UDC 004.8

USING AN ENSEMBLE OF NEURAL NETWORKS FOR DETERMINING THE DIAGNOSTIC PARAMETERS OF THE VERTEBRAE

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

Anatolii Pidhornyi Institute of Mechanical Engineering Problems of NAS of Ukraine, 2/10, Pozharskyi str., Kharkiv, 61046, Ukraine

Artificial intelligence opens up great prospects in many areas of human activity, primarily in medicine. One of the priority directions of using artificial intelligence in this field is the segmentation of medical images for the purpose of automatic diagnosis of common diseases. The application of neural network approaches to image analysis of medical images is becoming an increasingly promising direction in the field of medical diagnostics. In particular, this paper investigates the possibility of using an ensemble of neural networks for diagnosing osteoporosis. To achieve this goal, a study was conducted on the possibility of using machine learning methods to segment and determine the shape and size of certain vertebrae: Th8, Th9, Th10, Th11 of a human vertebra on X-ray images obtained in real conditions. Each network is configured and tested on different sets of medical images. Then, the two best networks were selected according to the accuracy and efficiency of the segmentation. One of the main results of the study was the selection of the two best neural networks that provide the most accurate segmentation of vertebrae. Next, the ensemble method was applied, based on the averaging of the predictions of the selected networks. This approach made it possible to improve the overall accuracy of determining the diagnostic parameters of the spine. The obtained results emphasize the effectiveness of using an ensemble of neural networks in the context of medical segmentation. Ensembles provide more stable and accurate predictions by reducing the impact of random errors of individual networks. Ensemble predictions of these networks lead to a statistically significant improvement in results compared to individual approaches. This is an important step in the direction of creating reliable systems of automated diagnostics capable of helping doctors in conducting more accurate and operative analyses.

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

Introduction

Health is the most important thing for every person because it affects all aspects of our existence, and forehanded highly qualified medical diagnosis is a guarantee of good health. One of the most promising areas, which can provide such a fruitful assistance for doctors, is the artificial intelligence. Artificial intelligence opens up great prospects in many fields, first of all in medicine. One of the priority areas of using artificial intelligence is the segmentation of medical images.

In particular, the spine is an important component of the human body, and its diagnosis with the use of various devices is important for the treatment of various diseases and illnesses.

In this paper, we will study the ways how neural networks can be used for automatic search for vertebrae in X-ray images. The main methods and algorithms for searching for four vertebrae will be analyzed. Finally, the obtained results will allow to increase the speed and to improve the accuracy of the analysis of the spine and vertebrae on medical images.

The spine role

Diseases of the spine are widespread in the world. According to WHO, osteoporosis ranks 4th in the list of causes of disability and death [1, 2]. There are a large number of diseases of the spine that can cause significant pain and lead to various symptoms. Back pain is the most common complaint that doctors hear. It can be caused by various injuries, muscle overload, sitting posture, etc. Osteochondrosis, herniated intervertebral discs, spondylosis, scoliosis, spondylolisthesis are only a small number of diseases of the spine.

To solve these medical problems, artificial intelligence algorithms [3], which are able to process large volumes of data faster and more accurately, are used. The main important applications of machine learning in medicine can be identified as follows: disease diagnosis, disease forecasting, medical image analysis, medical data management, medical robotics. This list demonstrates that the use of machine learning

in the medical field has great prospects and is an integral part of progress. Thanks to this, we can improve the diagnosis, processing, and management of medical data. Although on the other hand it should be understood that the use of machine learning in medicine should have a high degree of in-depth research, verification, and standardization.

The aim and objectives of the research

The main aim of this paper is the comparison of eight different neural networks and investigation of the method of simultaneous use of the best of them. We provide research and application of machine learning methods for segmentation and form determination of four vertebrae: Th8, Th9, Th10, Th11 on X-ray images.

