

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах рукопису

Сидорський Володимир Сергійович

УДК 004.932.2

ДИСЕРТАЦІЯ

ІНТЕЛЕКТУАЛЬНА СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ НА БАЗІ
НЕЙРОННИХ МЕРЕЖ ГЛИБОКОГО НАВЧАННЯ ДЛЯ ЗАДАЧ МЕДИЧНОЇ
ДІАГНОСТИКИ

124 - Системний аналіз

12 - Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ В. С. Сидорський

Науковий керівник Зайченко Юрій Петрович, доктор технічних наук, професор

АНОТАЦІЯ

Сидорський В. С. Інтелектуальна система підтримки прийняття рішень на базі нейронних мереж глибокого навчання для задач медичної діагностики. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 124 Системний аналіз (12 – Інформаційні технології). – Національний Технічний Університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, 2026.

Задачі медичної діагностики є критично важливими для збереження добробуту людства, однак сучасні демографічні тенденції, зокрема зростання чисельності населення, випереджають темпи збільшення кількості висококваліфікованих медичних працівників. Це призводить до ситуації, за якої потреба в якісному медичному догляді не може бути повністю задоволена, особливо в регіонах з обмеженими ресурсами. У зв'язку з цим актуальною є задача підвищення точності та адаптивності клінічних інтелектуальних систем підтримки прийняття рішень (ІСППР), а також їх узгодження з типовими структурами медичних даних. Такі системи здатні автоматизувати частину діагностичних процесів і стандартизувати лікарський висновок, що, своєю чергою, сприятиме суттєвому підвищенню продуктивності праці лікарів.

Значна частина діагностичних задач пов'язана з онкологічними захворюваннями, для яких процеси виявлення й оцінювання вже добре формалізовані та можуть бути використані як основа для побудови інтелектуальних систем, зокрема за критеріями Lung–RADS та Bosniak. Попри наявність значного масиву наукових напрацювань у сфері глибокого навчання для аналізу медичних зображень, кількість клінічних ІСППР, впроваджених у практику, залишається обмеженою. Це зумовлює актуальність розроблення таких систем з урахуванням вимог до форматів вхідних даних, способів представлення результатів і потреб кінцевих користувачів, а також із впровадженням інтерактивних механізмів,

що підвищують довіру лікарів. Особливий інтерес у цьому контексті становить діагностика раку легенів і нирок, водночас із закладанням методологічних засад для подальшої адаптації систем до інших медичних задач.

Для розв'язання задачі побудови сучасної та практично цінної клінічної ІСППР у роботі запропоновано методологію її створення, разом із низкою методів і підходів, спрямованих на підвищення точності діагностики та вирішення суміжних задач медицини й біології. В основі запропонованої методології закладено такі принципи:

- доступність і стандартизованість усіх проміжних результатів для кожної підсистеми;
- можливість простого розширення функціоналу ІСППР шляхом додавання нових підсистем;
- використання моделей глибокого навчання для розв'язання складно структурованих медичних задач.

Перевагами запропонованої методології є те, що:

- побудована ІСППР адаптована до протоколів обміну медичною інформацією (формату DICOM), приймає на вхід медичні зображення у типовому для клінічної практики вигляді та надає структуровану інформацію щодо кожного підозрілого утворення для прийняття остаточного рішення експертом;
- система є робастною до різних типів досліджень і можливих похибок окремих підсистем;
- система легко масштабується та може бути доповнена додатковими підсистемами;
- система має інтерактивну складову, яка дозволяє експерту з мінімальними зусиллями коригувати можливі похибки;
- система побудована на базі моделей глибокого навчання, що забезпечує високу точність розв'язання складно структурованих задач.

На основі запропонованої методології було розроблено ІСППР для діагностики раку легенів, яка без використання інтерактивної складової досягла

повноти 87% при середній кількості хибнопозитивних спрацювань 2 утвори на один КТ–скан, а з інтерактивною складовою – 94% повноти при середній кількості хибнопозитивних спрацювань 1 утвір на один КТ–скан.

Для сегментаційної підсистеми запропоновано удосконалену функцію втрат Boundary DoU++, яка дозволила підвищити значення метрики surface Dice similarity coefficient з 0.47 до 0.487. Запропоновані методи аугментації даних та використання багатоканального входу для моделі класифікації утворів покращили результати з 0.914 до 0.9383 за метрикою ROC AUC. Інтерактивну підсистему побудовано з урахуванням прогнозів статичної системи, що підвищило її адаптивність у складі загальної ІСППР.

Було розроблено клінічну ІСППР для виявлення раку нирок, яка продемонструвала покращення показників порівняно з попередньою версією, зокрема підвищення специфічності з 0.95 до 0.99 за збереження рівня чутливості 0.96. Дану ІСППР розроблено у співпраці з компанією Better Medicine; система визнана придатною до проведення клінічних випробувань, які наразі тривають у клініці.

У ході дисертаційного дослідження запропоновано методи доменної адаптації для сегментації функціональних тканинних одиниць, зокрема гістограмне вирівнювання між доменами, спеціально підібраний набір аугментацій, включаючи міждомений CutMix, а також напівкероване навчання для збільшення вибірки цільового домену. Запропоновані підходи дозволили підвищити точність сегментації з 0.156 до 0.834 за метрикою Dice на цільовому домені та зменшити розрив між доменами з 0.555 до 0.02, одночасно покращивши показники на суміжному домені з 0.711 до 0.854. Отримані результати перевищують показники низки сучасних передових рішень для розв’язання цієї задачі.

Було розроблено трирівневу систему моделей для класифікації раку шкіри на доброякісний та злоякісний. Особливістю цієї системи є поєднання інформації з фотографій утворів і табличних метаданих про пацієнта та всі його утвори. Система включає алгоритми класичного машинного навчання, глибокі згорткові нейронні мережі та гібридні нейронні моделі, для яких запропоновано

двоетапний підхід до оптимізації. Багаторівнева архітектура дозволяє агрегувати передбачення попередніх моделей, підвищуючи загальну точність. Досягнуто значення 0.17 за метрикою Partial ROC AUC (за максимального значення 0.2) для трирівневої системи та 0.974 за метрикою ROC AUC для дворівневої, що відповідає рівню найсучасніших підходів і перевищує результати попередніх моделей.

Отримані результати були впроваджені в навчальний процес, зокрема в магістерський курс «Інтелектуальна обробка голосової інформації», розроблений автором. Запропоновані методи доменної адаптації та гібридні нейронні мережі були інтегровані в лекційні та практичні матеріали розділу «Класифікація сигналів».

Підсумовуючи, наукова новизна дисертаційної роботи полягає в наступному.

Вперше розроблено методологію побудови клінічної інтелектуальної системи підтримки прийняття рішень на базі набору моделей глибокого навчання, що працюють із медичними зображеннями та метаданими основного клінічного протоколу обміну інформацією DICOM і надають на виході виявлені підозрілі утворення, їх радіометричні ознаки та оцінку ступеня ризику. На відміну від попередніх рішень, які реалізують лише окремі задачі (детекція, сегментація, класифікація), розроблена ІСППР охоплює повний процес діагностики та може бути безпосередньо використана в медичних закладах.

Вперше розроблено метод тренування згорткової нейронної мережі для класифікації утворень, який, на відміну від попередніх підходів, передбачає використання двох каналів із додатковим контекстом навколо утвору та з чітко виділеною рамкою утвору, а також застосування спеціалізованих аугментацій. Це дозволило підвищити точність моделі та її стійкість до похибок локалізації утвору.

Вперше розроблено комплекс методів доменної адаптації для сегментації функціональних тканинних одиниць, який, на відміну від попередніх робіт,

поєднує такі підходи: гістограмне вирівнювання між різними протоколами забарвлення; спеціалізований набір аугментацій, зокрема CutMix між доменами; напівкероване навчання на цільовому домені. Це дало змогу суттєво підвищити точність сегментації на цільовому домені.

Вперше розроблено комплекс методів для задачі класифікації раку шкіри, який, на відміну від підходів, заснованих на одиничних моделях або однорівневих ансамблях, передбачає: двоетапну оптимізацію гібридної нейронної мережі, що оперує зображеннями утворів і метаданими утвору та пацієнта; багаторівневу систему, яка поєднує моделі класичного та глибокого машинного навчання. Використання запропонованих методів дозволило суттєво підвищити точність виявлення раку шкіри на фотографічних знімках.

Дістала подальшого розвитку інтерактивна модель сегментації, яка, на відміну від існуючих рішень, оптимізується не лише за інтерактивним сигналом, а й з урахуванням результатів роботи статичної моделі сегментації, що дозволяє підвищити її точність.

Дістала подальшого розвитку сегментаційна функція втрат Boundary DOU шляхом її модифікації до Boundary DOU++, у якій, на відміну від базової версії, введено додатковий параметр γ , що забезпечує підвищення точності сегментації малих об'єктів, зокрема утворів у легенях.

Теоретичні та практичні результати дисертаційної роботи опубліковано у фахових вітчизняних і закордонних наукових виданнях та апробовано на міжнародних і національних наукових конференціях і семінарах.

За матеріалами дисертації опубліковано 5 наукових праць, з яких 2 – статті у журналах і збірниках наукових праць, що входять до переліку фахових видань, затверджених МОН України за спеціальністю дисертації (1 з них індексується у міжнародній наукометричній базі Scopus), 1 – стаття у періодичному науковому виданні, індексованому в міжнародній наукометричній базі Scopus, та 2 – тези міжнародних науково–практичних конференцій.

Ключові слова: медичні зображення, нейронні мережі, згорткові нейронні мережі, інтелектуальні системи підтримки прийняття рішень, глибоке навчання, семантична сегментація, діагностика раку, аналіз великих даних, інтерактивні моделі, напівкероване навчання, машинне навчання, оптимізаційна задача, ракові захворювання шкіри, КТ-зображення, штучний інтелект

ABSTRACT

Sydorskyi V. S. Intelligent Decision Support System Based on Deep Learning Neural Networks for Medical Diagnosis Tasks. – Qualifying scientific work submitted as a manuscript.

Dissertation for the degree of Doctor of Philosophy in specialty 124 System Analysis (12 – Information Technology). – National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, 2026.

Medical diagnostic tasks are critically important for maintaining human well-being; however, modern demographic trends, particularly global population growth, outpace the increase in the number of highly qualified medical professionals. This leads to a situation in which the demand for high-quality medical care cannot be fully satisfied, especially in resource-limited regions. In this context, the task of improving the accuracy and adaptability of clinical Intelligent Decision Support Systems (IDSS), as well as ensuring their alignment with standard medical data structures, becomes highly relevant. Such systems are capable of automating part of the diagnostic processes and standardizing clinical reports, which, in turn, will contribute to a significant increase in physicians' productivity.

A significant portion of diagnostic tasks is related to oncological diseases, for which detection and assessment procedures are already well formalized and can be used as a foundation for building intelligent systems, such as those based on the Lung-RADS and Bosniak criteria. Despite the substantial body of research in deep learning for medical image analysis, the number of clinical IDSS solutions deployed in real practice remains limited. This highlights the relevance of developing such systems with careful consideration of input data formats, result presentation, and end-user requirements, as well as the incorporation of interactive mechanisms that enhance physician trust. In this context, lung and kidney cancer diagnostics are of particular interest, while simultaneously establishing methodological foundations for adapting such systems to other medical tasks.

To address the challenge of building a modern and practically valuable clinical IDSS, this work proposes a methodology for its development, along with a set of methods and approaches aimed at improving diagnostic accuracy and solving related problems in medicine and biology. The proposed methodology is based on the following principles:

- availability and standardization of all intermediate results for each subsystem;
- ease of functional extension through the addition of new subsystems;
- use of deep learning models to solve complex, highly structured medical tasks.

The advantages of the proposed methodology include:

- the developed IDSS is adapted to medical information exchange protocols (DICOM format), accepts medical images in a form typical for clinical practice, and provides structured information for each suspicious lesion to support final decision-making by an expert;
- the system is robust to different types of studies and potential errors in individual subsystems;
- the system is easily scalable and can be extended with additional subsystems when needed;
- the system includes an interactive component that allows experts to correct potential errors with minimal effort;
- the system is built on deep learning models, enabling high-accuracy solutions for complex, structured tasks.

Based on the proposed methodology, an IDSS for lung cancer diagnosis was developed. Without an interactive component, the system achieved a sensitivity of 87% with an average of 2 false-positive detections per CT scan, while with the interactive component, sensitivity increased to 94% with an average of 1 false-positive detection per CT scan.

For the segmentation subsystem, an improved loss function, Boundary DoU++, was proposed, increasing the surface Dice similarity coefficient from 0.47 to 0.487. The proposed data augmentation methods and the use of multi-channel input for the

lesion classification model improved ROC AUC performance from 0.914 to 0.9383. The interactive subsystem was designed to utilize predictions from the static system, increasing its adaptability within the overall IDSS.

A clinical IDSS for kidney cancer detection was developed, demonstrating improved performance compared to its predecessor, including an increase in specificity from 0.95 to 0.99 while maintaining sensitivity at 0.96. This IDSS was developed in collaboration with the company Better Medicine and has been deemed suitable for clinical trials, which are currently ongoing.

The dissertation research proposes domain adaptation methods for segmentation of functional tissue units, including histogram matching between domains, a carefully selected set of augmentations incorporating inter-domain CutMix, and semi-supervised learning to expand the target-domain dataset. These approaches improved segmentation accuracy on the target domain from 0.156 to 0.834 according to the Dice metric and reduced the domain gap from 0.555 to 0.02, while also improving performance on a related domain from 0.711 to 0.854. These results are comparable to the best contemporary solutions for this task.

In addition, a three-level model system for skin cancer classification into benign and malignant categories was developed. A distinctive feature of this system is the integration of information from lesion photographs and tabular patient metadata covering all lesions. The system combines classical machine learning algorithms, deep convolutional neural networks, and hybrid neural models, for which a two-stage optimization approach was proposed. The multi-level architecture enables aggregation of predictions from previous models, improving overall accuracy. Performance of 0.17 was achieved for the Partial ROC AUC metric (with a maximum value of 0.2) for the three-level system and 0.974 ROC AUC for the two-level system, matching state-of-the-art results and surpassing previous approaches.

The obtained results were implemented in the educational process, in particular in the master's course "Intelligent Processing of Speech Information" developed by

the author. The proposed domain adaptation methods and hybrid neural networks were integrated into the lecture and practical materials of the “Signal Classification” section.

In summary, the scientific novelty of this dissertation includes.

For the first time, a methodology for constructing a clinical intelligent decision support system based on a set of deep learning models has been developed. These models operate on medical images and metadata from the primary clinical information exchange protocol DICOM and provide, as output, detected suspicious lesions, their radiomic features, and risk level assessment. Unlike previous solutions that address only individual tasks (detection, segmentation, classification), the developed IDSS covers the entire diagnostic process and can be directly used in medical institutions.

For the first time, a training method for a convolutional neural network for lesion classification has been developed. Unlike previous approaches, it involves the use of two channels with additional contextual information around the lesion and a clearly highlighted lesion boundary, as well as the application of specialized augmentations. This approach improves the model’s accuracy and robustness to localization errors.

For the first time, a comprehensive set of domain adaptation methods for the segmentation of functional tissue units has been developed. Unlike previous works, it combines the following approaches: histogram equalization across different staining protocols; a specialized set of augmentations, including cross-domain CutMix; and semi-supervised learning on the target domain. This made it possible to significantly improve segmentation accuracy on the target domain.

For the first time, a comprehensive set of methods for skin cancer classification has been developed. Unlike approaches based on single models or single-level ensembles, it involves: a two-stage optimization of a hybrid neural network that operates on lesion images and metadata of both the lesion and the patient; and a multi-level system that combines classical machine learning models with deep learning models. The use of the proposed methods made it possible to significantly improve the accuracy of skin cancer detection in photographic images.

The interactive segmentation model has been further developed. Unlike existing solutions, it is optimized not only with respect to the interactive signal but also taking into account the results of a static segmentation model, which allows improving its accuracy.

The Boundary DOU segmentation loss function has been further developed through its modification into Boundary DOU++, in which, unlike the base version, an additional parameter γ is introduced to improve the segmentation accuracy of small objects, in particular lung lesions.

Theoretical and practical results of the dissertation have been published in recognized national and international scientific journals and have been validated through presentations at international and national scientific conferences and seminars.

Based on the dissertation materials, 5 scientific works have been published, including 2 – journal articles published in journals and collections of scientific papers approved by the Ministry of Education and Science of Ukraine for the dissertation specialty (1 of which is indexed in the international scientific database Scopus), 1 – an article published in a periodical scientific journal indexed in the international scientific database Scopus, and 2 – proceedings of an international scientific and practical conferences.

Keywords: medical imaging, neural networks, convolutional neural networks, intelligent decision support systems, deep learning, semantic segmentation, cancer diagnosis, big data analysis, interactive models, semi-supervised learning, machine learning, optimization problem, skin cancer, CT scans, artificial intelligence

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

1. Sydorskyi, V. (2024). Interactive decision support system for lung cancer segmentation. *System Research and Information Technologies*, 2, 68–81. DOI: <https://doi.org/10.20535/SRIT.2308-8893.2024.2.05> ISSN: 1681–6048 (Scopus Q4)

2. Sydorskyi, V. (2025). Evaluating loss functions for 3D lung nodules segmentation. In *System analysis and data mining* (pp. 279–293). *Studies in Systems, Decision and Control* (Vol. 609) Cham: Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-97529-5_17 ISSN: 2198–4182 (Scopus)

3. Сидорський, В. С. (2025). Класифікація легеневих утворів на фрагментах КТ-зображень із використанням тривимірних згорткових нейронних мереж. *Таврійський науковий вісник. Серія: Технічні науки*, 1(5), 399–412. DOI: <https://doi.org/10.32782/tnv-tech.2025.5.1.42> ISSN: 2786–4596 (фахове видання категорії Б)

У матеріалах конференцій:

4. Sydorskyi, V., Krashenyi, I., Savka, D., & Zarichkovyi, O. (2022). Semi-supervised segmentation of functional tissue units at the cellular level. *CEUR Workshop Proceedings*, 3384, 11–23. https://ceur-ws.org/Vol-3384/Paper_2.pdf ISSN: 1613–0073 (Conference Paper). Внесок: Сидорський В.С. – розробка основних методів і фінальної системи; Крашений І.Е. – наукова редакція, дослідження архітектур нейронних мереж; Саква Д.Ю. – дослідження даних і різних методів забарвлення, часткова розробка стратегії аугментації; Зарічковий О.А. – огляд літератури та існуючих методів.

5. Sydorskyi, V. S. (2023). Interactive intelligent decision support system for lung cancer diagnosis (E. M. Kiseleva, Ed.). In *Mathematical support and software for intelligent systems (MSSIS–2023): Abstracts of the XXI International scientific and practical conference* (pp. 45–46). Dnipro: DNU. <http://mpzis.dnu.dp.ua/wp-content/uploads/2023/11/mpzis-2023.pdf>

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	17
ВСТУП	18
1 АНАЛІЗ СУЧАСНИХ ДОСЛІДЖЕНЬ ТА ВИКЛИКІВ У ПОБУ- ДОВІ ІСППР ДЛЯ ДІАГНОСТИКИ ОНКОЛОГІЧНИХ ЗАХВО- РЮВАНЬ	26
1.1 Медична діагностика в онкології: дані, задачі, метрики . . .	26
1.1.1 Формати представлення медичних зображень та специфіка їх використання	26
1.1.2 DICOM стандарт	28
1.1.3 NifTi стандарт	31
1.1.4 Основні задачі обробки медичних зображень . . .	32
1.1.5 Метрики в задачах обробки медичних зображень .	33
1.2 Клінічні ІСППР	35
1.3 Моделі глибокого навчання для класифікації, сегментації та детекції	37
1.4 Інтерактивна сегментація	40
1.5 Ключові виклики в побудові клінічних ІСППР	42
1.6 Суміжні задачі медицини й біології, які потребують залу- чення глибоких нейронних мереж	44
1.7 Висновки розділу 1	47
2 ІСППР ДЛЯ ДІАГНОСТИКИ РАКУ ЛЕГЕНЬ	49
2.1 Постановка задачі побудови ІСППР та вимоги до спроекто- ваної системи	49
2.2 Методологія побудови ІСППР	51
2.3 Загальна архітектура ІСППР	53
2.4 Підсистема обробки та підготовки медичних зображень . .	56

2.5	Підсистема сегментації області легенів	57
2.6	Підсистема сегментації утворів у легенях	60
2.7	Підсистема характеристики та фільтрації утворів	63
2.8	Підсистема класифікації утворів	65
2.9	Підсистема постобробки результатів	69
2.10	Підсистема інтерактивної сегментації утворів	70
2.11	Висновки розділу 2	78
3	ЗАСТОСУВАННЯ РОЗРОБЛЕНИХ МОДЕЛЕЙ І МЕТОДІВ ДЛЯ ІНШИХ ЗАДАЧ ДІАГНОСТИКИ, МЕДИЦИНИ ТА БІОЛОГІЇ . .	80
3.1	Адаптація системи для виявлення злоякісних та доброякісних пухлин у нирках	80
3.2	Доменна адаптація та напівкероване навчання для сегментації функціональних тканинних одиниць (FTU)	84
3.3	Багаторівнева система для виявлення раку шкіри	88
3.4	Висновки розділу 3	93
4	РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ	94
4.1	Опис вибірки для ІСППР діагностики раку легенів	94
4.2	Результати оцінювання моделі сегментації утворів у легенях	98
4.3	Результати оцінювання моделі класифікації утворів у легенях	103
4.4	Результати оцінювання загальної роботи ІСППР	111
4.5	Оцінка впливу інтерактивної підсистеми на результати роботи ІСППР	116
4.6	Результати оцінювання клінічної ІСППР для виявлення раку нирок	125
4.7	Результати сегментації функціональних тканинних одиниць (FTU)	130
4.8	Результати оцінки системи для виявлення раку шкіри	133
4.9	Висновки розділу 4	138
	ВИСНОВКИ	141

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	147
ДОДАТОК А. СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ	161
ДОДАТОК Б. АКТ ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОГО ДОСЛІДЖЕННЯ	162
ДОДАТОК В. АКТ ВПРОВАДЖЕННЯ В НАВЧАЛЬНИЙ ПРОЦЕС . .	163

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

ІСППР – Інтелектуальна Система Підтримки Прийняття Рішень

КТ – Комп'ютерна томографія

DICOM – Digital Imaging and Communications in Medicine

Lung RADS – Lung Imaging Reporting and Data System

NifTi – Neuroimaging Informatics Technology Initiative

JSON – JavaScript Object Notation

ВСТУП

Актуальність Тем. Задачі медичної діагностики є ключовими для підтримки добробуту людства. Проте сучасні демографічні тенденції, зокрема зростання чисельності населення Землі, є швидшими за темпи збільшення кількості медичних працівників, а особливо висококваліфікованих фахівців у галузі медицини [1]. Це означає, що в найближчому майбутньому потреба населення в якісному медичному догляді не зможе бути повністю забезпечена. Зазначена проблема вже є актуальною для країн і регіонів із низьким рівнем доходу, зокрема для окремих країн Африки [2].

У зв'язку з цим виникає нагальна задача підвищення продуктивності роботи медичних працівників з метою збільшення кількості пацієнтів, які можуть отримати якісні медичні послуги від одного лікаря. Одним із перспективних напрямів досліджень у цьому контексті є інтеграція інформаційних технологій, зокрема клінічних інтелектуальних систем підтримки прийняття рішень (ІСППР), для розв'язання широкого кола діагностичних задач [3].

Значна частина задач медичної діагностики припадає на виявлення та оцінювання онкологічних захворювань. Регулярні скринінгові обстеження суттєво підвищують імовірність виявлення злоякісних утворень на ранніх стадіях [4], однак водночас призводять до додаткового навантаження на медичну інфраструктуру. Важливо зазначити, що процеси діагностики багатьох онкологічних захворювань вже добре формалізовані, а для них розроблено низку стандартизованих методологій і критеріїв, зокрема Lung–RADS [5] для легенів та Bosniak [6] для нирок. Такі критерії можуть слугувати основою для формування як чітких, так і нечітких правил у побудові клінічних ІСППР.

Дослідження у сфері штучного інтелекту та глибокого навчання вже охопили широкий спектр задач, необхідних для створення систем медичної діагностики. Переважна більшість сучасних методів зосереджена на обробці медичних

зображень, зокрема на задачах класифікації [7], сегментації [8] та детекції [9]. Таким чином, на сьогодні існує значний масив наукових напрацювань, які можуть слугувати основою для побудови клінічних ІСППР.

Водночас наразі практично відсутні прикладні клінічні ІСППР, які широко використовуються в реальній медичній практиці, а кількість систем, побудованих на основі глибоких нейронних мереж, є вкрай обмеженою. Частина з існуючих рішень не підтримує необхідні для лікарів формати вхідних даних та представлення результатів. Отже, задача розроблення клінічної ІСППР для медичної діагностики, зокрема онкологічних захворювань, на основі глибоких нейронних мереж залишається вкрай актуальною. При цьому особливу увагу необхідно приділяти вимогам до формату вхідних даних і способів представлення результатів для кінцевих користувачів.

Крім того, для успішної інтеграції подібних систем у клінічну практику слід враховувати наявний рівень скептицизму та занепокоєння з боку медичних працівників щодо використання моделей із нечіткою або недостатньо інтерпретованою логікою прийняття рішень [10]. Одним із можливих шляхів подолання цієї проблеми є впровадження інтерактивної складової, яка уможлиблює корекцію результатів системи за мінімальних витрат часу та зусиль лікаря, водночас забезпечуючи прозорість і зрозумілість логіки її роботи.

Також необхідно приділяти особливу увагу підвищенню загальної точності системи та її підсистем, а також забезпеченню їх робастності до різних протоколів дослідження й типів медичних зображень, оскільки саме ці чинники є одними з ключових для успішного впровадження системи в реальну клінічну практику.

Одними з найпоширеніших онкологічних захворювань є рак легенів і рак нирок, тому розроблення клінічних ІСППР саме для діагностики цих онкологій є пріоритетним завданням. Водночас важливо не лише створити окрему ІСППР, а й закласти методичні засади та архітектурні рішення для побудови підсистем, які можуть бути повторно використані та адаптовані для діагностики інших видів раку, а також для розв'язання суміжних задач медицини та біології.

Мета і задачі дослідження. Метою дисертаційного дослідження є підвищення кінцевої точності систем виявлення раку легень, нирок і шкіри, зменшення впливу похибок, спричинених відмінностями у протоколах досліджень та різними джерелами походження даних, а також скорочення кількості інтеракцій користувача, необхідних для виправлення помилок системи.

Для досягнення поставленої мети були сформульовані такі завдання дослідження:

- підвищити точність сегментації малих об'єктів, якими є утвори в легенях;
- підвищити точність класифікації утворів за умови наявності похибок їх локалізації;
- підвищити точність роботи інтерактивної моделі сегментації утворів у легенях за наявності результатів попередньої статичної моделі;
- підвищити точність роботи системи виявлення раку нирок;
- підвищити точність роботи системи виявлення раку шкіри на фотографіях потенційних ділянок ураження;
- підвищити точність сегментації функціональних тканинних одиниць за різних протоколів дослідження мікроскопічних зображень;
- розробити методологію побудови ІСППР, здатну працювати з сирими медичними даними та надавати на виході дані, необхідні для встановлення діагнозу.

Об'єктом дослідження є процес автоматизованої та напіваавтоматизованої діагностики онкологічних захворювань.

Предметом дослідження є моделі, методи та алгоритми глибокого навчання для діагностики онкологічних захворювань.

Методи дослідження. Для проведення дослідження було використано такі методи:

- методи математичної статистики для аналізу великих наборів даних і оцінювання результатів моделювання;
- методи та моделі глибокого навчання для розробки й навчання нейронних мереж;

- чисельні методи оптимізації для налаштування параметрів моделей;
- методи інтелектуального аналізу даних для обробки та інтерпретації великих вибірок;
- методи й моделі комп'ютерного зору для обробки та аналізу зображень.

Наукова новизна одержаних результатів.

Вперше розроблено методологію побудови клінічної інтелектуальної системи підтримки прийняття рішень на базі набору моделей глибокого навчання, що працюють із медичними зображеннями та метаданими основного клінічного протоколу обміну інформацією DICOM [11] і надають на виході виявлені підозрілі утворення, їх радіометричні ознаки та оцінку ступеня ризику. На відміну від попередніх рішень, які реалізують лише окремі задачі (детекція, сегментація, класифікація), розроблена ІСППР охоплює повний процес діагностики та може бути безпосередньо використана в медичних закладах.

Вперше розроблено метод тренування згорткової нейронної мережі для класифікації утворень, який, на відміну від попередніх підходів, передбачає використання двох каналів із додатковим контекстом навколо утвору та з чітко виділеною рамкою утвору, а також застосування спеціалізованих аугментацій. Це дозволило підвищити точність моделі та її стійкість до похибок локалізації утвору.

Вперше розроблено комплекс методів доменної адаптації для сегментації функціональних тканинних одиниць, який, на відміну від попередніх робіт, поєднує такі підходи: гістограмне вирівнювання між різними протоколами забарвлення; спеціалізований набір аугментацій, зокрема CutMix між доменами; напівкероване навчання на цільовому домені. Це дало змогу суттєво підвищити точність сегментації на цільовому домені.

Вперше розроблено комплекс методів для задачі класифікації раку шкіри, який, на відміну від підходів, заснованих на одиничних моделях або однорівневих ансамблях, передбачає: двоетапну оптимізацію гібридної нейронної мережі, що оперує зображеннями утворів і метаданими утвору та пацієнта; багаторівневу

систему, яка поєднує моделі класичного та глибокого машинного навчання. Використання запропонованих методів дозволило суттєво підвищити точність виявлення раку шкіри на фотографічних знімках.

Дістала подальшого розвитку інтерактивна модель сегментації, яка, на відміну від існуючих рішень, оптимізується не лише за інтерактивним сигналом, а й з урахуванням результатів роботи статичної моделі сегментації, що дозволяє підвищити її точність.

Дістала подальшого розвитку сегментаційна функція втрат Boundary DOU шляхом її модифікації до Boundary DOU++, у якій, на відміну від базової версії, введено додатковий параметр γ , що забезпечує підвищення точності сегментації малих об'єктів, зокрема утворів у легенях.

Практичне значення одержаних результатів. У ході дисертаційного дослідження запропоновано методологію побудови клінічних інтелектуальних систем підтримки прийняття рішень (ІСППР), на основі якої реалізовано ІСППР для діагностики раку легень та ІСППР для виявлення раку нирок. Розроблена ІСППР для діагностики раку легень має високий потенціал практичного впровадження, оскільки спроектована з урахуванням сучасних діагностичних підходів і клінічних протоколів та продемонструвала високі показники точності. Зокрема, повнота виявлення утворів у легенях становить 87% при середній кількості хибнопозитивних спрацювань 2 утвори на один КТ-скан. За умови використання інтерактивної підсистеми повнота зростає до 94% при зменшенні середньої кількості хибнопозитивних спрацювань до 1 утвору на один КТ-скан. Водночас запропонована методологія може бути перевикористана для побудови інших клінічних ІСППР.

Запропоновані методи, системи та вдосконалення були оцінені за допомогою об'єктивних метрик і продемонстрували високу ефективність порівняно з існуючими підходами, а саме:

- запропонована функція втрат для сегментаційних моделей підвищила точність сегментації за метрикою surface Dice similarity coefficient [12] з 0.47 до 0.487;
- методи формування двоканалного входу та підходи до аугментації даних для класифікаційних моделей забезпечили зростання точності класифікації утворів за метрикою ROC AUC з 0.914 до 0.938;
- запропоновані методи доменної адаптації дозволили підвищити точність сегментації на цільовому домені за метрикою Dice з 0.156 до 0.834 та зменшити розрив між доменами з 0.555 до 0.02, одночасно покращивши показники на суміжному домені з 0.711 до 0.854;
- розроблена система виявлення раку шкіри досягла значення 0.17 за метрикою Partial ROC AUC [13] (за максимального значення 0.2) у трирівневій конфігурації та 0.974 за метрикою ROC AUC у дворівневій конфігурації, що відповідає рівню найсучасніших підходів і перевищує результати попередніх моделей.

Запропоновані методи та системи можуть бути перевикористані для розв'язання інших задач медичної діагностики та інтегровані в прикладні інформаційні системи, що використовуються лікарями у повсякденній практиці та під час проведення скринінгових досліджень.

Розроблені ІСППР та супутні інтелектуальні системи можуть стати важливим інструментом для вирішення проблеми дефіциту висококваліфікованих медичних кадрів і зростання кількості онкологічних захворювань.

Розроблена клінічна ІСППР для виявлення раку нирок, створена у співпраці з компанією Better Medicine, вже допущена до клінічних випробувань і наразі проходить їх у європейській клініці. Акт впровадження цієї ІСППР наведено в Додатку до роботи. За результатами оцінювання система продемонструвала покращення порівняно з попередньою версією за метрикою специфічності на рівні пацієнта – її підвищено з 0.95 до 0.99 при збереженні чутливості на рівні 0.96.

Також отримані результати були впроваджені в навчальний процес, зокрема в магістерський курс «Інтелектуальна обробка голосової інформації», розроблений автором. Запропоновані методи доменної адаптації та гібридні нейронні мережі були інтегровані в лекційні та практичні матеріали розділу «Класифікація сигналів».

Особистий внесок здобувача. Дисертаційна робота є самостійним науковим дослідженням автора. Одна з публікацій здобувача виконана у співавторстві. У роботі [4] здобувачем запропоновано підхід напівкерованого навчання для доменної адаптації, а також застосовано аугментації, зокрема CutMix, і гістограмне вирівнювання, для розв'язання цієї ж задачі. Крім того, здобувачем розроблено систему моделей, яка продемонструвала найкращі результати.

Апробація результатів дисертації. Результати та основні положення дисертаційної роботи були апробовані та представлені на таких наукових конференціях:

- IX International Scientific Conference «Information Technology and Implementation» (IT&I–2022), Kyiv, Ukraine, November 30 – December 02, 2022 [4];

- Міжнародна науково–практична конференція «Математичне та програмне забезпечення інтелектуальних систем» (МПЗІС-2023), Дніпро, 22–24 листопада 2023 р. [5];

Публікації. Основні теоретичні та практичні результати дисертаційної роботи відображено у 5 наукових працях, опублікованих у фахових вітчизняних і закордонних наукових виданнях, зокрема: в 2 статті у періодичних наукових виданнях, що індексується у наукометричній базі Scopus; 1 статті у фаховому виданні категорії Б; 2 тезах міжнародних наукових конференцій.

Використання штучного інтелекту. Під час підготовки цієї роботи автор використовував ChatGPT від OpenAI з метою підвищення зрозумілості, зв'язності та якості форматування тексту. Після використання цього інструменту автор здійснив перевірку й за потреби відредагував матеріал та несе повну відпо-

відальність за зміст рукопису. Використання інструментів штучного інтелекту повністю узгоджується з політикою застосування технологій штучного інтелекту в академічній діяльності в КПІ ім. Ігоря Сікорського (протокол № 11 Вченої ради КПІ ім. Ігоря Сікорського від 11 грудня 2023 р.).

Структура і обсяг роботи. Дисертаційна робота складається зі вступу, чотирьох розділів, загальних висновків, списку використаних джерел із 118 найменувань. Загальний обсяг дисертації становить 163 сторінок, з яких 129 сторінок основного тексту, 3 додатки, та містить 51 рисунків, 26 формул, 26 таблиць.

1 АНАЛІЗ СУЧАСНИХ ДОСЛІДЖЕНЬ ТА ВИКЛИКІВ У ПОБУДОВІ ІСППР ДЛЯ ДІАГНОСТИКИ ОНКОЛОГІЧНИХ ЗАХВОРЮВАНЬ

1.1 Медична діагностика в онкології: дані, задачі, метрики

1.1.1 Формати представлення медичних зображень та специфіка їх використання

Сучасна діагностика використовує широкий спектр медичних зображень залежно від типу захворювання, ділянки дослідження, стану здоров'я пацієнта або індивідуальних обмежень. Дослідження, що базуються на медичних зображеннях, необхідні для діагностики, стадіювання та моніторингу лікування. Серед найбільш поширених типів можна виділити наступні [14].

1. Комп'ютерна томографія (КТ) – тривимірні зображення (1.1). КТ є швидким і доступним методом із високою просторовою роздільною здатністю. Використовується для виявлення пухлин, оцінки поширення та планування променевої терапії. Обмеження: іонізуюче випромінювання та низький контраст м'яких тканин. Приклад поперечного зрізу наведено на рис 1.1.

2. Магнітно–резонансна томографія (МРТ) – так само, як і КТ, формує тривимірне зображення, але забезпечує значно кращий контраст м'яких тканин без опромінення. Є ключовою в діагностиці пухлин головного мозку, органів малого таза, простати та печінки. Широко застосовуються мультипараметричні підходи (наприклад, mpMRI у простаті). Недоліки: висока вартість і тривалість обстеження.

3. Позитронно–емісійна томографія (ПЕТ, зазвичай у комбінації з КТ або МРТ) – надає інформацію про метаболічну активність пухлини. FDG–ПЕТ дозволяє виявляти приховані метастази та оцінювати ранню відповідь на терапію. Нові радіофармпрепарати (наприклад, PSMA) суттєво розширюють застосування у простатичному та нейроендокринному раку.

4. Ультразвукова діагностика (УЗД) – безпечний і недорогий метод для первинної оцінки та наведення біопсій. Важливий у діагностиці вузлів щитоподібної залози, печінки, молочних залоз і гінекологічних пухлин. Ендобронхіальний та ендоскопічний УЗД допомагають у стадіюванні раку легень і шлунково-кишкового тракту.

5. Мікроскопічні зображення (патогістологія) – зазвичай класичні RGB-зображення, що є золотим стандартом у підтвердженні діагнозу. Включають світлову, флуоресцентну та електронну мікроскопію. Гістологічні зрізи дають змогу візуалізувати морфологію клітин і тканин, визначати злоякісність і тип пухлини.

$$\mathbf{V} \in \mathbb{R}^{H \times W \times D}, \quad I : \{0, \dots, H-1\} \times \{0, \dots, W-1\} \times \{0, \dots, D-1\} \rightarrow \mathbb{R}, \quad (1.1)$$

де $\mathbf{V}$ – тривимірний об'єм КТ-зображення;

H, W, D – висота, ширина та глибина об'єму відповідно;

$I(x, y, z)$ – інтенсивність вокселя у шкалі HU;

x, y, z – просторові координати вокселя.

Жоден із методів не є універсальним і зазвичай потребує доповнення іншими дослідженнями для підтвердження діагнозу. Наприклад, КТ-знімок може надати інформацію про розміщення потенційної злоякісної пухлини та її морфологічні характеристики (об'єм, форму). Цією інформацією радіолог може керуватися при прийнятті рішення щодо доцільності подальших досліджень чи спостережень за утворенням [15]. Проте для остаточного визначення характеру утворення (злоякісне чи доброякісне) необхідне дослідження тканини – біопсія та гістологічний аналіз [16].

У діагностиці й моніторингу онкологічних захворювань медичні зображення відіграють ключову роль на етапах раннього виявлення, локалізації, оцінки прогресії та визначення ризиків для пацієнта. Водночас остаточне рішення щодо

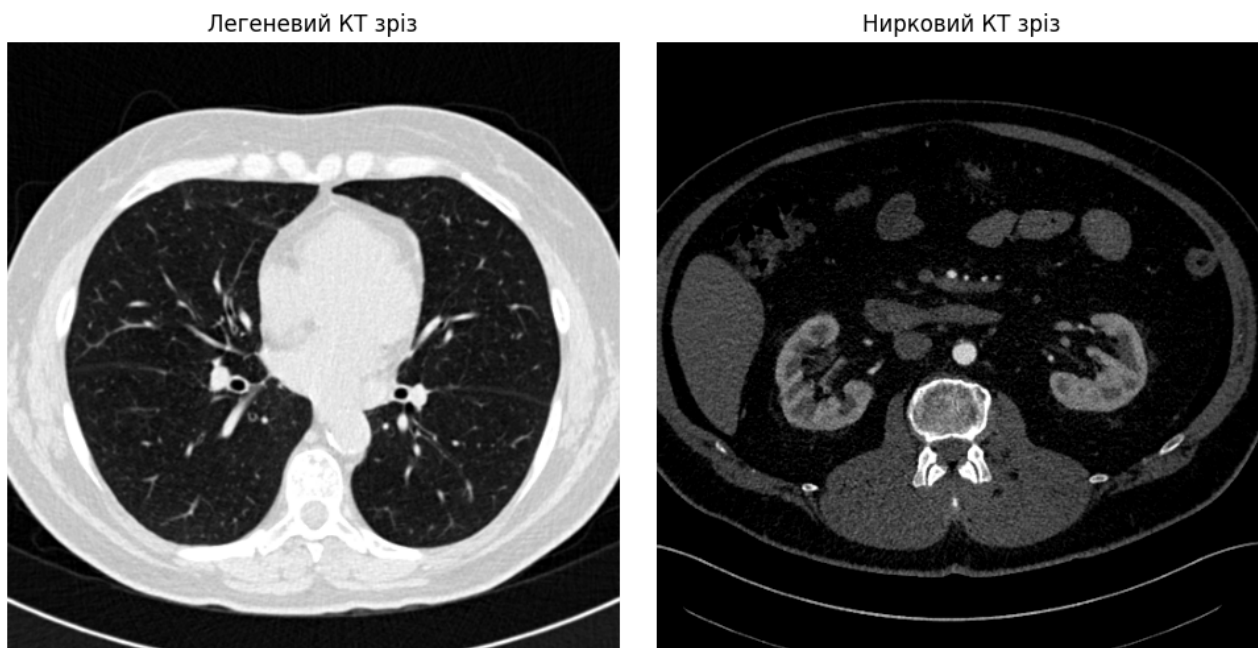


Рис. 1.1 – Приклад поперечного зрізу КТ скану легенів і нирок з відповідними параметрами відображення (WindowCenter й WindowWidth)

природи пухлини приймається виключно на основі морфологічного дослідження біопсійного матеріалу.

1.1.2 DICOM стандарт

Для зберігання, обміну та інтеграції медичних зображень у клінічних інформаційних системах використовується міжнародний стандарт DICOM (Digital Imaging and Communications in Medicine) [11].

