DOI: https://doi.org/10.15407/pmach2024.04.073

UDC 004.8

ADAPTIVE REFINEMENT OF SEGMENTED OBJECT CONTOUR BASED ON THE BRIGHTNESS OF NEIGHBORING PIXELS USING THE ENSEMBLE METHOD

Vladyslav D. Koniukhov
riggellllll@gmail.com
ORCID: 0009-0007-0256-1388

Anatolii Pidhornyi Institute
of Power Machines and Systems
of NAS of Ukraine,
2/10, Komunalnykiv str., Kharkiv,
61046, Ukraine

Improvement of the accuracy of computer vision algorithms plays a significant role in the tasks of medical image segmentation. After all, determining the boundaries of objects is a difficult task when using medical images, and especially X-ray images. The use of X-ray images in segmentation tasks is a complex process, since these images themselves can have a sufficient amount of noise and artifacts. Classical segmentation methods face significant challenges when segmenting X-ray images where there are objects with fuzzy boundaries. To solve such tasks, it is suggested to use segmentation with the help of machine learning, and to increase the accuracy of the objects' boundaries determination, it is necessary to use adaptive approaches. This paper proposes a new method to improve the accuracy of X-ray image segmentation, which analyzes the neighboring pixels of each contour element and adaptively reshapes it if necessary, and then combines all predictions using an ensemble method, which improves the previous version of the contour. The method was able to demonstrate an improvement in the quality of image segmentation on three datasets with different complexity of structures. Improvements in object boundary accuracy were obtained for all three sets.

***Keywords**: machine learning, neural networks, deep learning, image segmentation, medical image analysis.*

Introduction

In current conditions, image segmentation plays one of the most important roles in various fields, such as: industry, self-driving cars and medicine. The use of machine learning in image segmentation tasks has achieved significant success [1]. But despite the progress we have made in working with medical images, we still face significant challenges that need to be addressed. This is, in particular, because these images may contain noise and artifacts, which, in turn, complicate segmentation and cause problems in determining the boundaries. The heterogeneity of structures also plays a significant role in this, since different shapes and structures complicate the task. The huge number of pathologies that have different sizes, shapes, textures – all this greatly affects the final result.

The paper proposes a method that basically uses an algorithm for analyzing neighboring pixels. To increase the accuracy of segmentation, the ensemble method of combining predictions is used. The idea is that the segmentation accuracy is initially not able to reach 100% of the result and nearby pixels may be part of the contour of the segmented object. A kernel is applied to each element of the contour, after which the brightness values of the neighboring pixels are compared with the original value and if the neighboring pixel is brighter, it should be part of the new contour. This approach allows us to adaptively change the contour obtained as a result of using the ensemble. With this help, we can more accurately determine the boundaries of the object.

Literature review

Modern methods of segmentation can achieve good values in the performance of their tasks. They are developing very successfully and with every year we can see more and more methods aimed at increasing the accuracy of image segmentation. One of these methods is the method proposed in the paper [2]. This method is used to improve the segmentation of the object boundary region. The authors use contour-based boundary refinement. It is universal and does not require high computational costs. In the paper [3], the authors use texture analysis to improve segmentation. The method proposed by the authors analyzes each pixel using gray level feature extraction. They apply gradients to define the edges of the object. An adaptive active contour model based on local fitting for dividing the region of interest was presented in [4]. The method used by the authors assumes that an adaptive approach for moving the initial contour to the boundary of the object contributes to a better solution of the problem. In [5], the authors propose a method for adjusting boundaries for each boundary pixel. They made ad-

justments taking into account the structure, which shifted the contour to the edges in the direction of uniformity. The method of refining the contours taking into account the polygons of the vessel wall structures is presented in [6]. The authors use the approach of determining an anatomically plausible contour, after which they gradually refine it using a vector field to improve border alignment. The paper [7] proposed a method of combining the advantages of local and global information in the function of a set of image levels. Due to this, it was possible to obtain effective segmentation results even in cases of non-uniform intensity, bad edges and presence of noise.

Ensembles of neural networks are used to improve the quality of segmentation. Ensemble methods have an advantage over the use of single neural networks [8]. Combining a large number of neural networks provides an opportunity to minimize the impact of random errors and improve the overall efficiency of the segmentation.