Used neural networks

Eight different models of neural networks were selected for the study: Fcn8, Fcn32, Mobile-Net_Segnet, MobileNet_Unet, Segnet, U-Net, Vgg_Segnet, Vgg_Unet. These networks are used for various purposes, although some were specifically designed for use in medicine, such as U-Net [4]. The choice of neural networks was made according to the specific task and amount of data. For medical applications where accuracy, safety and reliability are most important, different combinations of models will be used for the best accuracy. Therefore, networks with an architecture based on a convolutional encoder-decoder structure are better suited for this. Each of the used models has its own advantages and disadvantages, so not only conventional architectures will be used, but also modifications and combinations of these architectures.

Data collection and preparation

The main part of all training is the data on which the model will later be trained. The final result and success will depend on their quality and quantity. At the first stage, X-ray images are collected. In the second stage, we need to remove duplicates, and bring all images to the same size, in some cases it will be necessary to crop or fill the image. At the third stage, mask coding takes place, where each vertebra pixel is assigned a value of 1, and each background pixel is assigned a value of 0. At the last stage, the data is divided into two groups – a training sample and a test sample. In our case a sample of 183 images with a size of 256×256 pixels was used for training of all models. All images were grayscale.

Fig. 1 demonstrates the fragment of the training sample that was used for the training of eight neural networks for each of the vertebra: Th8, Th9, Th10, Th11.