DICOM визначає не лише формат збереження зображень, але й структуру метаданих, які містять відомості про пацієнта, параметри дослідження, тип обладнання, часові мітки тощо. Це забезпечує сумісність між різними апаратними й програмними засобами, дозволяє інтегрувати зображення у системи архівації та передачі зображень (PACS), а також підтримує роботу медиків і радіологів на всіх етапах діагностичного процесу [17].

Ключові можливості DICOM.

1. Універсальність – підтримка різних модальностей: КТ, МРТ, ПЕТ, УЗД, мамографії тощо.

2. Інтероперабельність – забезпечує обмін зображеннями між обладнанням різних виробників.
3. Інтеграція з клінічними системами – можливість взаємодії з електронними медичними записами (EHR) та PACS.
4. Збереження історії – включення метаданих для відстеження параметрів дослідження та клінічного контексту.

У випадку комп'ютерної томографії формат DICOM використовується для збереження кожного окремого зрізу томографічного обстеження разом із супроводними метаданими. Самі зображення зберігаються у Hounsfield Units (HU) – числовій шкалі, що нормує коефіцієнт ослаблення рентгенівських променів: значення 0 відповідає воді, 1000 – повітрю, а більш високі значення (до +3000 і більше) характерні для кісткової тканини [18].

У DICOM-файлах КТ важливими є такі теги.

1. StudyInstanceUID – унікальний ідентифікатор дослідження. Під час одного дослідження зазвичай отримують декілька КТ знімків, наприклад різних частин тіла, або з різними типами реконструкції, або в різних площинах
2. SeriesInstanceUID – унікальний ідентифікатор серії (тобто КТ повного 3D КТ знімку) всередині дослідження.
3. PatientID – ідентифікатор пацієнта. Зазвичай відсутній або анонімізований задля збереження лікарської таємниці.
4. Modality – модальність візуалізації (наприклад, CT, MR, US).
5. ImageType – характеристики типу зображення (оригінальне, реконструйоване тощо).
6. BodyPartExamined – ділянка тіла, що досліджується (наприклад, Chest, Abdomen).
7. ConvolutionKernel – тип фільтра реконструкції (наприклад, B30f для «м'яких тканин» або B70f для «кісток»).
8. Manufacturer – виробник обладнання.
9. ContrastBolusAgent – назва контрастного засобу (якщо застосовувався).

10. ContrastBolusStartTime – час введення контрастного агента.
11. AcquisitionDateTime – дата та час отримання зображення.
12. RescaleIntercept та RescaleSlope – параметри для перетворення сирих піксельних значень у фізично значимі Hounsfield Units (HU) (1.2).
13. PixelSpacing – розмір пікселя у площині зображення (мм), тобто скільком реальним міліметрам відповідає один воксель (назва пікселю для КТ/МРТ сканів) зоображення в аксіальній площині.
14. SliceThickness – товщина зрізу (мм), тобто яка відстань між двома зрізами в КТ знімку.
15. ImagePositionPatient – координати лівого верхнього кута зображення у системі координат пацієнта.
16. ImageOrientationPatient – напрямні косинуси для визначення орієнтації зрізу у просторі.
17. WindowCenter та WindowWidth – параметри відображення для візуалізації певних тканин (наприклад, «вікно легенів», «вікно кістки»).

$$HU = SV \cdot \text{RescaleSlope} + \text{RescaleIntercept} \quad (1.2)$$

де SV – зображення у початкових піксельних значеннях;

HU – зображення у шкалі Hounsfield Units.

Приклад метаданих DICOM файлу зображений на рис. 1.2

(0008,0060) Modality	CS: 'CT'
(0008,0070) Manufacturer	LO: 'GE MEDICAL SYSTEMS'
(0008,0090) Referring Physician's Name	PN: ''
(0008,1090) Manufacturer's Model Name	LO: 'LightSpeed Plus'
(0008,1155) Referenced SOP Instance UID	UI: 1.3.6.1.4.1.14519.5.2.1.6279.6001.266980125019159618870110055006
(0010,0010) Patient's Name	PN: ''
(0010,0020) Patient ID	LO: 'LIDC-IDRI-0001'
(0010,0030) Patient's Birth Date	DA: ''
(0010,0040) Patient's Sex	CS: ''
(0010,1010) Patient's Age	AS: ''
(0010,2100) Last Menstrual Date	DA: '20000101'
(0012,0062) Patient Identity Removed	CS: 'YES'
(0012,0063) De-identification Method	LO: 'DCM:113100/113105/113107/113108/113109/113111'
(0013,0010) Private Creator	LO: 'CTP'
(0013,1010) Private tag data	LO: 'LIDC-IDRI'
(0013,1013) Private tag data	LO: '62796001'
(0018,0010) Contrast/Bolus Agent	LO: 'IV'
(0018,0015) Body Part Examined	CS: 'CHEST'
(0018,0022) Scan Options	CS: 'HELICAL MODE'
(0018,0050) Slice Thickness	DS: '2.500000'
(0018,0060) KVP	DS: '120'
(0018,0090) Data Collection Diameter	DS: '500.000000'
(0018,1020) Software Versions	LO: 'LightSpeedApps2.4.2_H2.4M5'
(0018,1100) Reconstruction Diameter	DS: '360.000000'
(0018,1110) Distance Source to Detector	DS: '949.075012'
(0018,1111) Distance Source to Patient	DS: '541.000000'
(0018,1120) Gantry/Detector Tilt	DS: '0.000000'
(0018,1130) Table Height	DS: '144.399994'
(0018,1140) Rotation Direction	CS: 'CW'
(0018,1150) Exposure Time	IS: '570'
(0018,1151) X-Ray Tube Current	IS: '400'
(0018,1152) Exposure	IS: '4684'
(0018,1160) Filter Type	SH: 'BODY FILTER'
(0018,1170) Generator Power	IS: '48000'
(0018,1190) Focal Spot(s)	DS: '1.200000'
(0018,1210) Convolution Kernel	SH: 'STANDARD'
(0018,5100) Patient Position	CS: 'FFS'
(0020,000D) Study Instance UID	UI: 1.3.6.1.4.1.14519.5.2.1.6279.6001.298806137288633453246975630178
(0020,000E) Series Instance UID	UI: 1.3.6.1.4.1.14519.5.2.1.6279.6001.179049373636438705059720603192
(0020,0010) Study ID	SH: ''
(0020,0011) Series Number	IS: '3000566'
(0020,0013) Instance Number	IS: '65'
(0020,0032) Image Position (Patient)	DS: [-166.000000, -171.699997, -170.000000]
(0020,0037) Image Orientation (Patient)	DS: [1.000000, 0.000000, 0.000000, 0.000000, 1.000000, 0.000000]
(0020,0052) Frame of Reference UID	UI: 1.3.6.1.4.1.14519.5.2.1.6279.6001.229925374658226729607867499499
(0020,1040) Position Reference Indicator	LO: 'SN'

Рис. 1.2 – Приклад DICOM тегів медичного зображення

Таким чином, DICOM у КТ містить не лише піксельні значення зрізів, але й усю необхідну мета інформацію для їх коректної інтерпретації, 3D–реконструкції та клінічного застосування.

1.1.3 NifTi стандарт

Формат NifTi (Neuroimaging Informatics Technology Initiative) [19] є де-факто стандартом у наукових дослідженнях, що базуються на МРТ, КТ та функціональній МРТ (fMRI). Його простота та відкрита специфікація зробили цей формат одним з основних у нейровізуалізації та медичній інформатиці.

Основні особливості формату.

1. Простота структури – файл складається з header, що містить метадані (розміри, воксельний розмір, афінне перетворення, тип даних), та image data, де зберігається сам масив інтенсивностей.
2. Підтримка багатовимірних даних – 3D (об’ємні дані) та 4D (часові ряди, наприклад для fMRI або динамічних досліджень).
3. Афінні матриці – у полях qform та sform зберігається інформація для відображення вокселів у фізичний простір пацієнта.
4. Стислість та сумісність – NifTi–файли легко інтегруються у наукові пайплайни, зокрема в Python (через бібліотеку nibabel [20]) та MATLAB.
5. Варіанти збереження – однокомпонентний файл (.nii) або розділений на .hdr (заголовок) та .img (дані).

Після конвертації КТ–зображень у формат NifTi можна застосовувати різні програмні пакети для просторових трансформацій: масштабування, обертання, трансляції та інші афінні перетворення. Використання афінних матриць дозволяє коректно переходити між різними системами координат, що критично для подальшої обробки медичних зображень, наприклад у задачах сегментації чи реєстрації.

Формат NifTi отримав широке поширення у медичній інформатиці та біоінженерії завдяки своїй гнучкості, відкритості та підтримці у численних наукових платформах. На відміну від DICOM, він не містить великої кількості клінічних метаданих, проте є оптимальним для машинного навчання та аналітичних досліджень, де основна увага зосереджується на структурі зображення та просторових трансформаціях.

1.1.4 Основні задачі обробки медичних зображень

Підходи машинного й глибокого навчання активно застосовуються у задачах обробки медичних зображень [21]. Основні напрями можна згрупувати у кілька категорій.

1. Попередня обробка (preprocessing) – нормалізація інтенсивностей, фільтрація шумів та артефактів, підвищення контрастності. Ці кроки необхідні для покращення якості даних перед подальшим аналізом.

2. Сегментація – виокремлення структур та органів (наприклад, легенів, нирок, пухлин) із томографічних даних. Сегментація дозволяє визначити форму, розмір і об'єм патологічних утворень.

3. Детекція та локалізація – виявлення підозрілих ділянок (пухлин, метастазів, аномалій) та визначення їх координат у просторі зображення. На відміну від сегментації, детекція зазвичай надає лише обмежувальні рамки (bounding boxes) без точної маски об'єкта. Це зменшує вимоги до навчальних вибірок та обчислювальних ресурсів, проте знижує деталізацію результатів.

4. Реєстрація – зіставлення кількох зображень одного пацієнта (наприклад, до і після лікування) або даних різних модальностей (КТ, МРТ, ПЕТ). Це дозволяє поєднувати структурну та функціональну інформацію.

5. Класифікація – віднесення утворень, ділянок зображення або цілого зображення до певного класу (наприклад, доброякісне чи злоякісне, тип пухлини, стадія захворювання).

6. Характеризація – отримання кількісних характеристик (об'єм пухлини, товщина тканини, метаболічна активність), що використовуються у клінічних протоколах та при оцінці ефективності лікування.

7. Візуалізація – побудова 2D/3D реконструкцій, багатоплощинних та інтерактивних відображень для клініцистів, які полегшують процес прийняття рішень.

Таким чином, задачі обробки медичних зображень охоплюють увесь цикл – від попередньої підготовки даних до отримання клінічно релевантної інформації, яку може використати лікар для прийняття рішень.

1.1.5 Метрики в задачах обробки медичних зображень

Для оцінки якості алгоритмів у задачах обробки медичних зображень застосовується широкий спектр метрик, вибір яких залежить від конкретної постанов-

ки задачі: сегментація, детекція, класифікація чи реєстрація. Метрики мають не лише технічне, але й клінічне значення, оскільки вони безпосередньо впливають на довіру до систем підтримки прийняття рішень [22].

1. Метрики сегментації: Dice Similarity Coefficient (DSC), Jaccard Index (IoU), Sensitivity (Recall), Specificity, Hausdorff Distance [23], Average Surface Distance [24]. Вони оцінюють точність виділення контурів органів чи патологічних утворень [25].

2. Метрики детекції: Precision, Recall, F1–score, Average Precision (AP), Area Under the Curve (AUC). Використовуються для оцінки здатності моделі виявляти об'єкти (наприклад, пухлини чи метастази) з урахуванням кількості хибних спрацьовувань.

3. Метрики класифікації: Accuracy, Balanced Accuracy, ROC AUC, Precision–Recall AUC, Cohen's Kappa [26], Area under the Free–Response ROC Curve (FROC) [27]. Ці показники важливі для задач диференціації доброякісних і злоякісних утворень або для класифікації стадій захворювання.

4. Метрики реєстрації: Root Mean Square Error (RMSE) між трансформованими та еталонними точками, Mutual Information (MI), Normalized Cross–Correlation (NCC). Вони оцінюють якість просторового узгодження між різними модальностями чи серіями.

5. Клінічно орієнтовані метрики: зміна об'єму пухлини, довжини найдовшого діаметра (за RECIST [15] або Lung–RADS [5]), оцінка збереження форми або відповідність протоколам лікування. Такі метрики дозволяють безпосередньо інтегрувати результати у практику лікарів.

Водночас важливо наголосити, що окрім вибору конкретного набору метрик, значення має і спосіб їх підрахунку. Наприклад, сегментаційні метрики можуть обчислюватися:

- для кожного пацієнта з подальшим усередненням;
- для кожного об'єкта (наприклад, пухлини чи органу) окремо з подальшим усередненням;

- по всій вибірці одразу.

Від обраного методу підрахунку залежить чутливість метрики до похибок у малих чи великих об'єктах, а також до інших їх характеристик. Крім того, сучасні модифікації класичних метрик (наприклад, Surface Dice [12]) краще відображають ті аспекти, які є критичними для клінічного застосування, зокрема точність опису контуру об'єкта, а не лише повну піксельну сегментацію.

Таким чином, використання метрик в обробці медичних зображень має дві ключові складові: технічну, що дозволяє порівнювати моделі між собою на основі стандартних наборів даних, та клінічну, яка орієнтована на прийняття практичних рішень у медицині.

1.2 Клінічні ІСППР

Інтелектуальна система підтримки прийняття рішень (Intelligent Decision–Maker Support System, iDMSS) – це інтерактивна комп'ютерна система, призначена для підтримки процесу прийняття рішень у різних сферах діяльності, зокрема у випадках слабкоструктурованих і неструктурованих проблем. Такі системи ґрунтуються на використанні моделей і процедур обробки даних та знань із залученням сучасних технологій штучного інтелекту [28].

У медичному контексті ІСППР застосовуються для діагностики, прогнозування та планування лікування. Вони можуть функціонувати на різних рівнях прийняття рішень.

1. На рівні окремого пацієнта – для аналізу результатів індивідуальних досліджень та надання лікарю повної, релевантної й ненадлишкової інформації для оцінки поточного стану пацієнта. Приклад наведено на рис. 1.3.

2. На рівні планування лікування – коли необхідно інтегрувати результати кількох досліджень, порівняти динаміку стану пацієнта та виробити оптимальну терапевтичну стратегію.

3. На рівні популяційного аналізу – для виявлення тенденцій розвитку певних захворювань у масштабах країни чи світу та обґрунтування рішень щодо впровадження заходів охорони здоров'я.

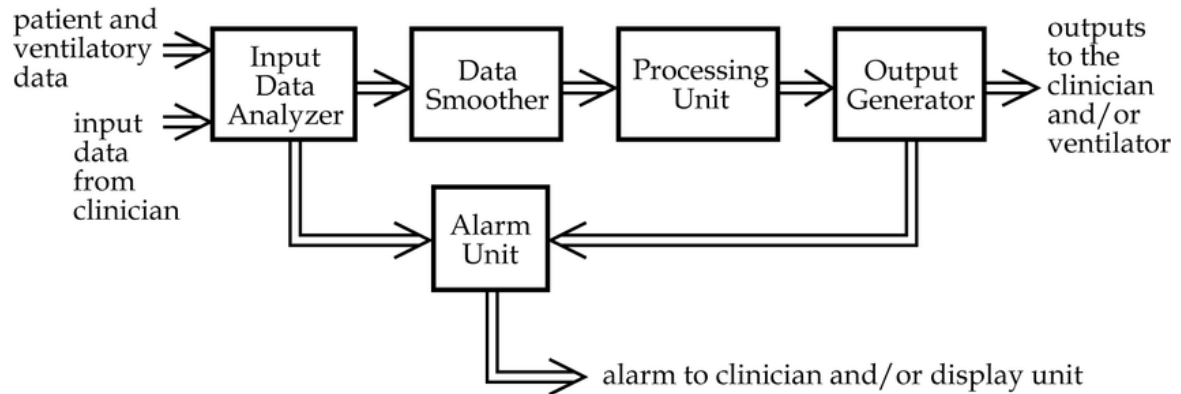


Рис. 1.3 – Схематична діаграма клінічної ІСППР для механічної вентиляції легенів [29]

У даному дослідженні розглядаються саме ІСППР першого рівня, орієнтовані на підтримку клінічної діагностики на основі медичних зображень.

В українському науковому просторі є низка робіт, присвячених розробці та розвитку ІСППР. Зокрема, дослідження українських науковців описують використання адаптивних онтологій для побудови інтелектуальних систем, що дозволяє динамічно оновлювати знання відповідно до нових даних і клінічного контексту [30].

Останніми роками спостерігається активний розвиток ІСППР, у яких ключову роль відіграють методи комп'ютерного зору, засновані на глибоких нейронних мережах. Такі системи поєднують автоматизований аналіз медичних зображень із модулями клінічної інтерпретації, формуючи основу сучасних інтелектуальних діагностичних підходів. Зокрема, у роботі [31] представлено архітектуру, що інтегрує згорткові нейронні мережі (CNN) для автоматичного виявлення патологій на томографічних зображеннях, що значно підвищує точність клінічних рішень при діагностиці пухлин головного мозку. Узагальнення підходів до застосування глибокого навчання в системах підтримки рішень наведено у [32], де

підкреслюється, що моделі типу CNN (ResNet, U-Net) стали стандартом для задач класифікації, сегментації та локалізації візуальних аномалій. Крім того, огляд [33] підкреслює, що саме комбінація глибоких моделей і СППР-компонентів формує нову парадигму автоматизованої підтримки прийняття рішень, де аналітичні можливості штучного інтелекту доповнюють експертні знання лікаря, підвищуючи ефективність та інтерпретованість діагностичного процесу.

1.3 Моделі глибокого навчання для класифікації, сегментації та детекції

Для розв'язання задач різного типу застосовують широкий спектр моделей глибокого навчання. Більшість із них базуються на архітектурах згорткових нейронних мереж, однак сучасні підходи дедалі частіше використовують моделі на основі механізмів уваги, зокрема візуальні трансформери.

У задачах обробки КТ-зображень виникає необхідність аналізу тривимірних даних. Існує кілька підходів до побудови систем класифікації таких даних. Найпростішим є використання двовимірних згорткових мереж із класифікацією кожного поперечного зрізу окремо [7]. Такий підхід дозволяє застосовувати добре відомі архітектури, як-от ResNet [34], DenseNet [35], EfficientNet [36]. Недоліком є обмеженість контексту вздовж третьої осі, що може суттєво знизити точність класифікації утворів, які займають декілька зрізів.

Модифікацією цього підходу є збільшення вхідної каналності 2D-згорткових мереж для врахування суміжних поперечних перерізів під час аналізу поточного (так званий підхід 2.5D) [7], [37]. Проте навіть такий підхід залишається обмеженим через двовимірну природу згортки, яка не дозволяє повноцінно захопити просторові закономірності вздовж третьої осі на глибших шарах мережі.

Альтернативним рішенням є використання гібридних нейронних мереж [7], у яких перша частина обробляє окремі поперечні зрізи, а наступний модуль агрегує інформацію з усіх зрізів по всьому КТ-скану. Наприклад, перший рівень може бути представлений 2D-згортковою мережею, а другий – рекурентною

мережею або трансформером. Такий підхід забезпечує охоплення всієї необхідної інформації для прийняття рішення, проте має певні обмеження: перший рівень не володіє глобальним уявленням про структуру об'єкта вздовж третьої осі, тоді як другий рівень оперує вже агрегованими ознаками без доступу до детальної інформації кожного окремого зрізу.

Саме такий підхід був розглянутий у роботі «Системи автоматичної медичної комп'ютерної діагностики з використанням методів штучного інтелекту» [38] та продемонстрував досить високу точність виявлення раку легенів за одночасно низького рівня необхідних обчислювальних ресурсів, особливо при великих розмірностях КТ-сканів. Проте, як було зазначено автором у контексті подальших досліджень, існує необхідність підвищення інтерпретованості системи, зокрема забезпечення можливості локалізації розміщення утворів, а також застосування методів напіваавтоматичного навчання для залучення додаткових даних до оптимізації моделі.

Найбільш комплексним з точки зору охоплення всієї просторової інформації є використання тривимірних згорткових мереж [39] або 3D-візуальних трансформерів. Вони дозволяють обробляти тривимірні об'єми як єдину структуру, проте вимагають значних обчислювальних ресурсів. Розмір одного КТ-скану може коливатись у межах від $[512, 512, 60]$ до $[512, 512, 800]$, і відповідно кількість параметрів моделі та обчислювальна складність зростають пропорційно розміру згортки. У більшості випадків обробка всього об'єму за один прохід є неможливою, тому застосовують методи «ковзного вікна» або аналізу цільових регіонів (Region of Interest, ROI).

Для задачі сегментації зазвичай використовують архітектури типу «Енкодер–Декодер». Енкодер виокремлює ключові ознаки з вхідного зображення, тоді як декодер відновлює просторову структуру, створюючи сегментаційну маску. Вибір енкодера і декодера залежить від підходу до обробки вхідних даних. У більшості випадків енкодер може бути перевикористаний із моделі класифікації,

тоді як декодер має дзеркальну архітектуру з використанням транспонованих згорток.

Класичною моделлю для медичних задач сегментації є U-Net [40], яка й досі демонструє високу точність. На основі її принципів було створено фреймворк nnU-Net [8], який автоматично адаптує гіперпараметри під конкретну задачу. Набувають популярності й архітектури, що використовують механізм уваги. Наприклад, Swin UNETR [41] поєднує концепцію *skip connections* із U-Net, але замість згорткових блоків застосовує Swin Transformer-блоки (рис 1.4). Архітектура SegFormer [42], побудована на концепції *deep supervision*, демонструє навіть вищу точність при меншій кількості параметрів.

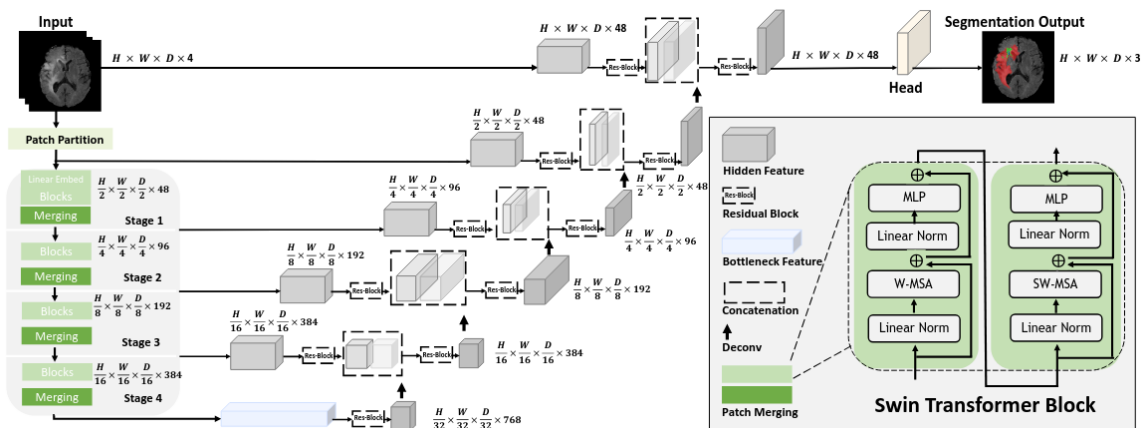


Рис. 1.4 – Архітектура моделі Swin UNETR для тривимірної медичної сегментації [41]

Для окремих задач аналізу медичних зображень доцільно використовувати моделі детекції, наприклад, коли потрібно ідентифікувати велику кількість малих об'єктів на одному зображенні. Прикладами є виокремлення нейронних клітин на мікроскопічних зображеннях [43] або детекція множинних утворень і метастазів на КТ-зображеннях. Перевагою використання детекторів є спрощений формат анотацій – *bounding box*, отримання якого є значно простішим і швидшим, ніж створення точної сегментаційної маски. Це дозволяє формувати більші навчальні вибірки за менших витрат часу експертів (радіологів, лікарів).

Серед найвідоміших архітектур для детекції можна відзначити RetinaNet [44], яку можна застосовувати як для 2D–, так і для 3D–зображень. Вона стала основою фреймворку nnDetection [9], що використовується для вирішення широкого кола задач детекції на медичних даних.

Отже, сучасні моделі демонструють високу точність і пропонують різноманіття підходів залежно від доступних обчислювальних ресурсів та обсягу навчальної вибірки. Водночас жоден із наведених методів окремо не забезпечує повного кінцевого результату, необхідного лікарю для прийняття клінічного рішення.

1.4 Інтерактивна сегментація

Сегментаційні моделі при створенні масок різних органів, утворів або пухлин завжди мають певний рівень похибки. Якщо задача полягає лише у локалізації об'єкта, такими похибками можна знехтувати, однак коли необхідно отримати радіологічні показники на основі сегментаційних масок (об'єм, площа найбільшого поперечного перерізу, максимальний діаметр тощо) – ці помилки стають критичними й унеможливають використання результатів клінічної ІСППР. Також система може пропустити певні утворення, які є важливими для постановки діагнозу. У такому випадку радіолог змушений повністю повторно проводити процес аналізу зображення вручну, що є надзвичайно трудомістким і ресурсозатратним. Альтернативним підходом є «корекція» результатів сегментаційної підсистеми за допомогою мінімального втручання лікаря для отримання задовільного результату всієї системи. Такі підходи можна узагальнити як інтерактивні системи (підсистеми) сегментації.

Моделі глибокого навчання на кшталт Segment Anything Model (SAM) [45] продемонстрували можливість роботи з додатковим вхідним сигналом, таким як точки, рамки, ініціалізаційні маски чи навіть текстові описи, що дозволяє сегментувати позначені об'єкти на зображеннях. Такі моделі відрізняються від класичної архітектури типу «Енкодер–Декодер» (розділ 1.3) наявністю додатко-

вого енкодера для обробки допоміжного сигналу. Ці підходи показали високу точність сегментації різноманітних об'єктів на багатьох вибірках, включно з тими, що значно відрізняються від тренувальної. Схожим напрямом є використання лише одного типу додаткового сигналу – наприклад, точок – з можливістю як додавання, так і видалення певних об'єктів із зображення [46]. Зазвичай такі моделі спеціалізовані під конкретний тип зображень, що забезпечує вищу точність у своїй галузі застосування, але обмежує їхню здатність узагальнення на інші модальності.

Моделі інтерактивної сегментації активно впроваджуються в аналіз медичних зображень. Для тривимірних томографічних даних виокремлюють кілька основних напрямів: «легкі» CNN-базовані підходи з поширенням маски між зрізами та модулями пам'яті/стабілізації; гібридні архітектури, які поєднують CNN і трансформери, приймають різні типи підказок і навчаються в парадигмі *human-in-the-loop*; а також класичні сегментаційні мережі, модифіковані для прийому додаткових вхідних каналів.

Представницькою моделлю першого типу є LIM-Net [47]: вона формує початкову 2D-маску за кліками користувача, потім поширює її крізь об'єм за допомогою ієрархічного модуля пам'яті та виконує *multi-round result fusion*, що зменшує накопичення помилок на далеких зрізах і водночас зберігає низькі вимоги до GPU-пам'яті. Прикладом другого напрямку є PRISM [48], який приймає як розріджені (точки, рамки, скрібли), так і щільні підказки (маски), використовує ітеративне навчання з вибором найнадійнішого прогнозу (*confidence learning*) та поверхневу мережу корекції для виправлення хибно маркованих вокселів. Гібридний енкодер (CNN+ViT) допомагає поєднувати локальні та глобальні ознаки в 3D-просторі. Представником третього напрямку є фреймворк nnInteractive [49], побудований на базі архітектури U-Net, який має додаткові канали для обробки кліків, скріблів і рамок користувача та використовує агентну модель навчання для симуляції дій користувача (рис 1.5).

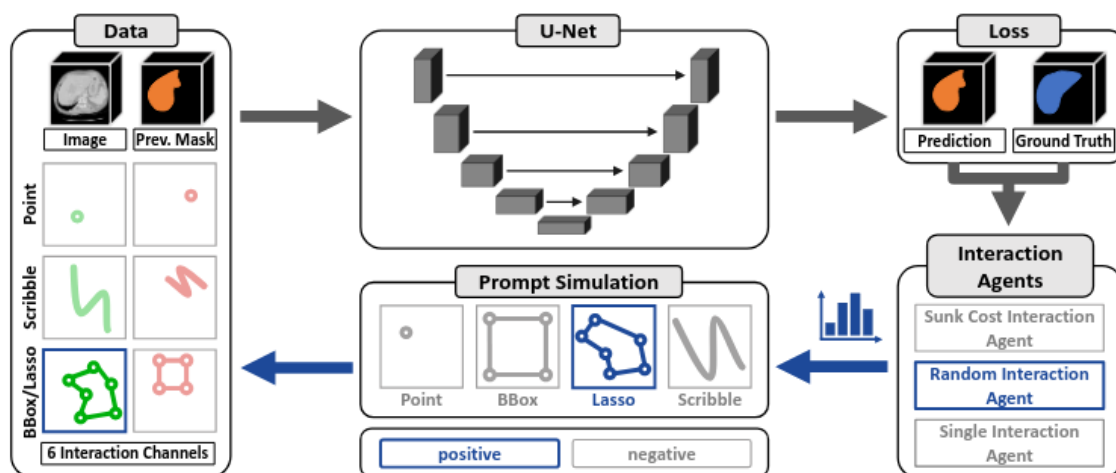


Рис. 1.5 – Схематичне зображення фреймворку nnInteractive [49]

Розглянуті підходи дозволяють локалізувати й сегментувати утворення за мінімального втручання лікаря. Проте більшість із них орієнтовані на отримання первинного (ініціалізуючого) сигналу від користувача, а не від попередньої моделі сегментації. Таким чином, перспективним і малодослідженим напрямом є розробка інтерактивних систем, здатних виправляти помилки базової моделі, використовуючи додаткову взаємодію з лікарем.

1.5 Ключові виклики в побудові клінічних ІСППР

Важливо виокремити основні виклики, які виникають як під час розробки окремих модулів, так і повноцінних клінічних ІСППР.

Фундаментальною проблемою є неповнота тренувальних і тестових вибірок. Цю проблему можна декомпозувати на такі аспекти.

1. Неповне охоплення різних протоколів досліджень (КТ із/без контрасту або з різними фазами контрастування), різних типів сканерів і їх технічних характеристик.
2. Відсутність деталізованої таксономії утворень або неточність визначення класів для певних вибірок. Наприклад, критерії Bosniak [6] для класифікації кіст нирок та Lung–RADS [5] для аналізу утворів легенового скрінінгу мають значно ширшу структуру, ніж проста класифікація «доброякісне/злаякісне».

3. Відсутність повних метаданих, наприклад, інформації про стать, вік, супутні захворювання пацієнта; тип пухлини – первинна чи метастатична; фазу контрасту дослідження тощо.

Усі наведені фактори ускладнюють побудову моделей для аналізу медичних зображень, здатних забезпечувати робастні результати в різних клініках та за різних протоколів дослідження. Більше того, вони створюють проблему об'єктивної оцінки системи для підтвердження її здатності працювати коректно в кожному окремому медичному закладі. Це безпосередньо пов'язано з явищем доменного зсуву (*domain shift*) [50] – відмінністю розподілів між тренувальною та тестовою вибірками, коли модель, оптимізована на даних певних клінік, демонструє значно гірші результати на даних з інших медичних центрів. Для подолання цього виклику необхідно формувати максимально різноманітні вибірки медичних зображень, що охоплюють широкий спектр типів утворень, протоколів і обладнання, які реально використовуються в клінічній практиці.

Ще одним важливим викликом у побудові клінічних ІСППР є невідповідність вихідних результатів потребам лікарів. Моделі зазвичай генерують надлишкову технічну інформацію (сегментаційні маски, ймовірності класів), тоді як у клінічній практиці необхідні чіткі морфологічні характеристики утворів (об'єм, діаметр, тип тканини) та оцінка рівня ризику для пацієнта. Додатковою проблемою є недовіра до систем і моделей, що функціонують як «чорні скриньки» без механізмів інтерпретації або корекції результатів. У цьому контексті інтерактивні підсистеми можуть суттєво полегшити інтеграцію таких систем, забезпечуючи зручний інтерфейс для взаємодії лікарів і радіологів з можливістю редагування або уточнення результатів.

Також варто відзначити нестачу повноцінних клінічних ІСППР – більшість сучасних досліджень зосереджені на розв'язанні окремих задач (сегментація, класифікація тощо), тоді як клінічним закладам потрібні комплексні системи, що охоплюють увесь цикл: від попередньої обробки даних і локалізації до класифікації та автоматичного формування структурованого клінічного звіту.

Отже, попри значний прогрес у застосуванні глибокого навчання для аналізу медичних зображень, наразі існує низка критичних факторів, які значно ускладнюють їхнє практичне впровадження в клінічну діяльність.

1.6 Суміжні задачі медицини й біології, які потребують залучення глибоких нейронних мереж

Окрім основної задачі детекції та класифікації небезпечних утворів у легенях та інших органах на основі КТ–зображень, у медицині й біології існує низка інших завдань, частина з яких безпосередньо не пов'язана з процесом діагностики. Проте їхнє вирішення дозволило б побудувати «структурні» мапи людського організму, що сприятиме пришвидшенню розробки ліків, діагностичних критеріїв і методологій, а також поглибленню знань про організм людини. Інша група задач безпосередньо стосується діагностики, проте використовує не КТ чи МРТ–скани, а мікроскопічні зображення або фотографії певних органів. У таких випадках застосування алгоритмів машинного та глибокого навчання може також відіграти ключову роль.

Задача виокремлення функціональних тканинних одиниць (ФТО) (наприклад, клубочків у нирці або альвеол у легенях) є однією з ключових для побудови еталонного атласу людини на клітинному рівні. Приклади різних ФТО наведені на рис. 1.6. Розв'язання цієї задачі дає змогу отримати глибше розуміння взаємозв'язків між клітинною та тканинною організацією, що, у свою чергу, сприятиме розробці нових лікарських засобів і протоколів лікування. Водночас процес виокремлення таких структур є надзвичайно ресурсозатратним – у середньому в одній нирці може міститися понад мільйон клубочків. Результати змагання з сегментації ФТО [51] продемонстрували, що алгоритми глибокого навчання здатні вирішувати подібні задачі з високою точністю.

Попри розвиток і поширення медичних технологій, значна частина населення світу досі не має доступу навіть до базового рівня медичної допомоги. У подоланні цієї проблеми потенційно можуть допомогти алгоритми штучного

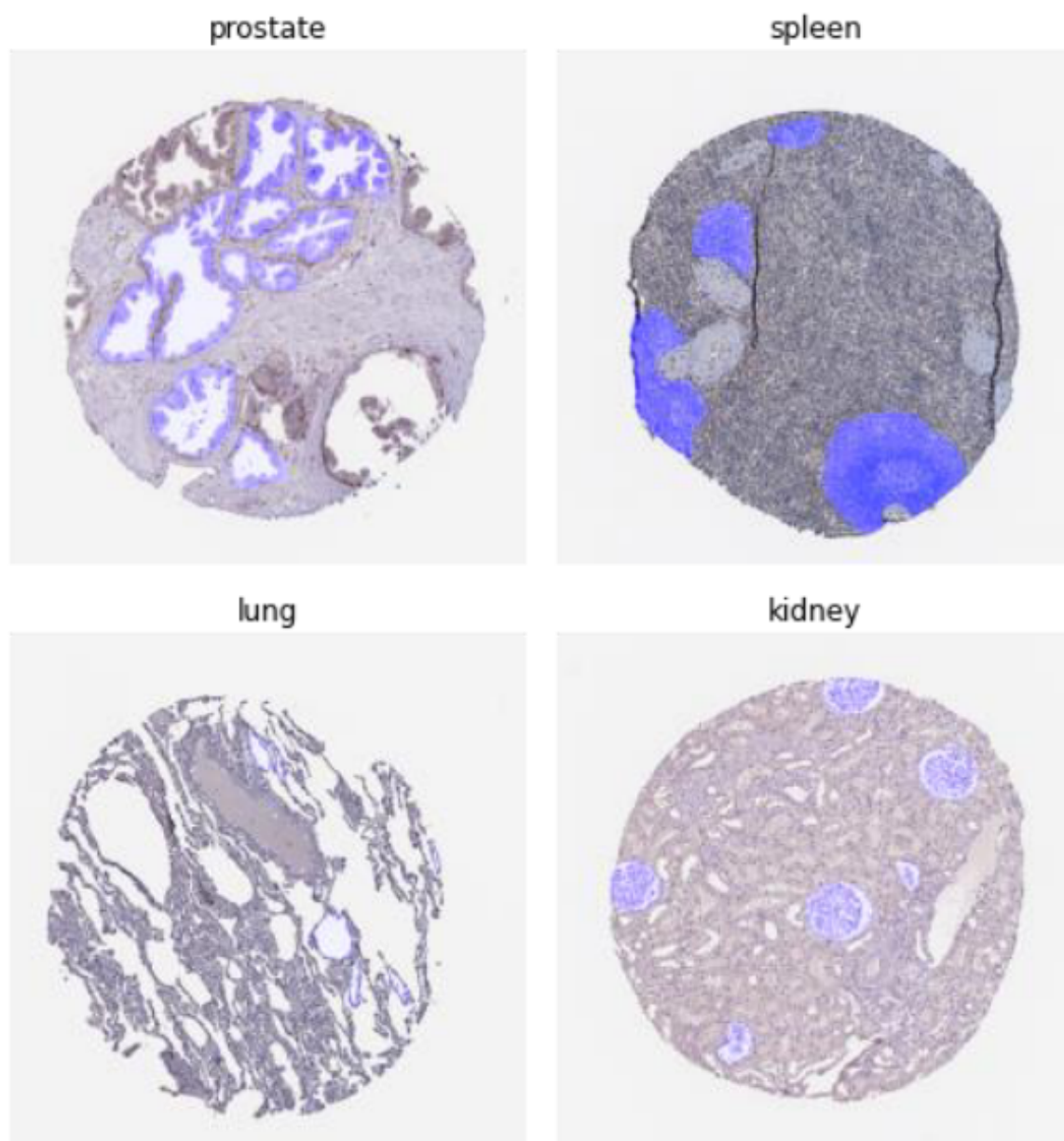


Рис. 1.6 – Мікроскопічне зображення функціональних тканинних одиниць [52]

інтелекту. Прикладом такої задачі є виявлення раку шкіри (меланому). Зазвичай для постановки діагнозу потрібні зображення з дерматоскопа та результати гістологічного аналізу, що є дорогою й трудомісткою процедурою. Альтернативою може бути 3D-фотографія всього тіла, яку можна замінити звичайними фотографіями ділянки шкіри з підозрою на меланому. Проте такі зображення містять значно менше інформації, тому діагностика за ними є складнішою. Однак результати змагання [53] показали, що використання ансамблів моделей різної

природи може забезпечити високу точність прогнозування навіть на звичайних фото для діагностики меланоми.

Іншим прикладом є застосування алгоритмів глибокого навчання для дослідження структури мозку та діагностики неврологічних розладів, зокрема нейродегенеративних захворювань, таких як хвороба Альцгеймера, а також пухлин головного мозку. Оцінка ефективності лікування таких патологій є надзвичайно складною. Одним із визнаних методів є аналіз нейронних клітин за допомогою світлової мікроскопії – доступного та неінвазивного підходу. Проте сегментація окремих нейронних клітин на мікроскопічних зображеннях залишається складним і трудомістким завданням. Результати змагання [43] демонструють, що методи детекції та інстанс-сегментації можуть доволі успішно розв'язувати такі задачі.

Алгоритми глибокого навчання також можуть бути адаптовані до нових типів медичних зображень, наприклад, HiP-CT (Hierarchical Phase-Contrast Tomography). Такі зображення застосовуються для побудови Єдиної координатної системи судин – Vasculature Common Coordinate Framework (VCCF). Ці координатні системи дають змогу досліджувати судинну систему людини та її вплив на різні типи клітин. Оскільки судинна система має складну розгалужену структуру, ручна сегментація її елементів на зображеннях різних органів є надзвичайно трудомістким процесом. Саме тому алгоритми глибокого навчання, здатні обробляти тривимірні медичні зображення, можуть ефективно застосовуватись для побудови таких мап у масштабних об'ємах даних [54].

Перелічені вище приклади є лише частиною спектра задач медицини та біології, де глибокі нейронні мережі можуть суттєво оптимізувати роботу висококваліфікованих експертів. Важливо зазначити, що в деяких випадках (наприклад, у задачах виокремлення мільйонів окремих тканинних структур або клітин) ручне виконання таких операцій практично неможливе, і саме штучний інтелект відкриває можливість їх реалізації.

1.7 Висновки розділу 1

Розробка, вдосконалення та інтеграція клінічних ІСППР дають змогу суттєво оптимізувати та підвищити точність процесу діагностики багатьох захворювань, зокрема онкологічних уражень різних органів людського організму. Водночас такі системи можуть прискорити розв’язання суміжних задач медицини й біології, що сприятиме подальшому розвитку медичної науки. З огляду на тенденцію зростання чисельності населення Землі, клінічні ІСППР у найближчому майбутньому можуть стати необхідною складовою для забезпечення належного рівня медичної допомоги.

Разом із тим, складність задач, пов’язаних із побудовою клінічних ІСППР, постійно зростає, і багато з них уже потребують використання моделей та методів нечіткої логіки й штучного інтелекту, зокрема глибоких нейронних мереж. Такі підходи дають змогу ефективно розв’язувати широкий спектр задач, зокрема сегментацію, детекцію, класифікацію та реєстрацію, які базуються на аналізі медичних зображень різних типів.

Провівши аналіз поточного стану досліджень у галузі обробки медичних зображень та побудови клінічних ІСППР, було виокремлено такі основні прогалини.

1. Відсутність повноцінних ІСППР, які на вхід отримують сирі медичні зображення (наприклад, DICOM-файли), а на вихід надають радіологу всю необхідну інформацію про всі виявлені утвори для прийняття рішень за критеріями RECIST [15], Bosniak [6] або системою Lung-RADS [5]. Такі системи мають також забезпечувати автоматизовану обробку даних із використанням сукупності моделей глибокого навчання.

2. Відсутність інтерактивної складової в клінічних ІСППР, яка дозволяла б ефективно виправляти похибки системи з мінімальним втручанням користувача, або ж їх низька інтегрованість із основною ІСППР.

