

Generative augmentation to improve lung nodules detection in resource-limited settings

N. F. Gusarova^a, PhD, Tech., Associate Professor, orcid.org/0000-0002-1361-6037, natfed@list.ru

A. P. Klochkov^a, Student, orcid.org/0000-0002-6843-7888

A. A. Lobantsev^a, Software Engineer, orcid.org/0000-0002-8314-5103

A. S. Vatian^a, PhD, Tech., Associate Professor, orcid.org/0000-0002-5483-716X

M. V. Kabyshev^a, Post-Graduate Student, orcid.org/0000-0002-1006-0408

A. A. Shalyto^a, Dr. Sc., Tech., Professor, orcid.org/0000-0002-2723-2077

A. A. Tatarinova^b, PhD, Med., Senior Researcher, orcid.org/0000-0001-5955-2529

T. V. Treshkur^b, PhD, Med., Associate Professor, orcid.org/0000-0001-5955-2529

Min Li^c, PhD, Professor, orcid.org/0000-0002-1361-6037

^aITMO University, 49, Kronverksky Pr., 197101, Saint-Petersburg, Russian Federation

^bAlmazov National Medical Research Centre, 2, Akkuratova St., 197341, Saint-Petersburg, Russian Federation

^cSchool of Computer Science and Engineering, Central South University, 932, South Lushan Road, Changsha, Hunan, 410083 P.R., China

Introduction: Lung cancer is one of the most formidable cancers. The use of neural network technologies in its diagnostics is promising, but the datasets collected from real clinical practice cannot cover various lung cancer manifestations. **Purpose:** Assessment of the possibility of improving pulmonary nodules classification quality utilizing generative augmentation of available datasets under resource constraints. **Methods:** The LIDC-IDRI dataset was used. We used the StyleGAN architecture, to generate artificial lung nodules and the VGG11 model as a classifier. **Results:** We generated pulmonary nodules using the proposed pipeline and invited four experts to evaluate them visually. Four experimental datasets with different types of augmentation were formed, including the use of synthesized data. We compared the effectiveness of the classification performed by the VGG11 network when training for each dataset. For an expert assessment, 10 generated nodules in each group of characteristics were presented: parietal nodules, ground-glass, sub-solid, solid nodules. In all cases, expert assessments of similarity with real nodules were obtained with a Fleiss's kappa coefficient $\kappa = 0.7-0.9$. We got the values of AUROC=0.9867 and AUPR=0.9873 with the proposed approach of a generative augmentation. **Discussion:** The obtained efficiency metrics are superior to the baseline results obtained using comparably small training datasets and slightly less than the best results achieved using much more powerful computational resources. We have shown that one can effectively use StyleGAN for augmenting an unbalanced dataset with a combination of VGG11 as a classifier, which does not require extensive computing resources and a sizeable initial dataset for training.

Keywords – lung nodules classification, data augmentation, generative adversarial networks, StyleGAN, CT image.

For citation: Gusarova N. F., Klochkov A. P., Lobantsev A. A., Vatian A. S., Kabyshev M. V., Shalyto A. A., Tatarinova A. A., Treshkur T. V., Li Min. Generative augmentation to improve lung nodules detection in resource-limited settings. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 6, pp. 60–69. doi:10.31799/1684-8853-2020-6-60-69

Introduction

Lung cancer is one of the most formidable cancers, both in terms of the development rate and the severity of the prognosis [1]. In this case, it is a vital necessity to get the earliest possible and accurate diagnosis. During the initial examination and screening of the population, procedures such as chest radiography and sputum cytology are widespread. However, when detecting the suspicions of lung cancer, the patient requires stronger diagnostic procedures, including bronchoscopic biopsies and computed tomography (CT) of the lungs. A bronchial biopsy is a highly invasive procedure. It involves the participation of proficient experts, is accompanied by complications and side effects, and cannot be used as a regular diagnostic procedure. Simultaneously, CT of the lungs is a non-invasive procedure, which

does not adversely affect the patient's health, and CT scanners are now widely used medical equipment. In this regard, increasing the efficiency of CT in the diagnosis of lung cancer is today one of the essential tasks of information technology.

The use of machine learning technologies and, above all, deep convolutional neural networks in this task has led to promising results in recent years. For instance, for the classification of malignant and benign nodes, the following results are given in literature: accuracy = 92.0%, sensitivity = 100% [2]; accuracy = 96.0%, sensitivity = 97% [3]. However, such high values of metrics are achieved, as a rule, on typical datasets (most often, the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset [4] is used). When moving to other datasets or real practice, the values of metrics drop dramatically. For example, the accu-

racy achieved by [3] on a relatively large proprietary dataset of 2054 images was 86% only.

The reason for such a fall in efficiency is mainly is as follows: benign and malignant nodes in lung cancer are very similar, both in objectively measurable features (such as diameter, optical density, etc.); in terms of integral visual assessment (such as smooth, lobulated, or irregular and spiculated margins, etc.) [5]. The existing datasets collected from real clinical practice cannot cover such a variety of lung cancer manifestations. Their size is not enough for full-fledged training of neural networks, which, as a result, leads to the explicit overfitting on a specific dataset and the drop in efficiency when switching to new datasets. An increase the dataset volume due to traditional augmentation methods, such as shifts, rotations, reflections, etc., does not give the desired improvement in the results of lung cancer classification. Therefore, generative adversarial networks (GAN) are considered a promising technology for solving this problem.

