

## АНОТАЦІЯ

*Радюк П. М.* Інформаційна технологія раннього діагностування пневмонії за індивідуальним підбором параметрів моделі класифікації медичних зображень легень. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії з галузі знань 12 Інформаційні технології за спеціальністю 122 – Комп'ютерні науки. – Хмельницький національний університет, Хмельницький, 2021.

Дисертаційна робота присвячена розв'язанню актуальної науково-прикладної задачі автоматизації процесу діагностування вірусного пневмонічного запалення за медичними зображеннями легень через розроблення інформаційної технології раннього діагностування пневмонії за індивідуальним підбором параметрів моделі класифікації медичних зображень легень. Застосування розробленої інформаційної технології раннього діагностування пневмонії в клінічній практиці дає змогу підвищити точність та надійність ідентифікації пневмонії на ранніх стадіях за медичними зображеннями грудної клітини людини.

На сьогодні клінічне діагностування за допомогою рентгенологічного випромінювання вважаються традиційним та ефективним засобом до виявлення вірусного пневмонічного запалення. У процесі діагностування рентгенологи зіставляють білі плями на медичному зображенні інфільтратам, а білі ділянки – пневмонічній рідині, ідентифікуючи в такий спосіб захворювання легень. Водночас фахівці-рентгенологи мусять мати достатньо навчені очі, щоби розрізнити неоднорідний колірний розподіл повітря в легенях, та мати широкий багаторічний досвід клінічного лікування. Рентгенологам вкрай важливо визначити чи відповідають білі плями на медичному зображенні пневмонічній рідині. Невчасно або неправильно поставлений діагноз може мати фатальні наслідки для пацієнтів із гострим пневмонічним запаленням легень. Отже, з огляду

на значну проблему ідентифікації вірусної пневмонії за медичними зображеннями легень, актуальним видається застосування обчислювальних методів комп'ютерного зору та машинного навчання до діагностування пневмонії.

Об'єктом дослідження є процес діагностування пневмонії за медичними зображеннями грудної клітини людини.

Предметом дослідження є моделі, методи та засоби інформаційної технології для раннього діагностування пневмонії за медичними зображеннями грудної клітини людини.

У дисертаційній роботі визначено актуальність застосування інформаційних технологій у галузі цифрового діагностування захворювань легень за медичними зображеннями грудної клітини. На основі проведено аналізу методів та підходів до виявлення пневмонії встановлено, що нейромережеві моделі є найкращим рішенням для розроблення інформаційної технології раннього діагностування. Досліджено методи для налаштування нейромережевої моделі та підходи до пояснення та інтерпретування результатів ідентифікації захворювання легень. За аналізом сучасних підходів, методів та інформаційних технологій для діагностування захворювання легень на ранніх стадіях за медичними зображеннями грудної клітини обґрунтовано потребу в створенні інформаційної технології раннього діагностування пневмонії.

Уперше розроблено нейромережеву модель медичного зображення легень з ознаками пневмонічного запалення для ідентифікації пневмонії на ранніх стадіях; особливістю запропонованої моделі є використання за навченою згортковою нейронною мережею гіперпараметрів для виділення карт пневмонічних ознак у випадках індивідуальних особливостей легень, які не входили в початкову навчальну вибірку; такий підхід дає змогу використовувати навчену нейронну мережу для прогнозування класів захворювання легень без перенавчання.

Удосконалено метод підбору квазіоптимальних гіперпараметрів нейромережевої моделі для ідентифікації пневмонії на ранніх стадіях за рентгенівськими зображеннями легень; удосконалення полягає у використанні в генетичному алгоритмові двох генів для формування набору квазіоптимальних гіперпараметрів згорткової нейронної мережі без зміни значень параметрів навчання та застосуванні нового алгоритму формування початкової популяції для створення наборів гіперпараметрів різного розміру; удосконалений метод підбору квазіоптимальних гіперпараметрів нейромережевої моделі за згортковою архітектурою призначений для локалізації та виявлення розсіяних ознак пневмонічного запалення, що відповідають індивідуальним особливостям легень людини; такий підхід дає змогу врахувати індивідуальні особливості легень, й отже, забезпечити високу точність ідентифікації захворювання легень.

