Noisy image enhancements using deep learning techniques

Kuanysh Daurenbekov[1], Ulzada Aitimova[2], Aigul Dauitbayeva[3], Arman Sankibayev[4], Elmira Tulegenova[3], Assel Yerzhan[5], Akbota Yerzhanova[6], Galiya Mukhamedrakhimova[7]

[1]Department for Student Affairs, S. Seifullin Kazakh AgroTechnical Research University, Astana, Kazakhstan
[2]Department of Information Systems, S. Seifullin Kazakh AgroTechnical Research University, Astana, Kazakhstan
[3]Department of Computer Science, Korkyt Ata Kyzylorda University, Kyzylorda, Kazakhstan
[4]Department of Artificial Intelligence Technologies, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
[5]Department of Telecommunications and Innovative Technologies, Almaty University of Power Engineering and Telecommunications named after G. Daukeev, Almaty, Kazakhstan
[6]Department of Technological Machines and Equipment, Faculty of Technology, S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan
[7]Department of Radio Engineering, Electronics and Telecommunications, Faculty of Physics and Technology, L. N. Gumilyov Eurasian National University, Astana, Kazakhstan

Article Info

Article history:

Received Jul 26, 2023
Revised Sep 25, 2023
Accepted Oct 9, 2023

Keywords:

Deep learning
Image processing
Machine learning
Multitasking learning model
Noisy image

ABSTRACT

This article explores the application of deep learning techniques to improve the accuracy of feature enhancements in noisy images. A multitasking convolutional neural network (CNN) learning model architecture has been proposed that is trained on a large set of annotated images. Various techniques have been used to process noisy images, including the use of data augmentation, the application of filters, and the use of image reconstruction techniques. As a result of the experiments, it was shown that the proposed model using deep learning methods significantly improves the accuracy of object recognition in noisy images. Compared to single-tasking models, the multi-tasking model showed the superiority of this approach in performing multiple tasks simultaneously and saving training time. This study confirms the effectiveness of using multitasking models using deep learning for object recognition in noisy images. The results obtained can be applied in various fields, including computer vision, robotics, automatic driving, and others, where accurate object recognition in noisy images is a critical component.

Corresponding Author:

Ulzada Aitimova
Department of Information Systems, S. Seifullin Kazakh AgroTechnical Research University
010000 Astana, Kazakhstan
Email: uaitimova@mail.ru

1. INTRODUCTION

Recognition of objects in images is an important task in the field of computer vision. However, real images often contain noise caused by various factors such as poor lighting, distortion, atmospheric effects, and other external noise. Noise can significantly complicate the process [1] of object recognition [2] and reduce model accuracy. In recent years, deep learning methods based on convolutional neural networks (CNNs) have received a lot of attention in the field of object recognition. These methods have shown high accuracy and the ability to extract complex features from images. However, the performance of deep learning on noisy images remains a challenge requiring further research. This paper proposes the use of multitasking models for object recognition in noisy images using deep learning methods. Multitasking models offer the solution to several related problems at the same time, including object recognition and noise filtering. This allows the model to learn and use noise information for more accurate object recognition. The paper considers various aspects of multitasking learning in the context of object recognition in noisy images.

Including the choice of model architecture, optimization of the loss function, adaptation of training data, and selection of suitable methods for noise filtering. This work aims to study the effectiveness of multitasking models using deep learning methods for object recognition in noisy images. It is assumed that this approach can significantly improve the process of object recognition, even in conditions of strong noise. Techniques that have been considered include the use of data augmentation, the use of filters to eliminate noise, and the development of custom CNN architectures that can account for noise and learn from noisy data. Also, the influence of various types of noise on the process of object recognition and optimal approaches for their elimination were studied. In the following, experiments and analysis of the results are presented that confirm or refute the hypothesis about the advantages of multitasking models in recognizing objects in noisy images. The results obtained can have significant practical applications, since the development of efficient models for object recognition in noisy images can be useful in many areas, including computer vision, automatic control, medical diagnostics, and industry.

