

УДК (UDC) 004.8

Конюхов Владислав Дмитрович	<i>аспірант Інститут проблем машинобудування ім. А.М.Підгорного НАН України адреса: вул. Пожарського 2/10, Харків, 61046, Україна e-mail: riggelllll@gmail.com https://orcid.org/0009-0007-0256-1388</i>
Моргун Олег Миколайович	<i>к.ф.-м.н., директор ТОВ «Лабораторія рентгенівської медичної техніки» адреса: вул. Достоевського, 1, м. Харків, 61102, Україна e-mail: lrmt@ukr.net https://orcid.org/0009-0005-6157-9110</i>
Немченко Костянтин Едуардович	<i>д.ф.-м.н., завідувач кафедри Харківський національний університет імені В. Н. Каразіна майдан Свободи, 4 м. Харків, 61022, Україна e-mail: nemchenko@karazin.ua https://orcid.org/0000-0002-0734-942X</i>

Багаторазове навчання нейронних мереж для автоматичної сегментації області хребта

Актуальність. Профілактичні та діагностичні дослідження наявності кісткових захворювань, потребують морфометричних досліджень рентгенівських знімків відділу грудної клітини. В теперішній час для розв'язування таких задач все частіше використовують методи штучного інтелекту. Зокрема в цій роботі досліджена можливість застосування штучного інтелекту при сегментації медичних зображень з ціллю автоматичного діагностування захворювань кісткової системи людини. Основні труднощі цієї задачі пов'язані з тим, що реальні рентгенівські зображення мають граничні характеристики якості, наприклад, за параметром відношення сигналу до шуму або контрасту. З цієї причини застосування стандартних методів розпізнавання зображень або автоматичної діагностики стає неможливим. Ці труднощі привели до того, що в теперішній час існує достатньо велика кількість робіт в цій галузі, але результати більшості з них демонструють недостатні для практичного використання результати. Наведена в цій роботі оригінальна методика використання ансамблю нейронних мереж дозволила узагальнити здобуті дотепер результати та поєднати переваги інших підходів.

Мета. Дослідити можливість використання штучного інтелекту в сегментації медичних зображень з метою автоматичної діагностики захворювань кісткової системи людини.

Методи дослідження. У даному дослідженні використовувався ансамблевий метод сегментації рентгенівських зображень. Основу навчальних даних було створено на базі рентгенівських знімків узятих з відкритих джерел. Всього кількість зображень становила 183 знімків. Початкові дані було модифіковано згідно з вимогами необхідними для навчання моделей. Всі зображення були переведені до градації сірого та змінені до розмірів 256x256 пікселів.

Результати. Завдяки використанню цього методу в двох тестових випадках було отримано покращення точності з 0,543 до 0,820 для першого знімку та з 0,725 до 0,923 для другого знімку.

Висновки. Було запропоновано і досліджено застосування методики використання ансамблю нейронних мереж, що багаторазово навчаються, для автоматичної сегментації певної області хребта, а саме ділянки хребта Th8-Th11. Застосування цього методу дозволило отримати більш стабільні та точні передбачення для шуканих ділянок хребта, навіть для знімків з високим рівнем шуму.

Ключові слова: штучний інтелект, машинне навчання, розпізнавання зображень, нейронна мережа, ансамбль нейронних мереж, морфометрія.

Як цитувати: Конюхов В. Д., Моргун О. М., Немченко К.Е. Багаторазове навчання нейронних мереж для автоматичної сегментації області хребта. *Вісник Харківського національного університету імені В.Н.Каразіна, серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»*. 2024. вип. 62. С.37-44.
<https://doi.org/10.26565/2304-6201-2024-62-04>

How to quote: V.D. Koniukhov, O.M. Morgun and K.E. Nemchenko, "Multiple training of neural networks for automatic spine segmentation". *Bulletin of V.N. Karazin Kharkiv National University, series Mathematical modelling. Information technology. Automated control systems*, vol. 62, pp. 37 - 44, 2024. [In Ukrainian] <https://doi.org/10.26565/2304-6201-2024-62-04>