Fig. 1. The fragment of training sample
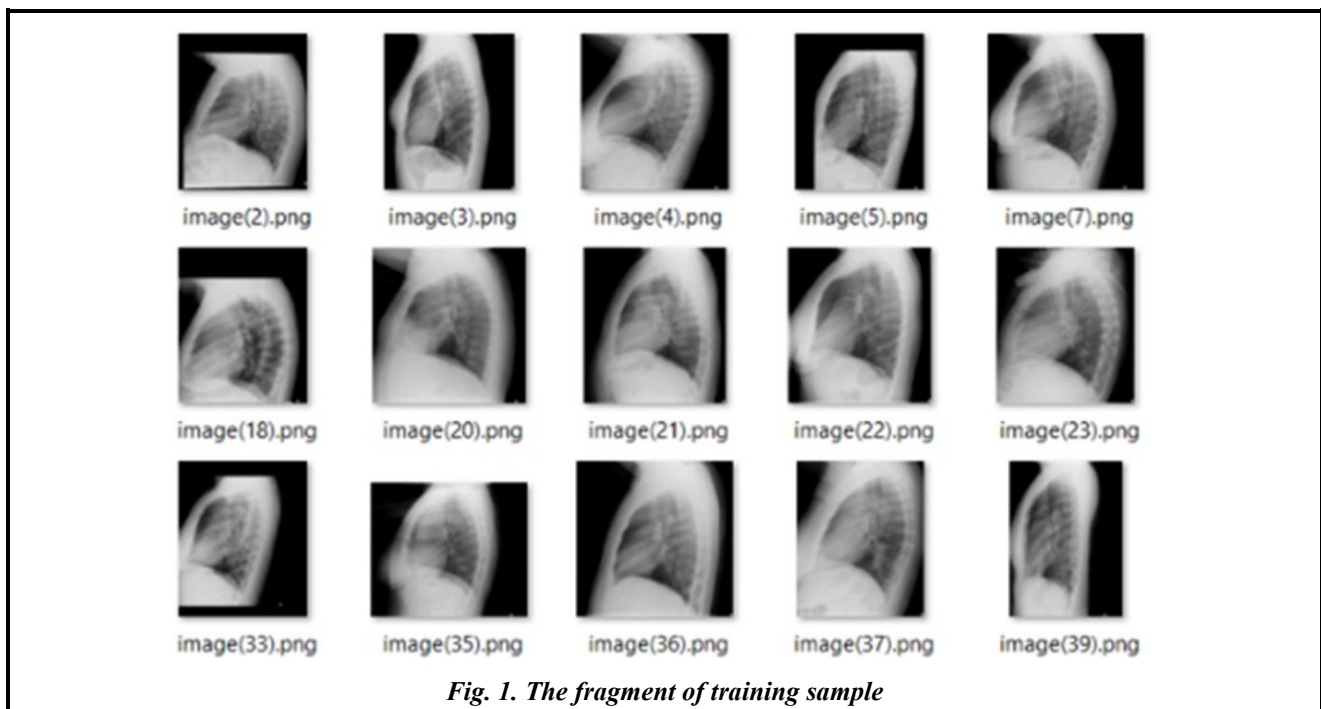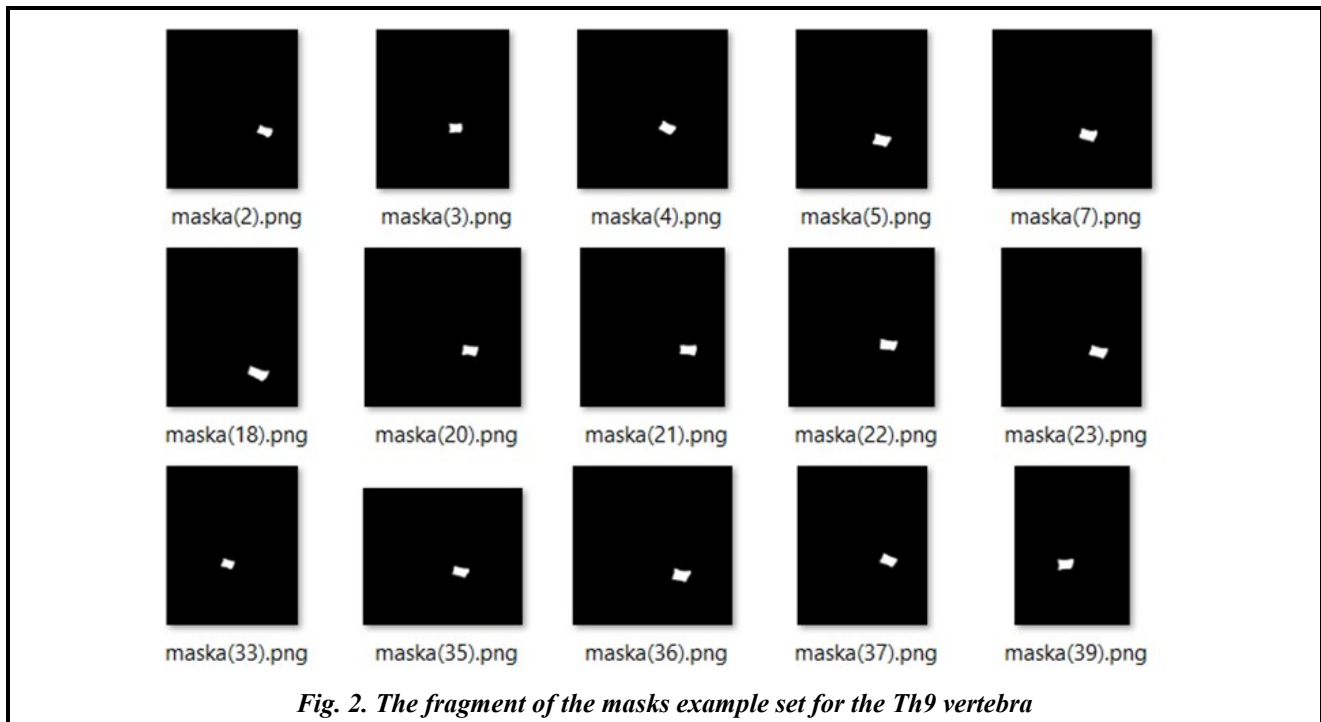
Fig. 2. The fragment of the masks example set for the Th9 vertebra

Moreover, a variety of masks were created for each vertebra. An example of such masks for the Th9 vertebra is shown in Fig. 2.