Generative adversarial network, proposed in [6], is a model to approximate an arbitrary distribution only by sampling from that distribution. The model consists of two parts — the generator and the discriminator. The generator aims to learn the sample distribution. It takes random noise and tries to generate the sample from the learned distribution. The discriminator tries to distinguish these generated objects from the real objects from the training sample with an arbitrary distribution and returns the results to the generator via gradient back-propagation. Thus, during training, the generator generates objects that are more and more similar to a sample.

Modern implementations of GAN technologies provide realistic-looking images, for example, pictures for an online store, avatars for games, and video clips. A significant advantage for the application of GAN in medicine is the fact that they provide the extraction of visual features by discovering the high dimensional latent distribution of the input data [7, 8]. Thus, in principle, it becomes possible to generate an unlimited number of images of benign and malignant nodes belonging to the same distribution as real nodes in a particular dataset and thereby augment this dataset to a size sufficient for effective training of the classifier. The article discusses the challenges of using GAN for generative augmentation of small datasets gathered in real clinical practice. To improve the availability of generative augmentation to the research community, this article focuses on the resource efficiency of the proposed methods.

Background and related works

As the literature review shows [8], today, GANs are actively used to analyze high-tech images in various medical applications, including diagnostics

and treatment of diseases of the brain, lungs, spine, cardiovascular system, etc. Within the framework of this article, we highlight the work related to the diagnosis of lung status. In the total flow of the publications, the share of works devoted to using GANs in pulmonology problems is relatively small. They can be divided into two groups — applying GANs for chest X-ray and CT images.

As concerning the first one, one should first of all mention the work [9], where the authors use a deep convolutional GAN (DCGAN) for mimicking common chest pathologies and then augment a labeled set of chest X-rays for the training of the deep CNN across five pathological classes. The authors in [10] use conditional GAN for improving lung segmentation in chest X-ray. But, as the authors themselves note, their construction is resource expensive and requires high computing power.

The authors [11] solve a complicated problem — predicting the dynamics of lung position during breathing. To do this, they use two CNNs, each of which is built according to the GAN scheme. The solution is very resource-intensive: to register images between any two breath phases, it uses a powerful NVIDIA Tesla V100 GPU within 1 min. The authors [12] also solve a dynamic problem — to visualize chronic obstructive pulmonary disease progress. They proposed a method of visualization for regression with GAN, which is also very resource-intensive.

The research [13] aims to balance dataset used in training CNN for pneumonia prediction via oversampling with CycleGAN [14] producing X-ray images with pneumonia from images with no pneumonia. The advantage of the proposed augmentation technique is that it does not require extensive computational resources (a single NVIDIA 1070 graphical card is enough).

Going to the overview of applying GANs for CT images, it is worth noting that the general concept behind GANs of [6] has been transformed here in various architectures. For example, to generate high-quality images of pulmonary nodes on CT scans, the authors [15] use DCGANs [16]. With a relatively simple architecture, their implementation required extensive computational resources (up to 110,000 iterations), while the generated images showed low results on the Turing test (58%). Attempts have been made to use a more complicated architecture for the same purpose: the authors of [17] used a 3D conditional GAN [18], and the approach in [19] suggest to apply a sophisticated variant called 3D multi-conditional GAN. In both works, the authors generate pulmonary nodes and their immediate environment (context) and then embed them into the general CT images. For this, they condition the GAN basing on a volume of interest whose central part containing the nodule has

been erased. Nevertheless, despite the sharp increase in the architecture complexity and a significant increase in computing resource requirements, the generated images are easily distinguished by qualified radiologists due to artifacts in synthetic samples [8].

Thus, as the analysis of existing achievements shows, it is hardly advisable today to set the task of creating realistic images of pulmonary nodes with the help of GAN having limited resources. Instead, researchers move on to modeling the characteristic components of the desired images. For pulmonary nodes, such a component is the maximum intensity projection (MIP) image. MIP [20, 21] is a post-processing method that projects 3D voxels with maximum intensity to the projection plane. The advantage of MIP is that its formation is a 2D task and therefore requires less computing resources. Besides, radiologists use MIP images as more easily interpretable concepts during the nodule screening stage in their routine clinical practice.

A typical pipeline for detecting neoplasms in the lungs using high-tech images consists of two independently solvable tasks — detecting nodules and their subsequent classification. When organizing research having only small proprietary datasets, the second task is more topical. As our analysis evidence, in the literature, works using GANs for detection tasks [17, 19, 22] are presented more widely and thoroughly than concerning classification tasks [23–26]. They used axial sections of the volume of interest centered on the pulmonary nodule as a generated characteristic component.

The authors [24] aiming to generate high-quality synthetic nodules images proposed a new GAN architecture named forward and backward GAN (F&BGAN) and formed a hierarchical learning framework based on multi-scale VGG11 network as a classifier. They tested different augmentation approaches including traditional methods, DCGAN generative augmentation, and proposed F&BGAN generative augmentation. The part of LIDC-IDRI [28] was used as the initial dataset. The accuracy from 88.09% up to 95.24% was obtained depending on the augmentation approach.

In [25], the authors used Wasserstein GAN (WGAN) to generate malignant nodes differing by the only feature — the presence of spicules. They used a relatively small proprietary CT dataset consisting of 60 cases for training. Due to the low resolution of the formed nodes, the authors obtained not very good classification accuracy values — up to 63.0% for benign nodes and 84.8% for malignant nodes. In [26], they improved these metrics by increasing the network complexity (moving to three CNNs).