1 Вступ

Діагностика та оцінка наявності кісткових захворювань, зокрема компресійних переломів, потребує морфометричних досліджень бічних рентгенівських знімків відділу грудної клітини (ВГК) [1], які вимагають від лікаря-рентгенолога дуже тривалої і трудомісткої роботи. Для полегшення цієї роботи, тобто автоматичної сегментації за допомогою комп'ютерної програми,

раніше розроблялася значна кількість різноманітних алгоритмів [2 – 4]. Проте труднощі реалізації цих алгоритмів пов'язані з низкою причин. При рентгенологічних дослідженнях повинен виконуватися принцип ALARA (*As Low As Reasonably Achievable*), тобто доза опромінення пацієнта має бути настільки низькою, наскільки це можливо. Це призводить до низького відношення сигнал/шум, де під сигналом розуміється радіаційний контраст, а під шумом квантові та структурні шуми. Структурний шум – це шум, який пов'язаний з накладенням на область хребта інших анатомічних органів при проектуванні тривимірного об'єкта (людського тіла) на двовимірний приймач рентгенівського зображення. Це проектування призводить також до розмиття границь зображення хребта, що відповідно погіршує його виявлення та сегментацію. У разі стандартного підходу до застосування, навіть спеціалізованих для медицини нейронних мереж (наприклад U-Net або аналогів) [5], виникають проблеми при сегментації реальних зашумлених і розмитих рентгенівських зображень бічної проекції ВГК. Для вирішення задачі автоматичної сегментації ділянки хребта навіть у таких зашумлених і розмитих рентгенівських знімках у цій статті запропоновано метод багаторазового навчання нейронних мереж.

Останнім часом завдяки прогресу в розвитку комп'ютерної техніки та алгоритмів штучного інтелекту з'явилася велика кількість досліджень з використанням згорткових нейронних мереж для автоматичної сегментації хребта [6 – 8]. Практично всі вони пов'язані з розробкою алгоритмів удосконалення стандартного застосування нейронних мереж до зображень хребта. До таких алгоритмів можна віднести гібридні методи, які поєднують раніше розроблені алгоритми машинного навчання та використання згорткових нейронних мереж з глибоким навчанням [7], або використання багатозадачних мультимодальних нейронних мереж [6].

У цій статті пропонується і досліджується застосування методики використання ансамблю нейронних мереж, що багаторазово навчаються, для автоматичної сегментації певної області хребта, а саме ділянки хребта Th8-Th11.

2 Методи

Задача автоматичної сегментації ділянки хребта з хребцями Th8-Th11 вирішувалася у два етапи, що дозволило досягти достатньо високого рівня визначення шуканої області.

2.1 Перший етап

На першому етапі, первинного передбачення шуканої області, досліджувалися повноформатні рентгенівські знімки бічної проекції ВГК людини. Процедура застосування нейронних мереж для визначення даної частини хребта була стандартною. Спочатку вибирався набір зображень за якими проводилося навчання нейронних мереж. В якості таких наборів, використовувалися загальнодоступні набори, які знаходяться у мережі Інтернет [10]. Потім вручну створювалися маски областей, тобто областей які містять шукану ділянку з хребцями Th8-Th11. Приклади вихідних зображень та відповідні маски показані на Рис. 2.1. У нашому випадку навчання всіх моделей нейронних мереж проводилося на наборі зі 183 знімків, які масштабувалися до розміру 256*256 пікселів. Усі зображення були виконані в градаціях сірого в діапазоні інтенсивності сигналу 0-255.