Sanqian *et al.* [3] proposes a method for translating noisy images into clean images using generative-adversarial networks. The method allows for improving the quality of noisy images before their further recognition. Li *et al.* [4] proposes a new convolutional neural network architecture for removing noise from images. The authors show that their method can be successfully used for preprocessing noisy images before recognizing objects. Shen *et al.* [5] propose a method for teaching object detection on noisy images. The method takes noise into account and applies augmentation to train the model to consistently recognize objects in noise conditions.

In study [6], a method is proposed to improve the recognition of objects in images with a low level of illumination. The method uses convolutional neural networks to improve image quality and increase the accuracy of object detection. In study [7], the problem of vascular segmentation on noisy retinal images is investigated. A method is proposed that takes into account noise and applies adaptive filters to improve the quality of vascular segmentation. The study [8] is an overview of deep learning methods for denoising images. The review covers various approaches used to train deep learning models for denoising and discusses their application in the context of object recognition. Soylu *et al.* [9] explores the problem of semantic object segmentation in noisy images. A method is proposed that uses fully convolutional networks that consider noise during training to improve the quality of segmentation. Zhang *et al.* [10] proposes a method for robust object detection in noisy images. The method uses deep neural networks with progressive feature extraction to account for noise and improve object detection accuracy. Zou *et al.* [11] investigates the problem of object tracking on noisy videos. A method is proposed that uses deep neural networks to understand and diagnose visual tracking systems under noisy conditions. Chen *et al.* [12] explores the problem of face recognition in noisy images. A method using generative adversarial networks (GANs) is proposed to improve image quality and increase the accuracy of face recognition in noisy conditions

2. METHOD

This paper proposes a multitasking learning architecture that solves two problems: image classification by context and image enhancement. Classification is performed for three classes: road camera, space, and medical images. The classification uses the ResNet 152 architecture [13]–[15] multitasking models [16]–[18] based on the ResNet-152 architecture can have multiple output layers corresponding to each task. Each output layer can be associated with a corresponding class or label for a classification problem. Shared layers and parameters between tasks can be shared, allowing the model to use common information to improve performance across all tasks. After image classification, additional classification for traffic images occurs. She divides them into subclasses: road images with raindrops, snowfall, and fog. After classification into road image subclasses, the AttentiveGAN submodel [19] is passed through, which detects and cleans images of noise such as snow marks, raindrops, and fog. Image classification [20] using the AttentiveGAN model requires two separate steps: training the AttentiveGAN model for image generation and feature extraction, and then image classification using a separate classifier model such as CNN.

Other medical and space imaging classes go through the super-resolution generative adversarial network (SRGAN) submodel [21]–[23]. It performs image enhancement using generative adversarial networks (GANs). Upon completion of training, the model combines both tasks: classification and image enhancement using GAN [24], [25]. The result is a single model that can perform both tasks: image context classification and image enhancement. Multi-tasking learning architecture Figure 1 performs two tasks, the first task is image classification by context, which contains three classes, such as road camera images, space images, and medical images.

ResNet 152 architecture is used for classification, and image enhancement is performed after classification. Also, after classifying the road image, it goes through the classification again, dividing it into sub-classes, such as road image with raindrops, road image with snowfall, and road image with fog. After classifying the road image into subclasses, it goes through the AttentiveGAN submodels, identifying and removing noise from the images, such as snow marks, raindrops, and fog. The rest of the classes related to medical and satellite imagery go through the SRGANs submodel, and after training, a single model is derived that performs two tasks, such as classifying and image enhancement using GANs.