Materials and methods

As a method of combining the predictions of the ensemble of neural networks, we will use the method of averaging the shape with alignment on the mean center. This method consists of two stages. Let us have n networks in the ensemble. For a test image I_i we get predictions $Masks_{pred}=\{P_1, P_2, \ldots, P_n\}$. At the first stage, the method determines the average value of the center for each object of all predictions $Masks_{pred}$. After that, all objects are shifted relative to the average center and we get $Masks_{shifted}$. The mean center is defined as $C_{mean}=(x_{mean}, y_{mean})$, where

$$x_{mean} = \frac{1}{n}\sum_{i=1}^{n} x_i \; ; \; y_{mean} = \frac{1}{n}\sum_{i=1}^{n} y_i \; .$$

At the second stage, it is necessary to combine all masks $Masks_{shifted}$. And that's why the distance conversion method is used. It averages the shapes of objects. For all masks, the distance must be obtained

$$d_{all} = \sum_{i=1}^{n} d(mask_{shifted,i}) \, ,$$

where $d(mask)$ is determined by the formula

$$d=dt(mask)–dt(\sim mask),$$

where $dt(mask)$ – is a distance transformation method and $\sim mask$ is an inverted mask.

In the end, an averaged image is obtained, where if the sum of the distances is greater than zero, then the pixels are colored white, otherwise they are colored black. A formula is used for this

$$I_{res} = d_{all} > 0 \, .$$

After the method of combining predictions is ready, the method of adaptive contour change can be defined. This method uses a kernel of size $k \times k$, which is superimposed on the contour element with its center, after which a line-by-line comparison of the brightness values of the contour element and neighboring pixels takes place. If the brightness of the neighboring pixel is greater, then the active contour element is replaced with a new value, with the coordinates where the new element was found, otherwise the element remains part of the contour. Let us have a contour $C = \{(x_i, y_i)\}_{i=1}^{N}$ which was obtained as a result of use of I_{res} and is a set of points. For each contour point (x_i, y_i) an analysis area with size $k \times k$ is used. Limits of the analysis area for a point (x_i, y_i) are given by the formula

$$x_{\min} = \max\left(0, x_i - \frac{k}{2}\right); \; x_{\max} = \min\left(W-1, x_i + \frac{k}{2}\right);$$

$$y_{\min} = \max\left(0, y_i - \frac{k}{2}\right); \; y_{\max} = \min\left(H-1, y_i + \frac{k}{2}\right),$$

where W and H – width and height of the image, respectively.

The search for a pixel that is brighter than the current element of the contour occurs as follows. For each point (x_i, y_i) it is necessary to perform a check of the brightness of the pixels in the given analysis window $(x, y) \in [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$. If for a point (x_i, y_i) pixel analysis found a pixel (x_i, y_i) such that

$$I(x_j, y_j) > I(x_i, y_i) \text{ та } (x_j, y_j) \notin C \, ,$$

then point (x_j, y_j) is added to C_{new}, otherwise in C_{new} is added (x_i, y_i).

After updating all the points belonging to the contour, the contour C_{new} is formed, which already contains more precise object boundaries.

Experiments

Images from open sources were used for experiments [9–11]. These were X-ray images that were reduced to a single size of 512×512 pixels, all images were in grayscale. Three data sets were created: the first included masks of the left lung; the second included right lung masks; the third consisted of masks of the spine region, the region consisted of four vertebrae. The first set included 703 images, the second set included 703 images, and the third set consisted of 190 images. The following distribution of images was used for training and testing: 90/10% (training/testing). Fig. 1 shows which images were used. From left to right: right lung, left lung, spine area.

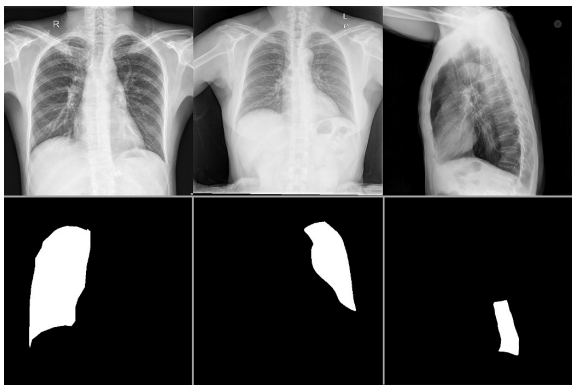

Fig. 1. Example of used images

FCN8-MobileNet was chosen as the model. The combination of FCN8 and MobileNet makes it possible to achieve a golden mean between segmentation accuracy and efficiency.