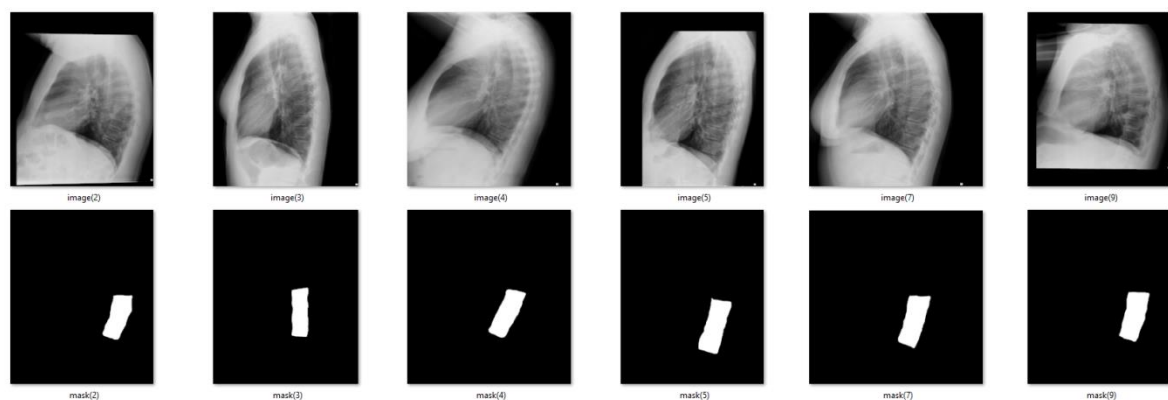


Рис.2.1 Приклади вхідних зображень та відповідні маски

Навчання здійснювалося на наступних нейронних мережах: FCN8, FCN32, MobileNet_Segnet, U-Net, MobileNet_Unet, VGG_Segnet. При цьому параметри навчання вибиралися однакові. Типовий результат передбачення - вихідний рентгенівський знімок, відповідна йому маска і результат визначення ділянки хребта, що шукається, показано на Рис.2.2.

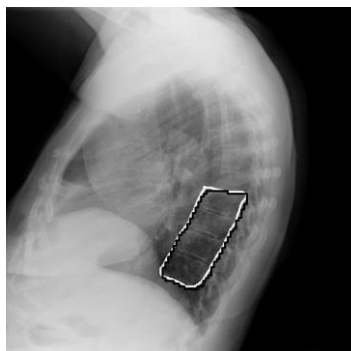


Рис.2.2 Приклад вихідного рентгенівського знімку, відповідна йому маска шуканої області (білий контур) та результат передбачення (чорний контур)

Відповідний цій моделі нейронної мережі графік втрат, а також графік зростання якості визначення шуканої області представлені на Рис.2.3.

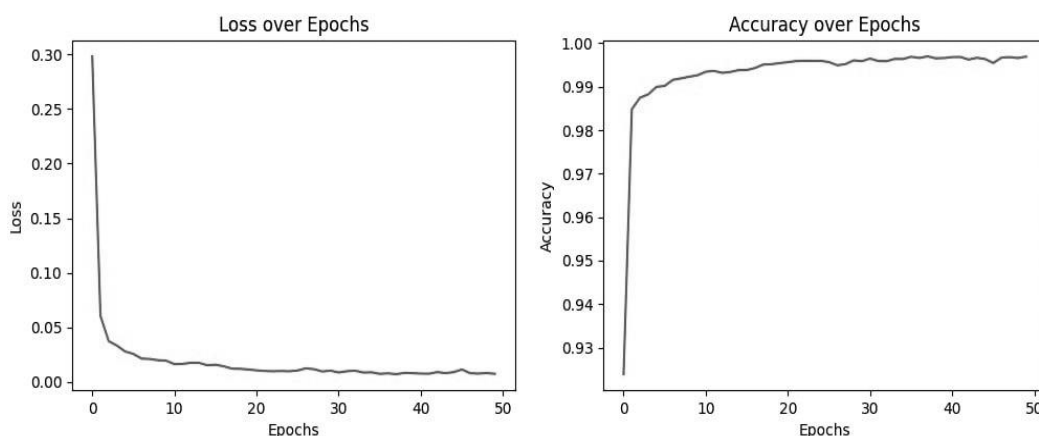


Рис.2.3 Залежність від кількості epoch втрат (а) та якості передбачення зображення (б)

Кількісне порівняння передбачень, отриманих за допомогою навчених моделей нейронних мереж та людини, здійснювалося за допомогою коефіцієнта Dice [9].