3. Відсутність спеціалізованих методів адаптації моделей глибокого навчання до різних типів сканерів, протоколів дослідження, доменних зсувів медичних даних, а також до похибок попередніх моделей у складі комплексних ІСППР.

4. Недостатня точність низки моделей глибокого навчання на окремих типах медичних зображень, зокрема для малих утворів у легенях або фотографічних знімків раку шкіри.

Виходячи з наведених вище положень, виокремлені основні завдання дисертаційної роботи.

1. Побудова повноцінної ІСППР для діагностики раку легень, яка прийматиме на вхід сирі медичні зображення та надаватиме на виході характеристики виявлених утворів і оцінку їх рівня ризику для пацієнта. В основі розробленої клінічної ІСППР передбачається використання комплексу моделей глибокого навчання, орієнтованих на розв'язання ключових задач – виявлення та характеристики утворів.

2. Інтеграція інтерактивної складової до ІСППР, яка враховуватиме результати статичної моделі з метою забезпечення можливості корекції результатів із мінімальною кількістю дій з боку радіолога.

3. Розробка методів і підходів для підвищення точності та адаптації моделей глибокого навчання до різних сценаріїв їх практичного застосування в задачах діагностики раку легень, а також для мінімізації впливу похибок попередніх моделей у межах системи.

4. Розробка методів для розв'язання задачі доменної адаптації під час сегментації функціональних тканинних одиниць, а також для побудови системи виявлення раку шкіри на основі фотографічних знімків.

2 ІСППР ДЛЯ ДІАГНОСТИКИ РАКУ ЛЕГЕНЬ

2.1 Постановка задачі побудови ІСППР та вимоги до спроектованої системи

На сьогодні рак легенів є одним із найбільш розповсюджених і смертельних видів раку, особливо якщо його не виявлено на ранніх стадіях [4]. Тому рання діагностика раку легенів є критично важливою для ефективного лікування. З огляду на це, багато країн ініціюють програми легеневого скринінгу на національному рівні [55], [56]. Проте такі масштабні програми охоплюють значну частину населення (десятки, а іноді й сотні тисяч людей), що потребує залучення великої кількості радіологів, яких часто бракує. Саме тому інтеграція клінічних ІСППР у подібні заходи національного рівня є вже критично необхідною.

Дослідження легеневого скринінгу зазвичай проводяться на основі КТ-знімків із низьким рівнем опромінення (low-dose CT) [56]. Після отримання КТ-зображень радіологи зазвичай керуються системою Lung-RADS [5] для виявлення потенційно небезпечних утворень, їх категоризації та подальшого відстеження прогресії. Для застосування цього критерію радіолог повинен:

- виокремити всі утворення;
- здійснити вимірювання діаметра кожного утворення;
- залишити утворення певного типу (підозрілі на злоякісні);
- відповідно до кількості утворень певного типу та їх діаметра віднести пацієнта до відповідної категорії ризику;
- визначити план подальшого обстеження відповідно до категорії Lung-RADS.

Критерії Lung-RADS охоплюють два основні сценарії застосування: коли дослідження проводиться вперше для конкретного пацієнта; коли наявні результати попереднього дослідження, які можна використати для уточнення категорії Lung-RADS.

У цій роботі пропонується ІСППР, сфокусована на автоматизації перших трьох етапів системи Lung–RADS, а також на роботі з випадками, коли відсутні попередні дослідження. Саме перші три кроки є найбільш трудомісткими для радіологів, а перший сценарій застосування відповідає масовим програмам легеневого скринінгу.

Отже, клінічна ІСППР повинна приймати на вхід КТ–знімок (у вигляді набору DICOM–файлів) і надавати на виході радіологу всі знайдені утворення в області легень разом із їхніми вимірами (зокрема, діаметром) та категорією, що відображає ступінь підозри на злоякісність. Крім того, важливо додавати координати центроїда кожного утворення у просторі вокселів для спрощення його подальшої локалізації радіологом.

Водночас розроблена система повинна бути інтерактивною, тобто надавати радіологу можливість за допомогою кліку на відповідний воксель КТ–скану вилучати певну область утворення або все утворення цілком, а також, за потреби, додавати нові області. Схематичне зображення вхідних і вихідних даних наведено на рис. 2.1.

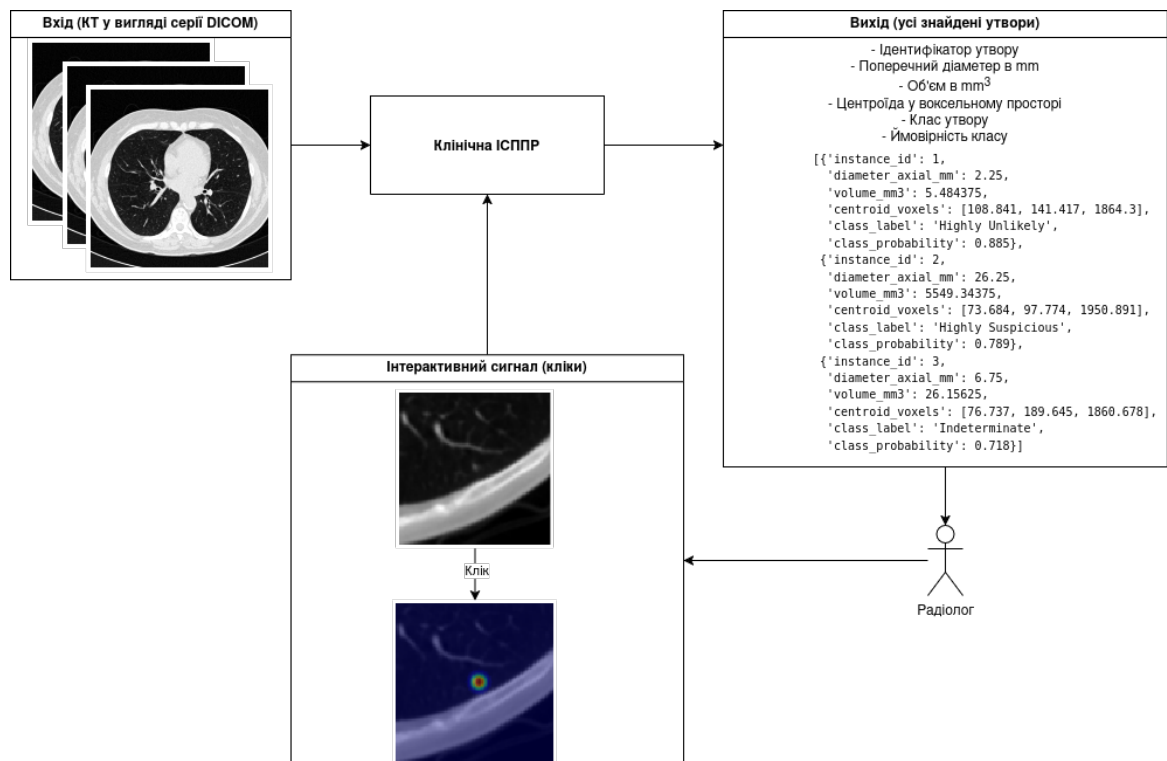


Рис. 2.1 – Схематичне зображення вхідних і вихідних даних клінічної ІСППР

Важливо зазначити, що підсистема класифікації у запропонованій ІСППР наразі не формує таксономію класів відповідно до критерію Lung–RADS через відсутність та складність отримання відповідної навчальної вибірки (ця проблема була детальніше описана в розділі 1.5). На даному етапі використовується спрощена таксономія, що описує «рівень» ризику утвору для пацієнта. Водночас, у разі появи необхідної вибірки для розширення або зміни таксономії, система може бути легко адаптована без потреби модифікації архітектури або методології оптимізації моделей.

2.2 Методологія побудови ІСППР

Керуючись поставленими задачами й вимогами з попереднього підрозділу, можна сформулювати основні засади методології побудови клінічної ІСППР.

1. Стандартизована робота з клінічними даними – ІСППР має приймати на вхід медичні зображення у стандартизованому форматі обміну медичною інформацією DICOM, водночас система повинна зберігати додаткову метаінформацію з цих файлів. За необхідності введення додаткової інформації це може бути реалізовано шляхом додавання на вхід JSON–файлу з відповідними даними. На виході ІСППР надає JSON–файл з характеристикою кожного підозрілого утвору та, за необхідності, NifTi–маски цих утворів для більш детального аналізу.

2. Стандартизованість і доступність даних у межах системи – кожна підсистема ІСППР має розширювати JSON–файл новою отриманою інформацією (наприклад, класом утвору або його радіометричними ознаками). Усі проміжні медичні маски та зображення повинні зберігатися у форматі NifTi з фізичним розміром вокселя та орієнтацією, ідентичними до вхідного зображення. ІСППР може надавати на вихід (за необхідності) всі проміжні результати для більш детального аналізу або аналізу помилок.

3. Масштабованість і розширюваність – під час проєктування підсистем необхідно враховувати два попередні пункти з метою коректного проєктування

схем вхідних і вихідних даних. Такий підхід забезпечує просту інтеграцію нових підсистем для подальшого розширення функціоналу ІСППР.

4. Перевикористання підсистем – за можливості підсистеми мають бути спроектовані максимально гнучко з метою спрощення їх адаптації до діагностики інших видів захворювань. У загальному випадку для адаптації підсистеми необхідно: оптимізувати модель глибокого навчання для нової задачі; змінити гіперпараметри системи відповідно до статистичних характеристик нової вибірки.

5. Інтерактивна складова – розроблена система має містити інтерактивну складову або, принаймні, можливість її подальшої інтеграції. При цьому інтерактивна підсистема повинна орієнтуватися на дві основні метрики: максимізацію точності виявлення після її застосування; мінімізацію кількості взаємодій з боку експерта.

6. Інтеграція моделей глибокого навчання – з урахуванням специфіки більшості медичних задач, а саме їх складної структурованості та високої складності навіть для експертів, підсистеми мають використовувати моделі глибокого навчання та ансамблі таких моделей для розв'язання більшості підзадач.

7. Оцінка ефективності та клінічна валідація – з огляду на високу ціну помилки клінічних ІСППР має бути розроблений протокол всебічної оцінки системи. Залежно від типу діагностованого захворювання він може відрізнятися, проте повинен включати такі основні етапи: оцінку кожної підсистеми окремо на відповідних тестових і валідаційних вибірках; оцінку системи в цілому; оцінку системи з використанням інтерактивної складової (за її наявності); проведення клінічних випробувань, у межах яких оцінюються як точність системи, так і зручність її використання.

Використовуючи наведену методологію, у подальшому буде побудовано ІСППР для діагностики раку легень.

2.3 Загальна архітектура ІСППР

Розроблена клінічна ІСППР складається з низки підсистем, кожна з яких отримує результати роботи попередніх модулів і формує вхід для наступних або фінальний вихід для радіолога.

1. **Підсистема обробки та підготовки медичних зображень** – перетворює серію DICOM-файлів у формат NifTi для подальшої роботи з КТ-сканом як із тривимірним зображенням.
2. **Підсистема сегментації області легенів** – виокремлює зону легень на КТ-скані для подальшої обробки лише цієї області.
3. **Підсистема сегментації утворів у легенях** – виконує сегментацію утворів у межах зони легень.
4. **Підсистема характеристики та фільтрації утворів** – виділяє об'єкти (утворів) з отриманих сегментаційних масок і розраховує їхні параметри. За потреби відфільтровує надто малі утворення.
5. **Підсистема класифікації утворів** – визначає клас утворів та ймовірність його належності до цього класу.
6. **Підсистема постобробки результатів** – виконує постобробку результатів, зокрема об'єднує об'єкти, які відповідають одному й тому ж утвору.
7. **Підсистема інтерактивної сегментації утворів** – у разі незадовільного результату отримує коригувальний сигнал від радіолога (у вигляді кліків), оновлює результати сегментації та активує всі наступні підсистеми для оновлення кінцевих результатів.

Схематичне представлення загальної архітектури системи наведено на рис. 2.2.

Важливо зазначити кілька ключових характеристик розробленої системи.

1. **Модульність і гнучкість.** Підсистеми можуть бути легко замінені за умови збереження формату вхідних та вихідних даних. Наприклад, підсистема класифікації може бути оновлена моделями з новою таксономією, а групу підсистем

«сегментація утворів у легенях» → «характеристика та фільтрація утворів»
можна замінити на єдину підсистему детекції.

2. Збереження проміжних результатів. Усі дані зберігаються в системі, що дає змогу кожній підсистемі мати доступ до результатів попередніх етапів. Це дозволяє, наприклад, уточнювати результати класифікаційної підсистеми, використовуючи додаткову інформацію про параметри утворів, отриману з підсистеми характеристики та фільтрації.

3. Ізольованість підсистем. Кожна підсистема може мати власну логіку попередньої обробки КТ-скану (у форматі NifTi) та постобробки результатів. Це дає можливість незалежно змінювати фізичний розмір вокселя, орієнтацію або просторове масштабування відповідно до вимог конкретних моделей глибокого навчання, що використовуються в різних підсистемах.

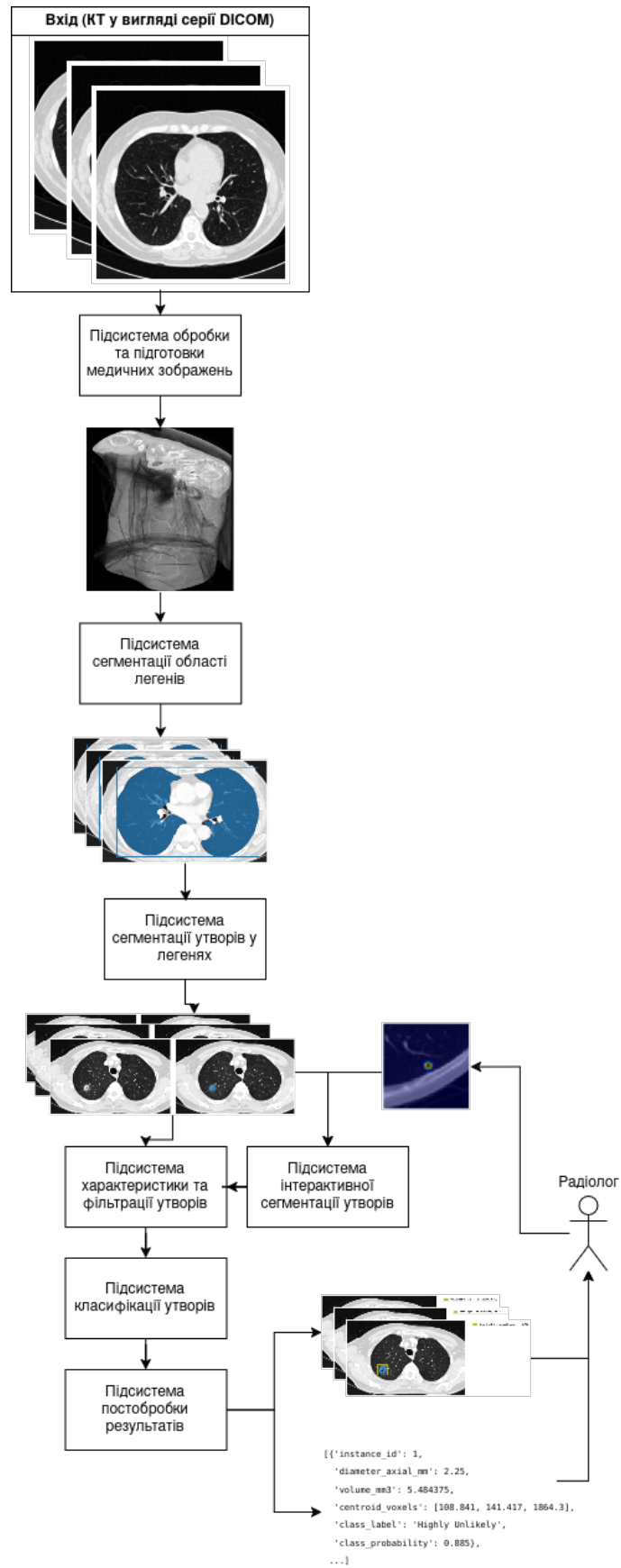


Рис. 2.2 – Схематичне зображення загальної архітектури клінічної ІСППР для діагностики утворів у легенях

2.4 Підсистема обробки та підготовки медичних зображень

Основна задача підсистеми обробки та підготовки медичних зображень – конвертувати серію DICOM–файлів у формат NifTi, який складається з тривимірного об’єму вокселів, а також метаінформації, що описує геометричні характеристики: фізичні розміри вокселя, орієнтацію КТ–скану відносно усталеної системи координат сканера, а також афінну матрицю зображення. Перетворення виконується з використанням пакета *dicom2nifti* [57], у результаті чого на виході отримується NifTi–об’єкт.

Окрім самої трансформації, ця підсистема виконує низку перевірок валідності отриманих даних, а також декомпресію за потреби.

1. Перевірка наявності таких тегів: SeriesInstanceUID, Modality, ImageOrientationPatient, TransferSyntaxUID.
2. Перевірка, що кількість DICOM–файлів більша за 1.
3. Перевірка, що модальність дослідження відповідає КТ.
4. Перевірка, що перше та передостаннє значення атрибуту ImageOrientationPatient дорівнює 1 або -1 . Атрибут Image Orientation (Patient) (0020,0037) визначає косинуси напрямків першого рядка та першого стовпця зображення відносно координат пацієнта. Ці атрибути мають бути задані у вигляді пари: спочатку значення косинусів напрямку рядка вздовж осей x , y та z відповідно, а потім – значення косинусів напрямку стовпця вздовж тих самих осей. Значення $[(1/\sqrt{2} \quad -1/\sqrt{2})^T, 0, 0, 0, (1/\sqrt{2} \quad -1/\sqrt{2})^T, 0]$ означають, що КТ–скан був реконструйований по поперечних зрізах.
5. Перевірка значення TransferSyntaxUID, щоб визначити тип компресії зображення. Якщо TransferSyntaxUID починається з 1.2.840.10008.1.2.4, це означає, що зображення у полі PixelData були стиснуті за допомогою JPEG, і їх необхідно декомпресувати.

Приклади деяких полів заголовка NifTi (header), а також афінної матриці наведено на рис. 2.3. Приклади різних зрізів КТ-скану, отриманих із NifTi-зображення, наведено на рис. 2.4.

```

<class 'nibabel.nifti1.Nifti1Header'> object, endian='<'
sizeof_hdr      : 348
data_type       : b''
db_name         : b''
extents         : 0
session_error   : 0
regular         : b''
dim_info        : 0
dim             : [ 3 399 250 374 1 1 1 1]
intent_p1       : 0.0
intent_p2       : 0.0
intent_p3       : 0.0
intent_code     : none
datatype        : int16
bitpix          : 16
slice_start     : 0
...
[[-7.500e-01  0.000e+00  0.000e+00  1.628e+02]
 [ 0.000e+00  7.500e-01  0.000e+00  3.521e+01]
 [ 0.000e+00  0.000e+00  7.500e-01  1.718e+03]
 [ 0.000e+00  0.000e+00  0.000e+00  1.000e+00]]

```

Рис. 2.3 – Приклад метаданих NifTi-зображення та афінної матриці, отриманих після перетворення з DICOM: зліва – афінна матриця; справа – метадані.

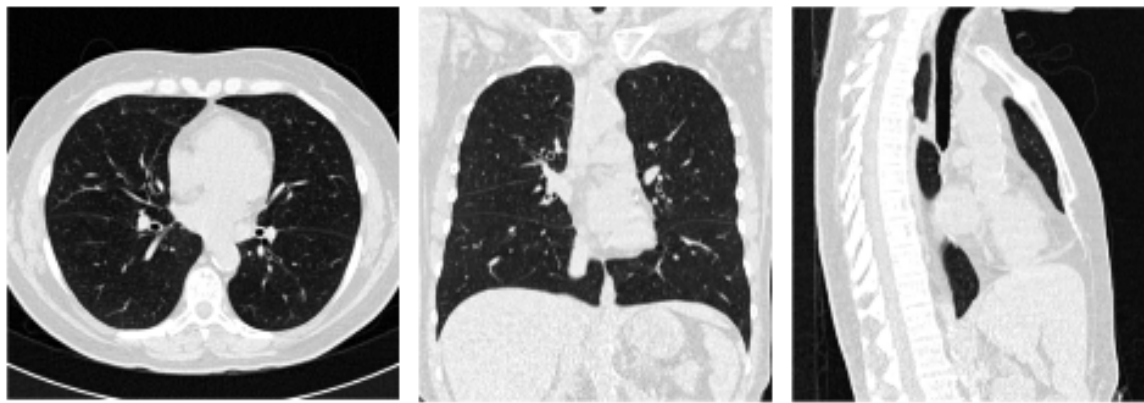


Рис. 2.4 – Різні зрізи одного КТ-скану (з ізотропним розміром вокселя) з NifTi-зображення: зліва направо – аксіальний, корональний та сагітальний

2.5 Підсистема сегментації області легенів

Для більш ефективної роботи подальших підсистем необхідно локалізувати регіон легенів. Під час КТ-дослідження грудної клітки легені займають більшу частину зображення, проте у випадку КТ-дослідження всього тіла розмір скану буде значно більший, і тому важливо виділити лише потрібну область. Навіть у випадку дослідження грудного відділу КТ-скан часто охоплює частини тіла, що

виходять за межі зони грудей (рис. 2.4), тому область подальшого аналізу можна суттєво зменшити.

Зменшення області дослідження має дві ключові переваги:

- зменшення кількості хибнопозитивних утворів, які можуть бути виявлені в інших анатомічних регіонах КТ-зображення;
- пришвидшення роботи подальших підсистем, оскільки обсяг даних для обробки зменшується.

В основі підсистеми сегментації області легенів використовується нейронна мережа *TotalSegmentator* [58]. Вона була оптимізована на вибірці з 1082 КТ-сканів для виділення 104 різних анатомічних структур (27 органів, 59 кісток, 10 м'язів і 8 судин). Архітектурно вона базується на фреймворку *nnU-Net* [8] і демонструє високу точність – середній показник Dice score на тестовій вибірці становить 0.943.

Після отримання маски на 104 класи з неї обираються класи, що відповідають зонам легенів: верхня частка лівої легені (*lung_upper_lobe_left*, id: 10), нижня частка лівої легені (*lung_lower_lobe_left*, id: 11), верхня частка правої легені (*lung_upper_lobe_right*, id: 12), середня частка правої легені (*lung_middle_lobe_right*, id: 13), нижня частка правої легені (*lung_lower_lobe_right*, id: 14). За належністю до цих класів формується бінарна маска. На основі цієї маски обчислюється рамка (bounding box) зони легенів, до якої додається 3% паддінгу для збільшення контексту та компенсації можливих неточностей сегментації *TotalSegmentator* на краях. По збільшеній рамці виконується обрізання тривимірного зображення. Повний алгоритм описано у формулі 2.1, а приклад результату наведено на рис. 2.5.

$$M_{\text{lungs}}(x, y, z) = \begin{cases} 1, & \text{якщо } L(x, y, z) \in \{10, 11, 12, 13, 14\}, \\ 0, & \text{інакше;} \end{cases}$$

$$x_{\min}, x_{\max} = \min / \max \{x \mid \exists y, z : M_{\text{lungs}}(x, y, z) = 1\},$$

$$y_{\min}, y_{\max} = \min / \max \{y \mid \exists x, z : M_{\text{lungs}}(x, y, z) = 1\},$$

$$z_{\min}, z_{\max} = \min / \max \{z \mid \exists x, y : M_{\text{lungs}}(x, y, z) = 1\},$$

$$I_{\text{crop}} = I[(x_{\min} - 0.03 W) : (x_{\max} + 0.03 W),$$

$$(y_{\min} - 0.03 H) : (y_{\max} + 0.03 H),$$

$$(z_{\min} - 0.03 D) : (z_{\max} + 0.03 D)]. \quad (2.1)$$

де $L(x, y, z)$ – початкова сегментаційна маска з 104 класами, отримана з виходу *TotalSegmentator*;

$M_{\text{lungs}}(x, y, z)$ – бінарна маска зон легенів, сформована на основі класів $\{10, 11, 12, 13, 14\}$;

I – початкове тривимірне КТ–зображення у форматі NifTi;

I_{crop} – обрізане зображення легеневої області.

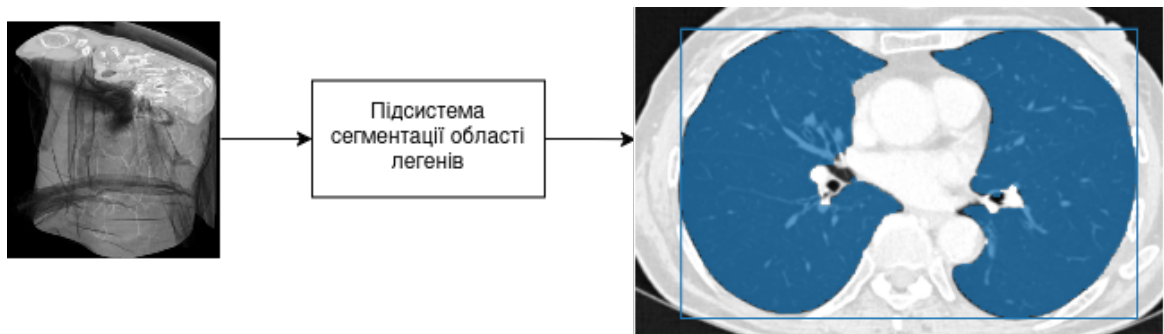


Рис. 2.5 – Результат роботи підсистеми сегментації області легенів: виділення зони легенів із повного КТ–зображення та обрізання за розрахованою рамкою

2.6 Підсистема сегментації утворів у легенях

Підсистема сегментації утворів у легенях є однією з ключових складових розробленої клінічної ІСППР. Її основна задача – побудова бінарної сегментаційної маски всього обрізаного КТ–скану, у якій кожен позитивний воксель відповідає регіону утвору.

Як базову архітектуру глибокої нейронної мережі було обрано згорткову модель на основі тривимірних згорток – *DynUNet* [59], створену за ідеями фреймворку *nnU-Net* [8]. Нейронна мережа була ініціалізована з параметрами, наведеними у таблиці 2.1. Як нормалізуючий шар використано *Instance Normalization*.

Таблиця 2.1 – Конфігурація моделі DynUNet 3D

Шар мережі	Розмір ядра	Крок (stride)
Шари зменшення розмірності (Downsampling)		
Вхідний блок	[3, 3, 3]	[1, 1, 1]
Downsample 0	[3, 3, 3]	[2, 2, 2]
Downsample 1	[3, 3, 3]	[2, 2, 2]
Downsample 2	[3, 3, 3]	[2, 2, 2]
Downsample 3	[3, 3, 3]	[2, 2, 2]
Вузьке місце (Bottleneck)	[3, 3, 3]	[2, 2, 2]
Шари збільшення розмірності (Upsampling)		
Upsample 0	[2, 2, 2]	–
Upsample 1	[2, 2, 2]	–
Upsample 2	[2, 2, 2]	–
Upsample 3	[2, 2, 2]	–
Upsample 4	[2, 2, 2]	–

Основне дослідження з оптимізації сегментаційної моделі було зосереджено на виборі функції втрат [60]. Було виявлено, що класичні функції втрат, такі

як Dice, Jaccard або Cross-Entropy, не забезпечують високої точності для даної задачі. Це пояснюється тим, що більшість легеневих утворів мають дуже малий розмір і займають незначну частку всього об'єму КТ-скану. Тому доцільно застосовувати функції втрат, спеціально адаптовані до задачі з сильним дисбалансом класів.

Серед сучасних підходів найкращі результати показали функція *Boundary DoU* [61] та її лінійна комбінація з *Focal loss* [44]. Проте в межах цього дослідження було запропоновано модифікацію – *Boundary DoU++*, яка вводить додатковий параметр γ у формулу функції втрат. Зміна цього параметра дозволяє контролювати чутливість функції до надмірно «впевнених» передбачень і підвищувати точність моделі на незбалансованих вибірках (2.4).

$$\begin{aligned} L_{\text{Boundary DoU}} &= 1 - \frac{(1 - \alpha) \text{TP}}{\text{FP} + \text{FN} + (1 - \alpha) \text{TP}} \\ &= 1 - \frac{(1 - \alpha) S_I}{S_D + (1 - \alpha) S_I} \\ &= \frac{S_D}{S_D + (1 - \alpha) S_I} \end{aligned} \quad (2.2)$$

$$S_D = \text{FP} + \text{FN} = \sum_{i=1}^N p_i \cdot (1 - y_i) + \sum_{i=1}^N (1 - p_i) \cdot y_i \quad (2.3)$$

$$L_{\text{Boundary DoU++}} = \frac{S_{Dfocal}}{S_{Dfocal} + (1 - \alpha) \cdot S_I} \quad (2.4)$$

$$S_{Dfocal} = \text{FP}_{focal} + \text{FN}_{focal} = \sum_{i=1}^N (p_i \cdot (1 - y_i))^\gamma + \sum_{i=1}^N ((1 - p_i) \cdot y_i)^\gamma \quad (2.5)$$

Окрім спеціальної функції втрат, було використано низку додаткових методів для покращення точності моделі.

1. Ресемплінг усіх зображень до одного фізичного розміру вокселя як під час тренування, так і під час інференсу, що дозволяє стандартизувати дані.

2. Використання більшого вікна по аксіальній площині, ніж по осі z , оскільки більшість утворів витягнуті саме в цій площині. Також під час інференсу застосовується ковзне вікно з 50% перекриттям і гаусівським згладжуванням, що зменшує вагу сегментаційних масок на краях вікна.

3. Спеціальний механізм вибірки 3D-патчів під час тренування: випадковим чином обирається число з біноміального розподілу з $p = 0.75$. Якщо результат дорівнює 1 – вибирається регіон із утвором, якщо 0 – без утвору. Самі регіони відбираються випадково та рівномірно з усього об'єму КТ-скану; у випадку 1 перевіряється до 10 кандидатів, поки не буде знайдено регіон, що містить вокселі утвору. Такий підхід дозволяє збільшити кількість позитивних вокселів і зменшити дисбаланс вибірки.

4. Використання різноманітних аугментацій – геометричних, інтенсивнісних та шумових. Це підвищує різноманітність тренувальних прикладів і робастність моделі до різних типів сканерів і протоколів дослідження. Додатково застосовується усереднення по аугментаціях під час інференсу (Test-Time Augmentation) [62].

Приклади результатів роботи підсистеми наведено на рис. 2.6.

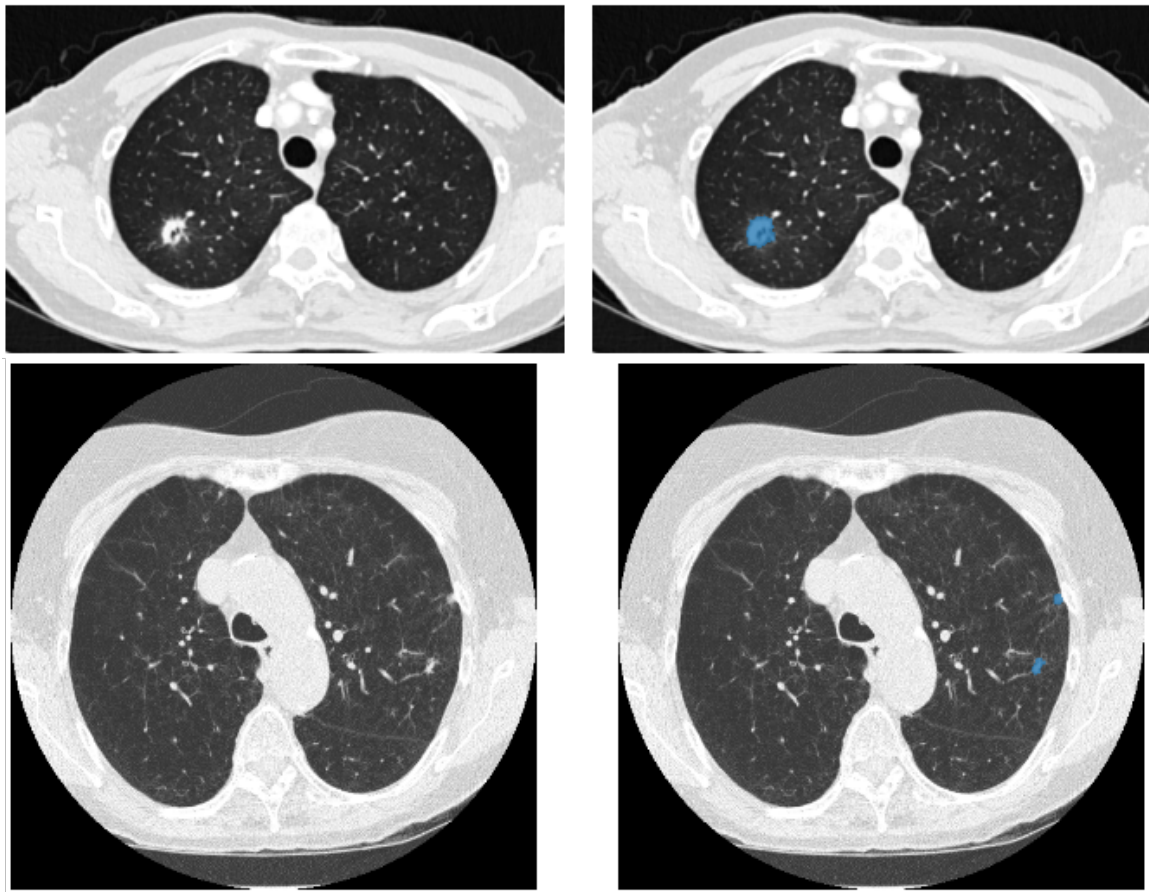


Рис. 2.6 – Результати роботи підсистеми сегментації утворів у легенях: зліва – оригінальні зрізи КТ, справа – сегментовані області утворів (позначені синім)

2.7 Підсистема характеристики та фільтрації утворів

Для подальшої роботи системи необхідно отримати виокремлені утвори із сегментаційних масок, створених підсистемою сегментації утворів у легенях. Також потрібно обчислити основні характеристики утворів, необхідні для застосування системи Lung-RADS, а за потреби – відкинути надто малі утвори, які не несуть потенційної загрози для пацієнта або є хибно-позитивними результатами сегментаційної моделі. Для цього використовується підсистема характеристики та фільтрації утворів.

Алгоритм роботи цієї підсистеми має такий вигляд.

1. Виокремлення тривимірних зв'язних компонентів за допомогою пакета `connected-components-3d` [63], спеціально оптимізованого для великих тривимір-

них обсягів даних. Для кожного зв'язного компонента одночасно обчислюються центроїд і рамка (bounding box).

2. Обчислення характеристик утворів і перехід від воксельних до фізичних координат. Центроїд та точки рамки переводяться у фізичну систему координат за допомогою афінної матриці (2.6). Після цього обчислюються: (1) фізичний об'єм утвору через множення кількості вокселів на фізичний об'єм одного вокселя; (2) найбільший діаметр ($d_{\max}$) – формула 2.7; (3) аксіальний діаметр (d_{axial}) – формула 2.8.

3. Фільтрація утворів, які виходять за допустимі інтервали розмірів за $d_{\max}$ або d_{axial} (за потреби).

$$p'_i = \mathbf{A}_{1:3, 1:3} p_i + \mathbf{A}_{1:3, 4}, \quad (2.6)$$

де p_i – координата точки у воксельній системі (індексах вокселів);
 p'_i – відповідна точка у фізичній системі координат (у міліметрах);
 $\mathbf{A}$ – афінна матриця КТ-скану.

$$d_{\max} = \max((b_{\max} - b_{\min}) \odot s), \quad (2.7)$$

$$d_{\text{axial}} = \max((b_{\max, \text{ax}} - b_{\min, \text{ax}}) \odot s_{1:2}), \quad (2.8)$$

де $d_{\max}$ – найбільший діаметр утвору у фізичних одиницях (мм);
 d_{axial} – найбільший аксіальний діаметр утвору;
 $b_{\min}, b_{\max}$ – координати мінімального та максимального кутів рамки у воксельній системі;
 $b_{\min, \text{ax}}, b_{\max, \text{ax}}$ – аналогічно, але лише для осей x та y (аксіальна площина);
 s – вектор розмірів вокселя у фізичних одиницях (мм/воксель);
 $\odot$ – покомпонентне множення.

На виході підсистема характеристики та фільтрації утворів формує список виокремлених утворів і обчислені для них розміри. Приклад результатів роботи показано на рис. 2.7.

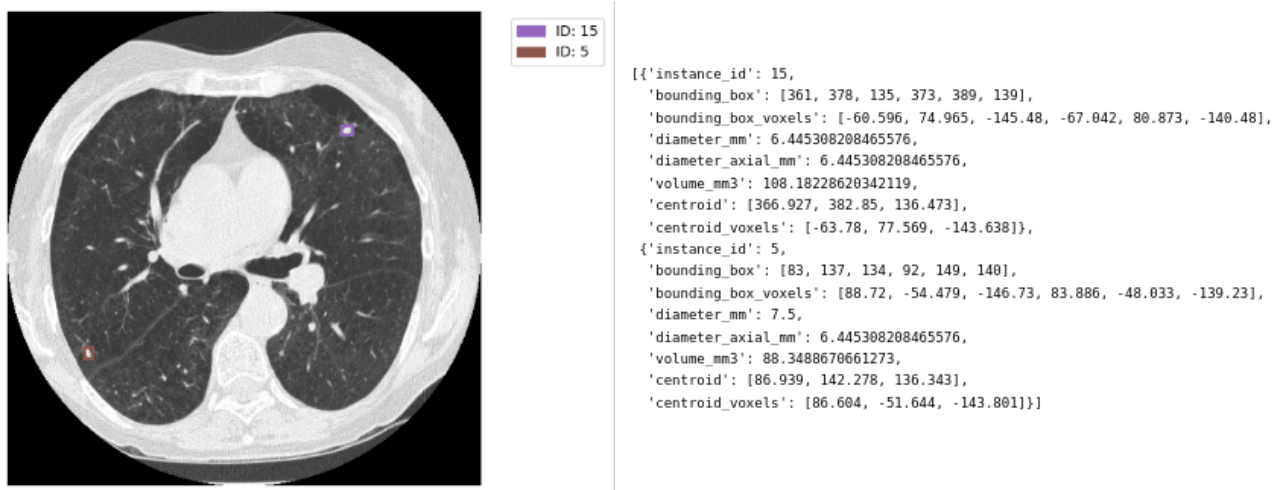


Рис. 2.7 – Приклад результатів роботи підсистеми характеристики та фільтрації утворів: зліва – КТ-зріз із виокремленими утворами, справа – інформація про їхні характеристики та розміри.

2.8 Підсистема класифікації утворів

Наступною ключовою складовою є підсистема класифікації утворів. Її основна мета – визначення класу утвору, який відображає рівень ризику для пацієнта, тобто ймовірність того, що утвір є злоякісним.

Дана підсистема реалізована у двох варіантах.

1. Мультикласова класифікація – утвори поділяються на чотири класи за рівнем ризику злоякісності: Дуже малоімовірно (Highly Unlikely), Помірно малоімовірно (Moderately Unlikely), Помірно підозріле (Moderately Suspicious), Дуже підозріле (Highly Suspicious). У датасеті LIDC-IDRI [64] також використовується додатковий клас Невизначене (Indeterminate), однак проведені дослідження [65] показали, що його застосування значно підвищує рівень шуму в даних і знижує точність класифікації.

2. Бінарна класифікація – визначається лише факт злоякісності утвору (malignant). Ймовірність доброякісного (benign) утвору розраховується як $1 - P(\text{malignant})$. Така таксономія утворюється шляхом об'єднання класів Помірно підозріле і Дуже підозріле в один клас – Злоякісний.

Розроблена класифікаційна модель отримує на вхід патчі КТ-зображень, обрізані за попередньо визначеними рамками. Для малих утворів рамки масштабуються вгору, а для надто великих – зменшуються, що дозволяє отримати фіксований розмір патча та використовувати механізм пакетної обробки (*batching*). Цей підхід суттєво прискорює обробку великої кількості утворів в одному скані без втрати точності класифікації, оскільки більшість утворів є меншими за розмір патча і містяться в ньому разом із контекстом навколишніх тканин.

Для класифікації використано модифікований варіант мережі ResNet50 із тривимірними згортками [39]. Замість одноканального зображення застосовано додатковий канал, який містить лише виділений утвір, що дозволяє сфокусувати увагу моделі на цільовій області (2.9). Приклад двоканального зображення наведено на рис. 2.8.

$$I = \text{concat}(X, M) \in \mathbb{R}^{2 \times H \times W \times D}, \quad (2.9)$$

де $X \in \mathbb{R}^{H \times W \times D}$ – патч КТ-скану.

$$M(x, y, z) = \begin{cases} X(x, y, z), & \text{якщо } (x, y, z) \in \text{рамці,} \\ 0, & \text{інакше.} \end{cases}$$

Для підвищення стійкості моделі до різних протоколів дослідження та типів сканерів було застосовано набір аугментацій: віддзеркалення по всіх осях, повороти на 90° в аксіальній площині, випадкові зміни масштабів, додавання гаусівського шуму та згладжування. Додатково використано дві спеціально розроблені техніки аугментації: випадковий зсув утвору (2.12) та випадкове видалення додаткового каналу. Обидві покликані підвищити робастність моделі до

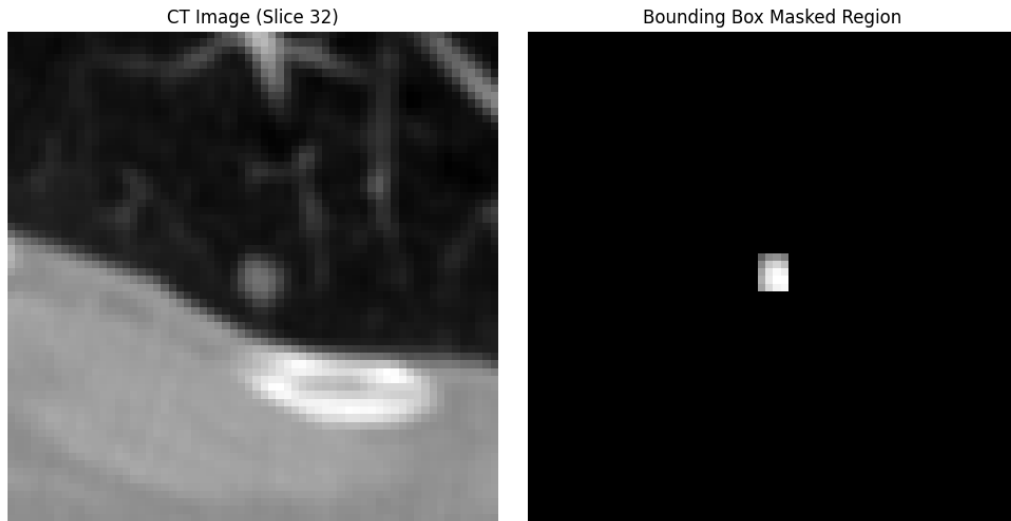


Рис. 2.8 – Зліва – патч КТ-зображення; справа – маскований регіон, що містить лише сам утвір [65].