In order to evaluate the quality of segmentation, the Dice-Sørensen coefficient was used, which has the following form:

$$DSC = \frac{2|A \cap B|}{|A| + |B|}.$$

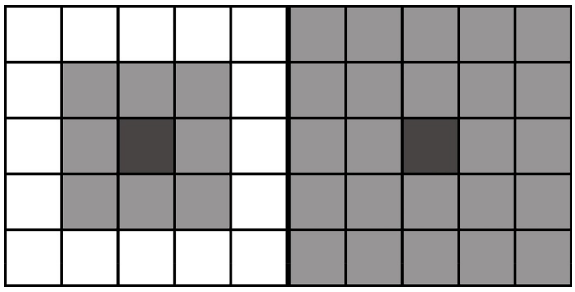

Fig. 2. Example of used kernels

Firstly, 10 neural networks were trained for each data set. Next, an ensemble consisting of these 10 networks will be used. For a more thorough and detailed study of the method, it was proposed to consider three variants of this method:

1. A neighboring pixel is outside or inside the output contour and is brighter than the current one.
2. The neighboring pixel is outside the output contour and is brighter than the current one.
3. The neighboring pixel is inside the output contour and is brighter than the current one.

The next task after using neural networks separately and in an ensemble is the application of three variants of the algorithm for each neural network, two variants of using kernel sizes were considered: 3×3 and 5×5. Visually used kernels can be seen in Fig. 2. On the left is a 3×3 kernel, and on the right is a 5×5 kernel. The dark gray square is the point of the contour to which the algorithm is applied, and the neighboring light gray squares are the neighbors that need to be analyzed.

At the end, the final version of the algorithm will be considered - combining the method of analyzing neighboring pixels and the ensemble approach. In the proposed method, the neighboring pixel analysis algorithm was first used, and then the ensemble method.

Results

Table 1 shows the training quality of the neural networks used. From the obtained data, it follows that the model used is stable, since the values for each network do not differ much. The average value for the left lung and the right lung is very close, but the value for the spine is smaller.

Table 2 shows the application of three variants of the neighboring pixel analysis algorithm for each neural network. For the left lung, the best result was given by the version 2 algorithm with a 5×5 kernel. For the right lung, the best result was also obtained using algorithm version 2 with a 5×5 kernel. For the last data set, algorithm version 1 with a 5×5 kernel was able to demonstrate the best result. As can be seen from the obtained results, in two cases out of three, the version 2 algorithm with a 5×5 kernel is leading. This result is due to the fact that this approach is based on the analysis of pixels that are outside the original contour. And with great confidence, we can say that the analysis of external pixels and the increase of the contour will lead to an improvement in the definition of the object boundaries. It is also worth noting that it was the use of a

5×5 kernel that helped to obtain the best result, again confirming that the contour improvement is based on expanding its shape outwards. For the left lung, the segmentation improvement was 0.56, for the right lung it was 0.34, and for the last dataset – 0.46. The comparison was made for the average values of single networks and the used algorithm.

The application of the neural network ensemble and the combination of the neighboring pixel analysis algorithm with the neural network ensemble are shown in Table 3. For the first set with the left lungs, the best result was demonstrated by the application of version 1 with a 5×5 kernel, version 2 with a 5×5 kernel was almost identical. For the second dataset with right lungs, version 1 with a 5×5 kernel and version 2 with a 5x5 kernel had identical results. In the third set, the best result was demonstrated by version 1 with a 5×5 kernel. Again, the 5×5 kernel gave a better result.

Summarizing the results, it is better to use version 1 and version 2 of the neighboring pixel analysis algorithm for the first data set. For the second data set, in one case version 2 was the best, and in the other case, version 1 and version 2 had the same result. For the third data set, it is better to use version 1 in both cases.