Для підвищення точності передбачення масок (а отже підвищення коефіцієнта Dice) було запропоновано навчати нейронну мережу 10 разів, а потім визначені області усереднити за всіма визначеннями.

2.2 Другий етап

Для покращення якості сегментації, на другому етапі було запропоновано використовувати ряд нейронних мереж ще раз, але з меншим розміром досліджуваної області рентгенівського знімку, де здійснюється сегментація.

Місце знаходження цієї області визначається за результатами першого етапу, в такий спосіб. Як цю зменшену область вибираємо прямокутник, що включає хребці один вище (Th7) і один нижче (Th12) шуканої ділянки хребта, що включає хребці Th8-Th11. Приклади рентгенівських знімків і відповідних їм, створених вручну масок, за якими здійснювалося навчання нейронних мереж, показано на Рис.2.4.

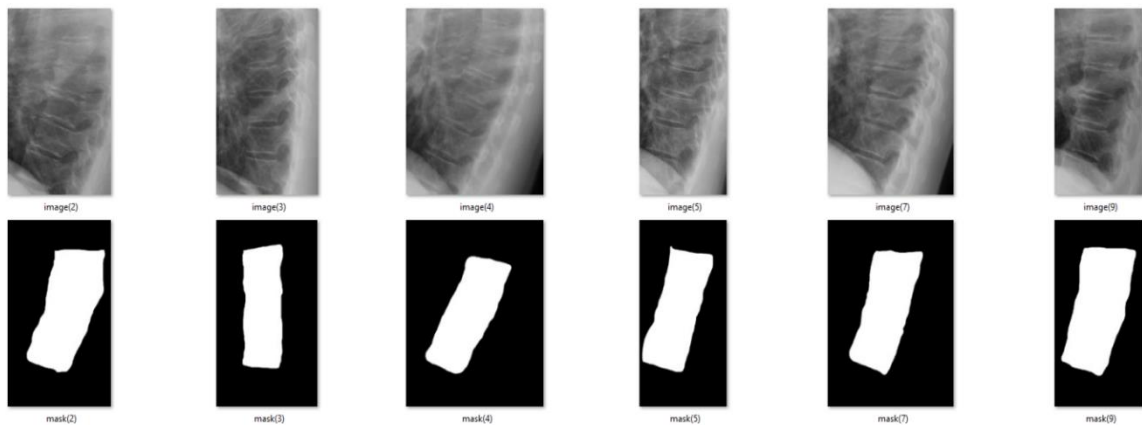


Рис.2.4 Приклади фрагментів вихідних зображень та відповідні маски шуканої області

Координати центру передбачуваної області зображення (X_c, Y_c) визначалися за допомогою передбаченої маски (отриманої на першому етапі) за формулою:

$$X_c = \frac{1}{2}(X_{min_{max}}) \quad (2.1)$$

та

$$Y_c = \frac{1}{2}(Y_{min_{max}}) \quad (2.2)$$

де X_{max} , X_{min} , Y_{max} , та Y_{min} – координати крайніх точок у зображенні передбаченої області.

Далі навколо цього центру вибиралася вирізана з повного рентгенівського зображення прямокутна ділянка з розмірами $L_0 \times H_0$, які визначаються за формулою:

$$L_0 = 0.33 \cdot L, \quad (2.3)$$

та

$$H_0 = 0.5 \cdot H \quad (2.4)$$

де L і H - повна ширина та висота вхідного рентгенівського зображення. Коефіцієнти 0.33 і 0.5 вибиралися з аналізу 183 рентгенівських зображень так, щоб у вирізану область вмістилася ділянка хребта з хребцями Th7-Th12.