можливих похибок попередніх підсистем – наприклад, при недосегментації або пересегментації утвору, коли він зміщується від центру патча або містить зайві пікселі. Використання таких аугментацій зменшує залежність мережі від точного центрованого положення утвору та його наявності у другому каналі. Приклад аугментації зсуву наведено на рис. 2.9. Для підвищення точності класифікації також застосовується техніка усереднення результатів по аугментаціях під час інференсу (аналогічно до підходу, описаного в розділі 2.6).

$$B = [x_{\min}, y_{\min}, z_{\min}, x_{\max}, y_{\max}, z_{\max}], \quad (2.10)$$

$$L = [L_x, L_y, L_z] = [x_{\max} - x_{\min}, y_{\max} - y_{\min}, z_{\max} - z_{\min}], \quad (2.11)$$

$$B' = [x_{\min} + \Delta_x, y_{\min} + \Delta_y, z_{\min} + \Delta_z, x_{\max} + \Delta_x, y_{\max} + \Delta_y, z_{\max} + \Delta_z], \quad (2.12)$$

$$\Delta_i \sim \mathcal{U}(L_i \cdot \alpha_i^{\min}, L_i \cdot \alpha_i^{\max}), \quad B' \subseteq [0, H] \times [0, W] \times [0, D]. \quad (2.13)$$

де B – вихідна рамка утвору;

L_x, L_y, L_z – довжини сторін рамки вздовж осей x, y, z ;

$\Delta_x, \Delta_y, \Delta_z$ – випадкові зсуви, вибрані з рівномірного розподілу;

$\alpha_i^{\min}, \alpha_i^{\max}$ – коефіцієнти діапазону зсуву для кожної осі;

B' – рамка після аугментації (з урахуванням зсуву);

H, W, D – висота, ширина та глибина КТ–скану.

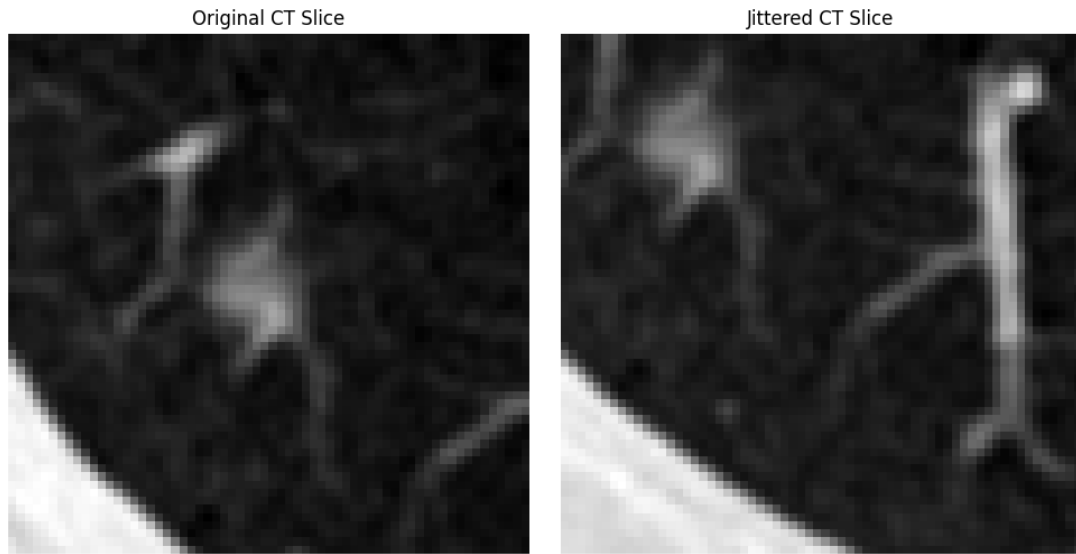


Рис. 2.9 – Зліва – оригінальний КТ–зріз; справа – результат після застосування аугментації зсуву [65].

На виході підсистема класифікації утворів формує розширений результат підсистеми характеристики та фільтрації – з визначеним класом утвору та ймовірністю належності до нього. Приклад роботи моделі мультикласової класифікації наведено на рис. 2.10, а бінарної – на рис. 2.11.

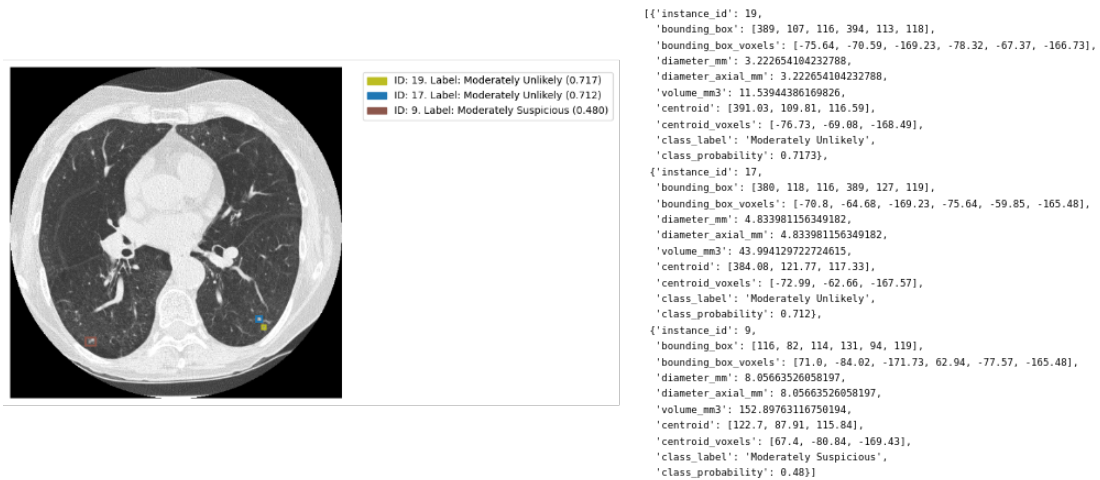


Рис. 2.10 – Зліва – КТ-зріз із виявленими та класифікованими утворами; справа – інформація про утвір, включаючи його клас та ймовірність належності до класу для мультикласової моделі.

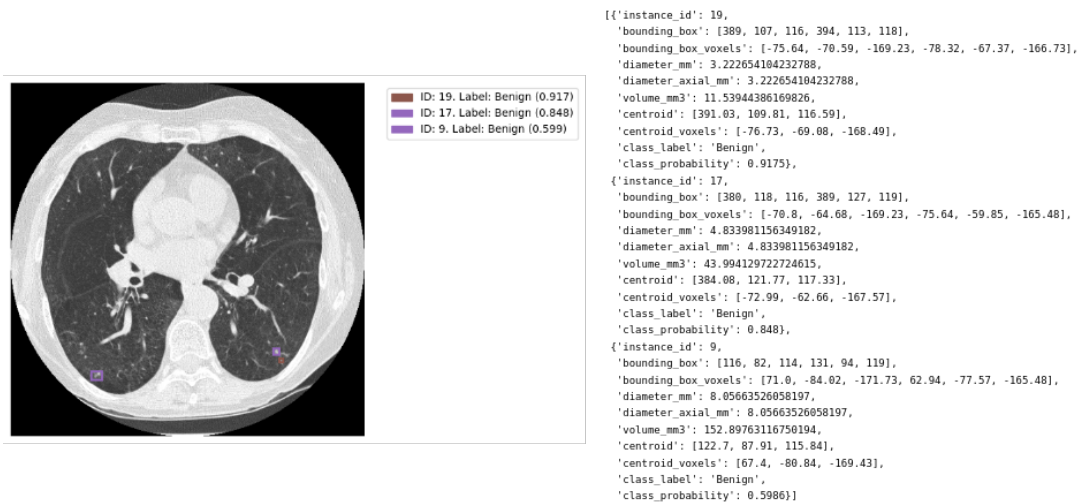


Рис. 2.11 – Зліва – КТ-зріз із виявленими та класифікованими утворами; справа – інформація про утвір, включаючи його клас та ймовірність належності до класу для бінарної моделі.

2.9 Підсистема постобробки результатів

Підсистема постобробки результатів призначена для усунення випадків, коли один і той самий утвір було помилково визначено як кілька окремих

об'єктів. Для цього застосовується алгоритм подавлення немаксимумів (*Non-Maximum Suppression, NMS*) [66], який із групи перетинних утворів залишає лише той, що має найвищу ймовірність належності до визначеного класу. Приклад застосування цього алгоритму наведено на рис. 2.12.

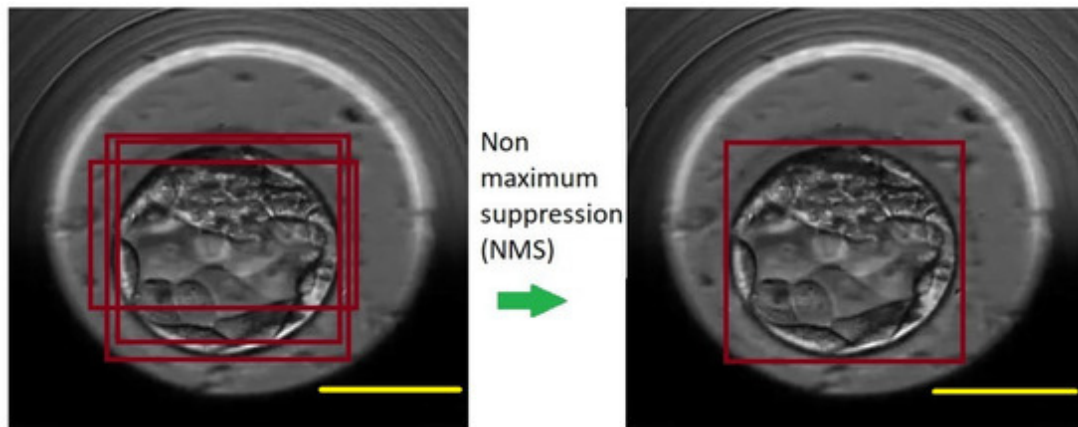


Рис. 2.12 – Приклад застосування алгоритму подавлення немаксимумів у задачах аналізу медичних зображень [67].

2.10 Підсистема інтерактивної сегментації утворів

Підсистема інтерактивної сегментації призначена для ефективного виправлення помилок, отриманих від статичної сегментаційної системи, таких як недосегментація, пересегментація, хибно-позитивне виділення або пропуск хибно-негативного утвору.

В основі системи використовується мережа *DynUNet* (так само, як у підсистемі сегментації – розділ 2.6), проте зменшено кількість шарів – видалено по одному шару зменшення та збільшення розмірності. Це забезпечує швидший час обчислення, необхідний для інтерактивної роботи. Водночас інтерактивний сигнал фокусує мережу саме на області корекції, тому глобальний контекст має менше значення при формуванні оновленої сегментаційної маски.

Інтерактивний механізм реалізовано через кліки, які радіолог проставляє у зонах некоректної сегментації: позитивний клік – у зоні пропущених (хибно-негативних) вокселів, які потрібно додати; негативний клік – у зоні хибно-

позитивних вокселів, які потрібно видалити. Зазвичай радіологи працюють з аксіальними зрізами, тому клік отримується у вигляді координати пікселя, але система може також працювати з координатою вокселя у 3D-просторі. Після отримання кліку формується його 3D-гаусова маска – формула 2.14. Схематичний приклад роботи підсистеми наведено на рис. 2.13.

$$M(y, x, z) = \text{clip} \left(\exp \left(-\frac{1}{2} \left[\left(\frac{y - y_c}{\sigma_y} \right)^2 + \left(\frac{x - x_c}{\sigma_x} \right)^2 + \left(\frac{z - z_c}{\sigma_z} \right)^2 \right] \right) \cdot 1 \left(\left(\frac{y - y_c}{r_y} \right)^2 + \left(\frac{x - x_c}{r_x} \right)^2 + \left(\frac{z - z_c}{r_z} \right)^2 \leq 1 \right), 0, 1 \right) \quad (2.14)$$

де $M(y, x, z)$ – 3D-гаусова маска;

(y_c, x_c, z_c) – координати кліку (центр гаусіани);

$\sigma_y, \sigma_x, \sigma_z$ – стандартні відхилення вздовж осей;

r_y, r_x, r_z – радіуси еліпсоїдного обрізання;

$1(\cdot)$ – індикатор еліпсоїдної області;

$\text{clip}(\cdot, 0, 1)$ – функція обрізання значень до діапазону $[0, 1]$.

Вхід у мережу формується як кубічний патч КТ-скану з чотирма каналами: саме зображення; попередня маска (до бінаризації); маска позитивного кліку; маска негативного кліку. Приклад продемонстровано на рис. 2.14. Кожен інференс обробляє один клік, але завдяки фіксованому розміру патча можливе ефективне пакування для паралельної обробки кількох кліків. Сам патч центрується навколо кліку.

Інтерактивна мережа використовує ті самі покращення, що й базова сегментаційна (Boundary DoU++, нормалізація фізичного розміру, аугментації), але оптимізується за спеціальною двоетапною методологією: підготовка вибірки патчів; семпсування кліків.

Спершу створюються маски на всій тренувальній вибірці за допомогою перехресної перевірки (*Cross Validation*). Для кожного прогнозу визначаються утвори через підсистему характеристики та фільтрації (розділ 2.7). Далі для кожного прогнозованого утвору обчислюється перетин з усіма анотованими

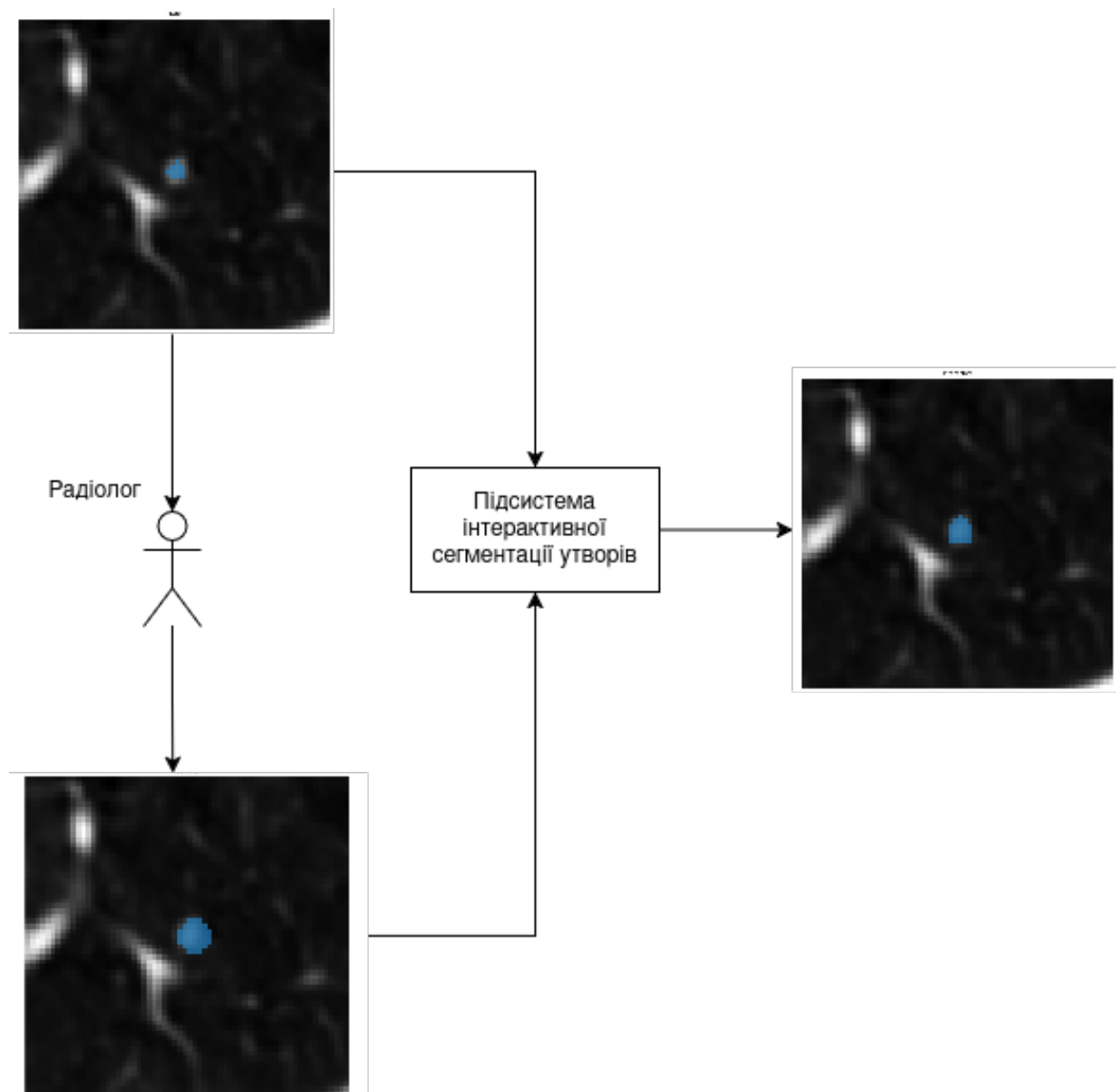


Рис. 2.13 – Схематичне зображення роботи підсистеми інтерактивної сегментації. Радіолог аналізує поперечний зріз і, якщо потрібно розширити область утвору, здійснює позитивний клік. Підсистема отримує на вхід попередню (статичну) маску та маску кліку, після чого оновлює сегментацію, розширюючи область утвору.

утворами за метрикою IoU на рівні рамок, і береться максимальне значення – формула 2.15. На основі цього утвору присвоюється один із трьох класів.

1. Якщо $\text{IoU}_{\max}(B_p) < 0.1$ – негативний (Negative), тобто хибно–позитивне спрацювання.

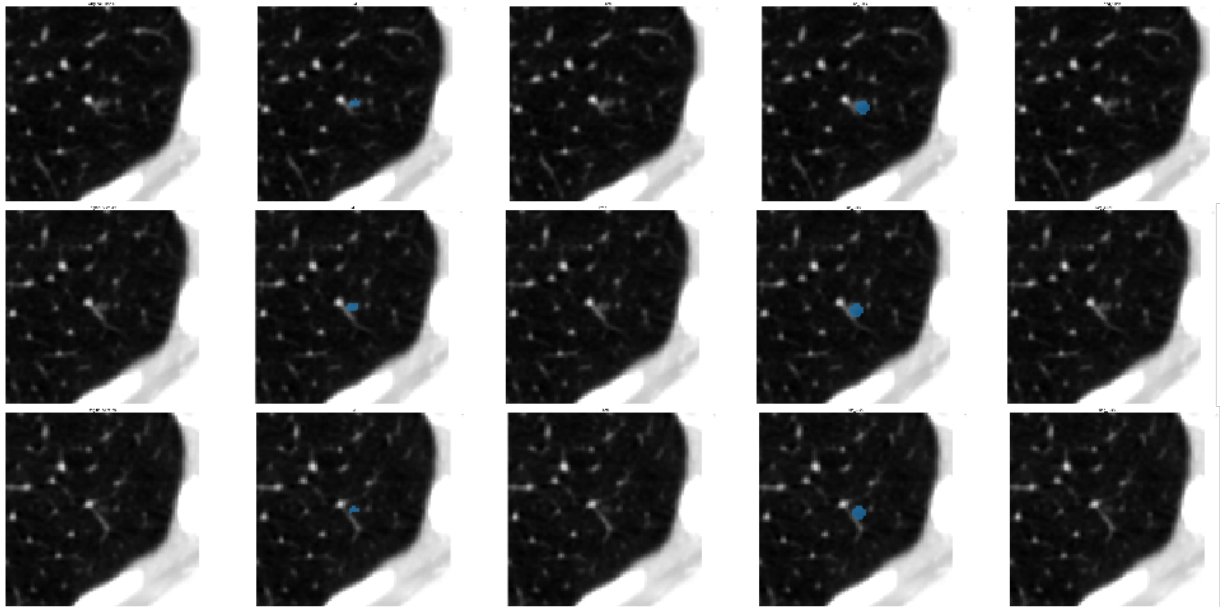


Рис. 2.14 – Приклад побудови багатоканального вхідного зображення для інтерактивної сегментаційної мережі (по рядках упорядковані три поперечні зрізи): перший стовпець – поперечні зрізи КТ-скану; другий – ановані області утвору; третій – прогнозована маска статичною моделлю; четвертий – гаусівська маска позитивного кліку; п'ятий – гаусівська маска негативного кліку.

2. Якщо $\text{IoU}_{\max}(B_p) > 0.2$ – позитивний (Positive), тобто коректно детектований утвір.

3. В іншому випадку – пограничний (Borderline), тобто утвір частково захоплений, проте значно пере– або недосегментований.

Такий самий підхід застосовується зворотно для кожного анованого утвору шляхом обчислення його перетину з кожним прогнозованим. Після цього визначаються відповідні класи.

1. Якщо $\text{IoU}_{\max}(B_{gt}) < 0.1$ – хибно–негативний (False Negative), тобто пропущений утвір.

2. Якщо $\text{IoU}_{\max}(B_{gt}) > 0.2$ – позитивний (True Positive), тобто правильно знайдений утвір.

3. В іншому випадку – пограничний (True Positive Borderline), тобто аналогічно до випадку Borderline для прогнозованого утвору.

$$\text{IoU}_{\max}(B_p) = \max_{j \in \{1, \dots, N_{\text{GT}}\}} \frac{\text{Vol}(B_p \cap B_{\text{GT}}^{(j)})}{\text{Vol}(B_p \cup B_{\text{GT}}^{(j)})}, \quad (2.15)$$

де B_p – рамка передбаченого утвору;
 $B_{\text{GT}}^{(j)}$ – j -та рамка анотованого утвору;
 N_{GT} – кількість анотованих утворів на скані;
 $\text{Vol}(\cdot)$ – оператор обчислення об'єму.

Отже, отримана тренувальна вибірка містить подвійний набір коректно детектованих, недосегментованих і пересегментованих утворів, а також набір хибно–негативних і хибно–позитивних утворів. Кожен утвір має відповідну рамку (B_{gt} або B_p), тому формат вхідних даних ідентичний до даних підсистеми класифікації (розділ 2.8).

Другий етап методології описує алгоритм семпсування кліку для кожного елемента вибірки. Залежно від класу утвору виконуються такі кроки.

1. Якщо клас Positive, True Positive або False Negative – обирається центральний об'єкт на анотованій масці. Для цього застосовується пакет `connected-components-3d`, після чого обирається об'єкт, центроїд якого найближчий до центру патча (за евклідовою відстанню). Потім на основі маски центрального об'єкта семплується позитивний клік із “обрізаного” Гаусового розподілу (2.16).
2. Якщо клас Negative – аналогічна процедура проводиться на передбаченій бінаризованій масці, але тип кліку змінюється на негативний.
3. Якщо клас Borderline або True Positive Borderline – обчислюються хибно–негативна маска M_{FN} (2.17) та хибно–позитивна маска M_{FP} (2.18). Якщо M_{FN} порожня, обирається M_{FP} ; якщо навпаки – M_{FN} ; якщо обидві непорожні – вибір робиться випадково. Потім із вибраної маски семплується клік, який буде позитивним у разі вибору M_{FN} або негативним у разі M_{FP} .

$$w_i = \exp\left(-\frac{1}{2} \left[\left(\frac{y_i - c_y}{\sigma_y}\right)^2 + \left(\frac{x_i - c_x}{\sigma_x}\right)^2 + \left(\frac{z_i - c_z}{\sigma_z}\right)^2 \right]\right) \cdot 1\left(\left(\frac{y_i - c_y}{r_y}\right)^2 + \left(\frac{x_i - c_x}{r_x}\right)^2 + \left(\frac{z_i - c_z}{r_z}\right)^2 \leq 1\right),$$

$$p_i = \frac{w_i}{\sum_{j \in \Omega} w_j}, \quad \text{де } \Omega = \{j \mid \text{mask}(j) = 1\},$$

$$(y^*, x^*, z^*) \sim \text{Categorical}(p_i), \quad (2.16)$$

де (y_i, x_i, z_i) — координати i -го вокселя, що належить масці Ω ;
 (c_y, c_x, c_z) — центроїд маски;
 $\sigma_y, \sigma_x, \sigma_z$ — параметри стандартного відхилення розподілу;
 r_y, r_x, r_z — радіуси еліпсоїдного відсікання;
 w_i — вага i -го вокселя, пропорційна гаусовому розподілу;
 p_i — нормалізована ймовірність вибору вокселя;
 (y^*, x^*, z^*) — випадково вибрана точка з розподілу p_i ;
 $1(\cdot)$ — індикатор, що обнуляє ваги поза еліпсоїдною областю.

$$M_{\text{FN}} = G \wedge \neg P \quad (2.17)$$

$$M_{\text{FP}} = \neg G \wedge P, \quad (2.18)$$

де M_{FN} — хибно–негативна маска;
 M_{FP} — хибно–позитивна маска;
 G — істинна маска;
 P — передбачена бінаризована маска.

Далі, на основі підготовленої вибірки та використовуючи алгоритм генерації кліку, можна сформуванати чотиріканальні зображення й оптимізувати інтерактивну мережу. Важливо зазначити, що алгоритм генерації кліку виконується на

кожному кроці навчання, що підвищує різноманітність можливих варіантів кліків користувача.

Під час використання вже оптимізованої моделі, тобто самої інтерактивної системи сегментації, користувач може надати набір кліків різного типу. Надалі вони групуються у пакети і обробляються за таким алгоритмом.

1. Навколо кожного кліку створюється патч КТ–скану того самого розміру, який використовувався під час оптимізації. Сам клік розташовується в центрі патча. Пакет даних формується з набору таких 3D–патчів.

2. Кожне зображення патча розширюється трьома додатковими каналами: масками позитивного й негативного кліків, побудованими за формулою 2.14, а також небінаризованою маскою статичної моделі сегментації (розділ 2.6).

3. До сформованих даних застосовується модель інтерактивної сегментації, у результаті чого отримується оновлена сегментаційна маска.

4. Для кожної маски з пакету проводиться бінаризація попередньої та оновленої масок за певним порогом. Якщо клік позитивний і оновлена маска непорожня, обирається її центральний об'єкт (за процедурою, описаною у першому випадку семплювання кліку), і лише пікселі маски цього центрального об'єкта оновлюються в попередній масці. Якщо ж клік негативний і попередня маска непорожня – центральний об'єкт обирається на основі початкової маски, після чого аналогічно проводиться оновлення ймовірностей пікселів цього об'єкта. Така процедура необхідна, оскільки, з урахуванням методології оптимізації інтерактивної мережі, покращення сегментації очікується саме для центральних об'єктів, що також узгоджується з поведінкою користувачів (радіологів).

5. Оновлені патчі сегментаційних масок (небінаризованих) інтегруються назад у глобальну маску всього 3D КТ–скану.

Результати роботи підсистеми наведено на рис. 2.15. Після отримання оновлених масок підсистема знову викликає модулі характеристики та фільтрації утворів, класифікації утворів і постобробки результатів для їх уточнення. Таким чином, інтерактивна підсистема не лише покращує якість сегментації та точність

геометричної характеристизації утворів, але й потенційно підвищує точність класифікації.

Під час дослідження спочатку була розроблена інтерактивна підсистема сегментації, що базувалася на 2.5D–мережах [68]. Ця система демонструвала хороші результати у покращенні сегментації, проте мала низку недоліків та обмежень.

1. Кожне інтерактивне оновлення впливало лише на один поперечний зріз, що вимагало значно більшої кількості кліків порівняно з одним кліком на утвір у 3D–мережі. У середньому кожен утвір займає близько 4.65 поперечних зрізів, отже система на базі 2.5D–мережі потребує у 4.65 разів більше взаємодій із радіологом.

2. Обмежений контекст по третій осі, що знижувало якість сегментації складних утворів, зокрема тих, які подібні до кровоносних судин у легенях.

Підсистема на основі 2.5D–мереж заклала підґрунтя подальших досліджень інтерактивної сегментації, результатом яких стала значно ефективніша та точніша 3D–підсистема інтерактивної сегментації.

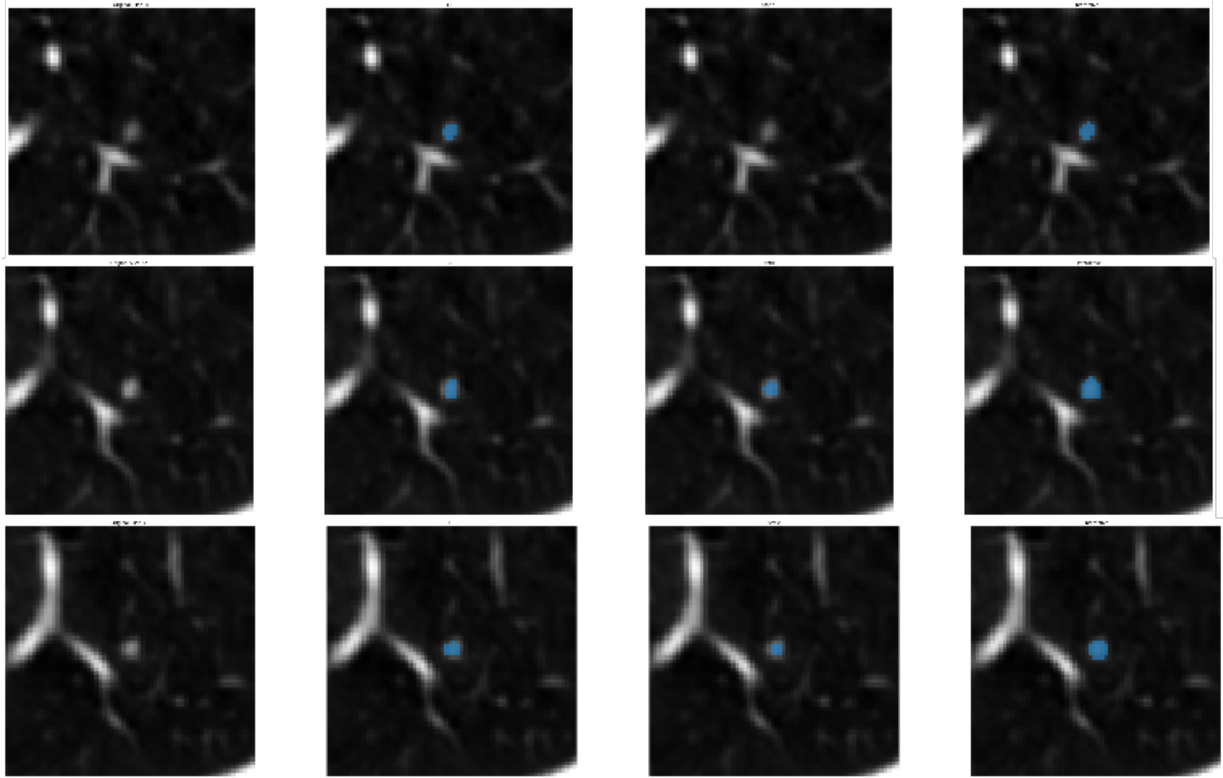


Рис. 2.15 – Результати роботи підсистеми інтерактивної сегментації (по рядках упорядковано три поперечних зрізи): перший стовпець – поперечні зрізи КТ-скану; другий – анотовані області утвору; третій – прогнозована маска статичною сегментаційною моделлю; четвертий – оновлені маски після інтерактивної сегментації.

2.11 Висновки розділу 2

Відповідно до поставлених у розділі 1.7 задач було розроблено клінічну ІСППР для діагностики раку легенів на основі низки моделей глибокого навчання. Розроблена система приймає на вхід «сирі» зображення – серію DICOM-файлів, що формують КТ-зображення, а на вихід надає всі знайдені підозрілі утвори з їхніми характеристиками, включно з геометричними ознаками та класом утвору, який відображає рівень ризику для пацієнта.

Система містить інтерактивну підсистему, що дозволяє радіологу за допомогою кількох кліків швидко скоригувати можливі похибки та отримати задовільний результат. Класифікаційна підсистема реалізована у двох варіантах:

бінарному (злаякісний / доброякісний) та мультикласовому (класи за рівнем ризику злоякісності). Результати, що генерує розроблена система, можуть бути безпосередньо використані радіологами та лікарями для застосування існуючих критеріїв і систем скринінгу та діагностики раку легенів.

Водночас важливо зазначити певні недоліки розробленої клінічної ІСППР.

1. Таксономія класифікації наразі не відповідає критеріям системи Lung–RADS.

2. Замість детекційних моделей застосовано сегментаційні глибокі нейронні мережі. Моделі детекції потенційно можуть бути обчислювально ефективнішими при збереженні або навіть підвищенні точності виявлення утворів.

Попри ці обмеження, завдяки модульній і гнучкій архітектурі розробленої ІСППР зазначені покращення можуть бути легко реалізовані надалі. Зокрема, у разі отримання навчальної вибірки з потрібною таксономією класифікаційна модель може бути оптимізована без зміни загальної методології. Інтеграція детекційної підсистеми вимагатиме більше зусиль, проте в рамках ІСППР вона може безпосередньо замінити підсистему сегментації утворів у легенях та підсистему характеристики й фільтрації утворів.

3 ЗАСТОСУВАННЯ РОЗРОБЛЕНИХ МОДЕЛЕЙ І МЕТОДІВ ДЛЯ ІНШИХ ЗАДАЧ ДІАГНОСТИКИ, МЕДИЦИНИ ТА БІОЛОГІЇ

3.1 Адаптація системи для виявлення злоякісних та доброякісних пухлин у нирках

Методологічні підходи, використані під час створення клінічної ІСППР для діагностики раку легенів, можуть бути застосовані й для побудови клінічних ІСППР для інших типів онкологічних захворювань, зокрема раку нирок.

Для діагностики раку нирок використовується класифікація Bosniak [6]. За цим критерієм виділяють п'ять категорій кістозних утворень: Bosniak I, Bosniak II, Bosniak IIF, Bosniak III, Bosniak IV. Окремо також розглядають злоякісні та доброякісні солідні пухлини. Категорії Bosniak I, II та IIF відносять до доброякісних утворень, тоді як Bosniak III та IV вважають потенційно небезпечними та відносять до злоякісних. Для спрощення задачі під час побудови клінічної ІСППР було вибрано три основні класи: ниркова тканина (ID: 1); доброякісні утворення (ID: 3), що включають доброякісні солідні пухлини та Bosniak I, II, IIF; злоякісні утворення (ID: 2), які охоплюють злоякісні солідні пухлини та Bosniak III і IV. Додатково введено клас «фону» (ID: 0), що відповідає всім іншим регіонам, які не належать до зазначених трьох класів. Надалі в цьому розділі ці класи називаються цільовими.

Для реального клінічного застосування від ІСППР необхідно отримати всі доброякісні та злоякісні утворення, а також їхні розміри (найдовший діаметр у 3D, діаметри на найбільшому аксіальному зрізі, висота в сагітальному зрізі, об'єм у міліметрах) і координати (номер аксіального зрізу з найбільшою площею утвору та координати для побудови діаметрів). Клас ниркової тканини не використовується як фінальний вихід ІСППР, проте він підвищує точність роботи системи та може бути корисним у майбутньому в разі зміни формату вихідних даних. На вхід системі подається серія DICOM-файлів.

Клінічна ІСППР для діагностики раку нирок складається з таких підсистем.

1. **Підсистема обробки та підготовки медичних зображень** – ідентична до підсистеми з розділу 2.4.

2. **Підсистема сегментації області нирок** – виокремлює зону нирок на КТ–скані для подальшої обробки лише цієї області. На відміну від підсистеми з розділу 2.5, тут використано сегментаційну мережу, спеціально оптимізовану на чотири цільові класи та на КТ–сканах зі зменшеним розміром для пришвидшення обчислень. Далі локалізація виконується через рамку трьох класів (1, 2, 3) відповідно до формули 2.1. Рішення замінити *TotalSegmentator* на оптимізовану мережу обумовлене тим, що *TotalSegmentator* призначений для сегментації типових анатомічних структур, тоді як пухлини в нирках можуть бути великими та виходити за межі області нирки, і така структура буде пропущена.

3. **Підсистема постобробки результатів сегментації області нирок** – видаляє сегментаційні області, що розташовані надто далеко по аксіальній осі від зони нирок. Область нирок визначається як найбільша зв’язна компонента. Підсистема призначена для усунення хибно позитивних об’єктів, особливо на КТ–сканах усього тіла.

4. **Перша підсистема сегментації утворень у нирках** – створює семантичні маски цільових класів. Архітектура нейронної мережі та методологія оптимізації відповідають тим, що наведено в розділі 2.6. Основна відмінність полягає в тому, що замість фізичного розміру вокселя використовується фіксований розмір 3D–зображення. Це забезпечує повний глобальний контекст та підвищує точність сегментації великих пухлин, які перевищують розміри ковзного вікна.

5. **Друга підсистема сегментації утворень у нирках** – аналогічна попередній, однак використовує класичну функцію втрат *Boundary IoU* замість модифікації *Boundary IoU++*. Підсистема спеціалізується на менших утвореннях, оскільки працює з ковзними вікнами. Небінаризовані маски обох підсистем усереднюються, а підсумковий клас кожного пікселя визначається за максимальною ймовірністю.

6. Підсистема постобробки результатів сегментації утворень – аналогічна першій частині фільтрації з розділу 2.7, але замість фільтрації за аксіальним діаметром застосовується фільтрація за об'ємом. Замалі утворення та нирки видаляються.

7. Підсистема характеристики утворень – працює аналогічно підсистемі з розділу 2.7. Після виокремлення утворень із сегментаційних масок визначаються їхні характеристики: розміри (найдовший діаметр у 3D, діаметри на найбільшому аксіальному зрізі, висота у сагітальному зрізі, об'єм у міліметрах) та координати (номер аксіального зрізу з максимальною площею та координати для побудови діаметрів).

Приклад результатів роботи ІСППР наведено на рис. 3.1, а архітектура – на рис. 3.2.



Рис. 3.1 – Зліва – локалізовані об'єкти нирок та пухлини на аксіальному зрізі КТ-скану. Справа – характеристика об'єктів (ниркова тканина – *tissue*; злоякісна пухлина – *malignant*). Важливо, що характеристики визначаються лише для злоякісних та доброякісних утворень.

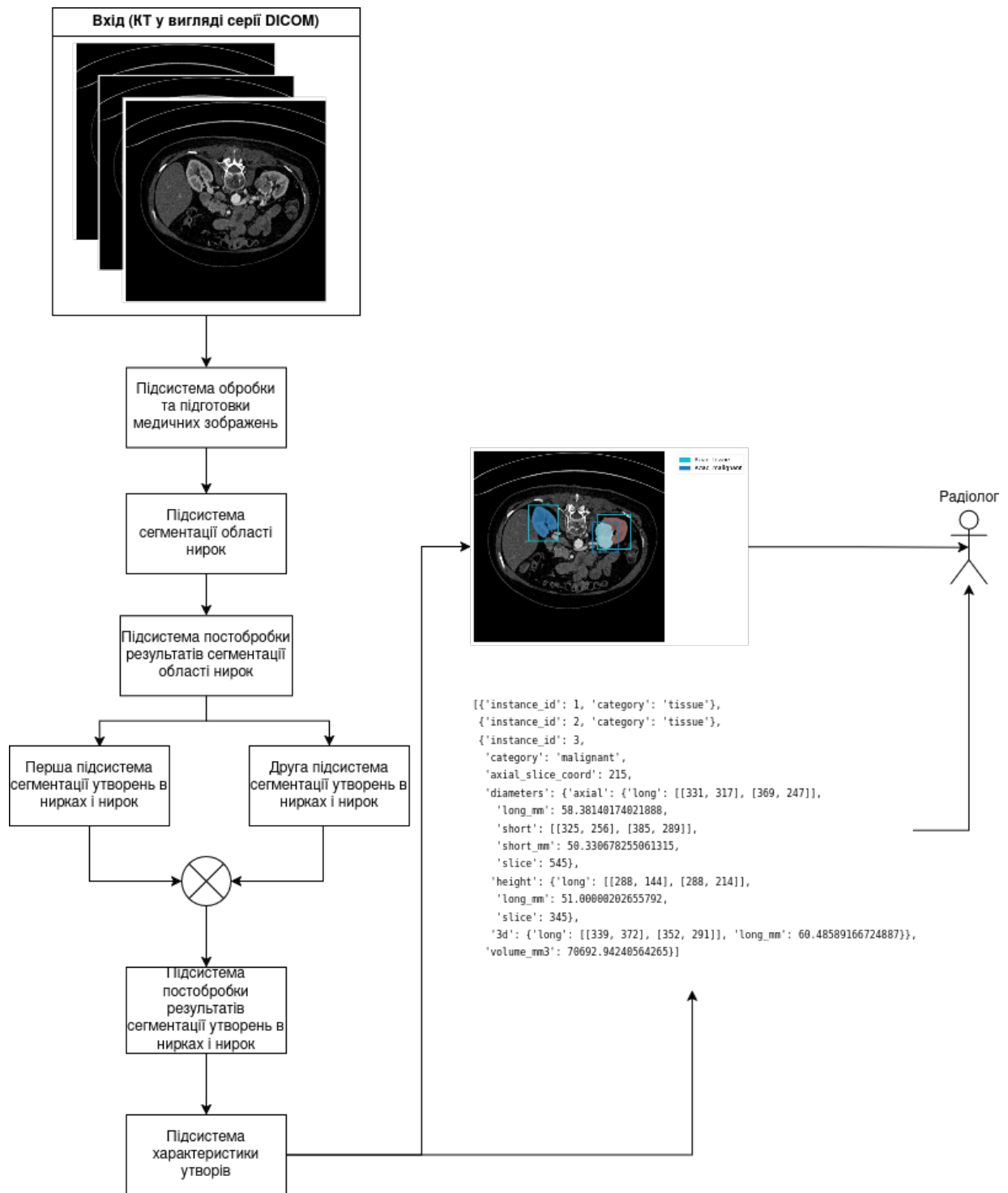


Рис. 3.2 – Схематичне зображення ІСППР для виявлення раку нирок

Важливо зазначити, що клінічна ІСППР для діагностики раку нирок має багато спільних або подібних підсистем до ІСППР для діагностики раку легенів, що демонструє можливість порівняно простої адаптації методології оптимізації глибоких нейронних мереж для інших типів онкологічних захворювань. Одночасно серед головних відмінностей можна виділити три.