Table 1. The value of the Dice-Sørensen coefficient % by neural networks

Data set	Neural network number										Statistics
	1	2	3	4	5	6	7	8	9	10	$\overline{DSC}$
Left lung	95.40	95.64	95.59	95.54	95.53	95.48	95.75	95.74	95.60	95.97	95.62
Right lung	94.85	95.53	95.19	95.32	94.88	95.61	95.80	95.67	95.68	95.58	95.41
Spine area	91.43	91.07	90.76	89.81	89.45	90.75	91.34	91.55	91.48	91.75	90.94

Table 2. The value of the Dice-Sørensen coefficient % for three variants of the neighboring pixel analysis algorithm

Data set	Version	$k\times k$	Neural network number										Statistics
			1	2	3	4	5	6	7	8	9	10	$\overline{DSC}$
Left lung	1	3×3	95.86	95.97	95.97	95.77	95.90	95.97	96.13	96.14	95.86	96.29	95.99
		5×5	96.01	96.11	96.16	95.85	96.11	96.33	96.33	96.33	95.95	96.42	96.16
	2	3×3	95.90	96.05	95.91	95.76	96.05	95.92	96.08	96.09	95.91	96.28	96.00
		5×5	96.09	96.22	96.07	95.86	96.34	96.28	96.26	96.28	96.04	96.41	96.18
	3	3×3	95.46	95.58	95.63	95.52	95.45	95.51	95.77	95.76	95.56	95.98	95.62
		5×5	95.43	95.53	95.65	95.50	95.35	95.52	95.79	95.77	95.51	95.95	95.60
Right lung	1	3×3	95.24	95.75	95.50	95.60	95.21	95.68	95.86	95.71	95.88	95.84	95.63
		5×5	95.51	95.82	95.70	95.76	95.47	95.70	95.82	95.72	95.95	96.00	95.74
	2	3×3	95.31	95.70	95.52	95.66	95.18	95.75	95.86	95.71	95.85	95.86	95.64
		5×5	95.56	95.76	95.70	95.81	95.38	95.80	95.84	95.71	95.91	96.01	95.75
	3	3×3	94.87	95.60	95.20	95.32	94.94	95.58	95.79	95.66	95.72	95.58	95.43
		5×5	94.87	95.61	95.21	95.31	94.98	95.55	95.77	95.65	95.73	95.58	95.43
Spine area	1	3×3	91.68	90.86	91.35	90.03	89.92	91.23	91.73	91.78	92.02	92.23	91.28
		5×5	91.70	90.36	91.79	90.05	90.22	91.65	91.88	91.80	92.26	92.25	91.40
	2	3×3	91.64	90.68	91.09	89.77	89.96	91.32	91.88	91.82	91.88	92.25	91.23
		5×5	91.70	90.28	91.40	89.69	90.30	91.65	92.12	91.86	92.01	92.43	91.34
	3	3×3	91.37	91.14	90.89	89.93	89.48	90.81	91.27	91.59	91.74	91.90	91.01
		5×5	91.27	91.04	91.03	90.00	89.41	90.81	91.07	91.45	91.74	91.68	90.95

Table 3. The value of the Dice-Sørensen coefficient % for the ensemble approach

Data set	Algorithm						
	Ensemble	Version 1, k=3+ +ensemble	Version 1, k=5+ +ensemble	Version 2, k=3+ +ensemble	Version 2, k=5+ +ensemble	Version 3, k=3+ +ensemble	Version 3, k=5+ +ensemble
Left lung	96.38	96.75	96.78	96.73	96.77	96.36	96.33
Right lung	96.10	96.28	96.37	96.32	96.37	96.08	96.10
Spine area	92.82	93.10	93.24	92.90	93.16	92.88	92.71

Conclusions

This paper proposed a method that improves segmentation. The proposed method, due to the combination of the neighboring pixel analysis method and the ensemble method, was able to demonstrate its effectiveness on three data sets. This was confirmed by an improvement in the Dice-Sørensen coefficient. For the first set of data, it was possible to increase the result by 1.16 compared to the use of separate neural networks, for the second set of data, the improvement was 0.96, and for the last set it was 2.3. The improvement of the segmentation quality for the three data sets shows the universality of the proposed method. A comparison of the method of analyzing pixels outside the contour with the method of analyzing pixels outside the contour and inside the contour showed a slight advantage of the second one. Thus, even with a small advantage, the best option would be to use a method with pixel analysis both outside and inside. Despite the results obtained, there are further prospects for improving the proposed method. Further research may include using different neural network architectures, changing kernel shapes, and adjusting the neighboring pixel analysis logic.