3 Експерименти та результати

Після проведення тренувань, проводилося тестування на іншому наборі зображень, що складаються з 58 рентгенівських знімків бічної проекції ВГК, взятих з інтернету [10], які отримані в різних умовах і на різних рентгенівських апаратах. При цьому слід зазначити, що попереднього відбору зображення за якістю не проводилося, а досліджувалися всі зображення, які були в цих базах.

Результати передбачення масок ділянки хребта, що включає хребці Th8-Th11, для двох різних навчань представлені на Рис.3.1. Для демонстрації ефективності багаторазового навчання вибрано знімки з високим рівнем шуму і розмитими межами хребта.

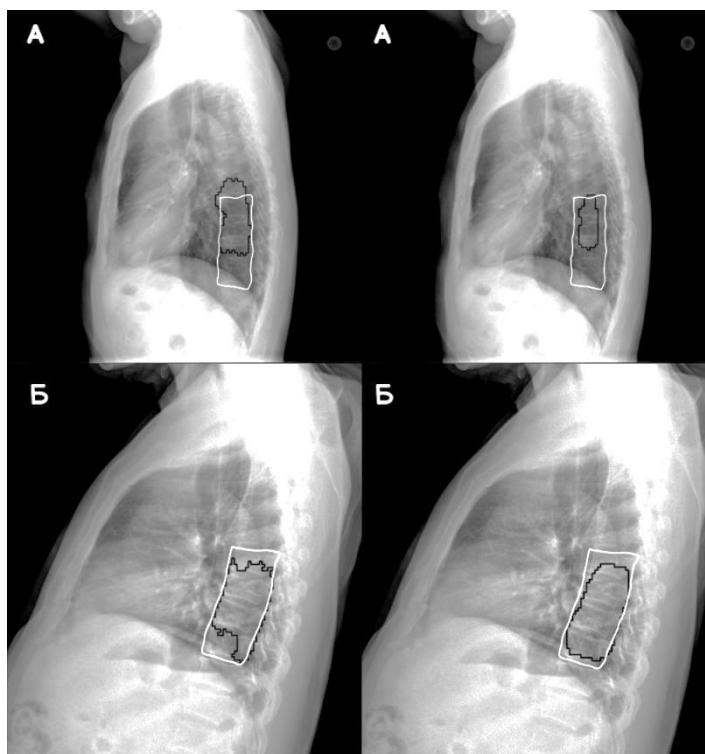


Рис.3.1 Приклади вихідних зображень, відповідних масок (білий контур) та передбачених областей (чорний контур) для двох різних навчань. А - зображення 14 з бази [10], Б - зображення 23 з бази [10]

Кількісні результати передбачень нейронних мереж, що окремо навчаються, а також усереднені по десяти навчаннях, наведені на Рис.3.2. Як видно з гістограм, наведених на Рис.3.2, окремі навчання можуть давати кращі результати за коефіцієнтом Дісе для якогось конкретного рентгенівського зображення, але для іншого рентгенівського зображення це окреме навчання дає гірший результат порівняно з усередненням. На Рис.3.2, також наводяться результати усереднення передбачень двох нейронних мереж. Результати передбачення інших нейронних мереж (FCN8, FCN32, U-net, VGG_Segnet) Рис.3.2 не наводяться, оскільки вони значно поступалися нейронним мережам MobileNet_Segnet і MobileNet_Unet.

