

Vladyslav KONIUKHOV

A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine, Kharkiv, Ukraine

IMPROVING THE SEGMENTATION OF THE VERTEBRAE USING A MULTI-STAGE MACHINE LEARNING ALGORITHM

*The health of the spine is an integral part of human health because the spine itself plays one of the key roles in human health, and diseases such as osteoporosis, vertebral injuries, herniated intervertebral discs, and other diseases can not only complicate a person's life but also have serious consequences. The use of X-ray images to diagnose spinal diseases plays a key role in diagnosis. Diagnosis of diseases with the help of X-rays is the most popular and cheapest option for patients to detect pathologies and diseases. **The subjects of this article** are algorithms for the segmentation of X-ray images of various qualities. **The aim** is to research the possibility of improving segmentation of vertebrae: Th8, Th9, Th10, Th11 using a multi-stage method of segmentation of the spine using machine learning to improve the accuracy of automation of vertebrae segmentation. **Task:** train a neural network that will segment the incoming X-ray image and produce a mask of the area of four vertebrae at the output; train a neural network that will segment each vertebra in the area found at the previous stage; cut out a section with one vertebra and train a neural network that will segment it; create an algorithm that, based on three previously trained neural networks, will segment vertebrae on an X-ray image. The following **methods** were used: a multi-stage approach using machine learning. The following **results** were obtained: thanks to segmentation in several stages, it was possible to reduce the region of interest, thereby removing unnecessary background when using segmentation. Using this algorithm for 48 vertebrae, an average improvement in segmentation accuracy of 4.83% was obtained. **Conclusions.** In this research, a multi-stage algorithm was proposed, and an improvement in the accuracy of segmentation of X-ray images in the lateral projection, namely the accuracy of all four vertebrae: Th8, Th9, Th10, Th11 - was obtained. The results demonstrate that the use of this method gives a better result than the usual segmentation of the input image.*

Keywords: artificial intelligence; machine learning; image recognition; neural network; image segmentation, computer vision.

1. Introduction

1.1. Motivation

The use of X-rays plays a key role in the diagnosis of spine diseases and pathologies. Since the use of this technology remains the most popular and cheapest way to obtain the necessary information about a patient's condition, automatic segmentation can be an important part of helping the physician. To reduce the burden on the doctor and improve the diagnosis process, an automation process is proposed, which makes it possible to identify pathologies, perform preoperative analysis, etc. However, due to the increasing use of X-ray images, there are also problems such as noise, artifacts, and incorrect exposure. All these factors can make the diagnostic process more difficult. It is for this purpose that it is proposed to consider and analyze the multi-stage method of segmentation of the spine region presented in this work, which includes the following vertebrae: Th8, Th9, Th10, and Th11.

This research used segmentation using neural networks, which nowadays have become a popular tool for solving such problems, however, despite their

advantages, they have some disadvantages. The main problem associated with using a convolutional neural network is the need for a large amount of data for training [1].

1.2. State of the Art

The use of classical segmentation methods prior to the advent of machine learning methods were critical to tasks related to medical images. A popular threshold segmentation method is the Otsu method [2]. Different algorithms have been built on this basis, such as the use of this method and multiple scaling [3]. The use of multiple scaling and the Otsu method was created to reduce the impact of noise and bad edges, resulting in a stable result. In [4], a multi-level approach for determining the interval iteration threshold was proposed based on the Otsu method. In the article [5], an effective method for determining the cancerous area using various types of segmentation of medical images was proposed. The cancer segmentation method based on fuzzy entropy with a threshold value of a set of levels improved the accuracy of cancer detection due to each segmentation of the medical image.



The interactive method, where the segmentation algorithm with the average shift is used in the first step, and the process of adaptive merging of regions based on the maximum similarity between regions takes place in the second step, was able to show a successful and effective method of separating bones from the background of the X-ray image due to manual interaction [6].

The study of neural networks has made it possible to create many algorithms capable of segmenting images. For example, in the article [7], the use of U-like architectures was investigated, and the article also noted that U-Net plays an important role in the segmentation of medical images due to its ability to scale and improve by adding or adjusting the number of layers or structural changes.