1. Відсутність спеціалізованої класифікаційної підсистеми – у нирковій ІСППР функція класифікації закладена безпосередньо в сегментаційні підсистеми. За потреби розширення чи зміни таксономії може бути додана окрема класифікаційна підсистема.

2. Відсутність інтерактивної підсистеми – через специфіку застосування ниркової ІСППР. Значна частина випадків раку нирок виявляється випадково під час дослідження інших органів (incidental findings) [69]. Тому ІСППР для нирок зазвичай повинна працювати повністю автономно, виконуючи роль «експерта», який перевіряє кожен КТ–скан на наявність підозрілих утворень. Водночас, за необхідності інтерактивну складову можна легко додати, використовуючи підхід із розділу 2.10.

3. Використання двох сегментаційних підсистем – зумовлене значною варіативністю розмірів пухлин нирок, порівняно з утворами в легенях, тому потрібні окремі підсистеми, спеціалізовані для малих та великих утворень.

3.2 Доменна адаптація та напівкероване навчання для сегментації функціональних тканинних одиниць (FTU)

Іншою прикладною задачею аналізу медичних зображень, де запропоновані методи можуть значно покращити якість роботи алгоритмів глибокого навчання й системи загалом, є задача сегментації функціональних тканинних одиниць (FTU) [52]. FTU визначається як тривимірний блок клітин, центрований навколо капіляра, причому кожна клітина цього блоку знаходиться в межах дифузійної відстані від будь-якої іншої клітини в тому ж блоці [70]. Такі клітинні композиції – локальні популяції клітин – відповідають за виконання основних фізіологічних функцій органів. Функціональні тканинні одиниці мають важливе патобіологічне значення, необхідне для моделювання та розуміння розвитку хвороб. Однак ручне визначення таких структур на мікроскопічних зображеннях потребує колосальних ресурсів, оскільки їх кількість може сягати сотень тисяч на одному

зразку. Тому раціональним є використання методів, моделей і систем штучного інтелекту.

Проте під час побудови робастної системи виникає низка проблем, здебільшого пов'язаних із різноманіттям протоколів отримання мікроскопічних зображень (приклад на рис. 3.3).

1. Різні протоколи забарвлення – зокрема забарвлення за допомогою 3,3'-діамінобензидину (DAB) [71] з контрзабарвленням гематоксиліном, або методом Періодичної кислоти-Шиффа (PAS) [72], або гематоксилін-еозином (H&E) [73].

2. Різні фізичні розміри пікселя.

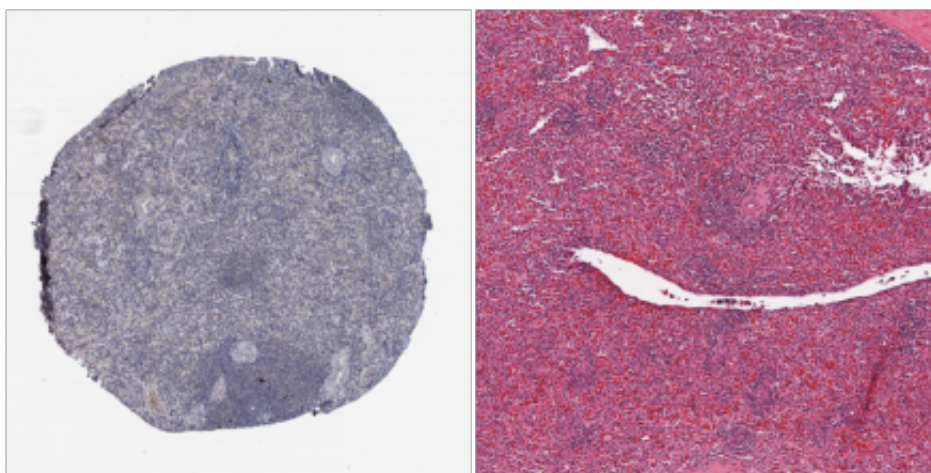


Рис. 3.3 – Зліва – зразки, забарвлені протоколом на основі 3,3'-діамінобензидину (DAB) з додатковим контрзабарвленням гематоксиліном; справа – зразки, забарвлені методом Періодичної кислоти-Шиффа (PAS) або гематоксилін-еозином (H&E).

Для вирішення задачі сегментації FTU на двовимірних медичних зображеннях були запропоновані такі методи та моделі.

1. Використання ансамбля сегментаційних моделей з енкодерами різного типу: згорткового енкодера EfficientNet B7 [36] та енкодера з механізмом уваги Mix Vision Transformer B5 [74]. Одночасно застосовано різні типи декодерів – Unet [40] для Mix Vision Transformer B5 та Unet++ [75] для EfficientNet B7.

2. Ресемплінг фізичного розміру пікселя до фіксованих значень для кожного органу. Значення були обрані так, щоб: стандартизувати розмір пікселя між різними протоколами, не змінювати надмірно розмір зображення. Для кожного органу був заздалегідь визначений набір значень, у які виконувалося ресемплювання відповідно до початкового розміру пікселя.

3. Використання гістограмного вирівнювання [76] між різними протоколами забарвлення, щоб зменшити різницю між ними.

4. Застосування великої кількості геометричних аугментацій, аугментацій у просторі забарвлення, спотворень та масштабування. Основна ідея – підвищити робастність моделей до різних типів забарвлення та масштабування. Окремо слід відзначити аугментацію CutMix [77], яка дозволяє поєднувати в одному зображенні фрагменти, отримані за різними протоколами забарвлення, що додатково об'єднує простір ознак глибоких мереж. Приклад наведено на рис. 3.4.

5. Використання методів напівкерованого навчання: залучення великої кількості неанотованих даних з іншого джерела з іншим протоколом отримання зображень. Отримані дані автоматично анотовано вже оптимізованими моделями, після чого вони використані в наступній ітерації оптимізації. Процес повторено двічі. Приклади автоматичних анотацій наведені на рис. 3.5.

6. Застосування постобробки для видалення надто малих регіонів з порогамі, визначеними відносно розміру зображення та окремо для кожного органу – формула 3.1. Додатково застосовано усереднення по аугментаціях під час інференсу (Test–Time Augmentation).

$$\frac{\text{RegionArea}}{\text{ImageArea}} < \text{OrganThresh} \quad (3.1)$$

де RegionArea – площа сегментованого регіону;

ImageArea – площа всього зображення;

OrganThresh – поріг для фільтрації регіонів, окремий для кожного органу.

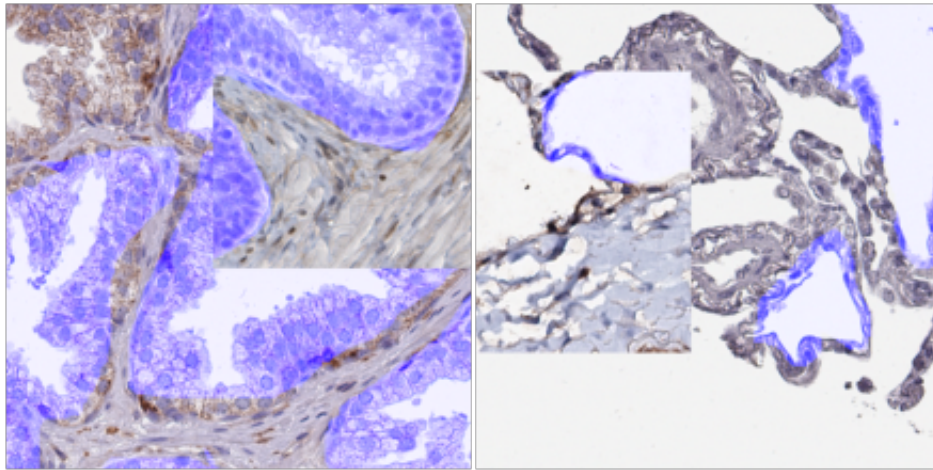


Рис. 3.4 – Приклад використання аугментації CutMix: зліва – для простати, справа – для легенів. Фіолетовим виділено зони сегментації.

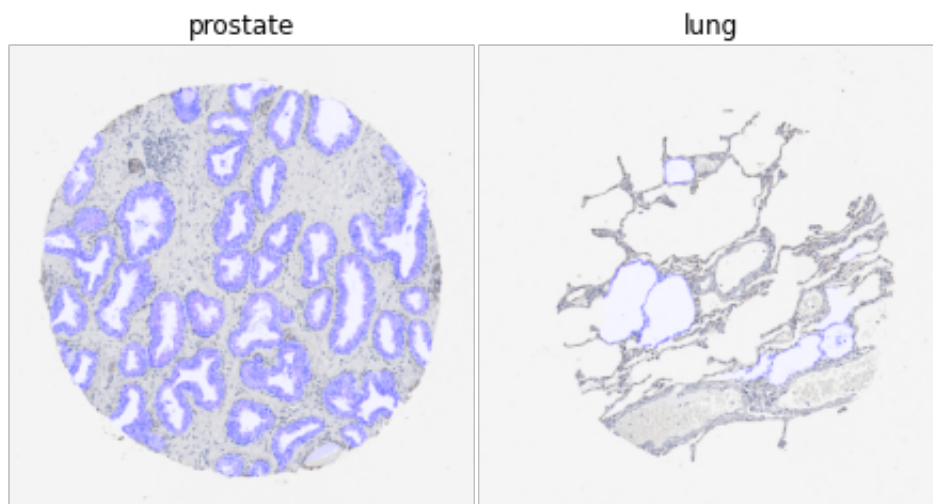


Рис. 3.5 – Результати, отримані під час автоматичної анотації додаткових даних, згенерованих у процесі застосування напівкерованого навчання. Фіолетовим виділено зони сегментації.

Розроблена система моделей продемонструвала високу якість сегментації (Dice Score на рівні 83–85%) і показала, що запропоновані попередні методики, такі як постобробка та спеціально підібрані аугментації, разом з новими підходами, зокрема адаптацією простору забарвлення, напівкерованим навчанням на вибірці, отриманій за іншими протоколами, а також використанням ансамбля

моделей різної природи, забезпечують високу якість виконання складної задачі сегментації FTU.

3.3 Багаторівнева система для виявлення раку шкіри

Клінічні ІСППР на базі нейронних мереж глибокого навчання можуть бути використані також для діагностики раку шкіри (меланоми). Рак шкіри є одним із найчастіше діагностованих видів раку та характеризується високими ризиками ускладнень у разі несвоєчасного лікування [78]. Сучасні методи діагностики значною мірою покладаються на експертизу дерматологів, яка може бути суб'єктивною, тому розробка об'єктивної клінічної ІСППР, здатної прискорити та підвищити точність діагностики, є актуальною задачею. Крім того, такі системи потенційно розширюють доступ до якісної діагностики для пацієнтів із низьким рівнем доходу, оскільки сучасні процедури часто потребують значних ресурсів [79].

Для вирішення задачі виявлення раку шкіри на основі медичних зображень та метаданих була запропонована клінічна система, що складається з багаторівневої системи алгоритмів класичного машинного навчання, глибоких згорткових мереж та гібридних нейронних мереж [80].

Основні виклики у побудові клінічної ІСППР для діагностики раку шкіри полягають у такому.

1. Різноманітність протоколів отримання медичних зображень: класичним стандартом є мікроскопічні зображення, однак сучасні дослідження фокусуються на розробці методів діагностики за фотографіями шкіри, зокрема 3D-фотографіями всього тіла (ТВР) [81]. Приклади наведені на рис. 3.6.
2. Значний дисбаланс позитивного класу у вибірці – зображення з діагностованою меланою зазвичай становлять менше 1%.
3. Відсутність прикладних мультимодальних систем, здатних водночас обробляти зображення та метадані (описи утворень, інформацію про пацієнта тощо) на рівні моделей глибокого навчання.

4. Значний шум анотацій у тренувальних даних – значна частина діагнозів була встановлена візуально, без гістологічного підтвердження.

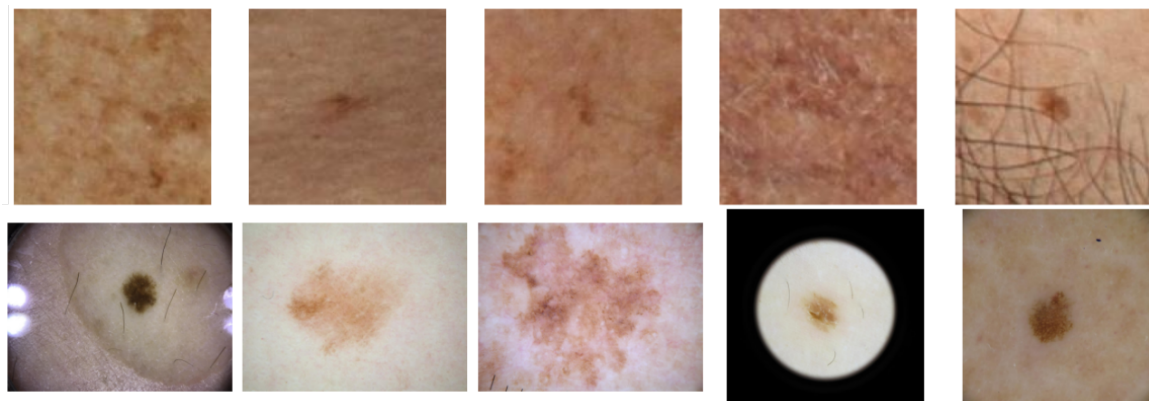


Рис. 3.6 – Перший ряд – приклади злоякісних утворень раку шкіри на фотографіях; другий ряд – приклади утворень на мікроскопічних знімках.

Розроблена система приймає на вхід медичні зображення (Image) – фотографії шкірних утворень, підозрілих на меланому, а також метадані про пацієнта, розташування утворення, характеристики зображення та додаткові ознаки, згенеровані самою системою (Meta). Отримана інформація обробляється ансамблем моделей, що складається з класичних алгоритмів машинного навчання на основі дерев рішень (Tabular Models), глибоких згорткових нейронних мереж (Vision Models), а також гібридних нейронних мереж (Vision/Tabular Models). Схематичне зображення системи наведено на рис. 3.7.

Для побудови системи були запропоновані такі моделі та методи.

1. Розширення тренувальної вибірки зображеннями з ISIC Archive [82], а також додатковими даними, згенерованими дифузійними моделями [83]. Таким чином вибірка була доповнена як фотографіями, так і мікроскопічними зображеннями, що збільшило частку позитивних прикладів та загальну різноманітність даних.

2. Використання двоетапної оптимізації згорткових нейронних мереж, а також застосування гібридних нейронних мереж, здатних одночасно опрацьовувати зображення та метадані; схематичне зображення відповідної архітектури наведено на рис. 3.8. Двоетапна оптимізація дозволяє адаптувати модель до цільових даних

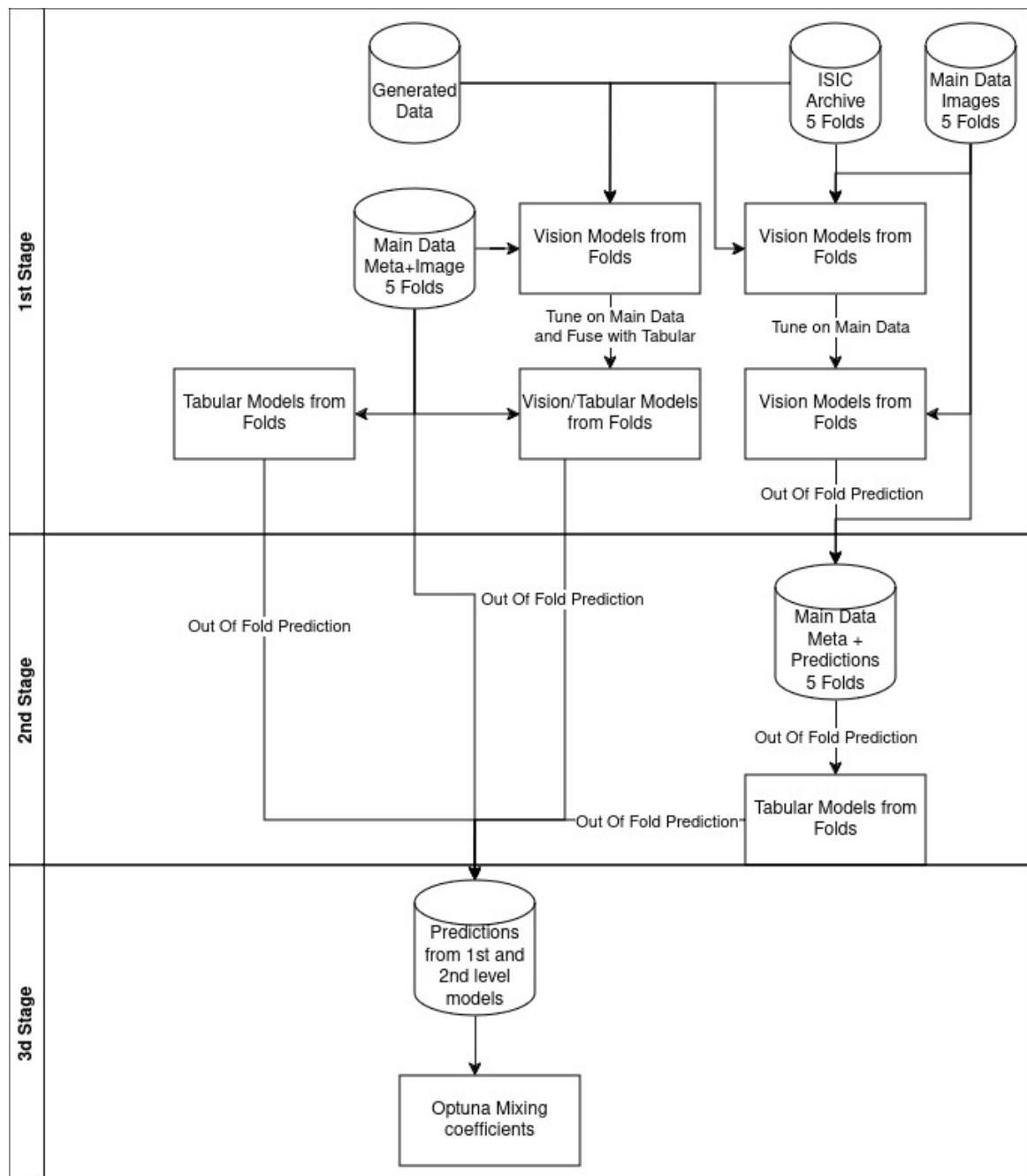


Рис. 3.7 – Схематичне зображення системи класифікації раку шкіри.

(фотографій) на другому етапі, а також отримати оптимізований енкодер для гібридних моделей. Схеми одноетапної (1st Approach) та двоетапних (2nd/3rd Approach) оптимізацій представлені на рис. 3.9.

3. Створення додаткових метаознак, зокрема: відносна площа утворення порівняно з найбільшим утворенням пацієнта, різниця забарвлення між центром та периферією утворення, інші відносні характеристики, обчислені для всіх утворень пацієнта.

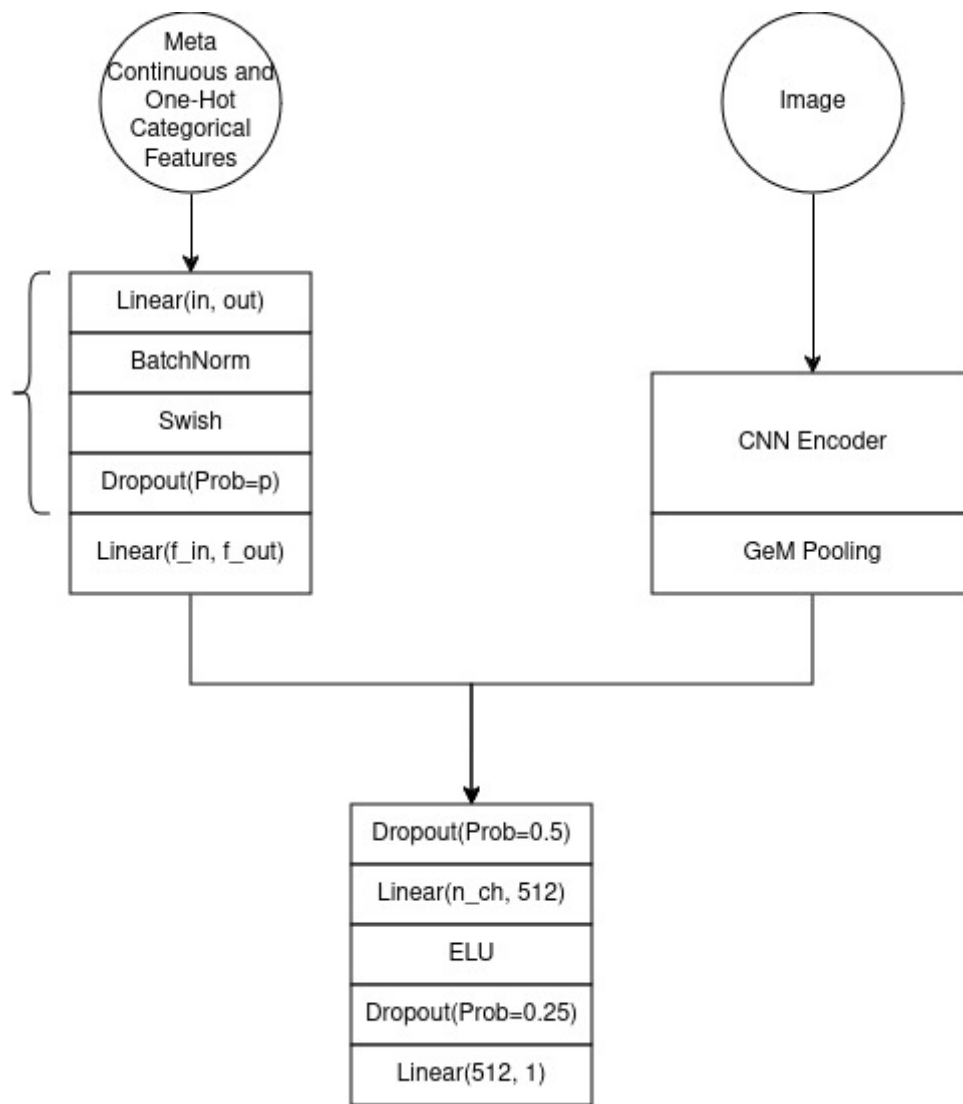


Рис. 3.8 – Схематичне зображення архітектури гібридної нейронної мережі, яка опрацьовує метадані (ліва гілка) та зображення (права гілка) й агрегує результати за допомогою фінального перцептрона.

4. Використання моделей різної природи. Для обробки зображень та побудови гібридних моделей застосовано низку згорткових архітектур: ConvNeXt V2 [84], EfficientNet V2 [85], EdgeNeXt [86]. Для обробки метаданих та побудови моделей другого рівня були використані методи на основі дерев рішень: LightGBM [87] та XGBoost [88]. Таке поєднання моделей підвищило робастність системи та дозволило ефективно працювати з різними модальностями даних.

5. Використання Optuna [89], неконвексного методу оптимізації гіперпараметрів, як моделі третього рівня для агрегування результатів моделей другого рівня.

Такий підхід дав змогу зважено усереднити передбачення моделей другого рівня для максимізації недиференційованої метрики ROC AUC, водночас уникаючи перенавчання системи;

6. Застосування трирівневої системи, де кожний наступний рівень отримує на вхід передбачення попереднього (а також метадані – лише на другому рівні). Така архітектура дозволяє калібрувати попередні передбачення та нівелювати їхні помилки для покращення фінального результату.

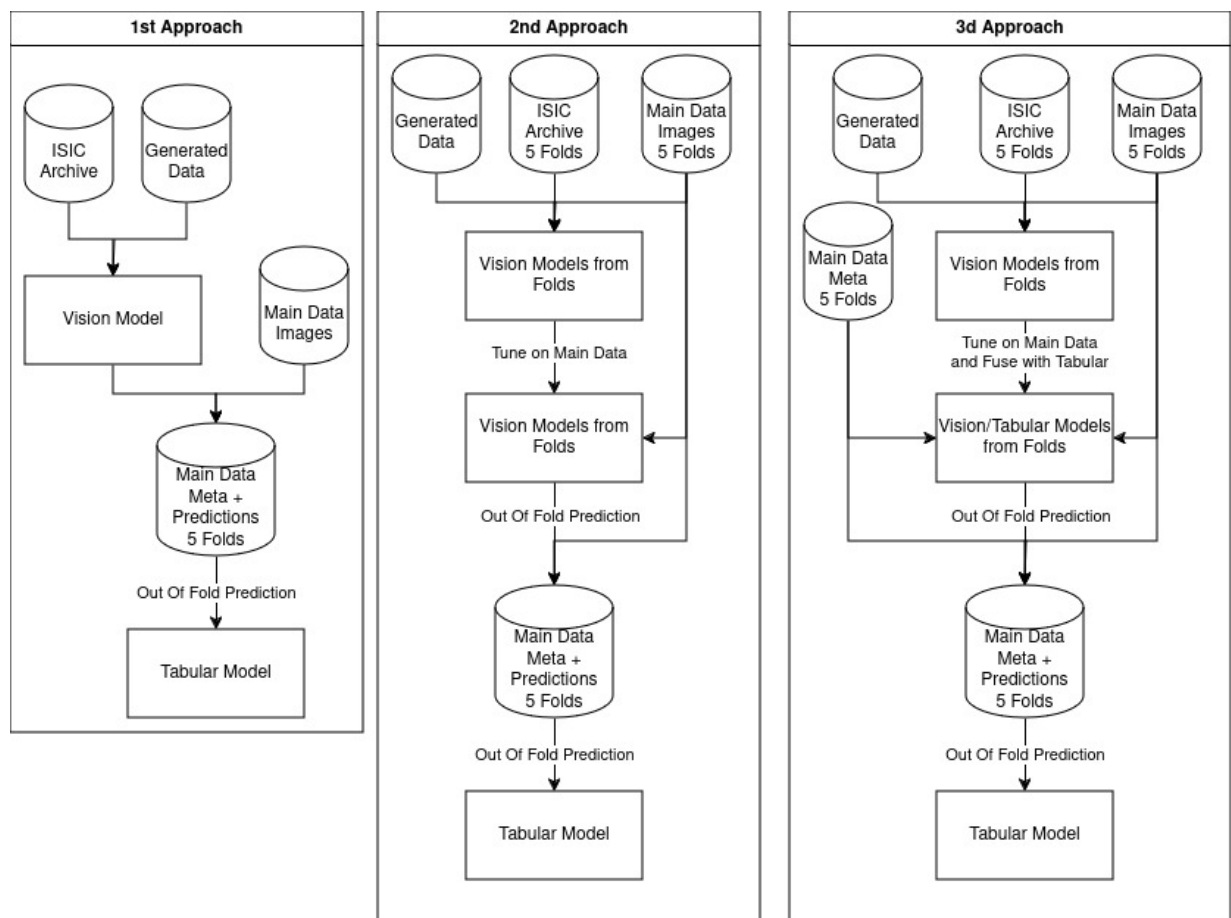


Рис. 3.9 – Схематичне зображення процесу оптимізації глибоких нейронних мереж для системи: перший підхід (1st Approach) – згорткова мережа з одноетапним процесом оптимізації; другий підхід (2nd Approach) – згорткова мережа з двоетапним процесом оптимізації; третій підхід (3rd Approach) – гібридна мережа з двоетапним процесом оптимізації.

Розроблена система продемонструвала високу точність прогнозу (ROC AUC на рівні 97%) на складних даних, отриманих із фотографій. Такий рівень

точності перевищує результати попередніх систем, створених для роботи з мікроскопічними зображеннями. Ці результати демонструють, що багаторівневі системи, які поєднують моделі глибокого навчання та класичного машинного навчання, ефективно обробляють мультимодальні дані та підвищують якість класифікації в умовах значного дисбалансу класів.

3.4 Висновки розділу 3

Розроблені та інтегровані системи, методи й моделі, представлені в Розділі 3, демонструють, що отримані напрацювання можуть бути застосовані не лише для діагностики раку легень, а й для інших завдань медицини та біології.

1. Продемонстровано, що використання загальної методології побудови клінічної ІСППР для діагностики раку легень, а також відповідних підсистем і моделей, дає змогу створити клінічну ІСППР для діагностики раку нирок, яка може бути впроваджена в медичну практику.

2. Запропоновані методи доменної адаптації для сегментації функціональних тканинних одиниць дозволили суттєво підвищити точність сегментації на цільовому домені.

3. Багаторівнева система моделей забезпечує можливість використання мультимодальних даних і моделей різної природи, зокрема гібридних нейронних мереж, для досягнення високої точності класифікації раку шкіри на складному типі медичних зображень – фотографіях шкіри.

Отже, запропоновані методи та системи не є специфічними лише для задачі діагностики раку легень, а можуть бути розширені й застосовані до інших прикладних задач, що значно підвищує їх наукову цінність.

4 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

4.1 Опис вибірки для ІСППР діагностики раку легенів

Для експериментальних досліджень розробленої системи, а також для оптимізації моделей глибокого навчання було обрано вибірку LIDC-IDRI [64]. Ця вибірка є стандартом для побудови моделей і систем для виявлення раку легенів. Вибірка складається з 1018 КТ-сканів від 1010 різних пацієнтів. Для анотації даних було залучено радіологів із 7 академічних центрів та 8 компаній, що працюють з медичними зображеннями. Було проведено двоетапну процедуру анотації [90].

1. Кожен радіолог переглядав кожен КТ-скан і відмічав одну з трьох категорій: утвір ≥ 3 мм; утвір < 3 мм; не утвір ≥ 3 мм.
2. Кожен радіолог переглядав утвори, які було відмічено ним (нею) самим, а також трьома іншими радіологами, щоб дійти до фінального висновку.

Вибірка надає доступ до анотацій усіх чотирьох радіологів, проте, на жаль, для деяких сканів дані щодо висновків певних радіологів відсутні. До фінальної вибірки включено лише утвори ≥ 3 мм, серед яких 928 утворів (34.7%) отримали маркування “утвір ≥ 3 мм” від усіх чотирьох радіологів.

Під час дослідження було перевірено валідність самих КТ-сканів, а саме: відсутність пошкоджень зображень, наявність необхідної метаінформації, коректність анотацій. Після валідації було відкинуто частину КТ-сканів, і в результаті фінальна вибірка для дослідження налічувала 850 КТ-сканів від 850 пацієнтів. Ця вибірка містить 2525 утворів з таким розподілом за класами ризику злоякісності.

1. Невизначене (Indeterminate) – 1177.
2. Дуже малоімовірно (Highly Unlikely) – 312.
3. Помірно малоімовірно (Moderately Unlikely) – 532.
4. Помірно підозріле (Moderately Suspicious) – 332.
5. Дуже підозріле (Highly Suspicious) – 172.

Об'єми утворів варіюються від 4.394 мм³ до 29067.2489 мм³, а загальна кількість утворів на одного пацієнта – від 1 до 23. Розподіли наведено на рис. 4.1.

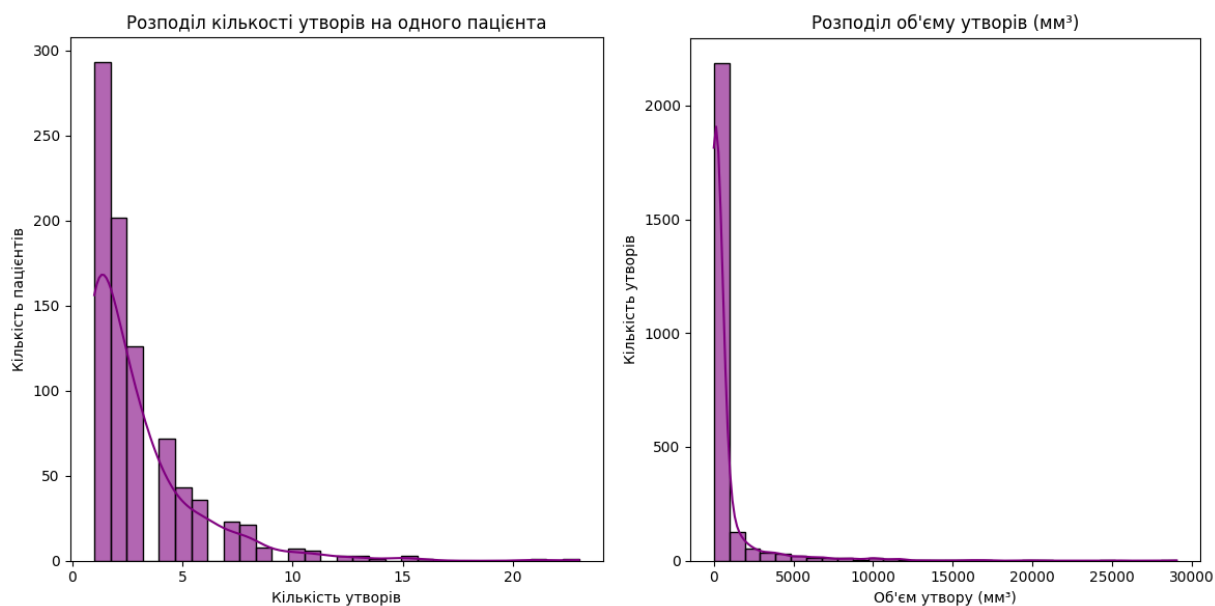


Рис. 4.1 – Зліва подано гістограму розподілу кількості утворів на пацієнта, справа – гістограму розподілу об'ємів утворів (мм³).

Описуючи характеристики КТ-зображень із розглянутої вибірки, важливо звернути увагу на такі ознаки.

1. Фізична величина сторони пікселя в аксіальній площині варіюється від 0.46 мм до 0.976 мм – розподіл наведено на рис. 4.2.
2. Товщина аксіальних зрізів змінюється від 0.45 мм до 5 мм – розподіл наведено на рис. 4.2.
3. КТ-скани отримано з чотирьох типів сканерів різних виробників – розподіл наведено на рис. 4.3.
4. Вибірка містить КТ-скани, що були зроблені з уведенням контрастної речовини та без неї; при цьому присутня значна різноманітність контрастних агентів – усього 66 видів. Розподіл КТ-сканів із та без контрасту наведено на рис. 4.3.

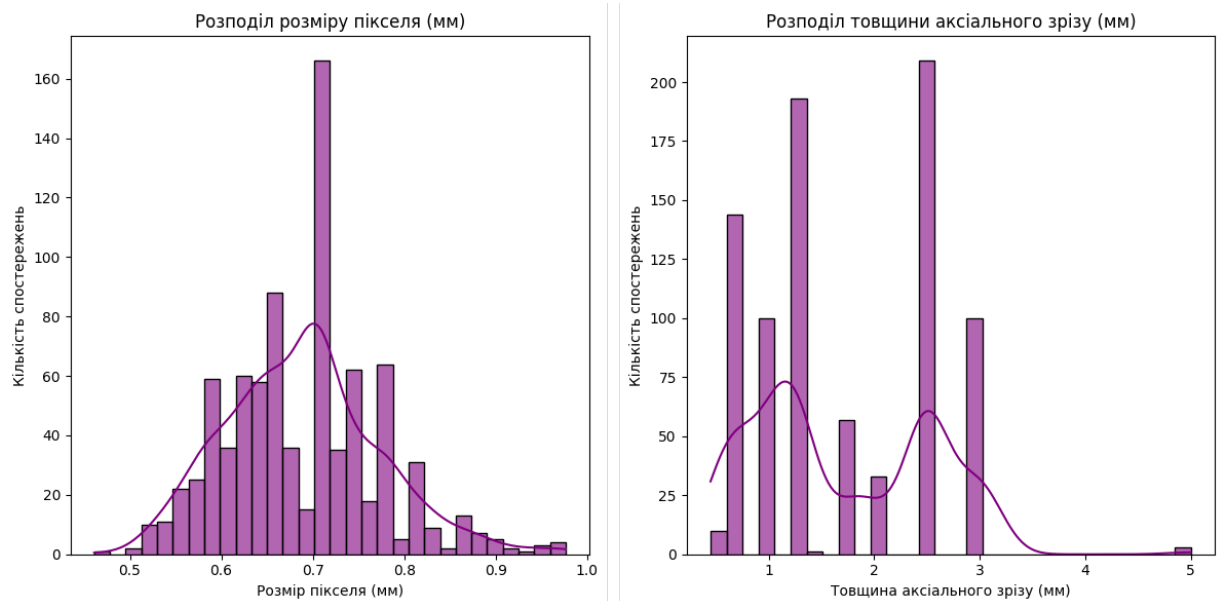


Рис. 4.2 – Зліва наведено гістограму розподілу фізичної величини сторони пікселя в аксіальній площині, справа – розподіл товщини аксіальних зрізів.

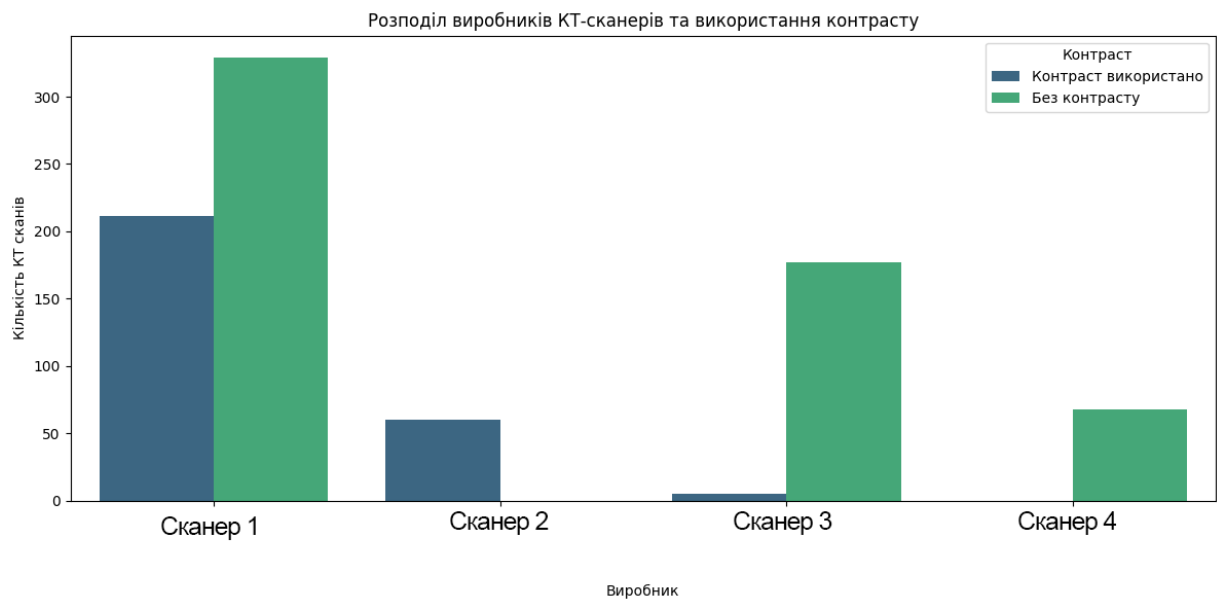


Рис. 4.3 – Розподіл сканів за виробниками КТ-сканерів та використанням контрастної речовини.

Загалом, обрана вибірка є досить репрезентативною і охоплює протоколи досліджень та типи сканерів, які найчастіше застосовуються в медичній практиці. Серед основних недоліків даної вибірки можна виділити.

1. Достатню, проте не дуже велику кількість КТ-сканів – найбільші вибірки містять не сотні, а тисячі зображень.
2. Не оптимальну таксономію, яка не відповідає системі Lung–RADS. [5].

Приклади зображень із вибірки наведено на рис. 4.4.

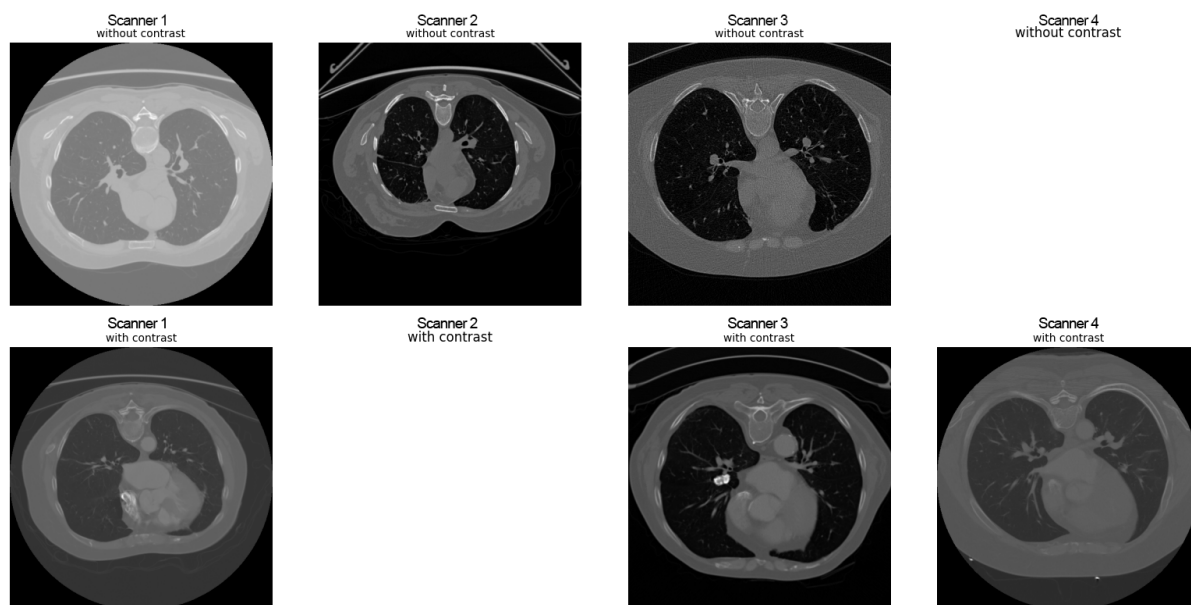


Рис. 4.4 – Приклади аксіальних зрізів, отриманих на різних КТ-сканерах, з використанням та без використання контрастної речовини. Порожні комірки означають відсутність відповідної комбінації виробника та контрастування у вибірці.