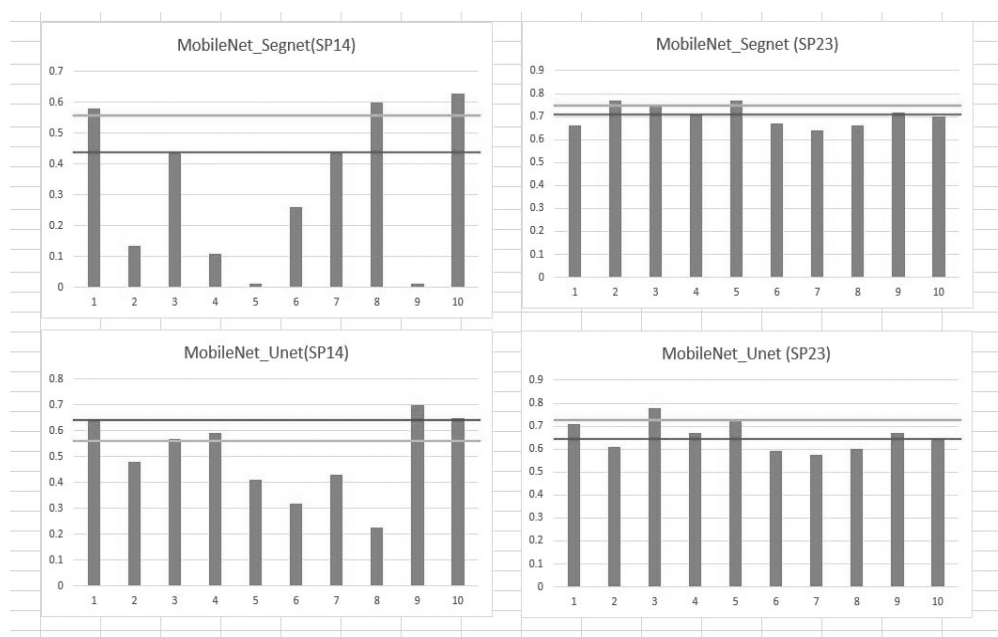


Рис.3.2 Кількісні результати десяти передбачень нейронних мереж, що окремо навчаються, а також усереднені по десяти навчаннях для двох зображень. Темна лінія – результат усереднення для певної мережі. Світла лінія – результат усереднення по обом мережам.

Однак, навіть результати цих найкращих нейронних мереж демонструють відносно низьку якість сегментації. Це зумовлено, як високим рівнем шуму, малим радіаційним контрастом, розмитими межами хребта, так й малим розміром шуканого зображення, проти повного зображення бічної проекції ВГК людини. Тому, як зазначалося вище, для поліпшення якості сегментації, на другому етапі, пропонувалося використовувати ряд нейронних мереж ще раз, але з меншим розміром досліджуваної області, де здійснюється сегментація.

Аналогічно першому етапу після проведення тренувань проводилося тестування на тому ж наборі зображень, що складаються з тих же 58 рентгеновських знімків бічної проекції ВГК, але зі зменшеною областю досліджень. Як і в першому етапі, для передбачення маски хребта, що включає хребці Th8-Th11, застосовувалися ті ж шість нейронних мереж з десятикратним навчанням і усередненням передбачених масок. Найкращі результати (за коефіцієнтом Dice) показали усереднення двох нейронних мереж MobileNet_Segnet та VGG_Segnet. Передбачені області для тих самих випадків, як і на Рис.5 показано на Рис.3.3. Коефіцієнт Dice цих двох випадків зріс з 0,543 до 0,820 для знімка № 14 і з 0,725 до 0,923 для знімка № 23.