References

1. Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2022). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523–3542. https://doi.org/10.1109/TPAMI.2021.3059968.
2. Zhu, C., Zhang, X., Li, Y., Qiu, L., Han, K., & Han, X. (2022). SharpContour: A contour-based boundary refinement approach for efficient and accurate instance segmentation. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (New Orleans, LA, USA), pp. 4382–4391. https://doi.org/10.1109/CVPR52688.2022.00435.
3. Aouat, S., Ait-hammi, I., & Hamouchene, I. (2021). A new approach for texture segmentation based on the Gray Level Co-occurrence Matrix. *Multimedia Tools and Applications*, vol. 80, pp. 24027–24052. https://doi.org/10.1007/s11042-021-10634-4.
4. Ma, D., Liao, Q., Chen, Z., Liao, R., & Ma, H. (2019). Adaptive local-fitting-based active contour model for medical image segmentation. *Signal Processing: Image Communication*, vol. 76, pp. 201–213. https://doi.org/10.1016/j.image.2019.05.006.
5. Judah, A., Hu, B., & Wang, J. (2014). An algorithm for boundary adjustment toward multi-scale adaptive segmentation of remotely sensed imagery. *Remote Sensing*, vol. 6, iss. 5, pp. 3583–3610. https://doi.org/10.3390/rs6053583.
6. Bransby, K. M., Bajaj, R., Ramasamy, A., Çap, M., Yap, N., Slabaugh, G., Bourantas, C., & Zhang, Q. (2024). POLYCORE: Polygon-based contour refinement for improved intravascular ultrasound segmentation. *Computers in Biology and Medicine*, vol. 182, article 109162. https://doi.org/10.1016/j.compbiomed.2024.109162.
7. Zia, H., Soomro, S., & Choi, K. N. (2024). Image segmentation using bias correction active contours. *IEEE Access*, vol. 12, pp. 60641–60655. https://doi.org/10.1109/ACCESS.2024.3391052.
8. Mohammed, A. & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University – Computer and Information Sciences*, vol. 35, iss. 2, pp. 757–774. https://doi.org/10.1016/j.jksuci.2023.01.014.
9. VinDr-SpineXR: An open dataset for spinal lesions detection and classification from radiographs. Vinbigdata: official site. Retrieved from https://vindr.ai/datasets/spinexr.
10. Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., Xue, Z., Palaniappan, K., Singh, R. K., Antani, S., Thoma, G., Wang, Y. X., Lu, P. X., & McDonald, C. J. (2014). Automatic tuberculosis screening using chest radiographs. *IEEE Transactions on Medical Imaging*, vol. 33, iss. 2, pp. 233–245. https://doi.org/10.1109/TMI.2013.2284099.
11. Candemir, S., Jaeger, S., Palaniappan, K., Musco, J. P., Singh, R. K., Xue, Z., Karargyris, A., Antani, S., Thoma, G., & McDonald, C. J. (2014). Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. *IEEE Transactions on Medical Imaging*, vol. 33, iss. 2, pp. 577–590. https://doi.org/10.1109/TMI.2013.2290491.

Received 26 September 2024

Адаптивне уточнення контуру сегментованого об’єкта на основі яскравості сусідніх пікселів із використанням ансамблевого методу

В. Д. Конюхов

Інститут енергетичних машин і систем ім. А. М. Підгорного НАН України,
61046, Україна, м. Харків, вул. Комунальників, 2/10