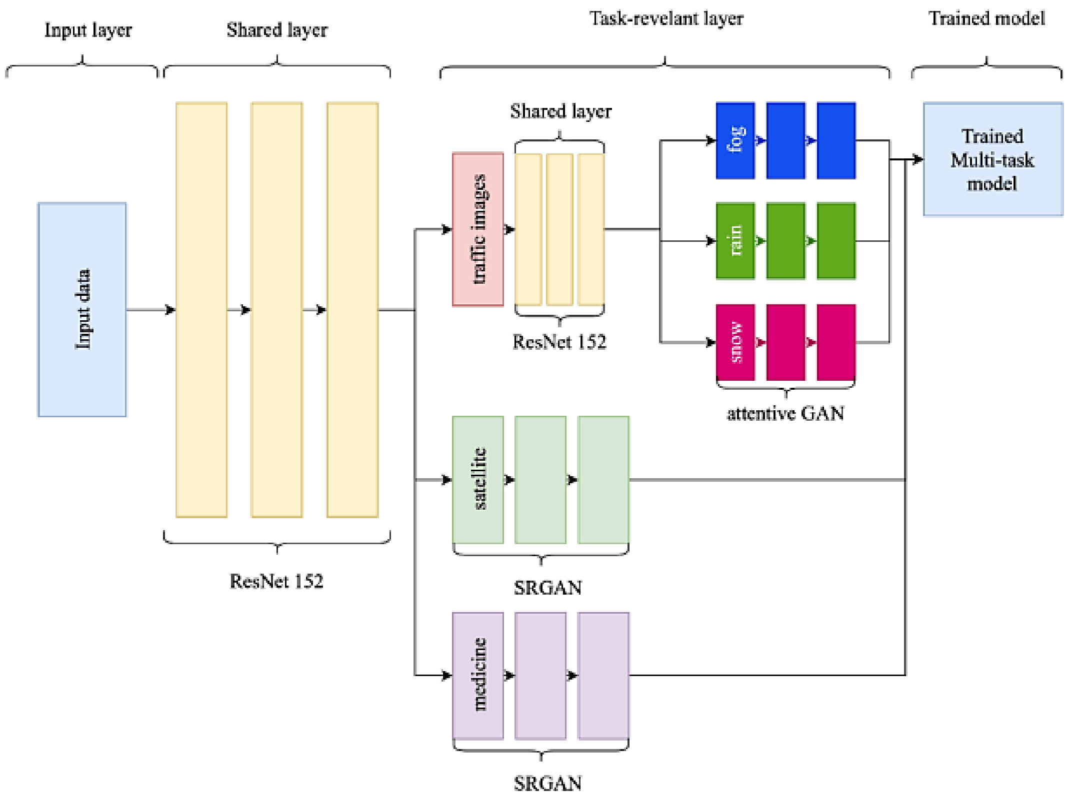


Figure 1. The architecture of the multitasking model

3. RESULTS AND DISCUSSION

The training of the multitasking model was carried out on a previously prepared database of 17,539 images as shown in Figure 2, containing 4,734 satellite images, 8,827 images with various types of noise (fog, snow marks, and raindrops), and 3,978 medical images from an open database as shown in Figure 2(a). The architecture was trained on a variant of the SRGAN model with extended CNN training. Based on the experiments and studies carried out, the following results were achieved:

a. Deep learning methods have increased the contrast, saturation, and other characteristics of the image, making it more attractive and informative. They enabled extracting complex features and training on large data sets, increasing recognition accuracy. As a result, CNN has detected and removed artifacts such as smudges, noise, flicker, and distortion to improve image quality and readability, as shown in Figure 2(b).
b. Data augmentation is an important tool for improving the generalizing ability of a model on noisy images. Scaling, rotating, adding noise, and other methods allow you to create various training examples that help the model better deal with real-world noise. These transformations provide data diversity and train the model to process and analyze noisy images with higher accuracy.
c. Applying filters and image restoration techniques is an effective approach to removing noise from noisy images. Low-pass filters were used to remove high-frequency noise from the image. They blur the image, retaining only low-frequency information, which reduces noise. This made it possible to improve the quality of images and facilitate the process of object recognition.

The multitasking learning architecture outlined above presents a holistic strategy for addressing image classification and enhancement challenges. This approach involves subcategorizing traffic images and applying AttentiveGAN and SRGANs submodels to enhance image quality based on contextual cues. To illustrate, Figure 3(a) demonstrates an original satellite image, while Figure 3(b) showcases the enhancement achieved through the chosen SRGAN learning model.