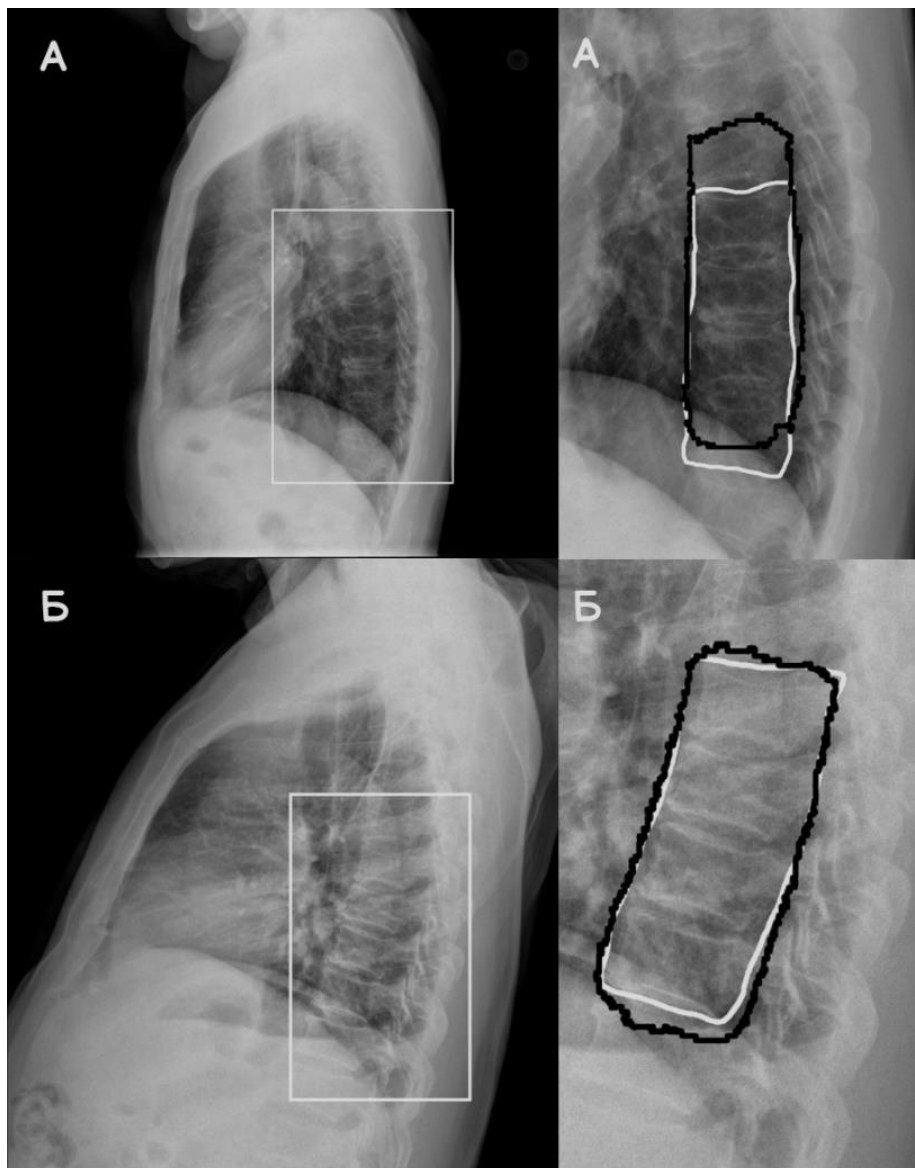


Рис.3.3 Приклади вихідних зображень, відповідних масок (білий контур) та передбачених областей (чорний контур). А - зображення 14 з бази [10], Б - зображення 23 з бази [10]

4 Висновки

Запропоновано метод підвищення точності передбачення нейронних мереж за рахунок двоетапного і багаторазового навчання низки нейронних мереж. Цей метод дозволив домогтися стабільності і точності передбачень шуканих ділянок хребта, навіть для знімків з високим рівнем шуму, низьким контрастом і розмитими межами шуканих об'єктів. Запропоновану методику можливо застосовувати як для дослідження інших органів людини, так і в не медичних застосуваннях, де зустрічаються неякісні зображення.