В силу невеликої кількості зображень у вибірці для подальшої оцінки моделей і системи загалом було використано механізм перехресної оцінки (Cross Validation) [91]. Вибірку було поділено на 5 підвибірок; також було застосовано групування за пацієнтом, щоб декілька сканів одного пацієнта не потрапили до різних підвибірок – це необхідно для коректного відтворення реальної медичної практики застосування подібних систем. Крім того, було збережено розподіл сумарного об'єму утворів у кожному скані для всіх підвибірок. Кожна підвибірка містить по 170 сканів.

4.2 Результати оцінювання моделі сегментації утворів у легенях

Для оптимізації моделі сегментації утворів використовувалися такі налаштування.

1. Розмір патча, вирізаного з КТ-скану – [256, 256, 32]. Такий вибір аргументований тим, що в більшості випадків утвори мають більші розміри в аксіальній площині, ніж уздовж товщини зрізів.
2. Фізичний розмір вокселя було приведено до значень [0.625, 0.625, 1.0], що відповідає медіанним значенням по вибірці.
3. Під час оптимізації використовувалося 12 патчів у пакеті (batch size).
4. КТ-скан обрізався за значеннями HU в межах [−1024, 700] та далі нормалізувався в діапазон [0, 1].

Були застосовані такі аугментації: відзеркалення по всіх трьох осях; повороти на 90° в аксіальній площині; повороти до 0.5236 радіан по всіх осях; масштабування в межах від 0.7 до 1.4; гаусівський шум та гаусівське згладжування; зміна інтенсивності та контрасту.

Для оптимізації використовувався стохастичний градієнтний спуск зі швидкістю навчання 0.01, моментом 0.99 та L2-регуляризацією $3e-5$. Швидкість навчання поступово змінювалася за половинним косинусним циклом [92] до $1e-6$, оновлення параметрів відбувалося раз на епоху. Оптимізація тривала 4500 епох; одна епоха відповідала проходженню всієї тренувальної вибірки (4 підвибірки перехресної оцінки з підрозділу 4.1), що становить 56 ітерацій оновлення.

Загалом модель містить 31 181 697 параметрів. Процес оптимізації пройшов стабільно, без значних флуктуацій значень функції втрат і метрик (Surface Dice 3D). Графіки наведено на рис. 4.5.

Після оптимізації було обрано три найкращі моделі за метрикою Surface Dice 3D і виконано усереднення їхніх ваг для отримання кращого оптимуму. Маски з усередненої моделі отримувалися шляхом усереднення інференсів на восьми можливих комбінаціях відзеркалень (Test-Time Augmentation) [62].

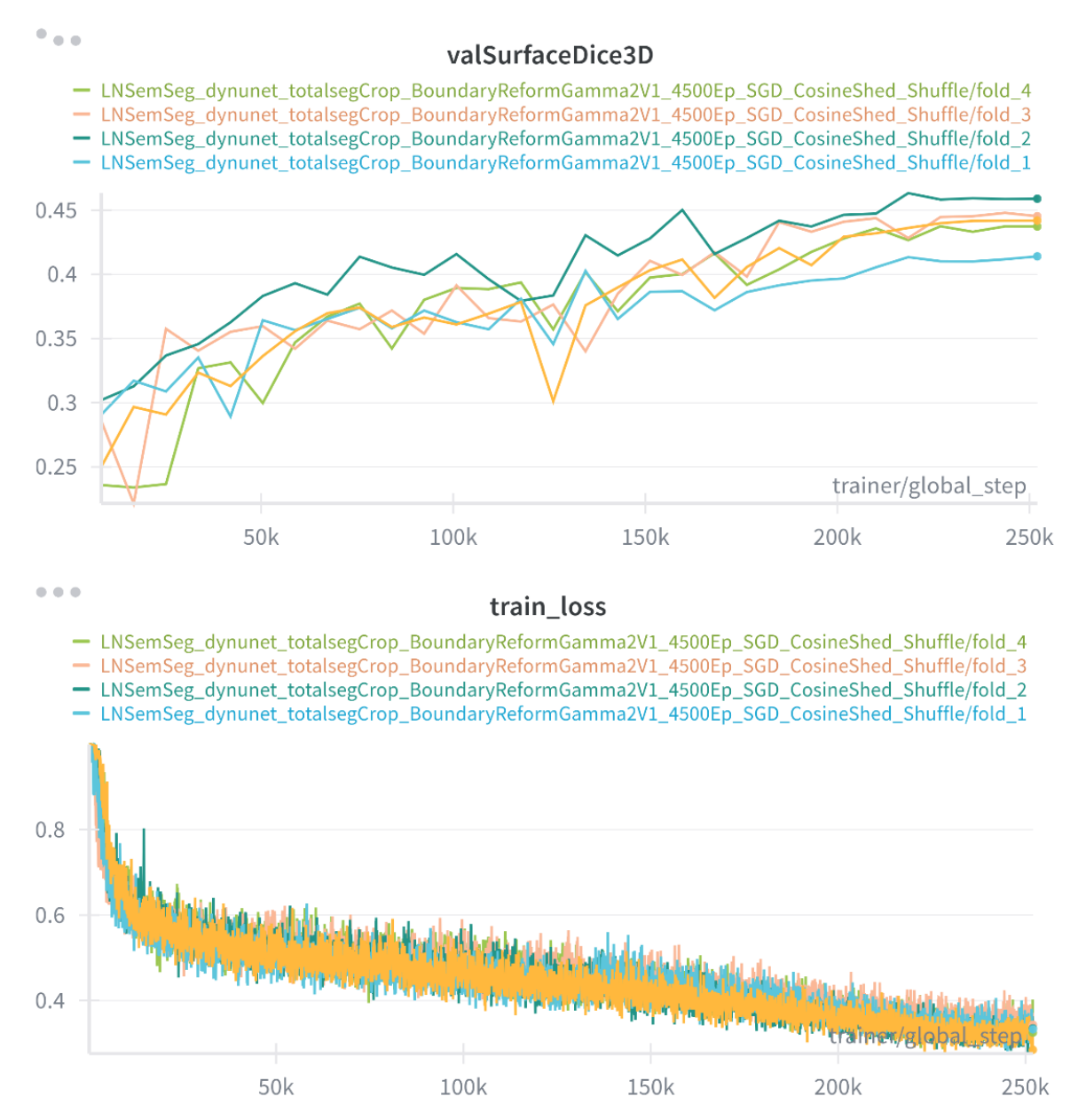


Рис. 4.5 – Графіки метрики Surface Dice (угорі) на валідаційних підвбірках та значення функції втрат Boundary Dou++ (унизу) на тренувальних підвбірках для всіх моделей перехресної перевірки під час їх оптимізації.

Оцінювання проводилося методом перехресної оцінки: для кожної підвбірки будувалася окрема модель, після чого здійснювалися передбачення на всій вибірці, а метрики обчислювалися для кожного КТ-скану в 3D-варіанті з подальшим усередненням по всіх сканах. Отримані результати перевершили базове дослідження [60] за всіма метриками – підсумки наведено в таблиці 4.10. Дослідження залежності метрик від порогу бінаризації (рис. 4.6) показує, що

оптимальні значення досягаються при дещо вищих порогах (0.5–0.7), що є логічним, адже метрика Dice фактично є F1–мірою й вимагає балансування похибок першого та другого роду, а сама задача має суттєво незбалансовану структуру.

Далі, аналізуючи помилки першого та другого роду на рівні всіх пікселів датасету, за метриками F1–міри, повноти (Recall) та точності (Precision), отримано оптимальне значення F1–міри 0.7034 на порозі 0.6; тенденцію наведено на рис. 4.7.

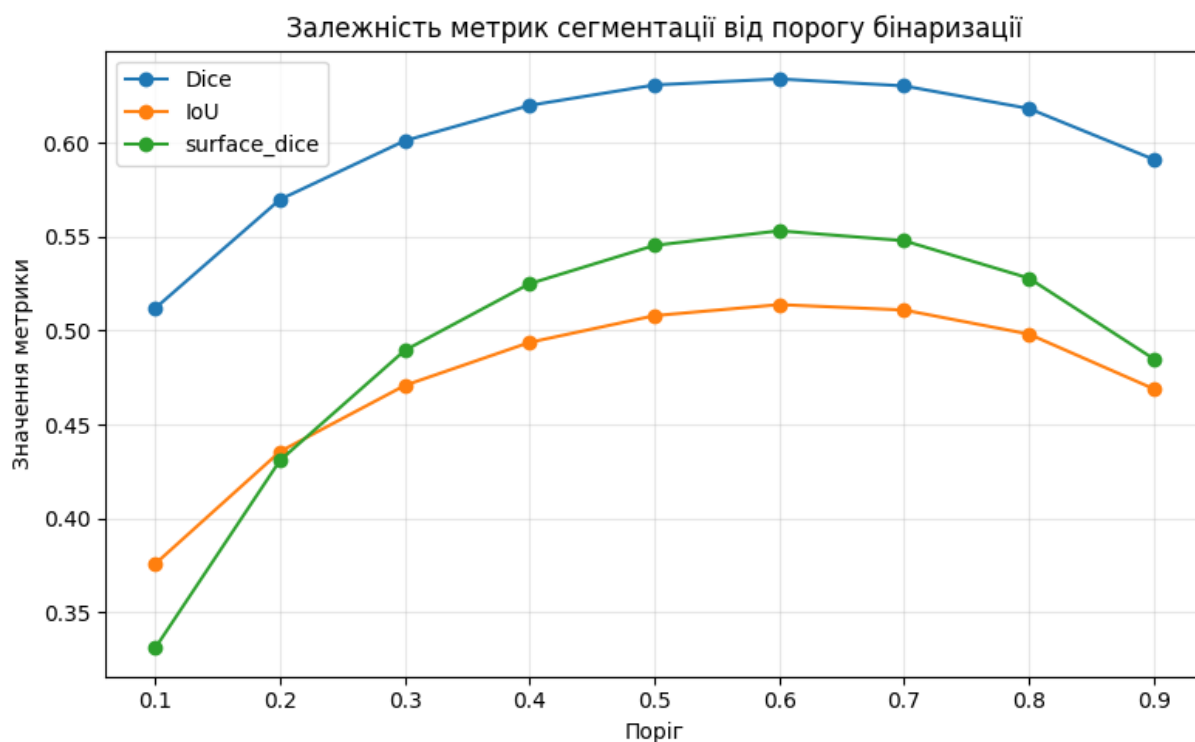


Рис. 4.6 – Залежність метрик сегментації від порогу бінарizaції.

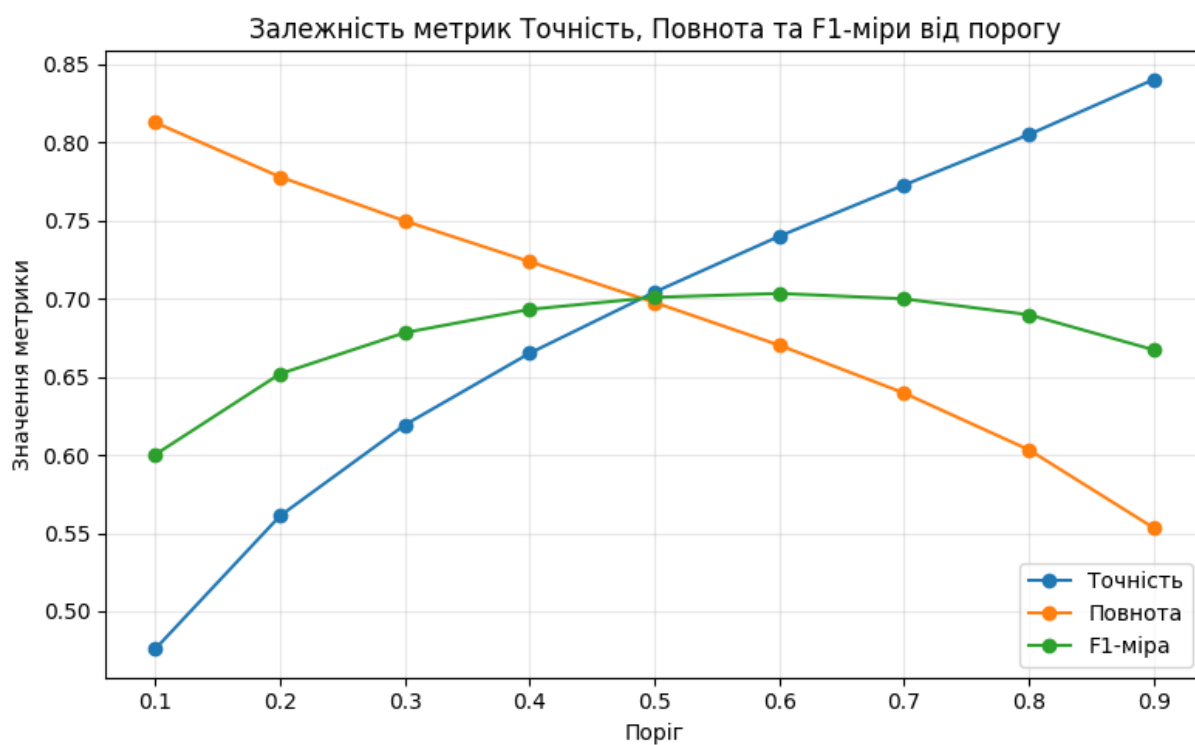


Рис. 4.7 – Залежність по-піксельної F1-міри, повноти та точності від порогу бінаризації.

Таблиця 4.1 – Метрики сегментаційної моделі

Експеримент	IoU	Surface Dice	Dice	Best Threshold
Попереднє дослідження [60]	0.488049	0.449673	–	0.6
Поточний результат	0.513877	0.553212	0.634158	0.6

Вивчаючи розподіл значень метрик по сканах (рис. 4.8), можна зробити висновок, що значна частина погіршень метрик зумовлена наявністю одного малого хибно–негативного або хибно–позитивного утвору на скан. Натомість ті скани, які містять кілька утворів або утвори більшого розміру, мають модальні значення метрик у межах 0.7–0.8.

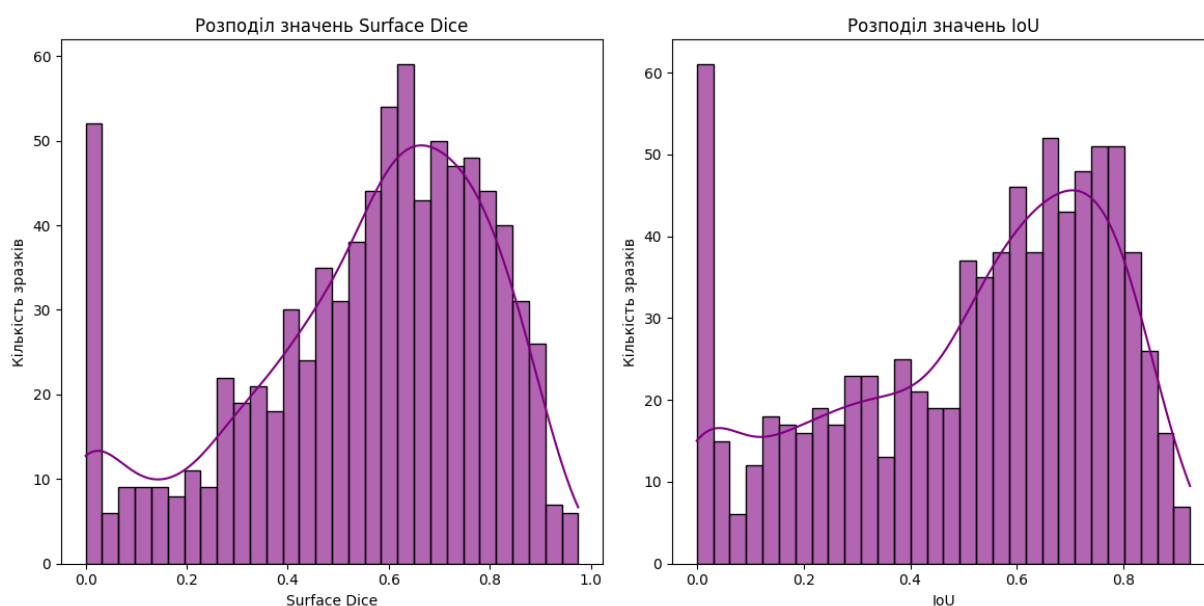


Рис. 4.8 – Розподіл значень метрик по сканам всієї вибірки.

Враховуючи складність задачі виявлення дуже малих утворів, а також алгоритм обчислення метрик на рівні кожного 3D–скану (коли один хибно–негативний або хибно–позитивний утвір може суттєво знизити значення метрики в усередненні по 850 сканах), отримані результати демонструють доволі високу точність моделі для задачі сегментації утворів у легенях.

4.3 Результати оцінювання моделі класифікації утворів у легенях

Для оптимізації моделі класифікації утворів використовувалися такі налаштування:

- розмір патча [64, 64, 64] та значення паддінгу у випадку надто великого утвору перед масштабуванням до розміру патча 16;
- фізичний розмір вокселя такий самий, як і для сегментації [0.625, 0.625, 1.0];
- розмір пакету (batch size) 64;
- нормалізація інтенсивності така сама, як і для сегментації.

Набір аугментацій був описаний у підрозділі 2.8, включаючи дві нові запропоновані аугментації. Параметри $\alpha_i^{\min}$, $\alpha_i^{\max}$ було визначено як -0.75 і 0.75 відповідно для всіх осей. Таке велике значення зумовлене тим, що більшість утворів є досить малими, і зміщення до 75% від сторін утвору може відповідати всього кільком (до 10) пікселям.

Для оптимізації використовувався оптимізатор Adam [93] зі силою навчання 0.001. Сила навчання поступово змінювалася за половинним косинусним циклом [92] до $1e-5$, оновлення сили навчання відбувалося раз на епоху. Оптимізація тривала 75 епох; одна епоха відповідала 16 ітераціям оновлень. Також використовувався механізм ранньої зупинки, якщо макро F1–міра по всіх класах не покращувалася протягом 20 епох.

Модель 3D ResNet–50 містить 46 144 772 параметри. Модель оптимізовувалася відповідно до функції втрат: крос–ентропія у мультикласовому формулюванні та бінарна крос–ентропія у бінарному формулюванні. Процес оптимізації був загалом стабільним відповідно до функції втрат і цільової метрики (макро F1–міри). Графіки наведено на рис. 4.9.

Для інференсу було обрано найкращі моделі за макро F1–мірою. Ймовірності усереднювалися на восьми можливих комбінаціях відзеркалень (Test–Time Augmentation) [62].

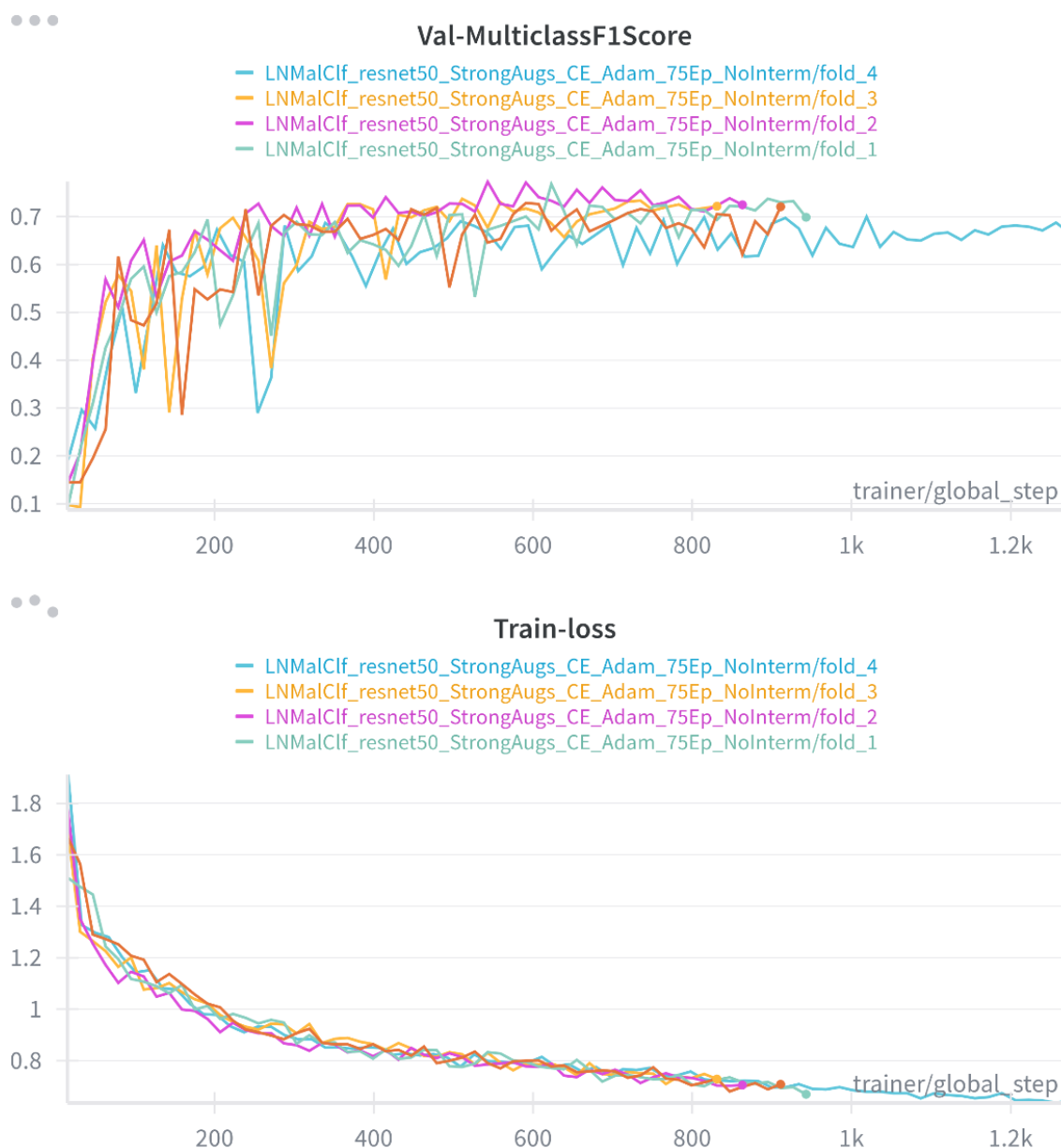


Рис. 4.9 – Графіки метрики Макро F1–міри (угорі) на валідаційних підвбірках та значення функції втрат Крос–Ентропія (унизу) на тренувальних підвбірках для всіх моделей перехресної перевірки під час їх оптимізації.

Оцінювання моделей виконувалося за тою самою методологією, що й у підрозділі 4.2. Було оптимізовано низку моделей у мультикласовому та бінарному формулюваннях задачі.

1. Мультикласова Модель 1 – оптимізована без класу Невизначене (Indeterminate) та з аугментацією випадкового зсуву.

2. Мультикласова Модель 2 – оптимізована з класом Невизначене (Indeterminate) та з аугментацією випадкового зсуву.

3. Мультикласова Модель 3 – оптимізована без класу Невизначене (Indeterminate) та без аугментації випадкового зсуву.

4. Бінарна Модель 1 – оптимізована без класу Невизначене (Indeterminate) та з аугментацією випадкового зсуву.

5. Бінарна Модель 3 – оптимізована без класу Невизначене (Indeterminate) та без аугментації випадкового зсуву.

Мультикласові моделі також були оцінені на утворах, сегментованих сегментаційною підсистемою, а потім локалізованих підсистемою характеристики та фільтрації утворів. Таким чином у процесі дослідження було оцінено точність роботи системи як на «ідеальній» локалізації, так і на «реалістичній», отриманій із попередніх підсистем.

Як було зазначено в розділі 4.1, кожен утвір був анотований чотирма радіологами (хоча для деяких утворів відсутні анотації частини з них, тобто кількість анотацій на один утвір варіюється від 1 до 4, медіана – 3). Поширеним підходом є використання найвищого рівня підозри серед усіх анотаторів для відображення найгіршого сценарію. Проте такий підхід часто призводить до появи шуму в мітках; тому класи були відображені на порядкову числову шкалу: *Дуже малоймовірно* $\rightarrow 0, \dots, \text{Дуже підозріле} \rightarrow 4$, після чого для кожного утвору обчислювалася медіана міток.

Для створення анотацій для бінарних моделей і подальшого порівняння їх із мультикласовими було запропоновано такі правила перетворення класів у бінарні:

- Дуже малоймовірно (Highly Unlikely) $\rightarrow$ Доброякісний (Benign);
- Помірно малоймовірно (Moderately Unlikely) $\rightarrow$ Доброякісний (Benign);
- Помірно підозріле (Moderately Suspicious) $\rightarrow$ Злоякісний (Malignant);
- Дуже підозріле (Highly Suspicious) $\rightarrow$ Злоякісний (Malignant).

Таке перетворення може бути не повністю медично коректним, проте в контексті моделювання воно адекватно відображає розподіл ймовірності «злоякісності»: нижчі значення відповідають малоймовірним класам, а вищі – підозрілим.

Для переходу від мультикласового прогнозу до бінарного застосовувалися два методи: (1) метод максимуму – формула 4.1; (2) метод суми – формула 4.2.

$$Prob_{\text{Malignant}} = \max (Prob_{\text{Moderately Suspicious}}, Prob_{\text{Highly Suspicious}}) . \quad (4.1)$$

$$Prob_{\text{Malignant}} = Prob_{\text{Moderately Suspicious}} + Prob_{\text{Highly Suspicious}} . \quad (4.2)$$

$$Prob_{\text{Benign}} = 1 - Prob_{\text{Malignant}} . \quad (4.3)$$

Було отримано такі результати: робота мультикласових моделей на «ідеальній» локалізації наведена у таблиці 4.2; на «реалістичній» локалізації – таблиця 4.3; робота бінарних та агрегованих мультикласових моделей – таблиця 4.4.

Таблиця 4.2 – Результати експериментів мультикласових моделей

Назва	Макро ROC AUC	Макро Повнота	Макро Точність	Макро F1
Мультикласова модель 1	0.9011	0.7333	0.7409	0.7346
Мультикласова модель 2	0.8255	0.5560	0.5560	0.5392
Мультикласова модель 3	0.8961	0.7213	0.7346	0.7249

Таблиця 4.3 – Результати експериментів мультикласових моделей на рамках, отриманих з моделі верхнього рівня

Назва моделі	Макро ROC AUC	Макро Повнота	Макро Точність	Макро F1
Мультикласова модель 1	0.9176	0.7644	0.7681	0.7658
Мультикласова модель 3	0.9173	0.7635	0.7685	0.7655

Таблиця 4.4 – Результати експериментів бінарних та агрегованих мультикласових моделей

Назва моделі	Бінарний ROC AUC	Бінарна Повнота	Бінарна Точність	Бінарна F1
Бінарна модель 1	0.9383	0.8651	0.8685	0.8668
Бінарна модель 3	0.9297	0.8373	0.8612	0.8491
Мультикласова модель 1 (сума)	0.9365	0.8552	0.8637	0.8594

Таблиця 4.4 – (продовження)

Назва моделі	Бінарний ROC AUC	Бінарна Повнота	Бінарна Точність	Бінарна F1
Мультикласова модель 1 (макс)	0.9300	0.7599	0.8764	0.8140
Мультикласова модель 2 (сума)	0.9273	0.8075	0.8772	0.8409
Мультикласова модель 2 (макс)	0.9221	0.7341	0.8937	0.8061

Також були обчислені матриці невірних класифікацій для різних моделей – для мультикласових (рис. 4.10) та для бінарної й агрегованої мультикласової (рис. 4.11).

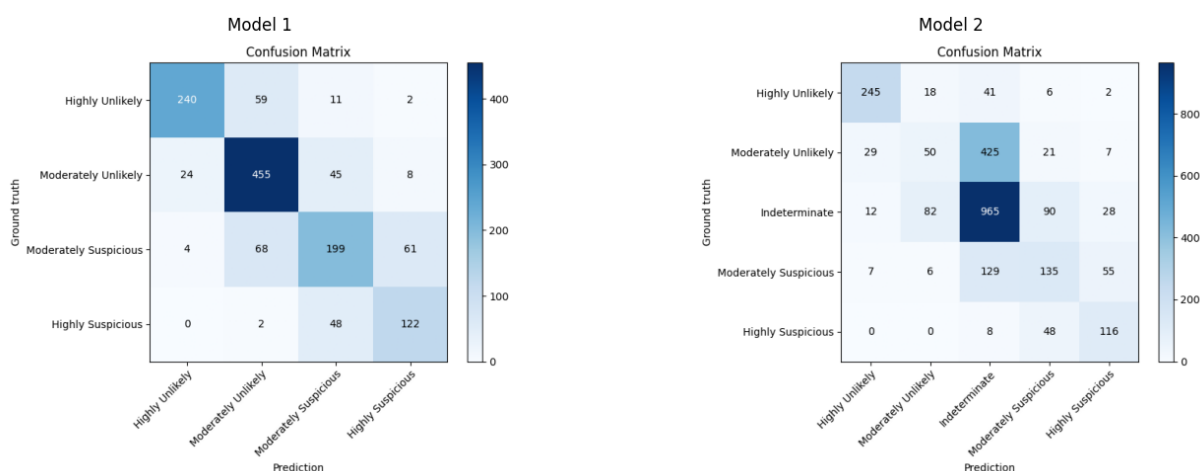


Рис. 4.10 – Матриця невірних класифікацій для Мультикласової Моделі 1 (зліва) та для Мультикласової Моделі 2 (справа).

На основі отриманих результатів можна сформулювати такі висновки.

1. Серед мультикласових моделей найбільший приріст метрик (+7.5% макро ROC AUC) забезпечує усунення класу Невизначене (Indeterminate), що логічно, адже цей клас є найбільш шумним. Дещо менший приріст дає додавання аугментації випадкового зсуву (+0.5% макро ROC AUC).

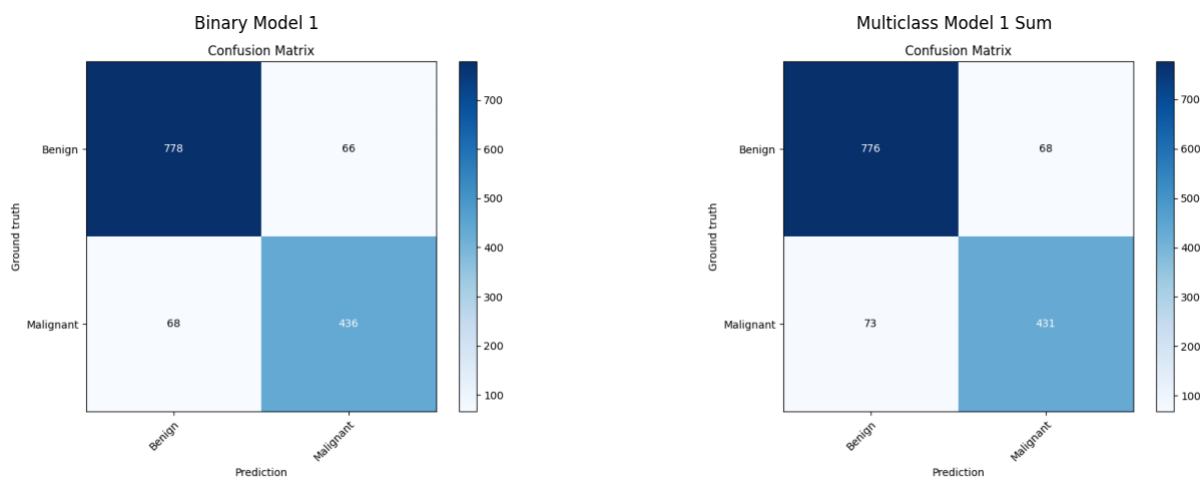


Рис. 4.11 – Матриця невірних класифікацій для Бінарної Моделі 1 (зліва) та для Мультикласової Моделі 1, агрегованої методом суми (справа).

2. Аналіз матриць невірних класифікацій демонструє, що найбільше помилок виникає між суміжними класами Помірно малоймовірно (Moderately Unlikely) та Помірно підозріле (Moderately Suspicious). У присутності класу Невизначене (Indeterminate) кількість помилок між усіма трьома класами суттєво зростає.

3. У випадку застосування моделі на «реалістичних» локалізаціях порівняно з «ідеальними» метрики класифікації навіть дещо покращуються. Це може пояснюватися тим, що модель локалізації пропускає найменші (а отже й найскладніші) утвори, що полегшує задачу для моделі класифікації. Також метрики Мультикласової Моделі 1 вищі за метрики Мультикласової Моделі 3, що підтверджує гіпотезу про позитивний вплив запропонованої аугментації на робастність моделі в умовах роботи у клінічній ІСППР.

4. Порівняння бінарних і агрегованих мультикласових моделей демонструє, що бінарна модель дає кращі результати. Аналіз двох методів агрегації показує: агрегація сумою пріоритизує зменшення помилок другого роду, а агрегація максимумом – помилок першого роду. Це логічно, зважаючи на формули 4.1, 4.2 та використання Softmax у мультикласовій моделі (імовірності сумуються до 1). Для максимізації ROC AUC та F1–міри агрегація сумою є оптимальною, що узгоджується з медичною практикою пріоритезації мінімізації помилок другого

роду. Загалом матриці похибок бінарної й агрегованої моделей демонструють близьку якість.

Порівнюючи розроблену підсистему з попередніми дослідженнями, можна побачити, що Мультикласова Модель 1 та Бінарна Модель 1 перевершують раніше запропоновані методи за метрикою ROC AUC (таблиця 4.5), що додатково підтверджує високу точність розробленої підсистеми.

Таблиця 4.5 – Порівняння з попередніми дослідженнями

Метод	ROC AUC
Attention pyramid pooling network [94]	0.914
Lung Nodule Classification [95]	0.800
Бінарна модель 1	0.9383
Мультикласова модель 1	0.9176

Приклади роботи мультикласової та бінарної моделей наведено на рис. 4.12.



Рис. 4.12 – Приклади роботи Бінарної Моделі 1 (зліва) та Мультикласової Моделі 1 (справа).

4.4 Результати оцінювання загальної роботи ІСППР

Для загальної оцінки роботи клінічної ІСППР були обчислені такі групи метрик.

1. Метрики локалізації – F1–міра, повнота та точність для рамок утворів, отриманих підсистемою сегментації та підсистемою характеристизації; також досліджено метрики для різних класів утворів та залежність похибок від розміру.
2. Метрика FROC [27] для локалізованих утворів, використовуючи ймовірності “злості”, отримані з Бінарної Моделі.
3. Метрики F1–міри, повноти та точності для локалізованих утворів, класифікованих Мультикласовою Моделлю.

Варто зазначити, що наведені в цьому розділі результати відповідають роботі статичної ІСППР. У наступному розділі ті самі метрики будуть розглянуті з урахуванням інтерактивної складової.

Визначення TP, FP та FN. Для обчислення метрик локалізації спершу необхідно визначити істинно позитивні (TP), хибно позитивні (FP) та хибно негативні (FN) об’єкти. Застосовано визначення з розділу 2.10.

1. Positive – відповідає істинно позитивним (TP).
2. Negative – хибно позитивні (FP).
3. False Negative – хибно негативні (FN).
4. Borderline (відповідає True Positive Borderline) – зазвичай зараховується до TP.

У таблиці 4.6 наведено метрики локалізації при двох варіантах трактування Borderline: Borderline $\rightarrow$ TP та Borderline $\rightarrow$ FN та FP. З отриманих результатів видно, що підсистема локалізації знаходить більшість справжніх утворів, але демонструє відносно високу кількість хибно позитивних спрацювань. Це є прийнятною поведінкою клінічних систем, оскільки пропуск злостісного утвору набагато небезпечніший, ніж зайве хибне попередження. Надалі, враховуючи не-

великі об'єми утворів (що може зумовлювати малі IoU), Borderline буде віднесено до TP.

Таблиця 4.6 – Метрики локалізації

Випадок	Точність (Precision)	Повнота (Recall)	F1–міра	F2–міра	F0.5–міра
Borderline у TP	0.4801	0.7965	0.5991	0.7037	0.5215
Borderline у FN та FP	0.4546	0.7541	0.5672	0.6663	0.4938

Похибки за класами утворів. У таблиці 4.7 наведено статистику TP/FN для всіх класів. Найбільші похибки спостерігаються для малоюмовірних класів та класу Невизначений, що узгоджується з тим, що система працює значно краще на утворах з високою підозрою на онкологію.

Таблиця 4.7 – Статистика TP/FN локалізації та відсоткові значення для кожного класу

Клас	TP	FN	TP%	FN%
Дуже підозрілий	156	16	0.906977	0.093023
Дуже малоюмовірний	257	53	0.829032	0.170968
Невизначений	910	256	0.780446	0.219554
Помірно підозрілий	283	44	0.865443	0.134557
Помірно малоюмовірний	386	140	0.733840	0.266160

На рис. 4.13 показано розподіл об'ємів рамок утворів для TP, FN та FP. Медіанні значення (в лінійній шкалі) становлять: 242.393 мм³ для TP; 85.070 мм³ для FN; 42.407 мм³ для FP. Це дозволяє зробити такі висновки.

1. Підсистеми локалізації добре виявляють середні та великі утвори – саме ті, що становлять найбільшу клінічну небезпеку.

2. Хибно позитивні спрацювання здебільшого стосуються малих утворів, що можуть бути як ранніми онкологічними проявами, так і артефактами анотації.

Загальна тенденція похибок відповідає патернам поведінки радіологів: найскладнішими є малі та малоймовірні утвори. Це дозволяє стверджувати, що попри невисокі значення метрик локалізації, характер похибок системи є логічним і порівнянним з експертними.

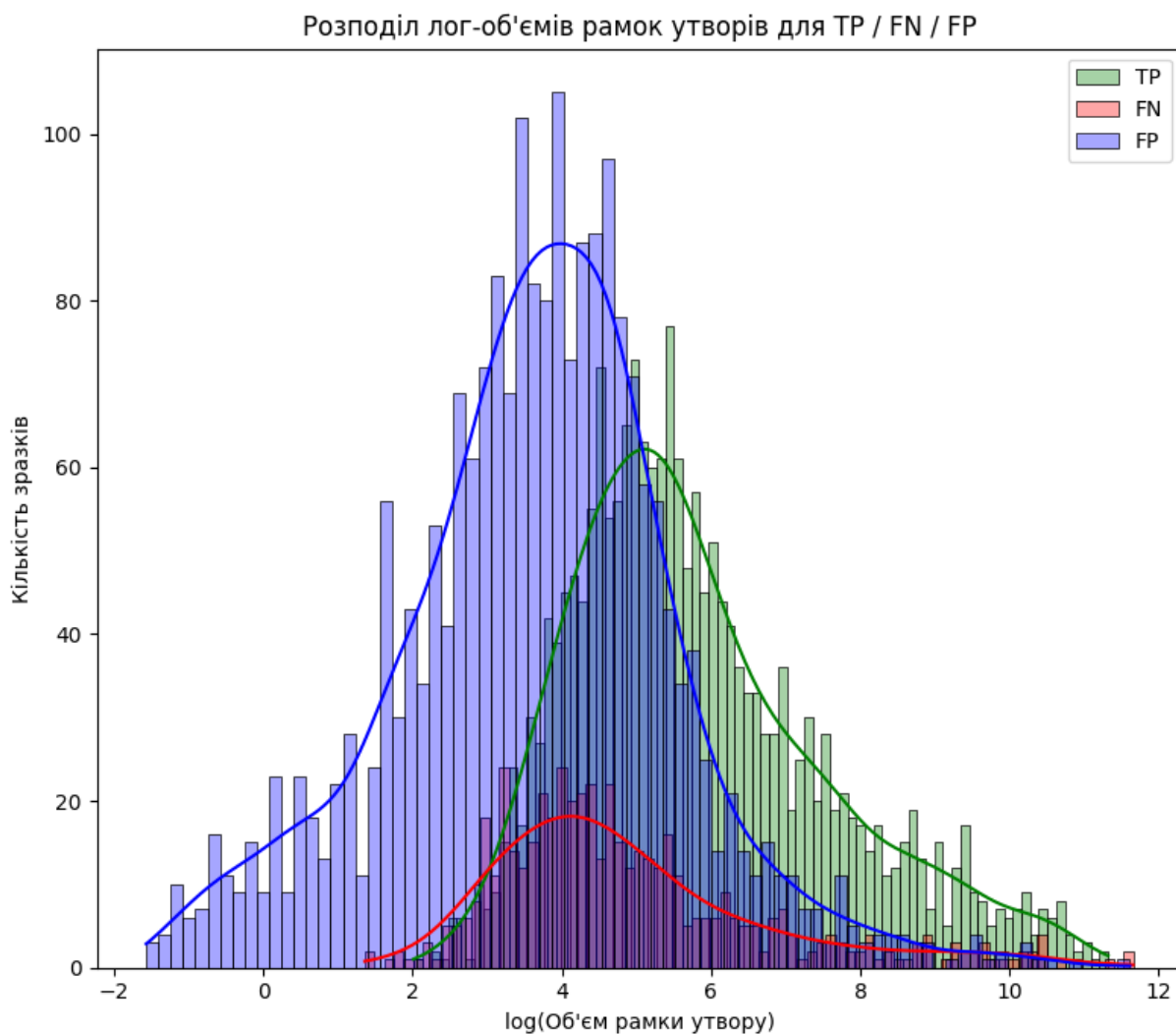


Рис. 4.13 – Розподіл об'ємів рамок утворів для істинно позитивних (TP), хибно негативних (FN) та хибно позитивних (FP) випадків локалізації утворів.

Аналіз з використанням Бінарної моделі. Далі розглянемо прогнози Бінарної Моделі 1 для локалізованих утворів. Ймовірність “злоякісності” може бути застосована як фільтр для зменшення кількості FP. Оптимальний поріг залежить

від клінічного сценарію, тому використовуємо метрику FROC, яка відображає залежність повноти від середньої кількості FP на один КТ-скан.

Метрику проаналізовано у двох варіантах.

1. Для всіх істинних утворів.
2. Після перетворення з підрозділу 4.3 у класи доброякісного та злоякісного утвору – істинними вважаються лише злоякісні утвори, усі Невизначені (Indeterminate) вилучено.