For a more thorough examination, an example of an enlarged fragment of the spine and the corresponding mask for vertebra Th11 will be shown in Fig. 3.
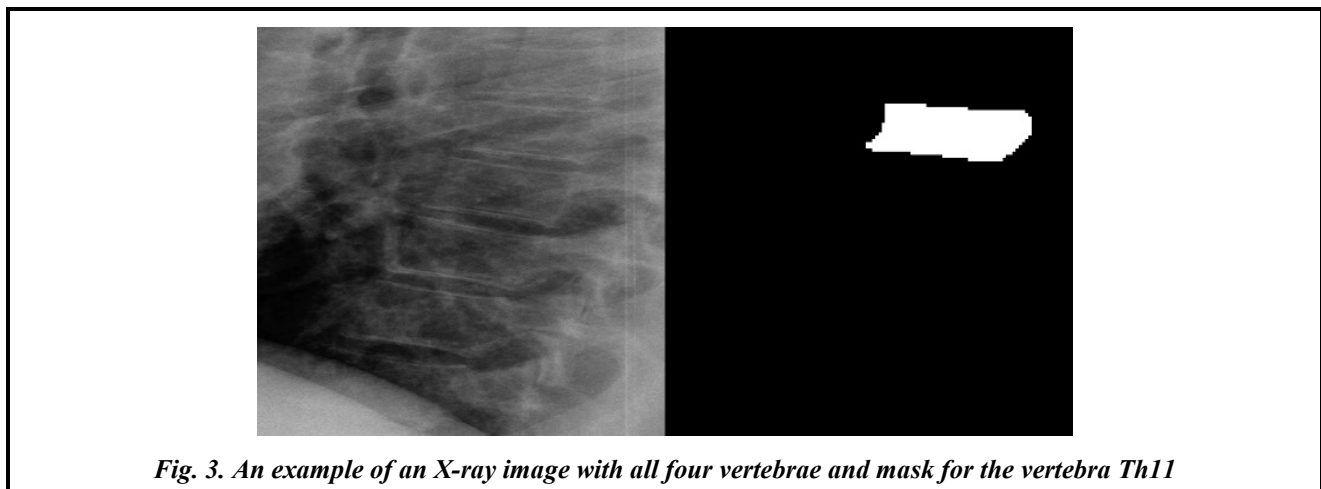
Fig. 3. An example of an X-ray image with all four vertebrae and mask for the vertebra Th11

Results of separate predictions

After the data preparation we provided the training of all the models. We use the same settings for each network.

The training process was divided into four stages – first, training was carried out on the Th8 vertebra, then Th9, Th10, and Th11. At each of the four stages, 10 trainings of each model were performed for each vertebra, after which the best result in each model was selected and tabulated.

To estimate the prediction quality, we used the following parameter: IoU (Intersection over Union). To determine this parameter, the concept of the correctly found part of the desired object (TP – True Positive), which is formally the intersection of the object and its predicate, is introduced:

$$TP = Object \cap Prediction.$$

The concept of the set of all points that relate to the desired object both in reality and from the point of view of prediction is also introduced. This set is the union of the desired object and the predicted image

$$Object \cup Prediction = TP + FP + FN ,$$

where FP (false positive) is the incorrectly predicted pixels, and FN (false negative) is the set of correct pixels that were not predicted.

The ratio of these values determines a parameter IoU (Intersection over Union) that is widely used to determine the quality of image recognition

$$IoU = \frac{TP}{TP + FP + FN} .$$

The results of the averaged over 10 predictions for all vertebrae are presented in Table 1. According to the training results, such networks as Vgg_Unet and MobileNet_Unet performed better than others. Thus, these networks will be selected at the second stage of prediction for their joint use.