New models began to replace U-Net, which were created for special tasks. According to the test results, the XNet neural network could distinguish bone and soft tissue areas well [8].

Finally, to improve the existing algorithms, it was proposed to use not one neural network, but two or more. This method was first called an ensemble of neural networks [9].

Considering the search for vertebrae on magnetic resonance images, an article [10] developed an algorithm for detecting and determining vertebrae only on magnetic resonance imaging. This method uses the idea of finding the angles of vertebrae after applying a convolutional neural network. The authors of the study [11] used random forest classifiers and selection of vertebra angles. The use of multiparameter ensemble learning based on the super pixels proposed in the article [12] was performed on a trained model of only 6 images. This study was performed to segment lumbar vertebrae on magnetic resonance images. Despite these positive results, the use of the super pixel algorithm for X-ray images may yield slightly worse results than for magnetic resonance images. The reason lies in the nature of X-ray images, the presence of noise and artifacts, and poor image contrast, which can significantly impair the usefulness of this method. A previous study [13] proposed a two-stage method for segmenting lumbar vertebrae on computed tomography images by cutting out the area of interest. Here, U-Net is used in the first stage, and XUNet in the second. The experimental results indicate that the method involving cutting out part of the spine provides good quality vertebral segmentation.

In the paper [14], the authors used a single convolutional network to segment all partially or fully visible vertebrae, as well as to estimate up to three vertebral center locations. The authors claim that their patch-based approach solves the problem of processing completely different volumes of computed tomography images.

The use of YOLO and 2D-U-Net was proposed in

a previous study [15]. The authors propose a two-stage algorithm for individual vertebral labeling and object detection followed by segmentation for computed tomography scan images. They used the YOLO algorithm to detect the vertebra, which was then cut out and used to train the 2D-U-Net.

A previous study [16] proposed an automatic segmentation algorithm that identifies anatomical structures using magnetic resonance imaging of the lumbar spine. The authors used the Residual U-Net network and developed a new rotation matrix approach for detecting disc protrusions.

A previous study [17] proposed segmentation of five lumbar vertebrae using a three-step approach. The main stages of which are: localization of the spine, segmentation of the lumbar region and fine-tuning of the segmentation.

1.3. Objectives and approach

In the conditions of the existence of a large number of X-ray machines and different experiences of a radiologist, it became possible to encounter significant difficulties in the diagnosis of diseases, as well as in treatment. Noises of various origins, artifacts, incorrect exposure, low-quality equipment, and unsatisfactory patient positioning can significantly affect the quality of the final result. It was for this that the main goal was set aside to develop an algorithm that can segment X-ray images regardless of their quality. For this purpose, it is proposed to consider the following tasks:

1. Develop an algorithm for vertebrae segmentation on X-ray images to improve the accuracy of the segmentation itself.
2. Prove the sense of using an approach with several regions of interest and multiple neural networks.
3. Analyze the causes of unsatisfactory cases.

To solve these tasks, it is necessary to develop a complex approach that consists of several stages. Based on the tasks, the following approaches were proposed:

1. Do not use preprocessing of images. Using such methods as a Gaussian filter or median filter cannot always improve image quality and remove noise, but, on the contrary, makes segmentation more difficult.
2. Use the modified Fcn8Resnet50 network to segment images of different qualities. Three such neural networks were implemented for image segmentation in three stages.

2. Materials and methods of research

2.1. Three-stage segmentation algorithm

The proposed method comprises three steps. To perform which we need to use a neural network.

Fcn8Resnet50, which is a fully convolutional network and has a good architecture for image segmentation tasks, was chosen as the neural network. Thanks to the use of a fully convolutional network with the idea of end-to-end learning, it is possible to directly match input images with masks, which provides more accurate segmentation [18]. Images from open sources were used for training [19]. Three samples of 182 images with a size of 512 by 512 pixels in grayscale were created. Each training sample consisted of 170 images, and the test sample consisted of 12. Augmentation was applied to all three trainings:

- 1) random rotation in degrees [-15, 15];
- 2) random shift in percentage vertically and horizontally [-10, 10];
- 3) random scaling in percent [0.8, 1.2];
- 4) random change in brightness [0.8, 1.2];
- 5) random change in contrast [0.75, 1.5].