REFERENCES

1. G. Guglielmi, D. Diacinti, C. van Kuijk, F. Aparisi, C. Krestan, J. E. Adams, and T. M. Link, "Vertebral morphometric: current methods and recent advances," *European Radiology*, p. 14, 2008. https://link.springer.com/article/10.1007/s00330-008-0899-8?error=cookies_not_supported&code=d445f129-d830-4a13-9590-5989bcf5141c
2. S. Li and J. Yao, *Spinal Imaging and Image Analysis*, vol. 18, p. 507, 2015. <https://link.springer.com/book/10.1007/978-3-319-12508-4>
3. S. Ebrahimi, L. Gajny, W. Skalli, and E. Angelini, "Vertebral corners detection on sagittal X-rays based on shape modelling, random forest classifiers and dedicated visual features," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 7, no. 2, pp. 132-144, 2018. https://hal.science/hal-02181802/file/IBHGC_CMBBEIV_2018_Ebrahimi.pdf
4. L. R. Long and G. R. Thoma, "Segmentation and feature extraction of cervical spine x-ray images," in *Proc. SPIE 3661, Medical Imaging 1999: Image Processing*, May 1999. <https://ui.adsabs.harvard.edu/abs/1999SPIE.3661.1037L/abstract>
5. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015. <https://arxiv.org/abs/1505.04597>
6. S. M. M. R. Al Arif, "Fully automatic image analysis framework for cervical vertebra in X-ray images," Doctoral thesis, City, University of London, 2018. <https://pubmed.ncbi.nlm.nih.gov/29477438/>
7. K. C. Kim, H. C. Cho, T. J. Jang, J. M. Choi, and J. K. Seo, "Automatic detection and segmentation of lumbar vertebrae from X-ray images for compression fracture evaluation," *arXiv:1904.07624v1 [physics.med-ph]*, Apr. 2019. <https://arxiv.org/abs/1904.07624>
8. Y. Chen, Y. Mo, A. Readie, G. Ligozio, T. Coroller, and B. W. Papie, "VertXNet: Automatic Segmentation and Identification of Lumbar and Cervical Vertebrae from Spinal X-ray Images," *arXiv:2207.05476v1 [eess.IV]*, Jul. 2022. <https://arxiv.org/abs/2207.05476>
9. L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, p. 297-302, 1945. <https://esajournals.onlinelibrary.wiley.com/doi/10.2307/1932409>
10. "Vindr.ai Datasets: SpineXR." [Online]. Available: <https://vindr.ai/datasets/spinexr>. [Accessed: 16-Jun-2024].

Koniukhov Vladyslav Dmytrovych *PhD student
Anatolii Pidhorneyi Institute of Mechanical Engineering Problems of NAS of Ukraine,
2/10, Pozharskyi str., Kharkiv, 61046, Ukraine*

Morgun Oleg Mykolayovych *Ph.D., director
"Laboratory of X-ray Medical Equipment" LTD
address: Dostoevsky str. 1, Kharkiv, 61102, Ukraine*

Nemchenko Kostyantyn Eduardovych *D. of Sc., head of the department
V. N. Karazin Kharkiv National University
Maydan Svobody, 4, Kharkiv, 61022, Ukraine*

Multiple training of neural networks for automatic spine segmentation

Actuality

Preventive and diagnostic studies of the presence of bone diseases require morphometric studies of X-ray images of the chest area. Nowadays, artificial intelligence methods are increasingly being used to solve such problems. The main difficulties of this task are related to the fact that X-ray images have quality limitations, for example, in terms of signal-to-noise ratio or contrast. For this reason, the application of standard methods of image recognition or automatic diagnosis becomes impossible. These difficulties have led to the fact that there is currently a fairly large number of works in this field, but the results of most of them are insufficient for practical use.

Goal

Investigate the possibility of using artificial intelligence in the segmentation of medical images for the purpose of automatic diagnosis of diseases of the human bone system.

Research methods

The ensemble method of X-ray image segmentation has been used in the study. The baseline of training data was created on the basis of X-ray images taken from open sources. The total number of images is 183. The initial data was modified according to the requirements necessary for model training. All images were converted to grayscale and resized to 256x256 pixels.

Results

Using this method in the two test cases resulted in an improvement in accuracy from 0.543 to 0.820 for the first snapshot and from 0.725 to 0.923 for the second snapshot.

Conclusions

We have proposed and investigated the application of the methodology of using an ensemble of reusable neural networks for automatic segmentation of a certain area of the spine, namely the Th8-Th11 spine region. The application of this method allowed obtaining more stable and accurate predictions for the desired spine regions, even for images with high noise levels.

Keywords: *artificial intelligence, machine learning, image recognition, neural network, ensemble of neural networks, morphometry.*