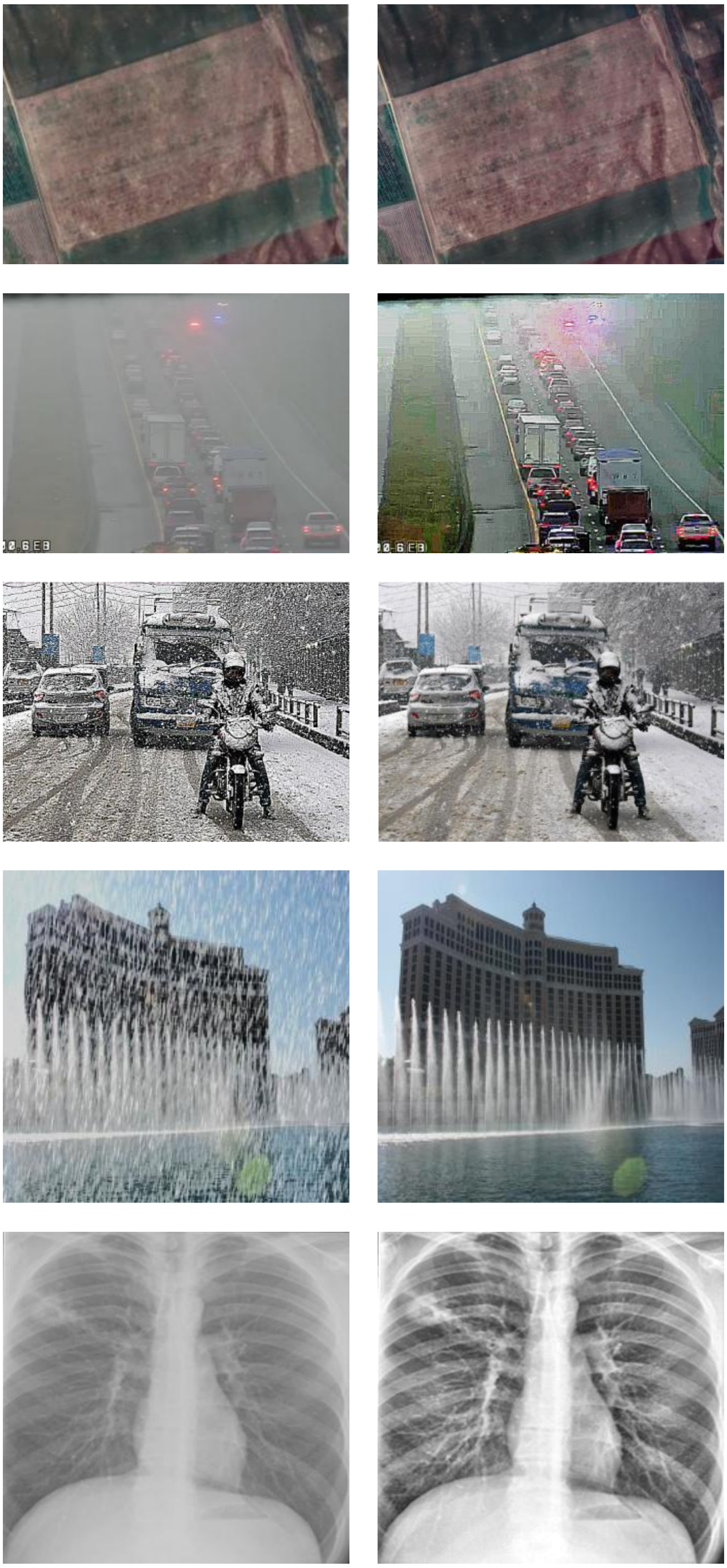
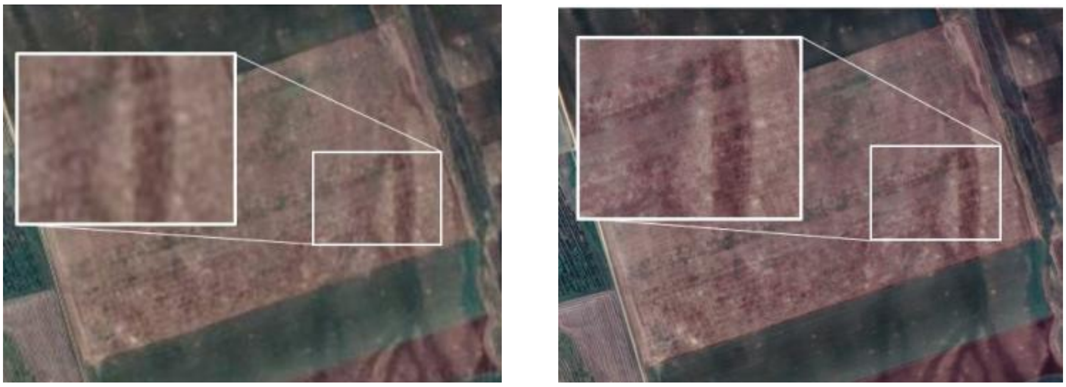