FROC-криві наведені на рис. 4.14, а детальна статистика – у таблиці 4.8. Результати демонструють, що підсистема класифікації може суттєво зменшити кількість FP, а якщо радіолога цікавлять лише злоякісні утвори, повнота усієї ІСППР може досягати 85–88% при лише 1–2 FP на скан. Якщо ж необхідно виявляти всі утвори, повнота становить близько 73% при 2 FP на скан.

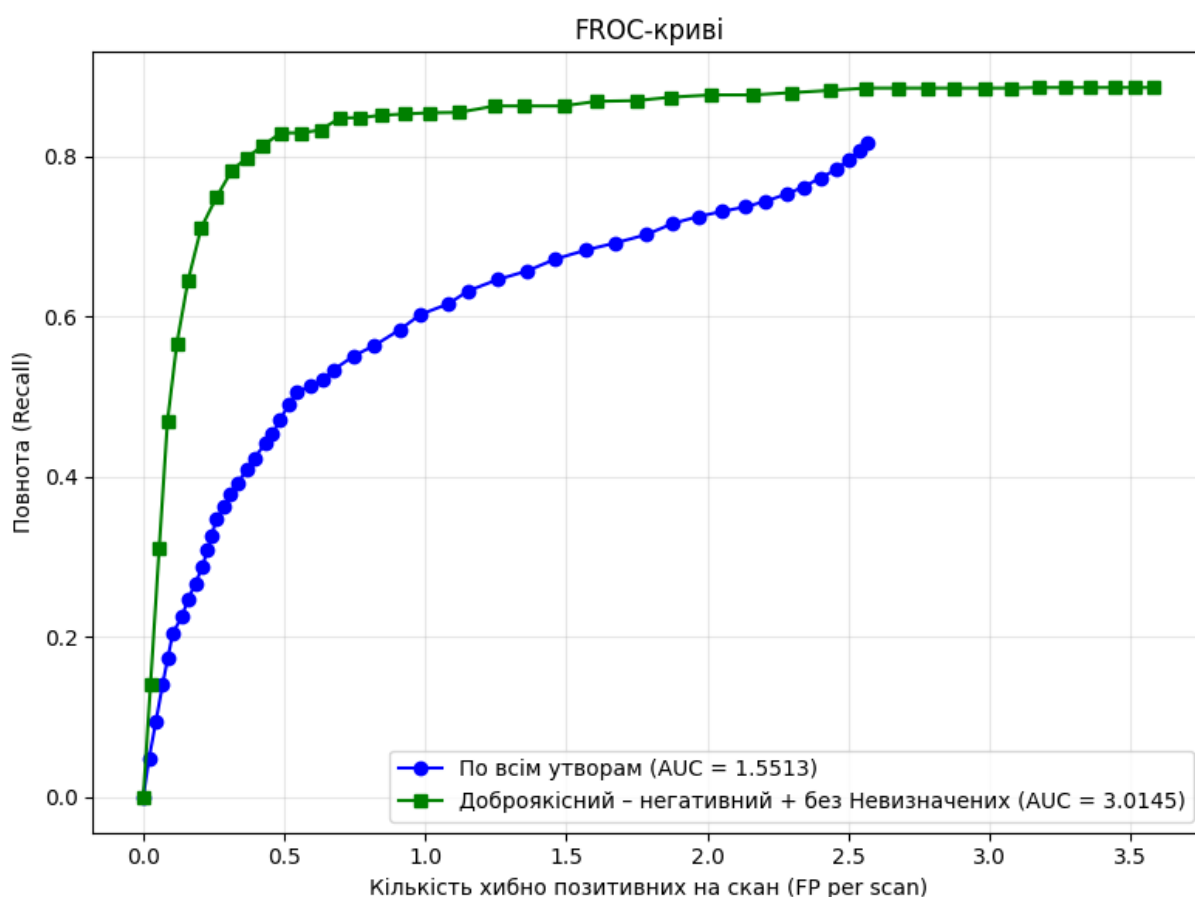


Рис. 4.14 – FROC-криві, побудовані для двох варіантів підрахунку істинних утворів

Таблиця 4.8 – Повнота та пороги для двох варіантів підрахунку істинних утворів

FP на скан	Усі утвори	Лише зложісні
Повнота		
0.125	0.219358	0.586756
0.25	0.336320	0.749710
0.5	0.484904	0.829617
1.0	0.604340	0.854813
2.0	0.729282	0.877111
Середнє	0.474841	0.779601
Поріг		
0.125	0.808105	0.844238
0.25	0.504395	0.661621
0.5	0.182373	0.315430
1.0	0.075317	0.113159
2.0	0.025131	0.044922
Середнє	0.319064	0.395874

Оцінка мультикласової конфігурації ІСППР. Для оцінки мультикласової моделі (Мультикласова Модель 1) обчислено метрики F1–міри, повноти, точності та ROC AUC для кожного класу (без класу Невизначений), а також макро–усереднені значення. Для хибно позитивних утворів встановлено істинний клас “Хибно позитивний”, для хибно негативних – прогноз “Хибно негативний” з “ймовірністю” –1 (для обчислення ROC AUC). Матриця невірних класифікацій наведена на рис. 4.15, а метрики – у таблиці 4.9.

Аналіз показує наявність виразної блочної структури в матриці невірних класифікацій: більшість помилок припадає на суміжні класи, що є ознакою коректної поведінки моделі. Наприклад, між класами “Дуже малоїмовірний” та

“Дуже підозрілий” помилки відсутні. Найбільша кількість хибних локалізацій припадає на малоймовірні класи, що логічно, оскільки ці утвори є найменшими й найскладнішими для локалізації. Таким чином, у мультикласовій конфігурації ІСППР демонструє адекватну поведінку, забезпечуючи найкращі метрики для підозрілих утворів, що мають найбільше клінічне значення.

Таблиця 4.9 – Метрики клінічної ІСППР для мультикласової конфігурації

Клас	Точність	Повнота	F1–міра	ROC AUC
Дуже малоймовірно	0.5047	0.7000	0.5865	0.7637
Помірно малоймовірно	0.2661	0.6065	0.3699	0.5287
Помірно підозріло	0.4737	0.5505	0.5092	0.7800
Дуже підозріло	0.6033	0.6453	0.6236	0.8851
Макро усереднено	0.4619	0.6256	0.5223	0.7394

4.5 Оцінка впливу інтерактивної підсистеми на результати роботи ІСППР

Налаштування моделі інтерактивної сегментації в більшості аспектів були аналогічними до моделі статичної сегментації (розділ 4.2). Основні відмінності полягають у наступному.

1. Розмір патча, вирізаного з КТ–скану, становив [96, 96, 96]. Такий вибір обумовлений тим, що інтерактивна модель отримує додатковий сильний сигнал у вигляді кліку та потребує менше контексту по аксіальному напрямку.

2. Механізм ковзного вікна був прибраний, оскільки майже всі утвори поміщаються в один патч; якщо ж утвір все–таки не поміщається, можливо додати наступний клік для досегментації. Під час тренування утвори, що не вміщалися у патч, масштабувалися до його розміру з падінгом у 16 пікселів перед масштабуванням.

3. Розмір пакету (batch size) збільшено до 32.



Рис. 4.15 – Матриця невірних класифікацій клінічної ІСППР для мультикласової конфігурації.

4. Сегментаційна нейронна мережа була зменшена на один шар в енкодері та декодері (розділ 2.10), а кількість входних каналів збільшено до чотирьох.

Для другого етапу методології оптимізації інтерактивної моделі були обрані такі гіперпараметри: стандартне відхилення для семплування кліків ($\sigma_{x/y/z}$) = 12.0; радіус еліпсоїдного відсікання ($r_{x/y/z}$) = 16.0; стандартне відхилення гаусової маски кліку ($\sigma_{x/y/z}$) = 3.0; радіус відсікання гаусової маски ($r_{x/y/z}$) = 8.0. Для утворів типу Borderline підібрано окремий набір гіперпараметрів, оскільки M_{FN} та M_{FP} зазвичай значно менші за повноцінні маски: стандартне відхилення = 1.0; радіус еліпсоїдного відсікання = 2.0.

Загальна кількість утворів у тренувальних підвибірках перехресної перевірки варіювалася від 5158 до 5754. Кожна модель тренувалася 200 епох, відповідно кількість кроків оптимізації становила від 32 200 до 35 800. Процес оптимізації, як і для статичних моделей, був стабільним – рис. 4.16.

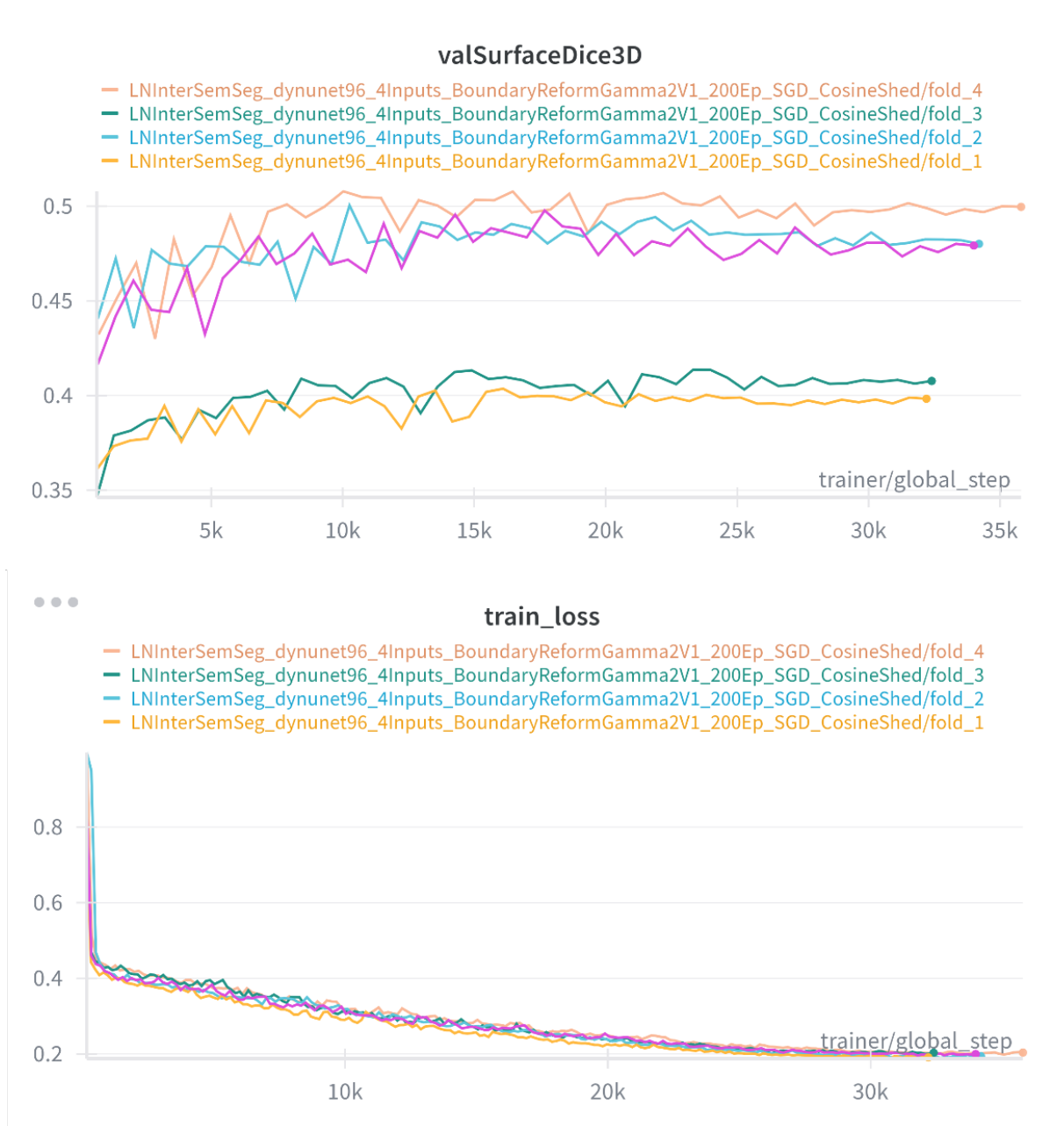


Рис. 4.16 – Графіки метрики Surface Dice (угорі) на валідаційних підвибірках та значення функції втрат Boundary Dou++ (унизу) на тренувальних підвибірках для всіх моделей перехресної перевірки під час їх оптимізації для моделі інтерактивної сегментації.

Для обчислення метрик інтерактивної моделі необхідно розробити алгоритм автоматичного розміщення інтерактивних кліків. Такий підхід демонструє “оптимістичний” сценарій роботи системи, оскільки припускає, що радіолог може ідеально визначити наявність/відсутність утвору та центроїду його локалізації. Повноцінне клінічне випробування з участю групи радіологів потребувало б значних ресурсів та часу, однак планується у майбутньому.

Для утворів із датасету істинних та прогнозованих утворів (розділ 2.10), які мали класи Negative та False Negative, було обчислено центроїд, що використовувався як координата кліку; тип кліку визначався за класом: позитивний для False Negative та негативний для Negative. Таким чином інтерактивна підсистема дозволяла коригувати хибно позитивні та хибно негативні локалізації. Результати моделі наведено в таблиці 4.10. Інтерактивна сегментація суттєво покращила якість: приріст IoU +0.1207; Surface Dice +0.1245; Dice +0.1202.

Таблиця 4.10 – Метрики статичної та інтерактивної сегментаційних моделей

Експеримент	IoU	Surface Dice	Dice	Best Threshold
Статична	0.513877	0.553212	0.634158	0.6
Статична → Інтерактивна	0.634610	0.677715	0.754336	0.6

Далі обчислимо метрики з розділу 4.4 після застосування інтерактивної підсистеми. Метрики локалізації наведено в таблиці 4.11, статистику TP/FN – у таблиці 4.12. Результати суттєво покращилися: для випадку врахування Borderline у TP приріст точності становив +0.4381; повноти +0.1409; F1–міри +0.3286. Аналогічне зростання спостерігається і для варіанта Borderline у FN та FP. Розподіл похибок між класами загалом повторює тенденції статичної моделі, за винятком того, що для Дуже малоїмовірного класу спостерігається найбільший приріст.

Таблиця 4.11 – Метрики локалізації після застосування інтерактивної підсистеми

Випадок	Точність (Precision)	Повнота (Recall)	F1–міра	F2–міра	F0.5–міра
Borderline у TP	0.9182	0.9374	0.9277	0.9335	0.9220
Borderline у FN та FP	0.8612	0.8792	0.8701	0.8756	0.8648

Таблиця 4.12 – Статистика TP/FN локалізації та відсоткові значення для кожного класу після застосування інтерактивної підсистеми

Клас	TP	FN	TP%	FN%
Дуже підозрілий	166	5	0.970760	0.029240
Дуже малоюмовірний	294	15	0.951456	0.048544
Невизначений	1089	72	0.937984	0.062016
Помірно підозрілий	308	19	0.941896	0.058104
Помірно малоюмовірний	479	45	0.914122	0.085878

Розподіл об’ємів рамок утворів після інтерактивної підсистеми наведено на рис. 4.17. Медіанні значення становлять: 214.638 мм³ для TP; 82.748 мм³ для FN; 44.131 мм³ для FP. Хоча медіанні значення залишилися близькими до статичної моделі, кількість TP суттєво зросла, а кількість FN та FP – помітно зменшилася.

FROC–криві після роботи інтерактивної підсистеми наведено на рис. 4.18, а деталізовані метрики – у таблиці 4.13. Порівнюючи зі статичною системою, додавання інтерактивної складової дозволяє досягти повноти 86% при лише 0.25 хибного спрацювання на скан (тобто одне хибне спрацювання на чотири дослідження). Для злоякісних утворів повнота сягає 94.5% при одному FP на дослідження. Це демонструє високу ефективність інтерактивної підсистеми.

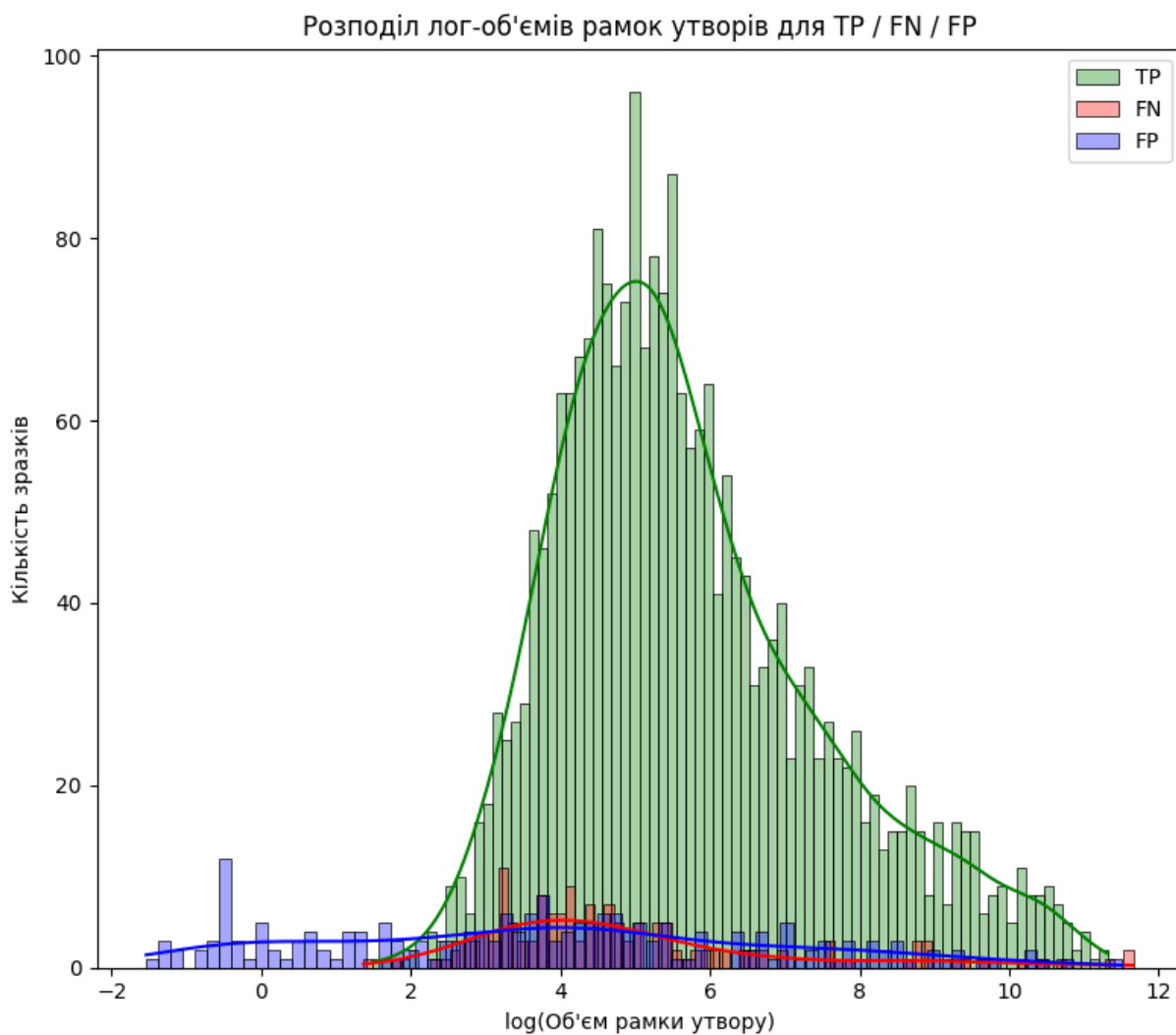


Рис. 4.17 – Розподіл об'ємів рамок утворів для істинно позитивних (TP), хибно негативних (FN) та хибно позитивних (FP) випадків локалізації утворів після застосування інтерактивної підсистеми.

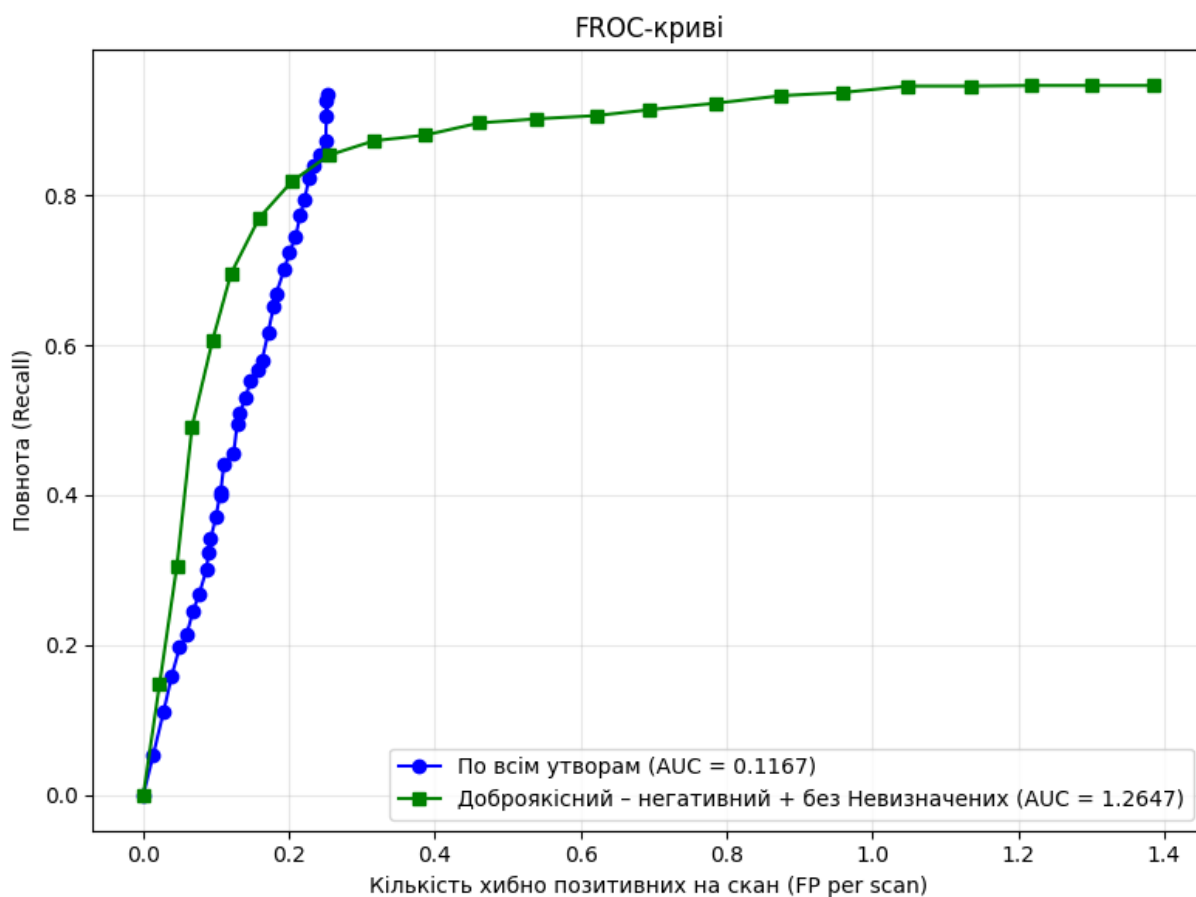


Рис. 4.18 – FROC-криві, побудовані для двох варіантів підрахунку істинних утворів після застосування підсистеми інтерактивної сегментації.

Таблиця 4.13 – Повнота та пороги для двох варіантів підрахунку істинних утворів після застосування інтерактивної підсистеми

FP на скан	Усі утвори	Лише зложкісні
Повнота		
0.125	0.456556	0.703061
0.25	0.864370	0.850310
0.5	—	0.900189
1.0	—	0.945875
2.0	—	—
Середнє	0.660463	0.849859
Поріг		
0.125	0.278564	0.772997
0.25	0.013329	0.440430
0.5	—	0.135254
1.0	—	0.030045
2.0	—	—
Середнє	0.145947	0.346256

Далі розглянемо вплив інтерактивної підсистеми для мультикласової конфігурації ІСППР. Метрики для кожного класу наведено в таблиці 4.14, а матрицю невірних класифікацій – на рис. 4.19. Метрики класифікації значно покращилися: F1–міра +0.1672; Точність +0.2331; Повнота +0.0641; ROC AUC +0.1105. Характер тенденцій також змінився: найкраще стали класифікуватися класи Дуже малоїмовірно та Помірно малоїмовірно, адже значна кількість утворів цих класів була правильно локалізована інтерактивною підсистемою. Водночас кількість хибно позитивних і хибно негативних утворів для підозрілих класів також суттєво зменшилася.

Таблиця 4.14 – Метрики клінічної ІСППР для мультикласової конфігурації після застосування інтерактивної підсистеми

Клас	Точність	Повнота	F1–міра	ROC AUC
Дуже малоймовірно	0.9020	0.7443	0.8156	0.8789
Помірно малоймовірно	0.6891	0.7615	0.7235	0.7949
Помірно підозріло	0.5778	0.5902	0.5840	0.8103
Дуже підозріло	0.6108	0.6608	0.6348	0.9155
Макро усереднено	0.6950	0.6897	0.6895	0.8499



Рис. 4.19 – Матриця невірних класифікацій клінічної ІСППР для мультикласової конфігурації після застосування підсистеми інтерактивної сегментації.

Підсумовуючи результати, можна зробити висновок, що інтерактивна підсистема суттєво покращує всі метрики роботи клінічної ІСППР. Очевидно, що отримані значення є певною верхньою межею ефективності, адже радіологи у реальних умовах також можуть припускатися помилок при взаємодії з інтерфейсом інтерактивної сегментації. Проте навіть за такого сценарію переваги інтерактивного механізму є беззаперечними.

4.6 Результати оцінювання клінічної ІСППР для виявлення раку нирок

Для задачі виявлення раку нирок була зібрана велика вибірка контрастованих КТ–сканів, що містить 1746 досліджень. Вибірка агрегована з двох публічних датасетів: TCGA–KIRC [96] та C4KC KiTS [97], а також з чотирьох клінік. Усі КТ–скани були анотовані на класи, визначені в розділі 3.1, згідно з чотирьохетапним процесом.

1. Створення пресегментацій за допомогою сегментаційної моделі на базі nnUNet.
2. Анотація радіологами на 5 категорій Bosniak та 2 класи солідних утворень відповідно до таксономії з розділу 3.1.
3. Перевірка та корекція результатів більш досвідченим радіологом.
4. Автоматична перевірка на наявність надто малих утворів та подальша корекція радіологами.

Після формування семантичних масок класи були агреговані відповідно до опису в розділі 3.1. Увесь процес проводився в компанії Better Medicine [98] за участю професійних радіологів з різних лікарень. Статистика щодо кількості доброякісних і злоякісних утворів та відповідних КТ–сканів наведена в таблиці 4.15. Із цього розподілу видно, що:

- вибірка є досить великою та містить значну кількість утворів;

– специфікою діагностики раку нирок є часта наявність кіст (доброякісних утворів), що ускладнює задачу не лише локалізації, а й диференціації злоякісних і доброякісних утворів.

Таблиця 4.15 – Кількість ниркових утворів на рівні сканів та окремих утворів у тренувальній вибірці

Рівень	Злоякісні	Доброякісні
КТ–скани	679	1273
Утвори	808	6364

Приклади КТ–зрізів з різних джерел тренувальної вибірки наведено на рис. 4.20.

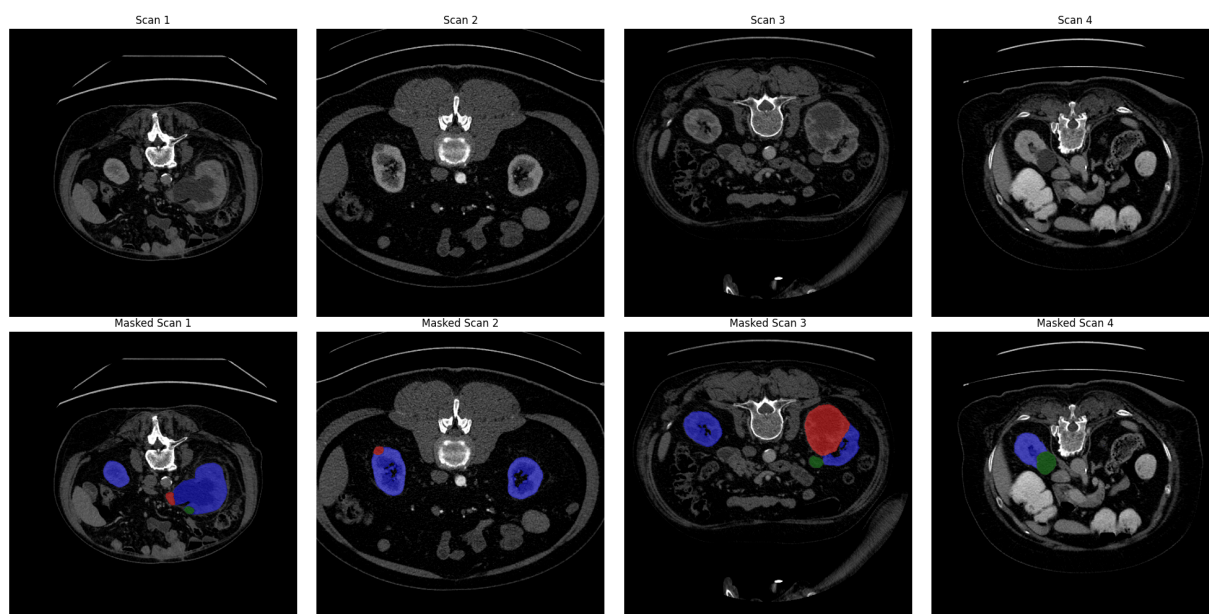


Рис. 4.20 – Приклади зрізів КТ–сканів з різних джерел ниркової вибірки у першому ряду; у другому ряду – візуалізація масок нирок (синій), злоякісних утворів (червоний) та доброякісних утворів (зелений).

Усі підсистеми сегментації – підсистема сегментації області нирок, а також перша та друга підсистеми сегментації утворень – побудовані аналогічно підсистемі сегментації утворень у легенях (розділи 2.6, 4.2). Відмінності стосувалися

лише приведення сканів до іншого фізичного розміру та параметрів патча згідно зі статистикою вибірки.

Для оцінювання системи використовувалися такі метрики.

1. Global Dice та Global IoU (4.4).
2. Average Dice та Average IoU (4.5).
3. Повнота (Sensitivity) та Точність (Precision) на рівні об'єктів – на основі IoU між істинними та прогнозованими об'єктами. Об'єкти з $\text{IoU} > 0.5$ вважалися істинно позитивними.
4. Повнота та специфічність на рівні КТ–сканів (еквівалентно рівню пацієнтів). Якщо модель знаходила хоча б один утвір відповідного класу з $\text{IoU} > 0.5$, скан вважався істинно позитивними.
5. Середня абсолютна похибка у відсотках (MAPE) по найдовшому діаметру (мм) та об'єму (мм^3) – лише для істинно позитивних утворів.

$$\begin{aligned}\text{Global Dice} &= \frac{2 |\text{Prediction} \cap \text{GroundTruth}|}{|\text{Prediction}| + |\text{GroundTruth}|}, \\ \text{Global IoU} &= \frac{|\text{Prediction} \cap \text{GroundTruth}|}{|\text{Prediction} \cup \text{GroundTruth}|}.\end{aligned}\quad (4.4)$$

де Prediction – усі вокселі, які модель віднесла до поточного класу у всій вибірці;

GroundTruth – усі вокселі, що насправді належать цьому класу у всій вибірці.

$$\begin{aligned}\text{Average Dice} &= \frac{1}{N} \sum_{i=1}^N \frac{2 |\text{Prediction}_i \cap \text{GroundTruth}_i|}{|\text{Prediction}_i| + |\text{GroundTruth}_i|}, \\ \text{Average IoU} &= \frac{1}{N} \sum_{i=1}^N \frac{|\text{Prediction}_i \cap \text{GroundTruth}_i|}{|\text{Prediction}_i \cup \text{GroundTruth}_i|}.\end{aligned}\quad (4.5)$$

де N – кількість КТ–сканів;

Prediction_i – усі вокселі, які модель віднесла до даного класу для i -го скану;

GroundTruth_i – усі вокселі, що насправді належать даному класу для i -го скану.

Тестова вибірка була отримана з однієї клініки та містила наступний розподіл утворів (табл. 4.16):

Таблиця 4.16 – Кількість ниркових утворів у тестовій вибірці

Рівень	Злоякісні	Доброякісні
КТ–скани	100	133
Утвори	106	518

Розроблена система була порівняна з попередньою версією, розробленою компанією Better Medicine [99]. Метрики наведено в таблиці 4.17. З отриманих результатів видно, що обидві системи демонструють високі показники, проте нова ІСППР значно покращує результати саме для злоякісних утворів, що є пріоритетом у клінічних сценаріях. Водночас система стала дещо гіршою у випадку доброякісних утворів, що є менш критично.

Окремо варто зазначити дуже високу специфічність на рівні КТ–сканів, що важливо для сценаріїв автоматичного перегляду великих масивів даних, коли виявлення ниркових утворів може не бути основною метою дослідження. У таких умовах велика кількість хибнопозитивних спрацювань є небажаною.

Таблиця 4.17 – Порівняння метрик для злоякісних та доброякісних утворів між попередньою та поточною версіями системи

Рівень деталізації	Метрика	Злоякісні (попер.)	Злоякісні (поточ.)	Доброякісні (попер.)	Доброякісні (поточ.)
Вокселі	Global Dice	0.93	0.94	0.88	0.86
	Global IoU	0.87	0.89	0.78	0.75
	Avg. Dice	0.91	0.91	0.80	0.76
	Avg. IoU	0.87	0.87	0.73	0.68
Об'єкти	Sensitivity	0.93	0.92	0.85	0.77
	Precision	0.78	0.95	0.86	0.84
	MAPE diam.	–	3.6115	–	8.1711
	MAPE vol.	6.75	7.3733	10.322	16.7922
Скани	Sensitivity	0.96	0.96	0.95	0.90
	Specificity	0.95	0.99	0.90	0.97

Наразі розроблена клінічна ІСППР проходить процес клінічних випробувань перед інтеграцією в медичні центри. Акт впровадження подано в додатках до роботи. Приклади прогнозів наведено на рис. 4.21.

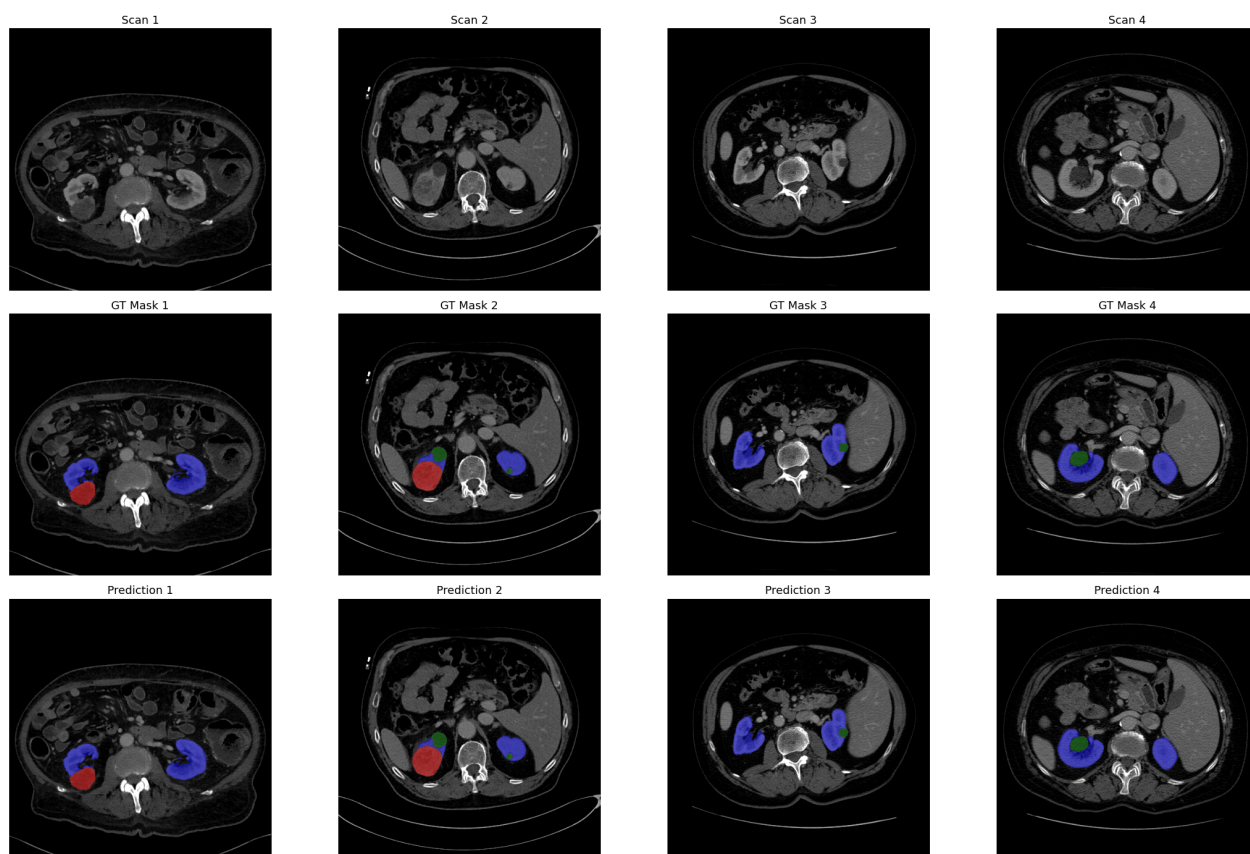


Рис. 4.21 – Приклади прогнозів розробленої клінічної ІСППР для діагностики раку нирок на тестовій вибірці. Перший ряд – зрізи КТ-сканів; другий – істинні сегментації: нирок (синій), злоякісних утворів (червоний), доброякісних утворів (зелений); третій – прогнози системи.

4.7 Результати сегментації функціональних тканинних одиниць (FTU)

Система моделей і методів, описана в розділі 3.2, була розроблена для сегментації функціональних тканин за умов відсутності даних одного з доменів, який характеризується іншим протоколом забарвлення тканин, а також відмінною фізичною величиною пікселя. На рис. 3.3 зображення зліва належить тренувальному домену, а справа – тестовому.

У контексті змагання HuBMAP + HPA – Hacking the Human Body [51], тренувальний домен відповідає датасету HPA [100], [101], [102], [103], [104],

[105], [106], [107], [108], а тестовий – NuBMAP [109]. Відповідно результати розроблених моделей і ансамблів оцінюються за двома критеріями:

- загальна якість сегментації на обох доменах, оцінена метрикою Dice;
- якість доменної адаптації, а саме різниця між метрикою Dice на доменах НРА та HubMAP.

Для оцінки на домені НРА застосовувалася перехресна перевірка, аналогічна до методики в розділі 4.2. Для оцінки на домені HubMAP було використано частину публічної тестової вибірки та всю приватну тестову вибірку зі змагання. Різницю метрик двох доменів визначали як різницю між перехресною оцінкою на НРА та приватною оцінкою на NuBMAP. У таблиці 4.18 наведено результати абляційного дослідження методів, запропонованих у розділі 3.2.

Таблиця 4.18 – Абляційне дослідження методів сегментації функціональних тканинних одиниць (FTU)

Конфігурація	НРА Dice	Публічний Dice	Приватний Dice	Різниця Dice
Базова модель	0.71172	0.12115	0.15632	0.55540
+ Фіксація розміру пікселя	0.71588	0.22653	0.30195	0.41393
+ Гістограмне вирівнювання	0.70283	0.38732	0.48486	0.21797
+ CutMix	0.75157	0.49483	0.64518	0.10639
+ Інші аугментації	0.76950	0.51506	0.71187	0.05763

Таблиця 4.18 – (продовження)

Конфігурація	НРА Dice	Публічний Dice	Приватний Dice	Різниця Dice
+ Додаткова адаптація розміру пікселя + напівкеро- ване навчання на GTEX [110]	0.82142	0.58037	0.78352	0.03790
+ Напівкероване навчання на додаткових НРА	0.83633	0.59800	0.81402	0.02231
Найкраща модель по пере- хресній перевірці	0.85405	0.60826	0.83332	0.02073
Ансамбль	0.85428	0.60931	0.83419	0.02009

Аналізуючи результати абляційного дослідження, можна зробити такі висновки.

1. Фіксація розміру пікселя та гістограмне вирівнювання незначно знижують метрику тренувального домену, проте суттєво підвищують метрики тестового домену, що робить їх ефективними методами доменної адаптації.
2. CutMix та додаткові аугментації значно покращують якість сегментації на обох доменах.
3. Доменна адаптація через напівкероване навчання також має суттєвий позитивний ефект, особливо на тестовому домені.
4. Вибір найкращої моделі за перехресною перевіркою та ансамблювання додатково підвищують робастність системи та покращують фінальні метрики.

Загалом запропоновані методи покращили якість на тренувальному домені на 0.14256 Dice, а на тестовому – на 0.67787, зменшивши різницю між доменами з 0.55540 до 0.02009. Це фактично означає досягнення однакової точності сегментації у двох різних доменах.

Таблиця 4.19 – Порівняння розробленої системи з іншими рішеннями

Найкращі результати	Приватний Dice
Перше місце	0.83562
Друге місце	0.83393
Розроблена система	0.83266
Четверте місце	0.82717
П'яте місце	0.82595

Порівнюючи розроблену систему з іншими рішеннями, представленими на змаганні (табл. 4.19), бачимо, що запропонований підхід маржинально поступається першим двом місцям, але суттєво перевершує наступні рішення. Це підтверджує ефективність розробленої методики доменної адаптації та сегментації функціональних тканинних одиниць.

4.8 Результати оцінки системи для виявлення раку шкіри

У розділі 3.3 була представлена багаторівнева мультимодальна система для виявлення (детекції) раку шкіри. Для оцінки розробленої системи були використані такі метрики.

1. ROC AUC.
2. Частковий ROC AUC (Partial ROC AUC) [13], який обчислюється на частині ROC-кривої, де рівень істинно позитивних спрацювань (True Positive Rate) перевищує 80%. Таким чином аналізується область, у якій системи демонструють високу повноту. Значення цієї метрики лежать у діапазоні $[0.0, 0.2]$.
3. Повнота за топ-15 обраними прикладами (Top-15 Retrieval Sensitivity) [111], яка характеризує кількість позитивних утворів, знайдених серед 15 зображень з найвищими оцінками системи. Далі результат нормується на загальну кількість істинних злоякісних утворів для пацієнта, після чого усереднюється за всіма пацієнтами зі злоякісними утворами. Таким чином метрика відображає середню кількість знайдених злоякісних утворів, зважену за кількістю утворів у пацієнта.