Table 1. Prediction results

Vertebrae	Fcn8	Fcn32	MobileNet_Segnet	MobileNet_Unet	Segnet	Unet	Vgg_Segnet	Vgg_Unet
Th8	91.82	89.72	93.70	95.51	79.32	91.87	86.24	95.09
Th9	91.36	86.18	93.77	95.26	88.01	92.71	54.42	93.39
Th10	91.26	86.55	93.04	94.33	90.86	92.55	82.88	95.68
Th11	91.46	88.97	93.35	95.89	73.92	87.30	83.92	95.45

Simultaneous usage of definite chosen neural networks

The previous analysis made it possible to implement the main idea of this paper. This idea consists in using not one network, but a whole ensemble of networks. After training each of the networks on the same data set, a certain number, for example, four (or two as in our case) out of eight that give the best result, are selected from the entire set. After that, only these selected networks are used when working with real images. Moreover, they are used to define objects independently and in parallel. After finding an object by each of these networks the procedure of averaging of the found image over this set is carried out. This averaging was done as follows.

At the first step, the so-called best network (exemplary), which demonstrated the best results on the entire sample of test images, is selected. The results of this network are the reference and the predictions of other networks are then compared with them. In our case, such a network turned out to be MobileNet_Unet.

In the next step, the predictions of another network, Vgg_Unet, which showed the second-best result, are compared with the predictions of the first network. Processing of the results showed that if the predictions of these two networks do not differ too much (less than a certain threshold), then taking them together gives predictions that are better than each of the networks separately. If their predictions differ by more than this threshold, then it is better to limit ourselves to the prediction of the first of the networks.

For a specific example, two predictions were used: $Pred_{Vgg}$, which was obtained from the Vgg_Unet network, and $Pred_{Mobile}$, obtained from MobileNet_Unet network. These masks are binary fields in which one indicates that the pixel belongs to a vertebra, and zero indicates that the pixel does not belong to a vertebra.

The joint use of predictions from these two networks is described by the relation

$$Pred_\Sigma = \begin{cases} \frac{1}{2}\left(Pred_{Mobile} + Pred_{Vgg}\right), if \ IoU_{pred} > r \\ Pred_{Mobile}, if \ IoU_{pred} \leq r \end{cases},$$

where r is the threshold above which the results of both networks are taken into account, and the ratio IoU_{pred} in this formula is the ratio of intersection to union for the masks of each network.

$$IoU_{pred} = \frac{Pred_{Mobile} \cap Pred_{Vgg}}{Pred_{Mobile} \cup Pred_{Vgg}} .$$

Informally, this algorithm can be described as follows. For each pixel, the number of times it belonged to the sought image according to the predictions of each of the networks was determined. Then this number was compared with a certain, predetermined limit, and if it exceeded this limit, then this pixel was marked as belonging to the sought object. As a result of this approach, it was possible to improve the quality of predicting the location of objects, their shape and size.

Table 2. Averaging result

Neural networks	Th8	Th9	Th10	Th11
Vgg_Unet	85.02	86.10	85.16	88.50
MobileNet_Unet	89.65	87.41	94.07	91.74
Result	91.17	90.75	94.07	92.24

In order to clearly demonstrate the result of the algorithm application, in Table 2 you can see the result of its work. Two networks, MobileNet_Unet and Vgg_Unet, were used and one image was predicted for each vertebra, after which the IoU metric was used to measure accuracy. At the end, the algorithm was applied and the above metric was also used for the resulting prediction.

Finally, it can be concluded that the use of an ensemble of networks definitely improves the quality of solving the problem, and allows to obtain better results than with each individual neural network.

Conclusions

This paper highlighted the relevance of using machine learning for segmentation of vertebrae. The use of this technology for medical purposes is undoubtedly a good step towards automating the diagnosis and treatment of spinal diseases.

The research carried out in the paper showed that in the process of using eight neural networks to determine the shape and size of vertebrae, two networks, which are Vgg_Unet and MobileNet_Unet, performed best. These networks are modifications of the U-Net network, which was specifically designed for medical research. Based on this, we see the advantage of a neural network specifically designed for medical imaging over other networks.