To compare two segmentation masks, the Jaccard coefficient was calculated as follows:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}. \quad (1)$$

In the first stage, a neural network was trained for spine segmentation, which consists of four vertebrae. This is visually indicated in Figure 1 as Step 1. The result of the first step is represented as a mask of the first step in Figure 1.

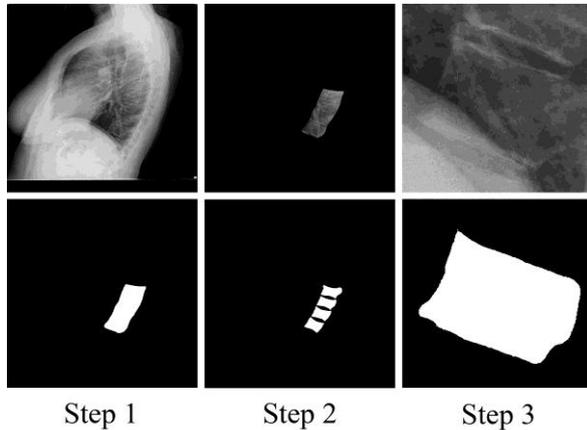


Fig. 1. Example image and mask for each training step

In the second stage, the prediction mask obtained in the previous stage must be taken, applied to the input image, and cut out. Because of this operation, only the region of the four vertebrae and a black background remain in the image. Figure 1 illustrates this in the second step. After the necessary part of the image is extracted using the second neural network, the four vertebrae are segmented as follows: Th8, Th9, Th10, and

Th11. The result of segmentation is shown in Figure 1, where it corresponds to the mask of the second step.

In the last - third step, the necessary vertebra obtained in the second step is selected, after which a bounding rectangle is drawn around the prediction mask of this vertebra. This rectangle is superimposed on the input image and cuts out only the part that is included in this rectangle. Everything outside the rectangle becomes black. Visually, the vertebra from the last stage is drawn as step 3 in Figure 1, with only one nuance-scaling performed for better understanding. Without scaling, the prediction mask from the third stage in Figure 2 can be visually observed.

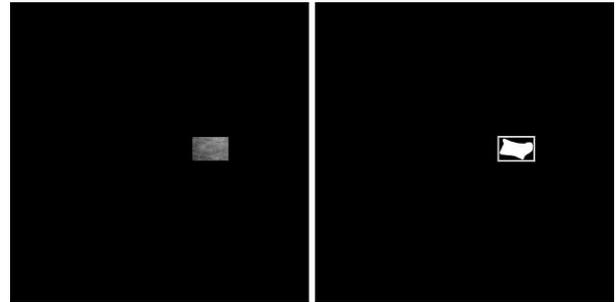


Fig. 2. Using a bounding rectangle in the third step

Figure 2 shows a bounding rectangle that was later cut out of the input image and on which segmentation occurred.

The main task of the third step was to determine the required size of the rectangle to be cut. The accuracy of vertebral segmentation and the number of positive results of improving vertebral segmentation depend on the dimensions of width and height.

First, the largest dimensions of the bounding rectangles were chosen for all four vertebrae. The largest width of one of the four bounding rectangles has the following form:

$$W_b = \max(W_{th8}, W_{th9}, W_{th10}, W_{th11}), \quad (2)$$

where W_b – the greatest width.

The largest height of one of the four bounding rectangles has the following form:

$$H_b = \max(H_{th8}, H_{th9}, H_{th10}, H_{th11}), \quad (3)$$

where H_b – the highest height.

The following formula was used to obtain the required width of the rectangle for cutting:

$$W_c = W_b \times K_w, \quad (4)$$

where W_c – width of the rectangle to cut out;

K_w – width coefficient of the rectangle.

The following formula was used to obtain the height of the cutting rectangle:

$$H_c = H_b \times K_h, \quad (5)$$

where H_c – height of the rectangle to cut;

K_h – height coefficient of the rectangle.