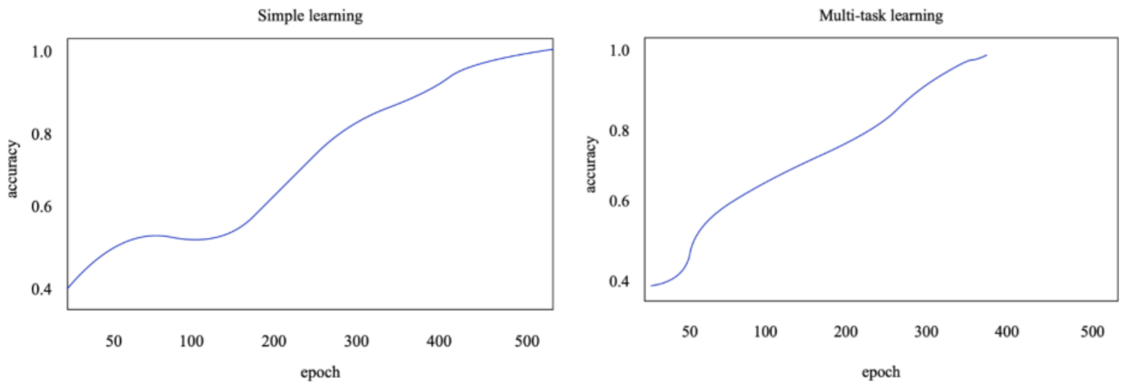
Figure 2. Overview of the multitasking model training on diverse image datasets: (a) Database distribution with satellite, noisy, and medical images, and (b) Improved image quality and readability after applying CNN techniques



(a) (b)

Figure 3. Illustration of image quality enhancement using the multitasking learning architecture: (a) Original satellite image, and (b) Enhanced image using the SRGAN model

After conducting experiments, the single-tasking model demonstrated its highest training performance, achieving a 99.5% accuracy after 530 epochs, as illustrated in Figure 4(a). However, this progress necessitated an extensive training duration of 18 hours. In contrast, the multi-task training model accomplished the same 99.5% accuracy but at a faster pace, reaching it in just 380 epochs, as illustrated in Figure 4(b), and requiring only 6 hours of training. Notably, the chosen model architecture can enhance image quality up to fourfold, making it valuable in diverse scientific domains.



(a) (b)

Figure 4. Comparative analysis of training accuracy between different models: (a) Single-tasking model's performance trajectory reaching 99.5% accuracy after 530 epochs, and (b) Multitasking model's accelerated performance, achieving the same accuracy in just 380 epochs