Удосконалено метод візуального подання та пояснення результатів медичного діагностування, який, на відміну від відомих підходів, ґрунтується на формуванні карт активації класів на основі середньозважених градієнтів, що дає можливість охопити всю область інтересу на зображенні та забезпечує виражену візуальну локалізацію аномальних зон на медичному зображенні; отримані в такий спосіб теплові карти на попередньо оброблених медичних зображеннях дають змогу ідентифікувати візуальні відмінності в «пневмонічних» зонах зображення, й у такий спосіб, забезпечують зручний інтерфейс для інтерпретування результатів ідентифікації захворювання легень.

Розроблено нову інформаційну технологію раннього діагностування пневмонії за індивідуальним підбором параметрів моделі класифікації медичних зображень легень, що на основі розробленої моделі згорткової нейронної мережі та методу підбору квазіоптимальних гіперпараметрів нейронних мереж, а також із використанням удосконаленого методу візуального подання результатів цифрового діагностування дає можливість ідентифікувати ранню пневмонію з високим показником точності та низьким рівнем обчислювальної складності.

Виконано експериментальну перевірку інформаційної технології раннього діагностування пневмонії через розроблення та використання інформаційної системи ідентифікації захворювання легень за рентгенівськими зображеннями. Результати експериментального тестування запропонованої інформаційної технології довели її спроможність розв'язувати поставлені задачі. Так, розроблена інформаційна технологія перевершила аналоги за показником точності класифікації на 0,58 % та 1,95 % за наборами даних CheXpert\_small та PadChest\_small, відповідно; за показником влучності на 0,02 % за PadChest\_small; за показником повноти на 0,69 % та 1,64 % за CheXpert\_small та PadChest\_small, відповідно; за показником оцінки площі під ROC-кривою на 1,08 % за PadChest\_small; за показником ймовірності помилки другого роду на 0,64 % та 1,64 % за CheXpert\_small та PadChest\_small, відповідно. Високі значення статистичних показників свідчать як про високу точність виявлення пневмонічного запалення, так і значну влучність моделі під час ідентифікації захворювання легень.

У результаті виконаного дослідження розроблено програмне забезпечення інформаційної технології раннього діагностування пневмонії за медичними зображеннями грудної клітини людини. Застосування інформаційної технології раннього діагностування пневмонії дає змогу виконувати ефективну ідентифікацію пневмонічного запалення за рентгенівськими зображеннями легень на малих обчислювальних пристроях; дає можливість враховувати індивідуальні особливості легень кожної людини на медичному зображенні; забезпечує користувача простим та зрозумілим інтерфейсом для виявлення слабо виражених пневмонічних ознак ранньої стадії захворювання. Результати експериментальних тестувань із використанням розробленого програмного забезпечення підтверджують вірність наукових положень запропонованої інформаційної технології, оскільки впровадження інформаційної технології дає змогу підвищити достовірність виявлення ознак пневмонічного запалення за

медичними зображеннями на 0,64 %-1,95 % та знизити ймовірність виникнення помилки під час ідентифікації захворювання легень на 0,64 %-1,64 % у порівнянні з відомими аналогами.

За результатами виконаних досліджень здобувачем розроблено методичне забезпечення, яке використане в навчальному процесі в Хмельницькому національному університеті на кафедрі комп'ютерних наук та інформаційних технологій у процесі викладання навчальних дисциплін «Моделювання систем», «Дослідження операцій та основи теорії прийняття рішень», «Інтелектуальний аналіз даних» та «Прикладні математичні пакети для аналізу результатів наукових досліджень». Розроблені в дисертації положення знайшли практичне застосування в роботі рентгенологічного відділення КП «Хмельницька міська лікарня», ТОВ «Науково-технічна фірма «Інфосервіс» та в навчальному процесі Хмельницького національного університету.

Ключові слова: інформаційна технологія, раннє діагностування, пневмонія, рентгенівське зображення грудної клітини, нейромережева модель ідентифікації, згортоква нейронна мережа, підбір гіперпараметрів, візуальне подання.

## ANNOTATION

*Radiuk Pavlo.* Information technology for early diagnosis of pneumonia by the individual selection of parameters of the classification model by medical images of the lungs. – Manuscript copyright.