After obtaining the cutting dimensions, it became possible to obtain an image where the size of the cut rectangle is $W_c \times H_c$. After obtaining a rectangle of the required size at the third stage, the rectangle can be cut out, and the next step is to segment the small vertebra. Using this image, we obtain the prediction mask $Pred_c$ for it. At the final stage, when the prediction mask for the second stage (vertebra mask from the input image) and the prediction mask for the third stage (vertebra mask from the cut rectangle), it remains to determine whether it is possible to improve accuracy using this approach or not. For this purpose, both prediction masks are compared with a manually created mask.

The next step is to obtain the Jaccard coefficient for the third-stage prediction and the ground truth mask of the second stage using formula (1):

$$J_c = J(Pred_c, Label). \quad (6)$$

The Jaccard coefficient for the second stage prediction was also obtained as follows:

$$J_{reg} = J(Pred_{reg}, Label). \quad (7)$$

The indicated coefficients of height K_h and width K_w were obtained from an experimental method. The final values of the coefficients were selected as follows – by enumerating all values in the specified range where $T_{min} = 1$ and $T_{max} = 2$ with a step $\Delta = 0.01$, for each case the number of positive results was calculated and those coefficients thanks to which it was possible to achieve the maximum number of positive results and became the resulting values. Obtain the number of positive results for test images where the Jaccard coefficient J_c more than J_{reg} :

$$Pos_n = \sum_{j=1}^{12} \sum_{i=1}^4 = \begin{cases} 1, & J_c > J_{reg}; \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

As a result, it was obtained $K_w = 1.32$, $K_h = 1.76$, the maximum number of positive vertebral accuracy improvement results for these coefficients was 43 out of 48.

For greater visibility of the comparison of masks, we did not use the pure Jaccard coefficient, but we used its percentage version, which has the following form:

$$J_p = J \times 100, \quad (9)$$

where J_p – Jaccard coefficient in percentage;
 J – Jaccard coefficient.

2.2. Analysis of unsatisfactory cases

It is not always possible to obtain excellent results in all cases. Then, we must analyze the failed results and understand the reason for this. In this case, two options were considered, one of which would allow us to understand the reason for undesirable results.

The contrast of the X-ray image is the main characteristic of the segmentation quality analysis. Depending on the difference in its values, a good or bad result can be obtained. In the first case, the image contrast was analyzed, which served as a measure of the standard deviation of the brightness of pixels near the average value of the brightness of the entire image. The formula for the average value of brightness looks like this:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i, \quad (10)$$

where N – the total number of pixels in the image;
 x_i – brightness i -th pixel.

Using formula (11), we obtain the standard deviation:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}. \quad (11)$$

With large values of this characteristic, the objects in the image are clearer and easier to identify.

Using the brightness range allows us to determine the difference between the maximum and minimum brightness values of the image. A large value of this characteristic indicates a large number of pixels with different brightness values, which, in turn, indicates a good image. Conversely, a small value indicates a low contrast value, which may be caused by poor exposure. Brightness range formula:

$$B = \max(I) - \min(I), \quad (12)$$

where $\max(I)$ – the maximum brightness value of a pixel in the image;

$\min(I)$ – the minimum brightness value of a pixel in the image.

3. Results

Applying the idea of using multiple regions of interest to segment a single object significantly improved

the Jaccard coefficient for the prediction mask. The use of the approach with the selection of the region from the four vertebrae made it possible to single out only the necessary part of the image, thereby all unnecessary information that was placed outside the zone of this part was painted in black. Thus, the removal of redundant information allowed the neural network to focus on only the necessary parameters, which made it impossible to obtain false predictions. Table 1 lists the similarities between the prediction and hand-drawn masks.

Table 1
Jaccard's percentage coefficient for the second stage

No.	Vertebrae			
	Th8	Th9	Th10	Th11
1	83.38	77.83	79.26	75.14
2	84.86	87.60	88.03	76.07
3	77.18	74.38	83.49	84.99
4	87.31	84.41	87.78	83.03
5	87.24	88.81	85.30	87.26
6	79.52	90.77	86.00	87.02
7	86.47	82.51	86.97	85.14
8	87.93	83.88	83.29	87.63
9	81.10	86.20	88.29	83.53
10	71.60	76.94	84.35	78.35
11	82.79	91.68	89.59	74.41
12	87.92	74.86	85.37	77.29

The average value of the Jaccard percentage coefficient for all vertebrae in Table 1 was 83.43%.