In [23], the authors used LUNA16 [27] database to extract candidate areas of images and nodules

and imitate them using GAN. The authors divided the nodules extracted from the dataset into three groups: large, medium, and small. Then, they generated artificial nodules using the GAN for each group separately. Synthesized nodules allowed them to change the distribution of node sizes in the generated dataset by weighting each group's share. As the authors write, they tested their method for 15 variants of the CNN of six different feature extraction types and classifiers on the newly generated images dataset. They received accuracy values with a wide scatter — from 78.21 to 95.13%. Unfortunately, no technical details of the development are provided, making it impossible to reproduce their results.

Considering all of the above, in our article, we set the task of experimentally testing the possibility of improving pulmonary node classification into malignant and benign utilizing generative augmentation of available datasets under resource constraints. The problem is solved in the 2D projection.

Method and materials

Initial dataset and data preprocessing

As the source of lung cancer nodules, we used the LIDC-IDRI dataset containing more than 1018 CT scans in DICOM format from about 1010 different patients. The characteristics of lesions in the dataset and specifics of the annotation of the data can be accessed in [28]. It should be noted that authors of the dataset are aware that the term nodule is more appropriately used for a spectrum of abnormalities in lung tissue, and according to that, they state that during the annotation procedure, each of four participating radiologists provided their own “noduleness” interpretation.

For the experiments, we selected only those DICOM series that contain tumor nodules. The DICOM series is a 3D scan of the lungs, and the tumor nodule is a 3D image. Therefore, by capturing part of the images from a series of images, we can extract the 3D nodule. The extraction of a nodule is performed as follows. We form a cube circumscribing the desired nodule. The bounding cube size is selected to capture the nodule completely and, if necessary, a small margin around the nodule. In our experiments, we used circumscribing cubes with a side length from 1 to 40 mm.

After extracting the bounding cube with a nodule, we resampled it to a size of $128 \times 128 \times 128$ pixels, and cut off the pixel values that are outside the range $[-1000, 800]$ on the Hounsfield scale [29] according to the formula

$$\text{pixel}_{\text{value}}^{\text{new}} = \min\left(800, \max\left(-1000, \text{pixel}_{\text{value}}^{\text{old}}\right)\right)$$

The resulting pixel values are scaled to the range $[-1, 1]$ by the formula

$$x = 2(x_{in} - in_{min}) / ((in_{max} - in_{min}) - 1),$$

where $in_{max} = 800$, $in_{min} = -1000$ are boundary values on the Hounsfield scale. As a result, we got 695 cubes, of which 294 have a malignant nodule. Examples of nodules which were used as input data to train the generative network are shown in Fig. 1, *a-e*.

As discussed in the previous section, there are two ways to go from 3D to a 2D image of a nodule: either to select a central slice (i. e., passing through the center of the nodule) in one of the projections or to build MIP for the central slice. As our experiments have shown, using MIP projection provides some advantages, namely:

- it shows the picture of the lungs in more detail;
- it rather clearly displays the nodules;
- it allows representing the nodules with the context around.

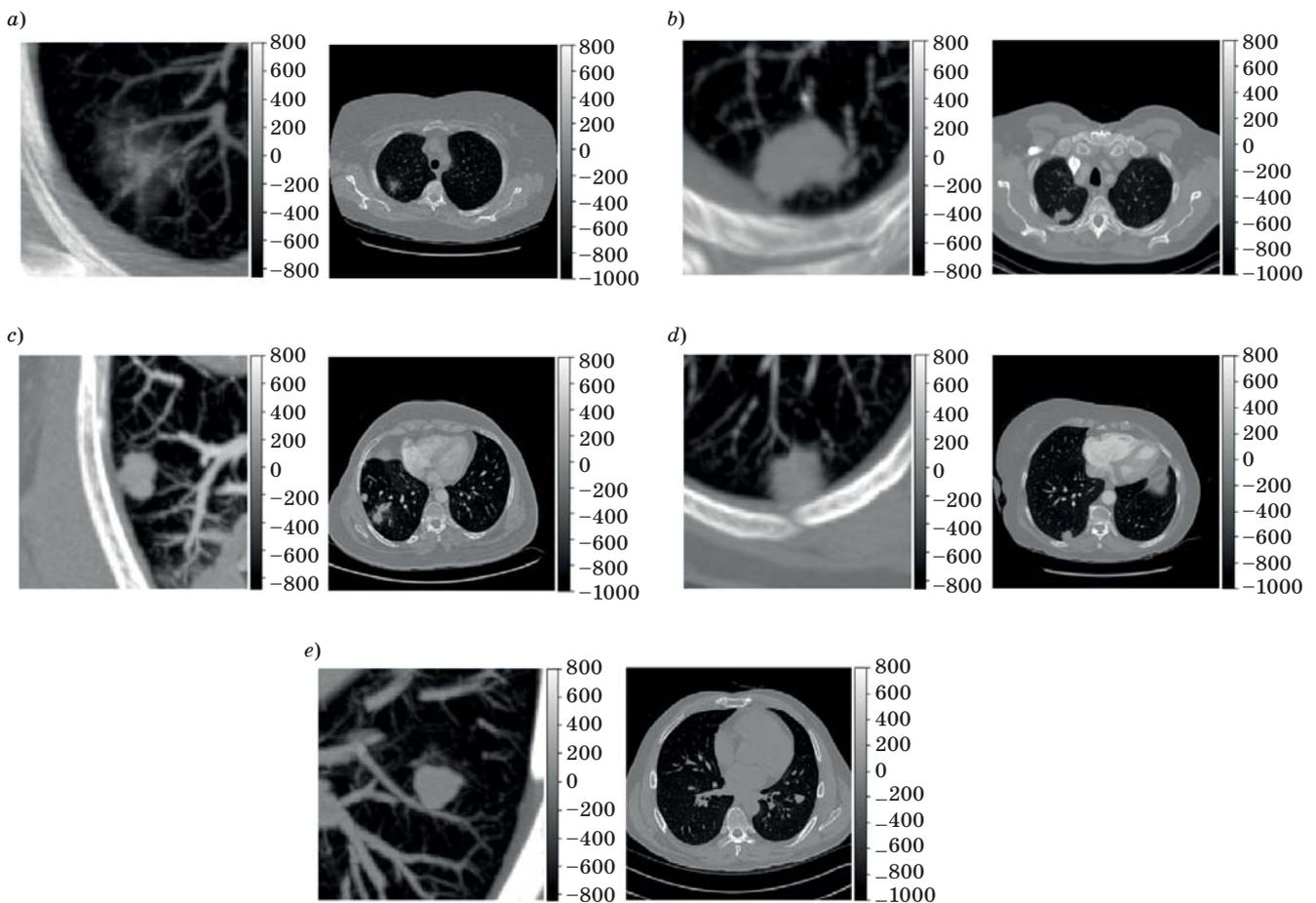
All of the above simplifies facilitates the task of classifying a cancerous tumor. We have performed a MIP lookup operation using the NumPy package.