At the second step the simultaneous usage of these selected networks was proposed. In addition, the algorithm of prediction that uses the independent predictions of each of these networks was proposed and applied.

As the result, it was shown that the use of such an ensemble of networks definitely improves the quality of solving the problem, and allows to obtain better results than with each individual neural network.

Finally, this paper highlighted the relevance of using machine learning for vertebral segmentation. The use of this technology for medical purposes is undoubtedly a good step towards automating the diagnosis and treatment of spinal diseases. Thanks to the use of machine learning, it is possible to automatically detect and highlight the necessary vertebrae in medical images, thereby ensuring good processing speed and accuracy.

References

1. Parashchuk, O. (2010). *Mizhnarodnyi den borotby z osteoporozom: razom do peremohy nad khvoroboiu!* [International day of fighting osteoporosis: Together to triumph over the disease!]. *Novyny medytsyny ta farmatsii – News of Medicine and Formation*, no. 18 (341), 6 p. [The application date is August 2, 2023] (in Ukrainian). Available at: http://urgent.mif-ua.com/archive/article/14406.
2. (2019). 80% UK adults experience back pain. Living Room Health Regenerative Treatments: official site. [The application date is August 2, 2023]. Available at: https://www.theregenerativeclinic.co.uk/news/80-uk-adults-experience-back-pain/.
3. (2023). What is machine learning? IBM: official site. [The application date is August 2, 2023]. Available at: https://www.ibm.com/topics/machine-learning.
4. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Lecture Notes in Computer Science. Cham: Springer, vol. 9351, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
5. Simonyan, K. & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations (ICLR 2015)*: Conference papers, pp. 1–14.
6. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. 9 p. https://doi.org/10.48550/arXiv.1704.04861.

Використання ансамблю нейронних мереж для визначення діагностичних параметрів хребців

В. Д. Конюхов

Інститут проблем машинобудування ім. А. М. Підгорного НАН України,
61046, Україна, м. Харків, вул. Пожарського, 2/10

Штучний інтелект відкриває великі перспективи в багатьох сферах діяльності людини, насамперед у медицині. Одним із пріоритетних напрямів використання штучного інтелекту в цій галузі є сегментація медичних зображень задля автоматичного діагностування поширених захворювань. У цій роботі досліджується, зокрема, можливість застосування ансамблю нейронних мереж для діагностування остеопорозу. Для досягнення вказаної мети вивчено можливості використання методів машинного навчання для сегментації і визначення форми й розмірів певних хребців (Th8, Th9, Th10, Th11 хребця людини) на рентгенівських зображеннях, отриманих у реальних умовах. Досліджено низку нейронних мереж, які застосовуються в обробці зображень, й запропоновано алгоритм одночасного їх використання для якісного визначення діагностичних параметрів.

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

Література

1. Паращук О. Міжнародний день боротьби з остеопорозом: разом до перемоги над хворобою! *Новини медицини та фармації*. 2010. № 18 (341). 6 с. [Дата звернення – 2 серпня 2023 р.]. URL: http://urgent.mif-ua.com/archive/article/14406.
2. 80% UK adults experience back pain. Living Room Health Regenerative Treatments: official site. 2019. [The application date is August 2, 2023]. URL: https://www.theregenerativeclinic.co.uk/news/80-uk-adults-experience-back-pain/.
3. What is machine learning? IBM: official site. 2023. [The application date is August 2, 2023]. URL: https://www.ibm.com/topics/machine-learning.
4. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Lecture Notes in Computer Science. Cham: Springer, 2015. Vol. 9351. P. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
5. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations (ICLR 2015)*: Conference papers. 2015. P. 1–14.
6. Howard A. G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. MobileNets: Efficient convolutional neural networks for mobile vision applications. 2017. 9 p. https://doi.org/10.48550/arXiv.1704.04861.