To improve the accuracy of segmentation of individual vertebra, a third stage was used, and the similarity of this stage with the original mask of the second stage is shown in Table 2.

Table 2
Jaccard percentage coefficient for the third stage

No.	Vertebrae			
	Th8	Th9	Th10	Th11
1	81.76	79.55	89.50	82.27
2	85.08	92.96	91.40	83.84
3	81.76	85.04	89.62	83.05
4	89.51	90.56	89.90	85.28
5	92.46	90.74	89.78	82.97
6	91.67	90.80	88.94	89.17
7	87.22	88.36	91.68	85.87
8	88.88	90.81	87.38	91.31
9	86.86	94.17	89.52	86.32
10	88.89	90.60	91.60	87.68
11	89.26	93.59	88.17	88.52
12	82.63	82.65	88.45	84.50

For the data presented in Table 2, the average value was 87.96%. However, if we analyze all 48 vertebrae listed in the tables, we can see that thanks to the use of this algorithm, only 43 vertebrae were improved. The

other five vertebrae showed no improvement in matching accuracy. The following cases: No. 1 Th8; No. 3 Th11; No. 5 Th11; No. 11 Th10; No. 12 Th8. In order to obtain improvement using this algorithm for 43 vertebrae, it was sufficient to use universal coefficients of width and height. However, universal coefficients are not suitable for five cases. The coefficients for five unsatisfactory cases were determined experimentally, and all of them are listed in Table 3.

Table 3
Separate coefficients for unsatisfactory cases

No.	Vertebra	K _w	K _h	J _p
1	Th8	1.62	1.9	85.73
3	Th11	1.6	1.7	88.54
5	Th11	1	1.9	87.28
11	Th10	1.9	1.5	91.27
12	Th8	1.8	1.7	89.54

If we calculate the average accuracy for the data in Table 2, replacing the five negative results with those obtained in Table 3, and selecting coefficients for each case, then the value for all 48 cases is 88.26%. This is a good result that made it possible to improve the segmentation of each vertebra by reducing the influence of artifacts and noise on the segmentation process.

The use of universal coefficients of height and width to cut the region of interest helped obtain 89.58% of positive cases. However, failed cases must be analyzed to understand poor segmentation. To perform this task, it was suggested to analyze image characteristics: image contrast, brightness range, and visual inspection. The values for these properties are given in Table 4.

Table 4
Comparison of image characteristics in unsatisfactory cases

No.	Properties	Vertebrae			
		Th8	Th9	Th10	Th11
1	Contrast	19.73	19.32	18.36	25.79
	Brightness range	105	115	133	161
3	Contrast	24.58	27.54	25.47	25.28
	Brightness range	120	140	131	147
5	Contrast	12.53	13.18	12.70	19.26
	Brightness range	80	85	86	116
11	Contrast	10.98	12.23	15.67	13.29
	Brightness range	81	81	104	88
12	Contrast	13.25	14.49	14.46	16.70
	Brightness range	78	94	87	100

4. Discussion

In the article [20], it was indicated that the trained neural network with the indication of the required region of interest showed higher productivity and surpassed traditional convolutional neural networks and even regression models taking into account clinical parameters. The use of the method proposed in this article included two regions of interest: the spine and vertebra. This made it possible to use two neural networks in parallel to improve the final segmentation results of the vertebra.

Using the first region of interest, which cut out the area with four vertebrae, made it possible to segment the vertebrae in a small area without redundant data. In this area, there were only vertebrae. The second region of interest made it possible to search for a vertebra in an area where there is only one vertebra. This approach, using three stages instead of two, improved the Jaccard percentage by 4.54% and improved 89.58% over the time. Undoubtedly, this approach unequivocally guarantees improvement due to its use because it focuses only on the necessary information.