Models and model training

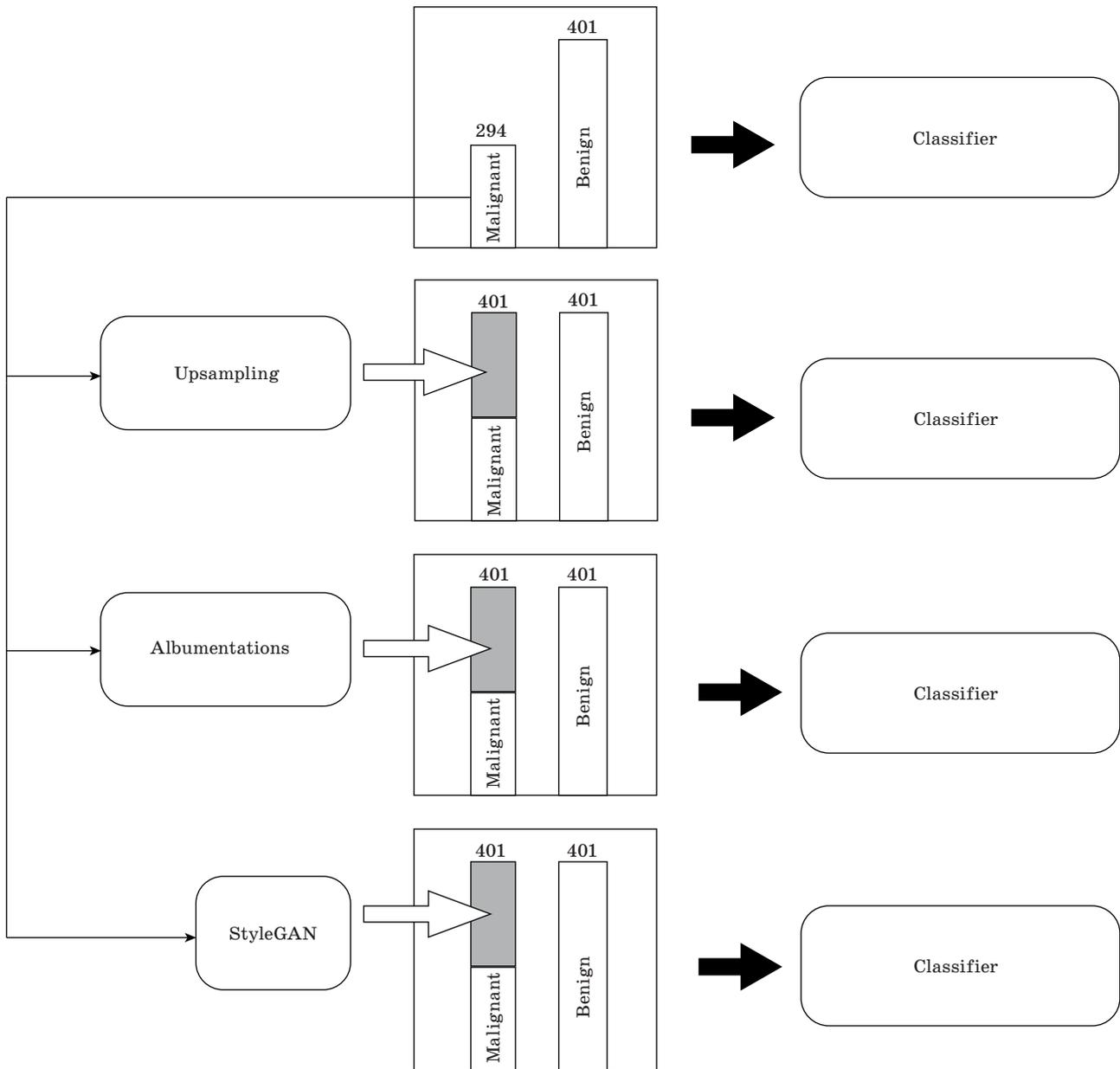
The complete pipeline of our experiments is shown in Fig. 2.