Підвищення точності алгоритмів комп'ютерного зору відіграє значну роль у сегментації медичних зображень, адже саме визначення меж об'єктів є складним завданням під час використання медичних зображень, а особливо рентгенівських знімків. Застосування рентгенівських знімків у сегментації є складним процесом, оскільки саме ці зображення можуть мати достатню кількість шумів й артефактів. Класичні методи сегментації стикаються з суттєвими викликами при сегментації рентгенівських зображень, де є об'єкти з нечіткими межами. Для вирішення таких завдань пропонується використовувати сегментацію за допомогою машинного навчання, а для підвищення точності визначення меж об'єктів необхідно застосовувати адаптивні підходи. У цій статті пропонується новий метод підвищення точності сегментації рентгенівських зображень, який аналізує сусідні пікселі кожного елемента контуру та, якщо потрібно, адаптивно змінює його форму, після чого комбінує всі передбачення за допомогою ансамблевого методу, що дає змогу покращити попередню версію контуру. Як демонструють дослідження, завдяки даному методу покращується якість сегментації зображень на трьох наборах даних із різною складністю структур. Для всіх трьох наборів отримано покращення точності меж об'єктів.

***Ключові слова**: машинне навчання, нейронні мережі, глибоке навчання, сегментація зображення, аналіз медичного зображення.*

Література

1. Minaee S., Boykov Y., Porikli F., Plaza A., Kehtarnavaz N., Terzopoulos D. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2022. Vol. 44. No. 7. P. 3523–3542. https://doi.org/10.1109/TPAMI.2021.3059968.
2. Zhu C., Zhang X., Li Y., Qiu L., Han K., Han X. SharpContour: A contour-based boundary refinement approach for efficient and accurate instance segmentation. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (New Orleans, LA, USA). 2022. P. 4382–4391. https://doi.org/10.1109/CVPR52688.2022.00435.
3. Aouat S., Ait-hammi I., Hamouchene I. A new approach for texture segmentation based on the Gray Level Co-occurrence Matrix. *Multimedia Tools and Applications*. 2021. Vol. 80. P. 24027–24052. https://doi.org/10.1007/s11042-021-10634-4.
4. Ma D., Liao Q., Chen Z., Liao R., Ma H. Adaptive local-fitting-based active contour model for medical image segmentation. *Signal Processing: Image Communication*. 2019. Vol. 76. P. 201–213. https://doi.org/10.1016/j.image.2019.05.006.
5. Judah A., Hu B., Wang J. An algorithm for boundary adjustment toward multi-scale adaptive segmentation of remotely sensed imagery. *Remote Sensing*. 2014. Vol. 6. Iss. 5. P. 3583–3610. https://doi.org/10.3390/rs6053583.
6. Bransby K. M., Bajaj R., Ramasamy A., Çap M., Yap N., Slabaugh G., Bourantas C., Zhang Q. POLYCORE: Polygon-based contour refinement for improved intravascular ultrasound segmentation. *Computers in Biology and Medicine*. 2024. Vol. 182. Article 109162. https://doi.org/10.1016/j.compbiomed.2024.109162.
7. Zia H., Soomro S., Choi K. N. Image segmentation using bias correction active contours. *IEEE Access*. 2024. Vol. 12. P. 60641–60655. https://doi.org/10.1109/ACCESS.2024.3391052.
8. Mohammed A., Kora R. A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University – Computer and Information Sciences*. 2023. Vol. 35. Iss. 2. P. 757–774. https://doi.org/10.1016/j.jksuci.2023.01.014.
9. VinDr-SpineXR: An open dataset for spinal lesions detection and classification from radiographs. Vinbigdata: official site. Retrieved from https://vindr.ai/datasets/spinexr.
10. Jaeger S., Karargyris A., Candemir S., Folio L., Siegelman J., Callaghan F., Xue Z., Palaniappan K., Singh R. K., Antani S., Thoma G., Wang Y. X., Lu P. X., McDonald C. J. Automatic tuberculosis screening using chest radiographs. *IEEE Transactions on Medical Imaging*. 2014. Vol. 33. Iss. 2. P. 233–245. https://doi.org/10.1109/TMI.2013.2284099.
11. Candemir S., Jaeger S., Palaniappan K., Musco J. P., Singh R. K., Xue Z., Karargyris A., Antani S., Thoma G., McDonald C. J. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. *IEEE Transactions on Medical Imaging*. 2014. Vol. 33. Iss. 2. P. 577–590. https://doi.org/10.1109/TMI.2013.2290491.