This study and the research conducted in [17] should be compared in more detail due to the partial similarity of some aspects. The first difference is the use of preprocessing methods in that article when no preprocessing is used. The reason for not including preprocessing in this paper was its experimental instability, which showed that the Gaussian method and median filter only worsened the quality. The next difference is the input image. The image used by Kim et al. (2021) [17] is immediately a cropped X-ray image, most of which is the desired part. In this study, a full image is used in a lateral projection. The proposed method is closer to practical application due to the use of the resulting images from an X-ray machine without processing. The method of cutting out the spine of the authors of the article with which the comparison is being made assumes a square, when in this research it was proven that using a rectangle where the width is greater than the height is the best option.

In this case, training was performed using the Fcn8Resnet50 network, which was not designed specifically to segment medical images. The network used in this study [17] was developed specifically for medical cases. This finding emphasizes the successful use of the Fcn8Resnet50 network for solving multidisciplinary problems and guarantees that not only networks developed specifically for the medical field can provide excellent results.

Using universal coefficients of width and height, it was possible to obtain an improvement not for all vertebrae, but only for 43 out of 48. To achieve the maximum number of positive cases, we must separately de-

termine the coefficients for bad images. The segmentation accuracy was improved only after assigning an eigenvalue to each of the five cases. Analyzing the cause of the five unsatisfactory cases, the key characteristics of any X-ray image were used: the brightness and contrast range of the image. By studying the data in Table 4, we can understand that the contrast of the image, which is given in the form of a standard deviation of brightness, is within the norm, which in turn indicates a normal exposure. In this case, the accuracy of the segmentation depends on the quality of the image, where it should have the following features: not blurred boundaries of the object over which the segmentation is performed, and low noise (quantum and structural). The size of the object over which segmentation is performed in relation to the size of the entire image.

The improvement for the five negative cases was obtained by separately defining the width and height coefficients. In these cases, the negative results could be caused by a non-ideal coincidence of the contours of the vertebrae and the size of the rectangle for cutting. In addition, in some cases, the presence of noise at the boundaries of the contour and rectangle can lead to unsatisfactory results. As shown in Table 3, the increase in the width coefficient in four out of five cases made it possible to transfer them to the positive class.

5. Conclusions

The comparison of the prediction masks at several stages made it possible to understand that the segmentation of vertebrae using a multi-stage algorithm is associated with better segmentation accuracy. Due to the use of several regions of interest, it was possible to minimize the influence of redundant and low-quality information on segmentation. Segmentation of vertebrae in several stages significantly improves the accuracy of finding the desired vertebra.

Due to the analysis of negative cases, it became clear that the use of universal coefficients of width and height for cutting a rectangle at the third stage of the algorithm is not sufficient to satisfy the requirement of 100% positive results. Separate coefficients were obtained for the negative cases, which improved the accuracy for the negative cases and improved the overall vertebral function.

In general, the algorithm was developed to perform segmentation of the X-ray image in three steps to improve the segmentation of the vertebrae. The implementation of this algorithm made it possible to obtain a positive result in 89.58% of cases, which is a good result. Further research on this approach and its application together with an ensemble of neural networks can lead to an increase in the number of positive cases that already use only universal coefficients of width and height.

Conflict of Interest

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

Financing

This study was conducted without financial support.

Data Availability

The manuscript contains no relevant data.

Use of Artificial Intelligence

The author confirm that they did not use artificial intelligence technologies in their work.