As an architecture for GAN, we chose StyleGAN [30] — one of the leading in the generation of photorealistic images. Considering that generating nodes in 2D is not more difficult than generating faces (on the example of StyleGAN was tested), this choice can be regarded as justified.

The paper aims to increase the cancer nodules classification quality under resource constraints. Hence the model with a simpler and more light-weight architecture suits better. Experiments show that VGG11 is the most appropriate model, which gives competitive results. We used the VGG11 model [31] as a classifier — a reasonably clear and easy-to-understand classifier model, which is often the baseline for research. We used the *torchvision* package to implement the model.



■ **Fig. 1.** Examples of extracted nodules (left column — nodules; right column — a corresponding lung slice): *a* — ground glass nodules; *b-d* — nodules of parietal localization; *e* — solid nodule



■ Fig. 2. Pipeline of our experiments

■ Table 1. Classifier training parameters

Name	Value
Model	VGG-11
batch_size	16
learning_rate	1e-5
optimizer	Adam
Loss function type	BinaryCrossEntropy
Training epochs	300

It should be emphasized that both models are relatively undemanding in terms of computing resources and can be applied in everyday machine learning practice [32].

To adapt the StyleGAN architecture following the prepared data, we made some changes to the model. The original implementation is configured to work with 3-channel images. To adjust the model in accordance with the prepared data, we have transformed the number of channels in the input and output layers in such a way as to be able to work with single-channel images. Besides, we made changes to the parameters of the model presented in Table 1.

Experimental datasets

To test the hypothesis that the binary classification of cancerous tumors in the lungs is better performed on a dataset augmented with synthesized GAN data, we formed four experimental datasets:

1) dataset A: original dataset with class imbalance;

2) dataset B: with the elimination of class imbalance by random copying of data of a smaller class (upsampling);

3) dataset C: with the elimination of class imbalance by transforming data of one class (vertical and horizontal reflection, and elastic transformation from the albumentations package [33]);

4) dataset D: balanced dataset using synthesized data. The imbalance was eliminated by generating new data using a pre-trained GAN model.

Results and discussion

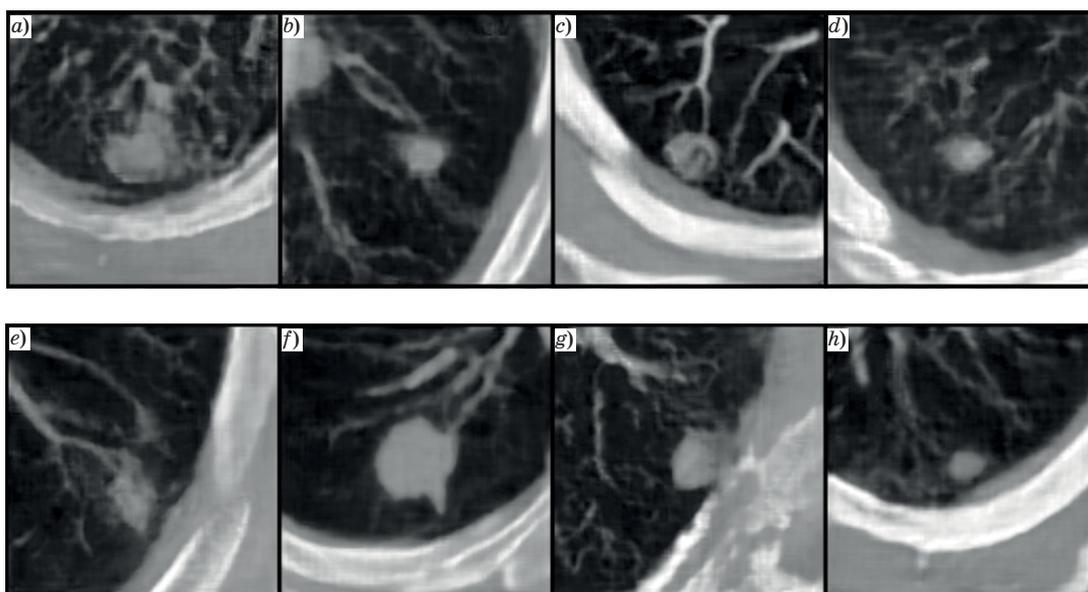
Examples of nodules generated by our GAN model are presented in Fig. 3, *a–h*. As shown in the previous section, it is not required to achieve an exhaustive execution of the Turing test on the generated nodules in the task under consideration. Therefore, we carried out an expert assessment of the “similarity” of the generated nodules for individual characteristics, including parietal nodule, solid nodule, subsolid nodule, ground-glass nodule. Four qualified radiologists participated in the examination, ten nodules in each group of characteristics were presented for assessment. In all cases,

positive expert assessments were obtained with a good Fleiss’s kappa coefficient $\kappa = 0,7–0,9$.