The thesis is on obtaining a scientific degree of Doctor of Philosophy in the field of knowledge 12 Information Technologies by specialty 122 – Computer Science. – Khmelnytskyi National University, Khmelnytskyi, 2021.

The present thesis is devoted to solving the topical scientific and applied problem of automating the process of diagnosing viral pneumonia by medical images of the lungs through the development of information technology for early diagnosis of pneumonia by the individual selection of parameters of the classification model by

medical images of the lungs. Applying the developed information technology for the early diagnosis of pneumonia in clinical practice by medical images of the human chest increases the accuracy and reliability of pneumonia identification in the early stages.

Nowadays, clinical diagnosis by X-ray is considered a traditional and effective means of detecting viral pneumonia. In the diagnosis process, radiologists compare white spots on the medical image to infiltrates and white areas - pneumonic fluid, thus identifying lung disease. Altogether, radiologists must have sufficiently trained eyes to distinguish the heterogeneous color distribution of air in the lungs and have extensive years of clinical treatment experience. Radiologists must determine whether the white spots on the medical image correspond to the pneumonic fluid. An untimely or incorrect diagnosis can be fatal for patients with acute pneumonia. Thus, given the significant problem of identifying viral pneumonia by medical images of the lungs, it is vital to use computational computer vision and machine learning methods to diagnose pneumonia.

The research object is the process of diagnosing pneumonia by medical images of the human chest.

The research subject is models, methods, and means of information technology for the early diagnosis of pneumonia by medical images of the human chest.

It was defined in the dissertation that applying information technologies was highly topical in digital diagnostics of lung diseases on medical images of a thorax. Based on the analysis of methods and approaches to detecting pneumonia, it was established that neural network models are the best solution for the development of information technology for early diagnosis. Methods for setting up a neural network model and approaches to explaining and interpreting the results of lung disease identification have been studied. According to the analysis of modern approaches, methods, and information technologies for diagnosing lung disease in the early stages based on medical images of the chest, the need to create information technology for the early diagnosis of pneumonia is substantiated.

For the first time, a neural network model of a medical image of the lungs with the pneumonic inflammation features was created to identify pneumonia in the early stages; the peculiarity of the proposed model is the use of hyperparameters on the pretrained convolutional neural network to select maps of pneumonic signs in cases of individual features of the lungs, which were not included in the initial training sample; this approach allows the use of a trained neural network to predict classes of lung disease without retraining.

The method of selection of quasi-optimal hyperparameters of the neural network model for identification of pneumonia in the early stages by X-ray images of the lungs was improved; improvement is the use in the genetic algorithm of two genes to form a set of quasi-optimal hyperparameters of the convolutional neural network without changing the values of learning parameters and the use of a new algorithm for forming the initial population to create sets of hyperparameters of different sizes; the improved method of selection of quasi-optimal hyperparameters of the neural network model according to the convolutional architecture is intended for localization and detection of the scattered signs of pneumonic inflammation corresponding to individual features of lungs of the person; this approach provides considering the individual characteristics of the lungs, and thus ensures high accuracy in identifying lung disease.

The method of visual presentation and explanation of medical diagnosis was improved, which, unlike known approaches, is based on the formation of activation maps of classes based on weighted average gradients; it allows covering the entire region of interest in the image and provides a pronounced visual localization of abnormal areas in the medical image; the heat maps obtained in this way on pre-treated medical images extract visual differences in the “pneumonic” areas of the image, and thus provide a convenient interface for interpreting the results of lung disease identification.

New information technology for early diagnosis of pneumonia by the individual selection of parameters of the classification model by medical images of the lungs was developed, that is based on the proposed model of convolutional neural network and

method of selection of quasi-optimal hyperparameters of neural networks, and using an advanced method of visual presentation ensures the identification of early pneumonia with high accuracy and low computational complexity.

Experimental validation of the information technology of early diagnosis of pneumonia was performed through the development and use of an information system to identify lung disease by X-ray images. The results of experimental testing of the proposed information technology proved its ability to solve the problem. Thus, the developed information technology surpassed the analogs in terms of classification accuracy by 0.58% and 1.95% for the CheXpert\_small and PadChest\_small datasets, respectively; in precision by 0.02% for PadChest\_small; in recall by 0.69% and 1.64% for CheXpert\_small and PadChest\_small, respectively; in area under the ROC-curve by 1.08% for PadChest\_small; in second type error by 0.64% and 1.64% for CheXpert\_small and PadChest\_small, respectively. High values of statistical indicators indicate both high accuracy of detection of pneumonic inflammation and significant accuracy of the model in the identification of lung disease.