References

1. Jamaludin, A., Kadir, T., & Zisserman, A. SpineNet: Automated classification and evidence visualization in spinal MRIs. *Medical Image Analysis*, 2017, no. 41, pp. 63-73. DOI: 10.1016/j.media.2017.07.002.
2. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, vol. 9, no. 1, pp. 62-66. DOI: 10.1109/TSMC.1979.4310076.
3. Feng, Y., Zhao, H., Li, X., Zhang, X. & Li, H. A multi-scale 3D Otsu thresholding algorithm for medical image segmentation. *Digital Signal Processing*, 2017, vol. 60, pp. 186-199. DOI: 10.1016/j.dsp.2016.08.003.
4. Feng, Y., Liu, W., Zhang, X., Liu, Z., Liu, Y., & Wang, G. An interval iteration based multilevel thresholding algorithm for brain MR image segmentation. *Entropy*, 2021, vol. 23, no. 11, pp. 22-51. DOI: 10.3390/e23111429.
5. Maalood, I., Al-Salhi, Y., & Lu, S. Thresholding for medical image segmentation for cancer using fuzzy entropy with level set algorithm. *Open Medicine*, 2018, vol. 13, pp. 374-383. DOI: 10.1515/med-2018-0056.
6. Stolojescu-Crisan, C., & Holban, S. An interactive X-ray image segmentation technique for bone extraction. *International Work-Conference on Bioinformatics and Biomedical Engineering*, Granada, Spain, 2014, vol. 2, pp. 1164-1171.
7. Yao, W., Bai, J., Liao, W., Chen, Y., Liu, M., & Xie, Y. From CNN to Transformer: A review of medical image segmentation models. *Journal of Imaging Informatics in Medicine*, 2023, pp. 1-18. DOI: 10.48550/arXiv.2308.05305.
8. Aylett-Bullock, J., Cuesta-Lázaro, C., & Quera-Bofarull, A. XNet: A convolutional neural network (CNN) implementation for medical X-ray image segmentation suitable for small datasets. *Proc. SPIE 10953, Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, San Diego, California, United States, 2019, vol. 10953, pp. 453-463. DOI: 10.1117/12.2512451.
9. Hansen, L. K. & Salamon, P. Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, vol. 12, iss. 10, pp. 993-1001. DOI: 10.1109/34.58871.
10. Windsor, R., Jamaludin, A., Kadir, T., & Zisserman, A. A convolutional approach to vertebrae detection and labelling in whole spine MRI. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Lima, Peru, 2020, pp. 712-722. DOI: 10.48550/arXiv.2007.02606.
11. Ebrahimi, S., Gajny, L., Skalli, W., & Angelini, E. D. Vertebral corners detection on sagittal X-rays based on shape modelling, random forest classifiers and dedicated visual features. *Computational Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2019, vol. 7, iss. 2, pp. 132-144. DOI: 10.1080/21681163.2018.1463174.
12. Gaonkar, B., Xia, Y., Villaroman, D. S., Ko, A., Attiah, M., Beckett, J. S., & Macyszyn, L. Multi-parameter ensemble learning for automated vertebral body segmentation in heterogeneously acquired clinical MR images. *IEEE Journal of Translational Engineering in Health and Medicine*, 2017, vol. 5, pp. 1-12. DOI: 10.1109/JTEHM.2017.2717982.
13. Lu, H., Li, M., Yu, K., Zhang, Y., & Yu, L. Lumbar spine segmentation method based on deep learning. *Journal of Applied Clinical Medical Physics*, 2023, vol. 24, iss. 6, pp. 1-12. DOI: 10.1002/acm2.13996.
14. Xiong, X., Graves, S.A., Gross, B.A., Buatti, J. M., Beichel, R. R. Lumbar and Thoracic Vertebrae Segmentation in CT Scans Using a 3D Multi-Object Localization and Segmentation CNN. *Tomography*, 2024, vol. 10, no. 5, pp. 738-760. DOI: 10.3390/tomography10050057.
15. Da Mutten, R., Zanier, O., Theiler, S., Ryu, S. J., Regli, L., Serra, C., & Staartjes, V. E. Whole Spine Segmentation Using Object Detection and Semantic Segmentation. *Neurospine*, 2024, vol. 21, iss. 1, pp. 57-67. DOI: 10.14245/ns.2347178.589.
16. Liang, Y. W., Fang, Y. T., Lin, T. C., Yang, C. R., Chang, C. C., Chang, H. K., Ko, C. C., Tu, T. H., Fay, L. Y., Wu, J. C., Huang, W. C., Hu, H. W., Chen, Y. Y., & Kuo, C. H. The Quantitative Evaluation of Automatic Segmentation in Lumbar Magnetic Resonance Images. *Neurospine*, 2024, vol. 21, iss. 2, pp. 665-675. DOI: 10.14245/ns.2448060.030.
17. Kim, K. C., Cho, H. C., Jang, T. J., Choi, J. M., & Seo, J. K. Automatic detection and segmentation of

lumbar vertebrae from X-ray images for compression fracture evaluation. *Computer Methods and Programs in Biomedicine*, 2021, vol. 200, pp. 1-18. DOI: 10.1016/j.cmpb.2020.105833.