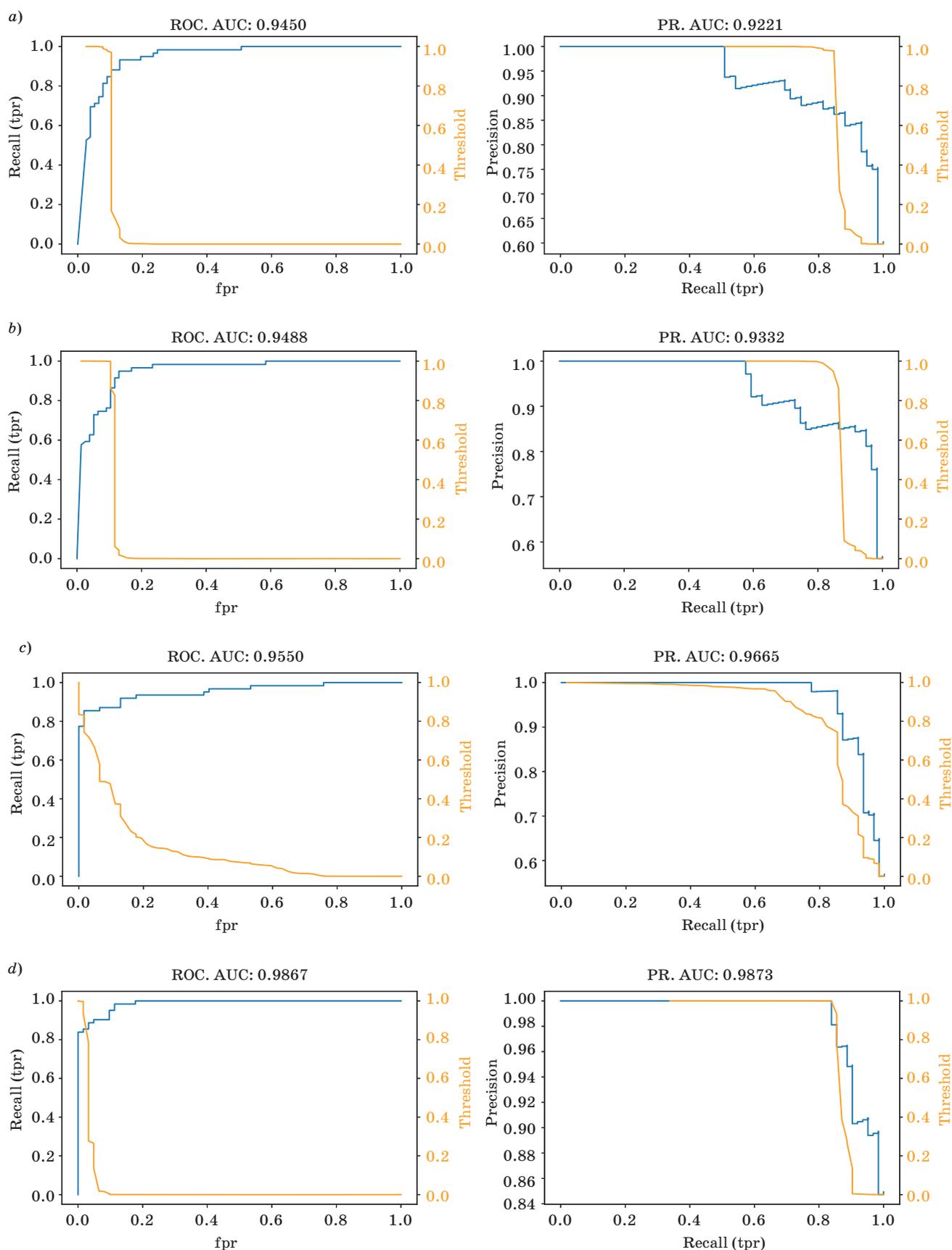
Figure 4 shows the ROC-curves for the best learning epochs for each experimental dataset described above, which makes it possible to compare the efficiency of the different augmentation techniques. As can be seen, the best values of AUROC: 0.9867, AUPR: 0.9873, accuracy: 94.35% were obtained with the proposed approach of a generative augmentation (see Fig. 3, *d*). Note that the obtained values are superior to the [25] results obtained using comparable training datasets (balanced accuracy: 81.7%), [26], (balanced accuracy: 85.6%), and comparable with the results of [24] (best accuracy: 95.24%, best AUROC: 0.984). Worth noting that the classifier model in [24] has approx-

■ *Table 2.* Results

Dataset	Accuracy, %	AUROC	AUPR
Original (A)	84.68	0.945	0.922
Upsampling (B)	87.50	0.949	0.933
Augmentation (C)	91.13	0.955	0.967
Synthetic (D)	94.35	0.987	0.987



■ *Fig. 3.* Examples of generated nodules: *a, g* — subsolid nodules of parietal localization nodules; *b–d, f, h* — solid nodules; *e* — subsolid nodule



■ Fig. 4. ROC- and PR-curves for the datasets: a — dataset A; b — dataset B; c — dataset C; d — dataset D

imately 143 million trainable parameters, while used VGG11 has 128 million trainable parameters. Hence, our proposed method has a lower GPU memory consumption with comparative quality results (Table 2).

Conclusion

It is not rare in common practice when a machine learning practitioner can encounter a lack of data to train the classification model properly. However, as our experiments have shown, augmenting an unbalanced dataset with synthetic data improves the classifier efficiency with comparatively no significant

effort. Regarding the classification of pulmonary nodules, we have shown that one can effectively use a combination of StyleGAN and VGG11, which does not require extensive computing resources and a sizeable initial dataset for training. We suggest that in future works, the use of StyleGAN in generative augmentation can be extended to conditional augmentation to synthesize the nodules with the specific parameters.

Financial support

This work was financially supported by Russian Science Foundation, Grant 19-19-00696.