As a result of the performed research, the software of information technology of early diagnosis of pneumonia on medical images of a human thorax was developed. The use of information technology for early diagnosis of pneumonia allows performing effective identification of pneumonic inflammation by X-ray images of the lungs on small computing devices; gives the chance to consider individual features of lungs of each person on the medical image; provides the user with a straightforward interface for detecting mild pneumonia in the early stages of the disease. The results of experimental testing using the developed software confirm the accuracy of the scientific provisions of the proposed information technology, as the introduction of information technology can increase the reliability of detection of signs of pneumonia by medical images by 0.64%-1.95% and reduce the likelihood of error in identifying the lung disease by 0.64%-1.64% compared with known analogs.



Based on research results, the applicant developed methodological support, which was used in the educational process at Khmelnytskyi National University at the Department of Computer Science and Information Technology in the teaching of disciplines “Systems Modeling,” “Operations Research and Fundamentals of Decision Theory,” “Intelligent Data Analysis” and “Applied mathematical packages for analysis of research results.” The provisions developed in the dissertation have found practical application in the work of the radiology department of KP “Khmelnytskyi City Hospital,” LLC Scientific and Technical Firm “Infoservice” and in the educational process of Khmelnytskyi National University.

Keywords: information technology, early diagnosis, pneumonia, chest X-ray image, neural network model of identification, convolutional neural network, selection of hyperparameters, visual representation.

Список опублікованих праць здобувача за темою дисертації

*Публікації в періодичних наукових виданнях, що індексуються в наукометричних базах Scopus, Web of Science:*

1. Radiuk P. M. Applying 3D U-Net architecture to the task of multi-organ segmentation in computed tomography. *Applied Computer Systems*. 2020. Vol. 25, No. 1. Pp. 43-50. (*WoS, Latvia*)

2. Радюк П. М., Бармак О. В., Крак Ю. В. Информационная технология ранней диагностики пневмонии с использованием сверточных нейронных сетей. *Международный научно-технический журнал «Проблемы управления и информатики»*. 2021. № 3. С 101-114. (*Scopus*)

3. Radiuk P., Barmak O., Krak Iu. An approach to early diagnosis of pneumonia on individual radiographs based on the CNN information technology. *The Open Bioinformatics Journal*. 2021. Vol. 14, No. 1, Pp. 92-105. (*Scopus, Netherlands*)

*Публікації в періодичних наукових виданнях інших держав, які входять до Організації економічного співробітництва та розвитку та/або Європейського Союзу:*

4. Radiuk P. M. Impact of training set batch size on the performance of convolutional neural networks for diverse datasets. *Information Technology and Management Science*. 2017. Vol. 20, No. 1. Pp. 20-24.

*Статті у фахових наукових виданнях України:*

5. Radiuk P. M. Neuroevolution of convolutional neural networks for the classification of lung cancer images. *Herald of Khmelnytskyi National University*. 2018. Vol. 267, No. 6(2). Pp. 188-192.

6. Radiuk P. M. Application of a genetic algorithm to search for the optimal convolutional neural network architecture with weight distribution. *Herald of Khmelnytskyi National University*. 2020. Vol. 281, No. 1. Pp. 7-11.

7. Radiuk P. M. Modular search space for automated design of neural architecture. *Proceedings of the O.S. Popov ONAT*. 2020. Vol. 1, No. 1. Pp. 37-44.

8. Бармак О. В., Радюк П. М. Інформаційна технологія візуального подання рентгенівських зображень для інтерпретації результатів діагностування пневмонії. *Вісник Хмельницького національного університету*. 2021. № 295(2). С. 52-55.

9. Barmak O., Radiuk P. Web-based information technology for classifying and interpreting early pneumonia based on fine-tuned convolutional neural network. *Computer systems and information technologies*. 2021. No. 1. Pp. 12-18.