In this research work, a multi-task model was selected and trained on a joint dataset containing a variety of images affected by various types of noise and degradation. This choice allowed the model to share information and knowledge between different tasks, which in turn greatly improved its ability to analyze and reconstruct images. One of the key factors to improve the accuracy of object recognition in noisy images is the selection of the optimal convolutional neural network (CNN) architecture. The architectural design of CNNs has been carefully considered to accommodate noisy images, resulting in significant improvements in analysis and reconstruction performance. In particular, the use of convolutional layers specifically tailored to deal with noise has proven critical to achieving outstanding results. In addition, the study explored and applied models based on the concept of attention, such as attention network models (AttentiveGAN). These models have proven to be particularly effective at removing noise and restoring detail in noisy images. Attention mechanisms in the models made it possible to focus more accurately on important parts of the image, which contributed to improving the reconstruction quality. The study also included a thorough analysis and comparison of various quality metrics used to evaluate the results to ensure the effectiveness of the selected architectures and methods. These results confirmed that the choice of a multi-task model optimized for the characteristics of noisy images, as well as the use of attention network architectures, made a significant contribution to improving the accuracy of object recognition and restoration in noisy images. In summary, this study highlights the importance of selecting optimal architectural solutions and deep learning methods when working with noisy images and confirms their significant impact on the quality of analysis and reconstruction in the context of multi-task models.

4. CONCLUSION

This work explored the application of multitasking models using deep learning methods to improve noisy images. The goal was to solve the problem of distortion and reduction in image quality due to the presence of noise. Research has shown that multitasking models can solve these problems. By solving several related problems at the same time, the model can then be trained and use information about noise for more accurate object recognition. Multitasking models allow information to be shared between tasks, resulting in better generalization and improved performance. During the experiments and analysis of the results, it was confirmed that the use of multitasking models by deep learning methods on noisy images leads to an increase in accuracy in image enhancement. This opens up prospects for the application of such models in various fields, including computer vision, automatic control, medical diagnostics, and industry. However, it should be noted that the performance of multitasking models may depend on the nature of the noise, the algorithms used, and the architecture of the model. Further research may be aimed at optimizing these aspects to achieve even more accurate results. In general, the use of multitasking models by deep learning methods to improve noisy images is a promising area of research. This opens up new opportunities for developing more efficient recognition systems and creating high-precision solutions in noisy environments.

REFERENCES

[1] M. Yessenova *et al.*, "Application of informative textural Law's masks methods for processing space images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 4, pp. 4557–4566, Aug. 2023, doi: 10.11591/ijece.v13i4.pp4557-4566.

[2] M. Yessenova *et al.*, "Features of growth of agricultural crops and factors negatively affecting their growth," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 30, no. 1, pp. 625–632, 2023, doi: 10.11591/ijeecs.v30.i1.pp625-632.

[3] L. Sanqian, H. Risa, F. Huazhu, L. Heng, N. Jingxuan, and L. Jiang, "Content-preserving diffusion model for unsupervised AS-OCT image despeckling," *arxiv.org/abs/2306.17717*, Jun. 2023.

[4] Y. Li *et al.*, "NTIRE 2023 challenge on image denoising: methods and results," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 1904-1920. Accessed: Oct. 25, 2023. [Online]. Available: https://cvlai.net/ntire/2023/

[5] Y. Shen *et al.*, "Noise-aware fully webly supervised object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2020, pp. 11323–11332, doi: 10.1109/CVPR42600.2020.01134.

[6] H. Liang, A. Yu, M. Shao, and Y. Tian, "Multi-feature guided low-light image enhancement," *Applied Sciences (Switzerland)*, vol. 11, no. 11, p. 5055, May 2021, doi: 10.3390/app11115055.

[7] H. Sheng *et al.*, "Autonomous stabilization of retinal videos for streamlining assessment of spontaneous venous pulsations," *arxiv.org/abs/2305.06043*, May 2023.

[8] S. Ghose, N. Singh, and P. Singh, "Image denoising using deep learning: Convolutional neural network," in *Proceedings of the Confluence 2020 - 10th International Conference on Cloud Computing, Data Science and Engineering*, Jan. 2020, pp. 511–517, doi: 10.1109/Confluence47617.2020.9057895.