References

1. *Cancer Facts & Figures 2020*. Atlanta, American Cancer Society, 2020. Available at: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf> (accessed 18 September 2020).
2. Makajua S., Prasad P. W. C., Alsadoona A., Singhb A. K., Elchouemic A. Lung cancer detection using CT scan images. *Procedia Computer Science*, 2018, vol. 125, pp. 107–114.
3. Wang S., Dong L., Wang X., and Wang X. Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy. *Open Medicine*, 2020, vol. 15, iss. 1, pp. 190–197.
4. *Lung Image Database Consortium image collection (LIDC-IDRI)*. Available at: <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI> (accessed 18 September 2020).
5. Purandare N. C., and Rangarajan V. Imaging of lung cancer: implications on staging and management. *The Indian Journal of Radiology and Imaging*, 2015, vol. 25, iss. 2, p. 109.
6. Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets. *Proceedings of Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680.
7. Creswell A., White T., Dumoulin V., Arulkumaran K., Sengupta B., and Bharath A. A. Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 2018, vol. 35, iss. 1, pp. 53–65.
8. Kazemina S., Baur C., Kuijper A., van Ginneken B., Navab N., Albarqouni S., and Mukhopadhyay A. GANs for medical image analysis. *Artificial Intelligence in Medicine*, 2020, p. 101938.
9. Salehinejad H., Valaee S., Dowdell T., Colak E., and Barfett J. Generalization of deep neural networks for chest pathology classification in X-rays using generative adversarial networks. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 990–994.
10. Munawar F., Azmat S., Iqbal T., Grönlund C., and Ali H. Segmentation of lungs in chest X-ray image using generative adversarial networks. *IEEE Access*, 2020, no. 8, pp. 153535–153545.
11. Fu Y., Lei Y., Wang T., Higgins K., Bradley J. D., Curran W. J., Liu T., Yanga X. LungRegNet: an unsupervised deformable image registration method for 4D-CT lung. *Medical Physics*, 2020, vol. 47, iss. 4, pp. 1763–1774.
12. Lanfredi R. B., Schroeder J. D., Vachet C., and Tadziden T. Adversarial regression training for visualizing the progression of chronic obstructive pulmonary disease with chest X-rays. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019, pp. 685–693.
13. Malygina T., Elicheva E., and Drokin I. GANs' N Lungs: improving pneumonia prediction. arXiv preprint arXiv:1908.00433, 2019. Available at: <https://arxiv.org/pdf/1908.00433> (accessed 18 September 2020).
14. Zhu J. Y., Park T., Isola P., and Efros A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2223–2232.
15. Chuquicusma M. J. M., Hussein S., Burt J. R., Bagci U. How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 240–244.
16. Radford A., Metz L., and Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015. Available at: <https://arxiv.org/pdf/1511.06434.pdf> (accessed 18 September 2020).

17. Jin D., Xu Z., Tang Y., Harrison A. P., and Mollura D. J. CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 732–740.
18. Mirza M., and Osindero S. Conditional generative adversarial nets. preprint arXiv:1411.1784, 2014. Available at: <https://arxiv.org/pdf/1411.1784.pdf> (accessed 18 September 2020).
19. Han C., Kitamura Y., Kudo A., Ichinose A., Rundo L., Furukawa Y., Umemoto K., Li Y., Nakayama H. Synthesizing diverse lung nodules wherever massively: 3D multi-conditional GAN-based CT image augmentation for object detection. *International Conference on 3D Vision (3DV)*, 2019, pp. 729–737.
20. Zhang J., Xia Y., Zeng H., and Zhang Y. NODULe: Combining constrained multi-scale LoG filters with densely dilated 3D deep convolutional neural network for pulmonary nodule detection. *Neurocomputing*, 2018, vol. 317, pp. 159–167.
21. Zheng S., Guo J., Cui X., Veldhuis R. N., Oudkerk M., and Van Ooijen P. M. Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection. *IEEE Transactions on Medical Imaging*, 2019, vol. 39, iss. 3, pp. 797–805.
22. Gao C., Clark S., Furst J., and Raicu D. *Augmenting LIDC dataset using 3D generative adversarial networks to improve lung nodule detection*. In: *Medical Imaging 2019: Computer-Aided Diagnosis*, 2019, vol. 10950, p. 109501K.
23. Esmailshahmirzadi N., Mortezaipoor H. A novel method for enhancing the classification of pulmonary data sets using generative adversarial networks. *Biomedical Research*, 2018, vol. 29, iss. 14, pp. 3022–3027.
24. Zhao D., Zhu D., Lu J., Luo Y., and Zhang G. Synthetic medical images using F&BGAN for improved lung nodules classification by multi-scale VGG16. *Symmetry*, 2018, vol. 10, iss. 10, p. 519.
25. Onishi Y., Teramoto A., Tsujimoto M., Tsukamoto T., Saito K., Toyama H., Imaizumi K., Fujita H. Automated pulmonary nodule classification in computed tomography images using a deep convolutional neural network trained by generative adversarial networks. *BioMed Research International*, 2019, vol. 2019, Article ID 6051939. <https://doi.org/10.1155/2019/6051939>
26. Onishi Y., Teramoto A., Tsujimoto M., Tsukamoto T., Saito K., Toyama H., Imaizumi K., Fujita H. Multi-planar analysis for pulmonary nodule classification in CT images using deep convolutional neural network and generative adversarial networks. *International Journal of Computer Assisted Radiology and Surgery*, 2020, vol. 15, iss. 1, pp. 173–178.
27. Setio A. A. A., Traverso A., De Bel T., Berens M. S., van den Bogaard C., Cerello P., ... and van der Gugten R. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. *Medical Image Analysis*, 2017, vol. 42, pp. 1–13.
28. Armato III S. G., McLennan G., Bidaut L., McNitt-Gray M. F., Meyer C. R., Reeves A. P., ... and Kazerooni E. A. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*, 2011, vol. 38, iss. 2, pp. 915–931.
29. Feeman T. G. *The Mathematics of Medical Imaging: A Beginner's Guide*. Springer Undergraduate Texts in Mathematics and Technology. Springer, 2015. 197 p.
30. Karras T., Laine S., and Aila T. A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4401–4410.
31. Simonyan K., and Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014. Available at: <https://arxiv.org/pdf/1409.1556.pdf> (accessed 18 September 2020).
32. Iglovikov V., and Shvets A. Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. arXiv preprint arXiv:1801.05746, 2018. Available at: <https://arxiv.org/pdf/1801.05746> (accessed 18 September 2020).
33. Buslaev A., Iglovikov V. I., Khvedchenya E., Parinov A., Druzhinin M., and Kalinin A. A. Albuementations: fast and flexible image augmentations. *Information*, 2020, vol. 11, iss. 2, p. 125.