18. Shelhamer, E., Long, J., & Darrell, T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 4, pp. 640-651. DOI: 10.1109/TPAMI.2016.2572683.

19. Vindr.ai. (n.d.). *Vindr.ai Datasets: SpineXR*. Available at: <https://vindr.ai/datasets/spinexr> (Accessed: 16.06.2024).

20. Wang, H., Zhang, T., Cheung, K. M., & Shea, G. K. Application of deep learning upon spinal radiographs to predict progression in adolescent idiopathic scoliosis at first clinic visit. *EClinicalMedicine*, 2021, vol. 42, pp. 1-9. DOI: 10.1016/j.eclinm.2021.101220.

Received 24.07.2024, Accepted 18.11.2024

ПОКРАЩЕННЯ СЕГМЕНТАЦІЇ ХРЕБЦІВ ЗА ДОПОМОГОЮ БАГАТОЕТАПНОГО АЛГОРИТМУ МАШИННОГО НАВЧАННЯ

В. Д. Конюхов

Здоров'я хребта є невід'ємною складовою здоров'я людини тому що саме хребет відіграє одну із ключових ролей у здоров'ї людини, а такі хвороби як – остеопороз, травми хребців, грижі міжхребцевих дисків та інші захворювання можуть не тільки ускладнювати життя людини, а й мати серйозні наслідки. Використання рентгенівських знімків для діагностування захворювань в області хребта відіграє ключову роль в діагностиці. Діагностування захворювань за допомогою рентгену є найпопулярнішою та найдешевшою можливістю для пацієнтів виявити патології та захворювання. **Предметом** вивчення в статті є алгоритми сегментації рентгенівських зображень різної якості. **Метою** є дослідження можливості покращення сегментації хребців: Th8, Th9, Th10, Th11 за допомогою багатоетапного методу сегментації хребта використовуючи машинне навчання з метою покращення точності автоматизації сегментації хребців. **Завдання:** провести навчання нейронної мережі, яка буде проводити сегментацію вхідного рентгенівського знімка і на виході видавати маску ділянки з чотирьох хребців; провести навчання нейронної мережі яка буде проводити сегментацію кожного хребця в знайденій ділянці на попередньому етапі; вирізати ділянку з одним хребцем та виконати навчання нейронної мережі яка буде проводити його сегментування; створити алгоритм який на основі попередньо навчених трьох нейронних мереж буде робити сегментацію хребців на рентгенівському знімку. Використовувались наступні **методи:** багатоетапний підхід з використанням машинного навчання. Було отримано такі **результати:** завдяки сегментації в декілька етапів стало можливим зменшити регіон інтересу, тим самим видалити непотрібний фон при використанні сегментації. Використовуючи даний алгоритм для 48 хребців було отримано середнє покращення точності сегментації на 4,83 %. **Висновки.** В цьому дослідженні було запропоновано багатоетапний алгоритм, завдяки якому було отримано покращення точності сегментації рентгенівських зображень в боковій проекції, а саме точності усіх чотирьох хребці: Th8, Th9, Th10, Th11. Було доведено, що використання даного методу дає кращий результат ніж звичайна сегментація на вхідному зображенні.

Ключові слова: штучний інтелект; машинне навчання; розпізнавання зображень; нейронна мережа; сегментація зображень, комп'ютерний зір.

Владислав Конюхов – аспірант, Інститут енергетичних машин і систем ім. А. М. Підгорного НАН України, Харків, Україна.

Vladyslav Koniukhov – PhD Student of A. Pidhornyi Institute of Power Machines and Systems of NAS of Ukraine, Kharkiv, Ukraine,
e-mail: rigelllll@gmail.com, ORCID: 0009-0007-0256-1388.