[9] B. Emek Soylu, M. S. Guzel, G. E. Bostanci, F. Ekinci, T. Asuroglu, and K. Acici, "Deep-learning-based approaches for semantic segmentation of natural scene images: A review," *Electronics (Switzerland)*, vol. 12, no. 12, p. 2730, Jun. 2023, doi: 10.3390/electronics12122730.

[10] Q. Zhang, S. Wang, X. Wang, Z. Sun, S. Kwong, and J. Jiang, "A multi-task collaborative network for light field salient object detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 5, pp. 1849–1861, May 2021, doi: 10.1109/TCSVT.2020.3013119.

[11] L. Zou, Y. Li, and F. Xu, "An adversarial denoising convolutional neural network for fault diagnosis of rotating machinery under noisy environment and limited sample size case," *Neurocomputing*, vol. 407, pp. 105–120, 2020, doi: 10.1016/j.neucom.2020.04.074.

[12] C. Chen, N. Hou, Y. Hu, S. Shirol, and E. S. Chng, "Noise-robust speech recognition with 10 minutes unparalleled in-domain data," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, May 2022, vol. 2022-May, pp. 8162–8166, doi: 10.1109/ICASSP43922.2022.9747755.

[13] B. Pishnamazi and E. Koushki, "Study of nonlinear optical diffraction patterns using machine learning models based on ResNet 152 architecture," *AIP Advances*, vol. 13, no. 1, Jan. 2023, doi: 10.1063/5.0135380.

[14] P. Nagpal, S. A. Bhinge, and A. Shitole, "A comparative analysis of ResNet architectures," in *2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, 2022, pp. 1-8, doi: 10.1109/SMARTGENCON56628.2022.10083966.

[15] S. Shanmugasundaram and N. Palaniappan, "Detection accuracy improvement on one-stage object detection using ap-loss-based ranking module and resnet-152 backbone," *International Journal of Image and Graphics*, 2023, doi: 10.1142/S021946782450030X.

[16] D. Bhattacharjee, T. Zhang, S. Susstrunk, and M. Salzmann, "MuIT: An end-to-end multitask learning transformer," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2022, vol. 2022-June, pp. 12021–12031, doi: 10.1109/CVPR52688.2022.01172.

[17] R. M. Samant, M. R. Bachute, S. Gite, and K. Kotecha, "Framework for deep learning-based language models using multi-task learning in natural language understanding: A systematic literature review and future directions," *IEEE Access*, vol. 10, pp. 17078–17097, 2022, doi: 10.1109/ACCESS.2022.3149798.

[18] S. H. Zou, X. Huang, X. D. Shen, and H. Liu, "Improving multimodal fusion with main modal transformer for emotion recognition in conversation," *Knowledge-Based Systems*, vol. 258, p. 109978, Dec. 2022, doi: 10.1016/j.knosys.2022.109978.

[19] G. Abdikerimova *et al.*, "Detection of chest pathologies using autocorrelation functions," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 4, pp. 4526–4534, Aug. 2023, doi: 10.11591/ijece.v13i4.pp4526-4534.

[20] J. Tussupov *et al.*, "Applying machine learning to improve a texture type image," *Eastern-European Journal of Enterprise Technologies*, vol. 2, no. 2–122, pp. 13–18, Apr. 2023, doi: 10.15587/1729-4061.2023.275984.

[21] R. S. Kushwaha, M. Rakhra, D. Singh, and A. Singh, "An overview: super-image resolution using generative adversarial network for image enhancement," in *Proceedings of 5th International Conference on Contemporary Computing and Informatics, IC3I 2022*, Dec. 2022, pp. 1243–1246, doi: 10.1109/IC3I56241.2022.10072862.

[22] K. ELKarazle, V. Raman, and P. Then, "Facial age estimation using machine learning techniques: An overview," *Big Data and Cognitive Computing*, vol. 6, no. 4, p. 128, Oct. 2022, doi: 10.3390/bdcc6040128.

[23] J. J. D. Khoo, K. H. Lim, and P. K. Pang, "Deep learning super resolution of sea surface temperature on South China Sea," in *2022 International Conference on Green Energy, Computing and Sustainable Technology, GECOST 2022*, Oct. 2022, pp. 176–180, doi: 10.1109/GECOST55694.2022.10010371.

