10.1007/978-3-030-50426-7_37
11. Dou, X., Li, C., Shi, Q., Liu, M. (2020). Super-Resolution for Hyperspectral Remote Sensing Images Based on the 3D Attention-SRGAN Network. Remote Sensing, 12 (7), 1204. doi: https://doi.org/10.3390/rs12071204
12. Zhou, S., Yu, L., Jin, M. (2022). Texture transformer super-resolution for low-dose computed tomography. Biomedical Physics & Engineering Express, 8 (6), 065024. doi: https://doi.org/10.1088/2057-1976/ac9da7
13. Kang, X., Liu, L., Ma, H. (2021). ESR-GAN: Environmental Signal Reconstruction Learning With Generative Adversarial Network. IEEE Internet of Things Journal, 8 (1), 636–646. doi: https://doi.org/10.1109/jiot.2020.3018621
14. Jaworek-Korjakowska, J., Kleczek, P., Gorgon, M. (2019). Melanoma Thickness Prediction Based on Convolutional Neural Network With VGG-19 Model Transfer Learning. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). doi: https://doi.org/10.1109/cvprw.2019.00333
15. Pan, H., Pang, Z., Wang, Y., Wang, Y., Chen, L. (2020). A New Image Recognition and Classification Method Combining Transfer Learning Algorithm and MobileNet Model for Welding Defects. IEEE Access, 8, 119951–119960. doi: https://doi.org/10.1109/access.2020.3005450
16. Alrashedy, H. H. N., Almansour, A. F., Ibrahim, D. M., Hammoudeh, M. A. A. (2022). BrainGAN: Brain MRI Image Generation and Classification Framework Using GAN Architectures and CNN Models. Sensors, 22 (11), 4297. doi: https://doi.org/10.3390/s22114297
17. Zhang, W., Liu, Y., Dong, C., Qiao, Y. (2019). RankSRGAN: Generative Adversarial Networks With Ranker for Image Super-Resolution. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). doi: https://doi.org/10.1109/iccv.2019.00319