*Статті в матеріалах конференцій, що індексуються в наукометричних базах*

*Scopus:*

10. Radiuk P. M., Hrypynska N. V. A framework for exploring and modelling neural architecture search methods. *The 4th International Conference on Computational Linguistics and Intelligent Systems (COLINS-2020) : CEUR-WS*. Vol. 2604. (Lviv, 23-24 April 2020). Lviv, 2020. Pp. 1060-1074. (*Scopus*)

11. Radiuk P. M., Kutucu H. Heuristic architecture search using network morphism for chest X-Ray classification. *The 1st International Workshop on Intelligent Information Technologies & Systems of Information Security (IntellITSIS-2020)* : CEUR-WS. Vol. 2623. (Khmelnyskyi, 10-12 June 2020). Khmelnyskyi, 2020. Pp. 107-121. (*Scopus*)

12. Krak Iu., Barmak O., Radiuk P. Information technology for early diagnosis of pneumonia on individual radiographs. *The 3rd International Conference on Informatics & Data-Driven Medicine (IDDM 2020)* : CEUR-WS. Vol. 2753. (Växjö, Sweden, November 19-21, 2020). Växjö, 2020. Pp. 11-21. (*Scopus*)

13. Krak Iu., Barmak O., Radiuk P. Detection of early pneumonia on individual CT scans with dilated convolutions. *The 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security (IntellITSIS-2021)* : CEUR-WS. Vol. 2853. (Khmelnyskyi, 24-26 March 2021). Khmelnyskyi, 2021. Pp. 214-227. (*Scopus*)

*Публікації в матеріалах конференцій (тези доповідей):*

14. Радюк П. М. Оптимізація розподілу розмірів фільтрів згорткових нейронних мереж у задачах класифікації. *Актуальні наукові дослідження в сучасному світі ISCIENCE.IN.UA* : тези міжнарод. науч. конф. (м. Переяслав, 21-22 груд. 2016). Переяслав, 2016. С. 48-51.

15. Радюк П. М., Праворська Н. І., Грипинська Н. В. Реалізація нейромережевого алгоритму пакета інструментів MatConvNet з використанням графічного процесора. *Science and Education 2017* : proceedings of X international conference of Science and Education (Rome, Italy, 27 April – 4 May, 2017). Khmelnyskyi, Council of Scientific and Engineering Union in Khmelnyskyi Region, 2017. С. 60-62.

16. Радюк П. М. Аналіз застосування методів нейромережевого моделювання для обробки медичних зображень. «ІНТЕЛЕКТУАЛЬНІ РІШЕННЯ». *Обчислювальний інтелект (результати, проблеми, перспективи)* :

матеріали V міжнар. наук.-практ. конф. (м. Ужгород, 15-20 квіт. 2019 р.). Ужгород, 2019. С. 273-274.

17. Радюк П. М. Поиск оптимальной архитектуры сверточной нейронной сети с использованием генетических алгоритмов. *Современные информационные и коммуникационные технологии на транспорте, в промышленности и образовании* : тезисы XIII междунар. науч.-практ. конф. (м. Днепр, 11-12 дек. 2019 г.). Д. : ДПТ, 2019. С. 117-118.

18. Радюк П. М. Аналітичний огляд архітектур згорткових нейронних мереж у задачах аналізу медичних зображень. *Science, Engineering and Technology: Global and Current Trends* : proceedings of inter. and pract. conf. (м. Прага, 27-28 груд. 2019). Прага, 2019. С. 32-25.

19. Радюк П. М. Стратегія пошуку оптимальної архітектури згорткової нейронної мережі. *VIII Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених з автоматичного управління присвяченої Дню космонавтики* : матеріали VIII всеукр. наук.-практ. конф. (м. Херсон, 8-10 квіт. 2020 р.). Херсон, 2020. С. 35-36.

20. Радюк П. М., Грипинська Н. В. Простір пошуку для задачі оптимізації архітектури нейронної мережі. *IX Міжнародна науково-практична конференція студентів, аспірантів та молодих вчених: Використання інформаційних та комунікаційних технологій в сучасному цифровому суспільстві* : матеріали IX міжнар. наук.-практ. конф. (м. Херсон, 4-5 черв. 2020 р.) / за заг. ред. Г.О. Райко. Херсон: Видавництво ФОП Вишемирський В.С., 2020. С. 308-311.