[24] R. A. Khan, Y. Luo, and F. X. Wu, "Multi-level GAN based enhanced CT scans for liver cancer diagnosis," *Biomedical Signal Processing and Control*, vol. 81, p. 104450, Mar. 2023, doi: 10.1016/j.bspc.2022.104450.

[25] M. Alauthman *et al.*, "Enhancing small medical dataset classification performance using GAN," *Informatics*, vol. 10, no. 1, Mar. 2023, doi: 10.3390/informatics10010028.

BIOGRAPHIES OF AUTHORS

Kuanysh Daurenbekov candidate of technical sciences. Currently working at the Eurasian National University by L.N. Gumilyov at the Department of Informatics. Has more than 15 years of scientific and pedagogical experience, publication of an article in the Scopus database. Professor of the Russian Academy of Natural Sciences. He can be contacted by email: dkuan@mail.ru, dkuankaz@gmail.com.

Ulzada Aitimova Candidate of physical and mathematical sciences, acting associate professor. Nowadays working at the Kazakh AgroTechnical Research University named after S. Seifullin in the Department of Information Systems. Has 30 years of scientific and pedagogical experience. Moreover, has 60 scientific articles, including 5 articles based on the Scopus base, and 7 textbooks with ISBN. She can be contacted by email: uaitimova@mail.ru.

Aigul Dauitbayeva candidate of technical sciences, associate professor. Currently, works at the Korkyt Ata Kyzylorda University at the Department of Computer Science. Has more than 20 years of scientific and pedagogical experience and more than 50 scientific papers, including 6 articles based on Scopus, 2 monographs with ISBN. You can contact her by email: aicos@mail.ru.

Arman Sankibayev master of technical sciences, senior teacher. Currently working at the Department of Artificial Intelligence Technologies, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan. Has 3 years of scientific and pedagogical experience. He can be contacted by email: armandos1980@mail.ru, armanjana1980@gmail.com.

Elmira Tulegenova candidate of economic sciences, associate professor. Currently, she works at the Korkyt Ata Kyzylorda University in the Department of Computer Science. She has more than 20 years of scientific and pedagogical experience and more than 40 scientific papers, including 3 articles based on Scopus, 1 textbook. You can contact her by e-mail: etulegenova80@mail.ru.

Assel Yerzhan In 2003, she graduated from the Almaty Institute of Power and Communications with a degree in industrial electronics. In 2004, he received a Master's degree in radio electronics and telecommunications. In 2013, she graduated from the doctoral program "Kazakh National Technical University named after K.I. Satbayev", specialty 6D071900 - "radio engineering, electronics and telecommunications". From 2014 to the present, he has been a doctor of philosophy PhD in the specialty 6D071900 - "radio engineering, electronics and telecommunications" of the Gumarbek Daukeev Almaty University of Power and Communications. She is the author of more than 40 works. Her research interests include circuit engineering in telecommunications, wireless communications, mobile communication systems, GSM and mobile systems management, sensor networks, 5G and mobile communication technologies. You can contact her by e-mail: erjanasel@gmai.com, a.erzhan@aues.kz.

Akbota Yerzhanova defended her doctor of philosophy (Ph.D.) in the specialty 6D070300 - “information systems" at the L.N. Gumilyov Eurasian National University in Astana. Currently, she works at the Kazakh Agrotechnical Research University named after S. Seifullin. She is the author or co-author of more than 15 publications. The Hirsch index is 2. Her research interests include geoinformation technologies, pattern recognition, and artificial intelligence. You can contact her at erjanova_akbota@mail.ru.

Galiya Mukhamedrakhimova candidate of pedagogical sciences. Currently, she works at the Department of Radio Engineering, Electronics and Telecommunications, Faculty of Physics and Technology, Eurasian National University by L.N. Gumilyov. Research interests: methods of professional education in higher education, machine learning, pattern recognition. You can contact her by e-mail: isatai-07@mail.ru.