УДК 004.895

doi:10.31799/1684-8853-2020-6-60-69

Генеративная аугментация для улучшения обнаружения узелков в легких в условиях ограниченных ресурсовН. Ф. Гусарова^а, канд. техн. наук, доцент, orcid.org/0000-0002-1361-6037, natfed@list.ruА. П. Клочков^а, студент, orcid.org/0000-0002-6843-7888А. А. Лобанцев^а, инженер-программист, orcid.org/0000-0002-8314-5103А. С. Ватьян^а, канд. техн. наук, доцент, orcid.org/0000-0002-5483-716XМ. В. Кабышев^а, аспирант, orcid.org/0000-0002-1006-0408А. А. Шалыто^а, доктор техн. наук, профессор, orcid.org/0000-0002-2723-2077А. А. Татарина^б, канд. мед. наук, старший научный сотрудник, orcid.org/0000-0001-5955-2529Т. В. Трешкур^б, канд. мед. наук, доцент, orcid.org/0000-0001-5955-2529Мин Ли^в, доктор наук, профессор, orcid.org/0000-0002-1361-6037^аУниверситет ИТМО, Кронверкский пр., 49, Санкт-Петербург, 197101, РФ^бНМИЦ им. В. А. Алмазова, Аккуратова ул., 2, Санкт-Петербург, 97341, РФ^вШкола компьютерных наук и инженерии, Центральный Южный университет, 932, Союф Люшан роуд, Чанша, Хунан, 410083 P.R., Китай

Введение: рак легкого — один из самых опасных видов рака. Использование технологий нейронных сетей для его диагностики является многообещающим, но датасеты, собранные из реальной клинической практики, не могут охватить различные проявления рака легких. **Цель:** оценка возможности улучшить классификацию легочных узлов посредством генеративной аугментации доступных датасетов при ограниченных ресурсах. **Методы:** использован датасет LIDC-IDRI, архитектура StyleGAN для создания искусственных изображений легочных узлов и модель VGG11 в качестве классификатора. **Результаты:** проведены генерация изображений легочных узлов с помощью предложенной схемы и их визуальная оценка с привлечением четырех экспертов. Сформированы четыре экспериментальных датасета с различными типами аугментации, включая использование синтезированных данных, и проведено сравнение эффективности классификации, выполняемой сетью VGG11 при обучении на каждом датасете. Для экспертизы отобраны по 10 генерированных изображений легочных узлов в каждой группе характеристик. Во всех случаях получены экспертные оценки схожести с реальными экземплярами с коэффициентом каппа Флейса $\kappa = 0,7-0,9$. Предложенный подход генеративной аугментации позволил получить значения AUROC = 0,9867 и AUPR = 0,9873. **Обсуждение:** полученные показатели эффективности превосходят результаты бейзлайна с использованием сравнительно небольших обучающих датасетов и немного уступают лучшим результатам, достигнутому с применением гораздо более мощных вычислительных ресурсов. Тем самым показано, что для аугментации несбалансированного датасета можно эффективно использовать StyleGAN в комбинации с VGG11 классификатором, которая не требует больших вычислительных ресурсов, а также большого начального датасета для обучения.

Ключевые слова — классификация легочных узлов, аугментация данных, генеративные состязательные сети, StyleGAN, КТ-изображение.

Для цитирования: Gusarova N. F., Klochkov A. P., Lobantsev A. A., Vatian A. S., Kabyshev M. V., Shalyto A. A., Tatarinova A. A., Treshkur T. V., Li Min. Generative augmentation to improve lung nodules detection in resource-limited settings. *Информационно-управляющие системы*, 2020, № 6, с. 60–69. doi:10.31799/1684-8853-2020-6-60-69

For citation: Gusarova N. F., Klochkov A. P., Lobantsev A. A., Vatian A. S., Kabyshev M. V., Shalyto A. A., Tatarinova A. A., Treshkur T. V., Li Min. Generative augmentation to improve lung nodules detection in resource-limited settings. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 6, pp. 60–69. doi:10.31799/1684-8853-2020-6-60-69