Зазначені метрики були обчислені для вибірок перехресного оцінювання шляхом об'єднання прогнозів (OOF) та усереднення метрик (Mean), а також для публічної (Public) і приватної (Private) тестових вибірок змагання ISIC 2024 – Skin Cancer Detection with 3D–TBP [53]. Метрики Partial ROC AUC для базових моделей і дворівневих систем наведені в таблиця 4.20, метрики ROC AUC – у таблиця 4.21, а значення Top–15 Retrieval Sensitivity – у таблиця 4.22.

Таблиця 4.20 – Partial ROC AUC для базових моделей й дворівневих систем

Модель / Система	OOF	Mean	Private	Public
Дворівнева v2	0.17666	0.17862	0.16941	0.18608
Гібридний ConvNeXt	0.17497	0.17698	0.16090	0.17714
XGBoost v2	0.17348	0.17460	–	–
XGBoost v1	0.17252	0.17351	–	–
EarlyStop LightGBM v4	0.17225	0.17305	–	–
LightGBM v3	0.17173	0.17266	0.16107	0.18400
LightGBM v4	0.17105	0.17210	–	–
EarlyStop LightGBM v1	0.17029	0.17183	–	–
EarlyStop LightGBM v2	0.17024	0.16187	0.16187	0.18336
Дворівнева v1	0.17010	0.17116	0.16173	0.18283
LightGBM v1	0.17005	0.17165	–	–
Гібридний EdgeNeXt	0.15892	0.17410	0.16082	0.17481

Таблиця 4.21 – ROC AUC для базових моделей і дворівневих систем

Модель	OOF	Mean
Дворівнева v2	0.97234	0.97395
Гібридний ConvNeXt	0.97082	0.97244
XGBoost v2	0.96826	0.96916
XGBoost v1	0.96741	0.96817
EarlyStop LightGBM v4	0.96719	0.96765
LightGBM v3	0.96599	0.96682
LightGBM v4	0.96586	0.96666
Дворівнева v1	0.96518	0.96600
EarlyStop LightGBM v2	0.96502	0.96605
EarlyStop LightGBM v1	0.96491	0.96629
LightGBM v1	0.96485	0.96632
Гібридний EdgeNeXt	0.95110	0.96701

Таблиця 4.22 – Top–15 Retrieval Sensitivity для базових моделей і дворівневих систем

Модель	OOF	Mean
Гібридний ConvNeXt	0.76081	0.75995
Дворівнева v2	0.74809	0.75375
XGBoost v2	0.73791	0.73919
LightGBM v1	0.73537	0.73769
EarlyStop LightGBM v2	0.72774	0.73095
LightGBM v4	0.72774	0.72939
XGBoost v1	0.72519	0.72621
Дворівнева v1	0.72265	0.72304
EarlyStop LightGBM v4	0.72265	0.72658
EarlyStop LightGBM v1	0.72010	0.72296
LightGBM v3	0.71501	0.71614
Гібридний EdgeNeXt	0.71247	0.71542

Різні конфігурації дворівневих систем проілюстровані на схемі 3.9: v1 відповідає лівій частині схеми (1st Approach); v2 – центральній частині (2nd Approach); права частина (3rd Approach) ілюструє підхід до тренування гібридної моделі. Запропонована дворівнева система v2 демонструє найкращі результати на всіх наборах даних за метриками Partial ROC AUC та ROC AUC. Водночас гібридна модель ConvNeXt досягає найвищих значень метрики Top–15 Retrieval Sensitivity.

У таблиці 4.23 наведені результати оцінювання дворівневих і трирівневих систем. Отримані результати показують, що трирівневі системи перевершують дворівневі та базові моделі майже на всіх вибірках за метриками Partial ROC AUC та Top–15 Retrieval Sensitivity. Позначення “mc” означає, що коефіцієнти агрегації були підібрані вручну, а не за допомогою алгоритму Optuna.

Таблиця 4.23 – Метрики дворівневих і трирівневих систем

Система	OOF Partial ROC AUC	Private Partial ROC AUC	Public Partial ROC AUC	OOF Top–15 Retrieval Sensitivity
Дворівнева v2	0.17666	0.16941	0.18608	0.74809
Дворівнева v1	0.17010	0.16173	0.18283	0.72265
Трирівнева v2	0.18068	0.17042	0.18528	0.78371
Трирівнева v1	0.18014	0.16982	0.18449	0.77608
Трирівнева v1 mc	0.17939	0.17039	0.18527	0.78117

Порівняння розробленого рішення з найкращими результатами змагання (таблиця 4.24) показує, що запропонована система лише маржинально поступається лідерам. Аналіз сучасного стану галузі (таблиця 4.25) свідчить, що запропоновані моделі й системи демонструють значно кращі результати порівняно з попередніми дослідженнями, навіть з урахуванням того, що попередні аналоги оцінювалися на мікроскопічних зображеннях, які є більш сприятливими для задачі детекції. Слід зазначити, що трирівнева система не включена до таблиці 4.25, оскільки алгоритм Optuna оптимізує параметри на всій тренувальній вибірці, що призвело б до надто оптимістичної оцінки на тій самій вибірці.

Таблиця 4.24 – Порівняння з найкращими результатами змагання ISIC 2024

Рішення	Private	Public
1–ше місце (Private)	0.17264	0.18611
1–ше місце (Public)	0.17051	0.18800
Трирівнева v1 mc	0.17039	0.18527
Трирівнева v2	0.17042	0.18528

Таблиця 4.25 – Порівняння з іншими системами виявлення раку шкіри

Система	Вибірка	ROC AUC
2020 Best Solution [112]	ISIC 2020 Competition [113]	0.9490
Saranya N. et al. [114]	PH2 [115]	0.87
Saranya N. et al. [114]	Derm7pt [116]	0.76
Jojoa Acosta et al. [117]	ISIC 2017 Challenge [118]	0.91
Запропонована (Гібридний ConvNeXt)	ISIC 2024 Kaggle Challenge	0.97244
Запропонована (XGBoost v2)	ISIC 2024 Kaggle Challenge	0.96916
Запропонована (Дворівнева система v2)	ISIC 2024 Kaggle Challenge	0.97395

Отже, розроблена мультимодальна дворівнева та трирівнева система демонструє високі показники якості детекції, досягаючи значень ROC AUC 0.97395, Partial ROC AUC 0.18528 та Top-15 Retrieval Sensitivity 0.78371. З урахуванням того, що система працює зі складним типом даних – фотографіями ділянок шкіри, отримані результати є дуже високими. Крім того, порівняно з попередніми дослідженнями, запропоноване рішення перевершує їх.

4.9 Висновки розділу 4

У розділі було описано вибірку, на якій оптимізувалися моделі глибокого навчання, на основі яких побудовано клінічну ІСППР для виявлення раку легенів, а також наведено результати оцінювання окремих підсистем і системи загалом. Запропонована система продемонструвала досить високу точність, зокрема повноту на рівні 0.877 за середнього значення двох хибно позитивних спрацювань на один КТ-скан. З огляду на складність задачі, а саме велику кількість малих утворів і потенційну зашумленість анотацій, можна вважати, що розроблена ІСППР успішно впоралася з поставленою задачею. Водночас інтеграція інтерактивної

складової навіть за мінімальної взаємодії з боку радіолога дозволяє досягти повноти 0.945 лише за одного хибно позитивного утвору на КТ–скан.

У мультикласовій конфігурації система без інтерактивної складової досягає значення F1–міри 0.623 для найбільш підозрілих утворів, тоді як із використанням інтерактивної підсистеми це значення зростає до 0.634. Водночас важливо зауважити, що згідно з матрицею невірних класифікацій похибки мультикласової системи мають «блочну структуру», тобто більшість помилок виникає під час розрізнення суміжних ступенів підозрілості. Це призводить до того, що майже всі підозрілі утвори потрапляють до підозрілих класів, хоча інколи й нижчого ступеня підозрілості.

Оцінюючи клінічну ІСППР для виявлення раку нирок, побудовану на методологічних засадах попередньої системи та таку, що перевикористовує низку її підсистем, було використано вибірку, отриману з клініки. Крім того, моделі глибокого навчання, що входять до складу ІСППР, оптимізувалися як на даних з відкритих джерел, так і на даних із кількох клінік. Розроблена система продемонструвала високу точність: зокрема, повнота на рівні пацієнта досягла значення 0.96, а специфічність – 0.99, що свідчить про перевагу системи над попередньою версією за низкою ключових метрик. Наразі клінічна ІСППР проходить етап клінічних випробувань і в разі їх успішного завершення буде інтегрована в низку лікарень і медичних центрів.

Окрім цього, було розроблено методи доменної адаптації для суміжної медичної задачі – сегментації функціональних тканинних одиниць, а також побудовано систему моделей для її розв’язання. Запропонована система продемонструвала високу точність (Dice на рівні 0.85) та дозволила усунути розрив у метриках між різними протоколами забарвлення мікроскопічних зображень. Також була розроблена мультимодальна багаторівнева система виявлення раку шкіри, яка поєднує глибокі згорткові мережі, гібридні архітектури та алгоритми дерев рішень для детекції меланоми. Запропонована система досягла високих показників якості виявлення раку шкіри (ROC AUC на рівні 0.97) на фотографіях

ділянок шкіри, що перевищує метрики низки попередніх систем, побудованих на мікроскопічних зображеннях.

ВИСНОВКИ

Дисертаційним дослідженням розв'язано задачу побудови Клінічної Інтелектуальної Системи Підтримки Прийняття Рішень для діагностики раку легенів, яка адаптована до поточних потреб радіологів. Дана задача була розв'язана шляхом інтеграції низки моделей глибокого навчання в цілісну інформаційну систему, а також розробки методів для кращої адаптації моделей до різних протоколів досліджень і похибок попередніх моделей у системі, разом з інтерактивною складовою для зручної та ефективної корекції можливих помилок системи. Водночас загальна методологія побудови ІСППР була перевикористана для розробки іншої клінічної ІСППР для виявлення раку нирок, яка продемонструвала високу ефективність і наразі проходить клінічні випробування в Європейській лікарні. Також було запропоновано низку методів доменної адаптації в задачі сегментації мікроскопічних зображень, що ґрунтуються на аугментаціях у просторі кольорів та методах напівкерованого навчання, а також систему виявлення раку шкіри, яка використовує гібридні нейронні мережі та багаторівневий каскад моделей для обробки даних різної модальності.

У ході цього дисертаційного дослідження було побудовано клінічні ІСППР для діагностики раку легенів та нирок, які відповідають вимогам прикладного застосування в медичній практиці. Розроблені методологічні засади та підсистеми можуть бути перевикористані й адаптовані для діагностики онкологічних захворювань інших органів. Розроблені системи та методи підвищили точність виявлення і класифікації утворень у різних умовах досліджень і за наявності похибок попередніх моделей.

За результатами дисертаційної роботи виконано наступні задачі.

1. Розроблено клінічну ІСППР для діагностики раку легенів, яка оброблює КТ-скани в DICOM-форматі та надає радіологу (експерту) локалізацію кожного з підозрілих утворів, їхні радіологічні ознаки, а також ступінь ризику для пацієнта

і впевненість системи в діагнозі. Розроблена система продемонструвала повноту на рівні 87% при середній кількості хибнопозитивних спрацювань 2 утвори на один КТ-скан.

2. У розроблену клінічну ІСППР інтегровано інтерактивну підсистему на базі глибоких нейронних мереж, що дозволило підвищити точність системи до 94% повноти при середній кількості хибнопозитивних спрацювань 1 утвір на один КТ-скан.

3. Розроблено специфічні методи для підвищення точності системи в задачах сегментації та класифікації утворів, зокрема запропоновано функцію втрат Boundary DoU++, аугментації зсуву для тренування класифікаційної моделі, багатоканальний вхід для класифікаційної мережі та аугментацію маскуванню допоміжного каналу. Запропоновані методи підвищили точність системи та її стійкість до різноманітних вхідних даних. Також ці методи були перевикористані в системі виявлення раку нирок, де продемонстрували свою ефективність.

4. Розроблено методику оцінки ІСППР для виявлення раку легенів, яка включає три основні етапи: (1) оцінка кожної підсистеми окремо, (2) оцінка всієї ІСППР без інтерактивної складової, (3) оцінка всієї ІСППР з інтерактивною складовою. Було використано низку сегментаційних метрик, зокрема IoU, Surface Dice, Dice, а також класифікаційні метрики ROC AUC, повнота, точність, F1/F2/F0.5-міри та спеціалізовану метрику для оцінки клінічних систем — FROC-криву. Метрики було досліджено залежно від класу (небезпечності) утворів і їхніх розмірів.

5. Розроблено ІСППР для виявлення раку нирок на базі методології та підсистем попередньої ІСППР, яка досягла значень чутливості (на рівні сканів) 0.96 і специфічності 0.99, що є достатнім для початку клінічних випробувань.

6. Розроблено метод доменної адаптації для сегментації функціональних тканинних одиниць, за допомогою якого досягнуто значення метрики Dice 0.832 при зменшенні розриву між доменами до 0.02.

7. Розроблено багаторівневу мультимодальну систему виявлення раку шкіри, яка досягла значення метрики ROC AUC 0.973.

Наукова новизна даної роботи полягає в наступному.

1. Вперше розроблено методологію побудови клінічної інтелектуальної системи підтримки прийняття рішень на базі набору моделей глибокого навчання, що працюють із медичними зображеннями та метаданими основного клінічного протоколу обміну інформацією DICOM [11] і надають на виході виявлені підозрілі утворення, їх радіометричні ознаки та оцінку ступеня ризику. На відміну від попередніх рішень, які реалізують лише окремі задачі (детекція, сегментація, класифікація), розроблена ІСППР охоплює повний процес діагностики та може бути безпосередньо використана в медичних закладах.

2. Вперше розроблено метод тренування згорткової нейронної мережі для класифікації утворень, який, на відміну від попередніх підходів, передбачає використання двох каналів із додатковим контекстом навколо утвору та з чітко виділеною рамкою утвору, а також застосування спеціалізованих аугментацій. Це дозволило підвищити точність моделі та її стійкість до похибок локалізації утвору.

3. Вперше розроблено комплекс методів доменної адаптації для сегментації функціональних тканинних одиниць, який, на відміну від попередніх робіт, поєднує такі підходи: гістограмне вирівнювання між різними протоколами забарвлення; спеціалізований набір аугментацій, зокрема CutMix між доменами; напівкероване навчання на цільовому домені. Це дало змогу суттєво підвищити точність сегментації на цільовому домені.

4. Вперше розроблено комплекс методів для задачі класифікації раку шкіри, який, на відміну від підходів, заснованих на одиничних моделях або однорівневих ансамблях, передбачає: двоетапну оптимізацію гібридної нейронної мережі, що оперує зображеннями утворів і метаданими утвору та пацієнта; багаторівневу систему, яка поєднує моделі класичного та глибокого машинного навчання. Використання запропонованих методів дозволило суттєво підвищити точність виявлення раку шкіри на фотографічних знімках.

5. Дістала подальшого розвитку інтерактивна модель сегментації, яка, на відміну від існуючих рішень, оптимізується не лише за інтерактивним сигналом, а й з урахуванням результатів роботи статичної моделі сегментації, що дозволяє підвищити її точність.

6. Дістала подальшого розвитку сегментаційна функція втрат Boundary DOU шляхом її модифікації до Boundary DOU++, у якій, на відміну від базової версії, введено додатковий параметр γ , що забезпечує підвищення точності сегментації малих об'єктів, зокрема утворів у легенях.

Також важливо підкреслити практичну значимість результатів дослідження, яка полягає в наступному.

1. Запропоновано та реалізовано методологію побудови клінічних інтелектуальних систем підтримки прийняття рішень, на основі якої створено ІСППР для діагностики раку легень та ІСППР для виявлення раку нирок.

2. Розроблена ІСППР для діагностики раку легень має високий потенціал практичного застосування, оскільки спроектована з урахуванням сучасних діагностичних практик і чинних клінічних протоколів.

3. Запропонована методологія побудови ІСППР є універсальною та може бути перевикористана для створення інших клінічних систем підтримки прийняття рішень.

4. Розроблені технічні покращення, зокрема нова сегментаційна функція втрат, методи формування двоканального вхідного представлення та підходи до аугментації даних для класифікаційних моделей, можуть бути застосовані в інших задачах обробки медичних зображень.

5. Запропоновані методи доменної адаптації можуть бути використані та розширені для адаптації сегментаційних моделей до кількох доменів (типів забарвлення) мікроскопічних знімків, що підвищує їх робастність.

6. Система виявлення раку шкіри може бути інтегрована в повноцінну інформаційну систему та використана для проведення скринінгових досліджень, зокрема в країнах з низьким рівнем доходу.

7. Розроблені ІСППР та супутні інтелектуальні системи можуть бути використані як інструмент для зменшення наслідків дефіциту висококваліфікованих медичних кадрів та зростання кількості онкологічних захворювань.

Результати роботи були впроваджені в діючу клінічну ІСППР у компанії Better Medicine, яка наразі проходить клінічні випробування для подальшого практичного застосування.

Також отримані результати були впроваджені в навчальний процес, зокрема в магістерський курс «Інтелектуальна обробка голосової інформації», розроблений автором. Запропоновані методи доменної адаптації та гібридні нейронні мережі були інтегровані в лекційні та практичні матеріали розділу «Класифікація сигналів».

Водночас важливо зазначити певні обмеження даної роботи, а саме: обмеженість таксономії класифікації утворів у легенях, відсутність інтерактивної складової та окремої класифікаційної підсистеми в ІСППР для виявлення раку нирок. Крім того, підсистема класифікації легеневих утворів може бути покращена шляхом побудови гібридної глибокої нейронної мережі, яка здатна оперувати всіма утворами пацієнта та враховувати радіометричні ознаки. Також розроблені ІСППР наразі оперують одним КТ-зображенням, тоді як для діагностики окремих онкологічних захворювань або складних клінічних випадків радіологам необхідно аналізувати кілька зображень одночасно (наприклад, отриманих у різних фазах контрастування).

Отже, з наведених вище обмежень впливають такі напрями подальших досліджень:

- створення вибірки для класифікації легеневих утворів, яка відповідала б системі Lung-RADS [5];
- інтеграція інтерактивної підсистеми та підсистеми класифікації в ІСППР для виявлення раку нирок;
- розробка гібридної класифікаційної мережі, здатної оперувати всіма утворами пацієнта та враховувати радіометричні ознаки;

- модифікація ІСППР з метою уможливлення одночасної обробки кількох КТ-сканів з одного дослідження пацієнта.

Усі результати дисертаційного дослідження доведено до високого практичного інженерного рівня. Відповідні наукові результати опубліковано у фахових вітчизняних та закордонних виданнях, що входять до міжнародно визнаних наукометричних баз даних Scopus, а також апробовано на вітчизняних і міжнародних наукових конференціях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] J. X. Liu, Y. Goryakin, A. Maeda, T. Bruckner and R. Scheffler, “Global health workforce labor market projections for 2030”, *Human resources for health*, т. 15, № 1, с. 11, 2017.
- [2] P. Pieterse, “Health Workers in Sub-Saharan Africa: Concurrent Skilled Health Worker Shortages and Under-Employment”, *The International Journal of Health Planning and Management*, 2025.
- [3] N. Bienefeld, E. Keller and G. Grote, “AI interventions to alleviate healthcare shortages and enhance work conditions in critical care: qualitative analysis”, *Journal of Medical Internet Research*, т. 27, e50852, 2025.
- [4] Cancer Research UK. “Lung cancer statistics”. Accessed: 2025-10-29. url: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer>.
- [5] American College of Radiology Committee on Lung-RADS®. “Lung-RADS Assessment Categories 2022”. Accessed: 2023-01-01. url: <https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/Lung-RADS-2022.pdf>.
- [6] S. G. Silverman et al., “Bosniak classification of cystic renal masses, version 2019: an update proposal and needs assessment”, *Radiology*, т. 292, № 2, с. 475–488, 2019.
- [7] Y. Zaychenko, G. Hamidov and B. Chapaliuk, *The application of CNN and hybrid networks in medical images processing and cancer classification*. Cambridge Scholars Publishing, 2023.
- [8] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen and K. H. Maier-Hein, “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”, *Nature methods*, т. 18, № 2, с. 203–211, 2021.

- [9] M. Baumgartner, P. F. Jäger, F. Isensee and K. H. Maier–Hein, “nnDetection: a self–configuring method for medical object detection”, в *International conference on medical image computing and computer–assisted intervention*, Springer, 2021, с. 530–539.
- [10] W. K. Diprose, N. Buist, N. Hua, Q. Thurier, G. Shand and R. Robinson, “Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator”, *Journal of the American Medical Informatics Association*, т. 27, № 4, с. 592–600, 2020.
- [11] National Electrical Manufacturers Association (NEMA), *Digital Imaging and Communications in Medicine (DICOM®) Standard*, Accessed: 2025–10–01, 2023. url: <https://www.dicomstandard.org>.
- [12] S. Nikolov et al., “Clinically applicable segmentation of head and neck anatomy for radiotherapy: deep learning algorithm development and validation study”, *Journal of medical Internet research*, т. 23, № 7, e26151, 2021.
- [13] N. R. Kurtansky et al., “Automated triage of cancer–suspicious skin lesions with 3D total–body photography”, *npj Digital Medicine*, т. 8, № 1, с. 708, 2025.
- [14] P. K. Singh, P. K. Gupta, V. Gupta, S. Singh and S. Akhtar, “Making Cancer Imaging Smarter: Emerging Techniques and Computational Outlook to Guide Precision Diagnostics”, *The Open Bioinformatics Journal*, т. 17, № 1, 2024.
- [15] E. A. Eisenhauer et al., “New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)”, *European journal of cancer*, т. 45, № 2, с. 228–247, 2009.
- [16] L.-J. Tseng, A. Matsuyama and V. MacDonald–Dickinson, “Histology: The gold standard for diagnosis?”, *The Canadian Veterinary Journal*, т. 64, № 4, с. 389, 2023.
- [17] O. S. Pianykh, *Digital imaging and communications in medicine (DICOM) a practical introduction and survival guide*. Springer, 2012.

- [18] National Electrical Manufacturers Association (NEMA), *Digital Imaging and Communications in Medicine (DICOM) Standard: CT Pixel Value Transformation Macro*, Accessed: 2025–10–01, 2023. url: https://dicom.nema.org/medical/dicom/current/output/chtml/part03/sect_c.8.15.3.10.html.
- [19] National Institute of Mental Health (NIMH), *Neuroimaging Informatics Technology Initiative (NifTI)*, Accessed: 2025–10–01, 2025. url: <https://nifti.nimh.nih.gov/>.
- [20] M. Brett, C. J. Markiewicz et al., *NiBabel: Neuroimaging data I/O in Python*, <https://nipy.org/nibabel/>, 2022. DOI: 10.5281/zenodo.591597.
- [21] B. Sistaninejhad, H. Rasi and P. Nayeri, “A review paper about deep learning for medical image analysis”, *Computational and Mathematical Methods in Medicine*, т. 2023, № 1, с. 7 091 301, 2023.
- [22] L. Maier–Hein et al., “Metrics reloaded: Recommendations for image analysis validation. arXiv 2023”, *arXiv preprint arXiv:2206.01653*, 2023.
- [23] D. P. Huttenlocher, G. A. Klanderman and W. J. Rucklidge, “Comparing images using the Hausdorff distance”, *IEEE Transactions on pattern analysis and machine intelligence*, т. 15, № 9, с. 850–863, 2002.
- [24] A. A. Taha and A. Hanbury, “Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool”, *BMC medical imaging*, т. 15, № 1, с. 29, 2015.
- [25] D. Müller, I. Soto–Rey and F. Kramer, “Towards a guideline for evaluation metrics in medical image segmentation”, *BMC Research Notes*, т. 15, № 1, с. 210, 2022.
- [26] J. Kohen, “A coefficient of agreement for nominal scale”, *Educ Psychol Meas*, т. 20, с. 37–46, 1960.

- [27] A. I. Bandos, H. E. Rockette, T. Song and D. Gur, “Area under the free–response ROC curve (FROC) and a related summary index”, *Biometrics*, т. 65, № 1, с. 247–256, 2009.
- [28] В. В. Литвин, Р. Р. Даревич та Д. Г. Досин, *Інтелектуальні системи підтримки прийняття рішень*. Київ: Національна академія наук України, Інститут проблем штучного інтелекту, 2021, ISBN 978–966–8406–94–2, доступ 2025–10–04.
- [29] F. T. Tehrani, “Computerized decision support systems for mechanical ventilation”, *Efficient decision support systems, practice and challenges in biomedical related domain* (In–Tech publishers, London, UK, 2011), с. 227–238, 2011.
- [30] В. В. Литвин, Р. Р. Даревич та Д. Г. Досин, “Інтелектуальні системи підтримки прийняття рішень на основі адаптивних онтологій”, *Штучний інтелект*, № 3, с. 388–395, 2011.
- [31] Z. Ullah et al., “A deep learning based intelligent decision support system for automatic detection of brain tumor”, *Biomedical Engineering and Computational Biology*, т. 15, с. 11 795 972 241 277 322, 2024.
- [32] A. Masood, U. Naseem, J. Rashid, J. Kim and I. Razzak, “Review on enhancing clinical decision support system using machine learning”, *CAAI Transactions on Intelligence Technology*, 2024.
- [33] M. Li, Y. Jiang, Y. Zhang and H. Zhu, “Medical image analysis using deep learning algorithms”, *Frontiers in public health*, т. 11, с. 1 273 253, 2023.
- [34] K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition”, в *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, с. 770–778.
- [35] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, “Densely connected convolutional networks”, в *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, с. 4700–4708.

- [36] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”, в *Proceedings of the 36th International Conference on Machine Learning (ICML)*, PMLR, 2019, с. 6105–6114. DOI: 10.48550/arXiv.1905.11946. url: <https://doi.org/10.48550/arXiv.1905.11946>.
- [37] W. Shen, M. Zhou, F. Yang, C. Yang and J. Tian, “Multi-scale convolutional neural networks for lung nodule classification”, в *International conference on information processing in medical imaging*, Springer, 2015, с. 588–599.
- [38] Б. В. Чапалюк, “Системи автоматичної медичної комп’ютерної діагностики з використанням методів штучного інтелекту”, Дисертація доктора філософії, Київ, Київ, 2020, с. 139.
- [39] S. Chen, K. Ma and Y. Zheng, “Med3d: Transfer learning for 3d medical image analysis”, *arXiv preprint arXiv:1904.00625*, 2019.
- [40] O. Ronneberger, P. Fischer and T. Brox, “U-net: Convolutional networks for biomedical image segmentation”, в *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, с. 234–241.
- [41] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth and D. Xu, “Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images”, в *International MICCAI brainlesion workshop*, Springer, 2021, с. 272–284.
- [42] S. Perera, P. Navard and A. Yilmaz, “Segformer3d: an efficient transformer for 3d medical image segmentation”, в *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, с. 4981–4988.
- [43] A. Howard, A. Chow, CorporateResearchSartorius, M. Ca, P. Culliton and T. Jackson, *Sartorius – Cell Instance Segmentation*, <https://kaggle.com/competitions/sartorius-cell-instance-segmentation>, Kaggle, 2021.
- [44] T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, “Focal loss for dense object detection”, в *Proceedings of the IEEE international conference on computer vision*, 2017, с. 2980–2988.

- [45] A. Kirillov et al., “Segment anything”, в *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, c. 4015–4026.
- [46] S. Mahadevan, P. Voigtlaender and B. Leibe, “Iteratively trained interactive segmentation”, *arXiv preprint arXiv:1805.04398*, 2018.
- [47] B. Shen, L. Chang, S. Chen, S. Guo and H. Liu, “Lightweight Method for Interactive 3D Medical Image Segmentation with Multi–Round Result Fusion”, *arXiv preprint arXiv:2412.08315*, 2024.
- [48] H. Li, H. Liu, D. Hu, J. Wang and I. Oguz, “Prism: A promptable and robust interactive segmentation model with visual prompts”, в *International Conference on Medical Image Computing and Computer–Assisted Intervention*, Springer, 2024, c. 389–399.
- [49] F. Isensee et al., “nninteractive: Redefining 3d promptable segmentation”, *arXiv preprint arXiv:2503.08373*, 2025.
- [50] H. Guan and M. Liu, “Domain adaptation for medical image analysis: a survey”, *IEEE Transactions on Biomedical Engineering*, т. 69, № 3, c. 1173–1185, 2021.
- [51] A. Howard et al., *HuBMAP + HPA – Hacking the Human Body*, <https://kaggle.com/competitions/hubmap-organ-segmentation>, Kaggle, 2022.
- [52] V. Sydorskyi, I. Krashenyi, D. Savka and O. Zarichkovyi, “Semi–Supervised Segmentation of Functional Tissue Units at the Cellular Level”, в *IT&I Workshops*, 2022.
- [53] N. Kurtansky, V. Rotemberg, M. Gillis, K. Kose, W. Reade and A. Chow, *ISIC 2024 – Skin Cancer Detection with 3D–TBP*, <https://kaggle.com/competitions/isic-2024-challenge>, Kaggle, 2024.
- [54] Y. Jain et al., *SenNet + HOA – Hacking the Human Vasculature in 3D*, <https://kaggle.com/competitions/blood-vessel-segmentation>, Kaggle, 2023.

- [55] Lung Cancer Policy Network. “Lung cancer screening: Factsheets”. Accessed: 2025–10–29. url: <https://www.lungcancerpolicynetwork.com/app/uploads/Lung-cancer-screening-factsheets.pdf>.
- [56] Lung Cancer Policy Network. “Lung cancer screening: Learning from implementation”. Accessed: 2025–10–29. url: <https://www.lungcancerpolicynetwork.com/app/uploads/Lung-cancer-screening-learning-from-implementation.pdf>.
- [57] icometrix. “dicom2nifti: Convert DICOM files to Nifti format”. Accessed: 2025–10–29. url: <https://icometrix.github.io/dicom2nifti/dicom2nifti.html>.
- [58] J. Wasserthal et al., “TotalSegmentator: robust segmentation of 104 anatomic structures in CT images”, *Radiology: Artificial Intelligence*, т. 5, № 5, e230024, 2023.
- [59] F. Isensee, P. F. Jäger, S. A. Kohl, J. Petersen and K. H. Maier-Hein, “Automated design of deep learning methods for biomedical image segmentation”, *arXiv preprint arXiv:1904.08128*, 2019.
- [60] V. Sydorskyi, “Evaluating Loss Functions for 3D Lung Nodules Segmentation”, в *System Analysis and Data Mining*, M. Zgurovsky and N. Pankratova, ред. Cham: Springer Nature Switzerland, 2026, с. 279–293, ISBN: 978-3-031-97529-5. DOI: 10.1007/978-3-031-97529-5_17. url: https://doi.org/10.1007/978-3-031-97529-5_17.
- [61] F. Sun, Z. Luo and S. Li, “Boundary difference over union loss for medical image segmentation”, в *International conference on medical image computing and computer-assisted intervention*, Springer, 2023, с. 292–301.
- [62] C. Szegedy et al., “Going deeper with convolutions”, в *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, с. 1–9.
- [63] Seung Lab, *connected-components-3d: 3D connected component labeling for NumPy arrays*, <https://github.com/seung-lab/connected-components-3d>, Accessed: 2025–10–29, 2021.

- [64] S. G. I. Armato et al., *Data from LIDC-IDRI*, [Data set], The Cancer Imaging Archive (TCIA), 2015. DOI: 10.7937/K9/TCIA.2015.LO9QL9SX. url: <https://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX>.
- [65] В. С. Сидорський, “Класифікація легеневих утворів на фрагментах КТ-зображень із використанням тривимірних згорткових нейронних мереж”, *Таврійський науковий вісник. Серія: Технічні науки*, т. 1, № 5, с. 399–412, 2025. DOI: 10.32782/tnv-tech.2025.5.1.42.
- [66] A. Neubeck and L. Van Gool, “Efficient non-maximum suppression”, в *18th international conference on pattern recognition (ICPR'06)*, IEEE, т. 3, 2006, с. 850–855.
- [67] A. Sharma, A. Z. Ansari, R. Kakulavarapu, M. H. Stensen, M. A. Riegler and H. L. Hammer, “Predicting cell cleavage timings from time-lapse videos of human embryos”, *Big Data and Cognitive Computing*, т. 7, № 2, с. 91, 2023.
- [68] V. Sydorskyi, “Interactive decision support system for lung cancer segmentation”, *System research and information technologies*, № 2, с. 68–81, 2024.
- [69] M. S. Al-Marhoon, A. M. Osman, M. M. Kamal and A. A. Shokeir, “Incidental vs symptomatic renal tumours: Survival outcomes”, *Arab Journal of Urology*, т. 9, № 1, с. 17–21, 2011.
- [70] B. de Bono, P. Grenon, R. Baldock and P. Hunter, “Functional tissue units and their primary tissue motifs in multi-scale physiology”, *Journal of biomedical semantics*, т. 4, № 1, с. 22, 2013.
- [71] J. Litwin, “Histochemistry and cytochemistry of 3, 3'-diaminobenzidine. A review.”, *Folia Histochemica et Cytochemica*, т. 17, № 1, с. 3–28, 1979.
- [72] G. Gomori, “The Periodic-Acid Schiff Stain”, *American Journal of Clinical Pathology*, т. 22, № 3, с. 277–281, 1952, ISSN: 0002-9173. DOI: 10.1093/ajcp/22.3_ts.277. url: https://doi.org/10.1093/ajcp/22.3_ts.277.

- [73] A. T. Feldman and D. Wolfe, “Tissue processing and hematoxylin and eosin staining”, в *Histopathology: methods and protocols*, Springer, 2014, c. 31–43.
- [74] J.-N. Chen, S. Sun, J. He, P. H. Torr, A. Yuille and S. Bai, “Transmix: Attend to mix for vision transformers”, в *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, c. 12135–12144.
- [75] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh and J. Liang, “Unet++: Redesigning skip connections to exploit multiscale features in image segmentation”, *IEEE transactions on medical imaging*, т. 39, № 6, c. 1856–1867, 2019.
- [76] S. van der Walt, J. L. Schonberger and J. e. a. Nunez-Iglesias, *Scikit-image: Histogram Matching*, https://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_histogram_matching.html, Accessed: 2025–02–22.
- [77] S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe and Y. Yoo, “Cutmix: Regularization strategy to train strong classifiers with localizable features”, в *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, c. 6023–6032.
- [78] N. C. Sathe and P. M. Zito, *Skin Cancer*. Treasure Island (FL): StatPearls Publishing, 2025, PMID: 28722978. url: <https://www.ncbi.nlm.nih.gov/books/NBK441949/>.
- [79] K. M. Hosny, M. A. Kassem and M. M. Foaud, “Classification of skin lesions using transfer learning and augmentation with Alex-net”, *PloS one*, т. 14, № 5, e0217293, 2019.
- [80] V. Sydorskyi, I. Krashenyi and O. Yakubenko, *Multimodal system for skin cancer detection*, arXiv preprint, 2026. DOI: 10.48550/arXiv.2601.14822. arXiv: 2601.14822 [cs.CV]. url: <https://arxiv.org/abs/2601.14822>.
- [81] I. Canfield Scientific, *VECTRA® WB360 Imaging System*, <https://www.canfieldsci.com>, Accessed: 2024–10–21, 2024.

- [82] International Skin Imaging Collaboration (ISIC), *ISIC Archive*, <https://www.isic-archive.com/>, Accessed: 2025-02-22, 2025.
- [83] R. Rombach, A. Blattmann, D. Lorenz, P. Esser and B. Ommer, “High-resolution image synthesis with latent diffusion models”, в *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, c. 10684–10695.
- [84] S. Woo et al., “Convnext v2: Co-designing and scaling convnets with masked autoencoders”, в *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, c. 16133–16142.
- [85] M. Tan and Q. Le, “Efficientnetv2: Smaller models and faster training”, в *International conference on machine learning*, PMLR, 2021, c. 10096–10106.
- [86] M. Maaz et al., “Edgenext: efficiently amalgamated cnn-transformer architecture for mobile vision applications”, в *European conference on computer vision*, Springer, 2022, c. 3–20.
- [87] G. Ke et al., “Lightgbm: A highly efficient gradient boosting decision tree”, *Advances in neural information processing systems*, т. 30, 2017.
- [88] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system”, в *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, c. 785–794.
- [89] T. Akiba, S. Sano, T. Yanase, T. Ohta and M. Koyama, “Optuna: A next-generation hyperparameter optimization framework”, в *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019, c. 2623–2631.
- [90] S. G. Armato et al., “The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans”, *Medical Physics*, т. 38, c. 915–931, 2011. DOI: 10.1118/1.3528204.

- [91] M. Stone, “Cross–validatory choice and assessment of statistical predictions”, *Journal of the royal statistical society: Series B (Methodological)*, т. 36, № 2, с. 111–133, 1974.
- [92] I. Loshchilov and F. Hutter, “SGDR: Stochastic Gradient Descent with Warm Restarts”, в *International Conference on Learning Representations*, 2017.
- [93] D. P. Kingma, “Adam: A method for stochastic optimization”, *arXiv preprint arXiv:1412.6980*, 2014.
- [94] H. Wang, H. Zhu, L. Ding and K. Yang, “Attention pyramid pooling network for artificial diagnosis on pulmonary nodules”, *Plos one*, т. 19, № 5, e0302641, 2024.
- [95] K. Mehta, A. Jain, J. Mangalagiri, S. Menon, P. Nguyen and D. R. Chapman, “Lung nodule classification using biomarkers, volumetric radiomics, and 3D CNNs”, *Journal of digital imaging*, т. 34, № 3, с. 647–666, 2021.
- [96] O. Akin et al., *The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma Collection (TCGA–KIRC)*, The Cancer Imaging Archive, Version 3, 2016. DOI: 10.7937/K9/TCIA.2016.V6PBVTDR. url: <https://doi.org/10.7937/K9/TCIA.2016.V6PBVTDR>.
- [97] N. Heller et al., *C4KC KiTS Challenge Kidney Tumor Segmentation Dataset*, The Cancer Imaging Archive, Version 3, 2019. DOI: 10.7937/TCIA.2019.IX49E8NX. url: <https://doi.org/10.7937/TCIA.2019.IX49E8NX>.
- [98] *Better Medicine*, <https://bettermedicine.ai/>, Accessed: 2025–02–07, 2025.
- [99] B. Petryshak et al., “A clinically validated AI framework for kidney cancer detection and characterization”, *Communications Medicine*, 2025.
- [100] M. Uhlén et al., “Tissue–based map of the human proteome”, *Science*, т. 347, № 6220, с. 1 260 419, 2015.
- [101] P. J. Thul et al., “A subcellular map of the human proteome”, *Science*, т. 356, № 6340, eaal3321, 2017.

- [102] E. Sjöstedt et al., “An atlas of the protein–coding genes in the human, pig, and mouse brain”, *Science*, т. 367, № 6482, eaay5947, 2020.
- [103] M. Karlsson et al., “A singletype transcriptomics map of human tissues”, *Science advances*, т. 7, № 31, eabh2169, 2021.
- [104] M. Uhlen et al., “A pathology atlas of the human cancer transcriptome”, *Science*, т. 357, № 6352, eaan2507, 2017.
- [105] M. B. Álvarez et al., “A human pan–disease blood atlas of the circulating proteome”, *Science*, eadx2678, 2025.
- [106] H. Jin et al., “Systematic transcriptional analysis of human cell lines for gene expression landscape and tumor representation”, *Nature communications*, т. 14, № 1, c. 5417, 2023.
- [107] M. Uhlen et al., “A genome–wide transcriptomic analysis of protein–coding genes in human blood cells”, *Science*, т. 366, № 6472, eaax9198, 2019.
- [108] M. Uhlén et al., “The human secretome”, *Science signaling*, т. 12, № 609, eaaz0274, 2019.
- [109] C.-.-U. T. C. L. lcai@ caltech. edu 21 b Shendure Jay 9 Trapnell Cole 9 Lin Shin shinlin@ uw. edu 2 e Jackson Dana 9, U. T. Z. K. kzhang@ bioeng. ucsd. edu 15 b Sun Xin 15 Jain Sanjay 24 Hagood James 25 Pryhuber Gloria 26 Kharchenko Peter 8, C. I. of Technology TTD Cai Long lcai@ caltech. edu 21 b Yuan Guo–Cheng 35 Zhu Qian 35 Dries Ruben 35, H. T. Y. P. peng_yin@ hms. harvard. edu 36 37 b Saka Sinem K. 36 37 Kishi Jocelyn Y. 36 37 Wang Yu 36 37 Goldaracena Isabel 36 37, P. T. L. J. jlaskin@ purdue. edu 10 b Ye DongHye 10 38 Burnum–Johnson Kristin E. 39 Piehowski Paul D. 39 Ansong Charles 39 Zhu Ying 39, S. T. H. P. harbury@ stanford. edu 11 b Desai Tushar 40 Mulye Jay 11 Chou Peter 11 Nagendran Monica 40 et al., “The human body at cellular resolution: the NIH Human Biomolecular Atlas Program”, *Nature*, т. 574, № 7777, c. 187–192, 2019.

- [110] GTEx Consortium, *GTEx Portal: Histology Images*, <https://www.gtexportal.org/home/histologyPage>, Accessed: 2025–02–14.
- [111] ISIC Research, *Challenge 2024 Metrics*, <https://github.com/ISIC-Research/Challenge-2024-Metrics>, GitHub repository, 2024.
- [112] K. Community, *SIIM–ISIC Melanoma Classification — 1st Place Solution (All Data Are Ext)*, <https://www.kaggle.com/competitions/siim-isic-melanoma-classification/writeups/all-data-are-ext-1st-place-solution>, Kaggle competition solution write–up, accessed 2025–03–XX, 2020.
- [113] A. Zawacki et al., *SIIM–ISIC Melanoma Classification*, <https://kaggle.com/competitions/siim-isic-melanoma-classification>, Kaggle, 2020.
- [114] S. N, J. A. C, R. RR and G. I. I, “Analysis of GAN for Melanoma Skin CancerClassification with Dermatologist Recommendation”, в *2024 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, 2024, c. 1—8. DOI: 10.1109/RAEEUCCI61380.2024.10547727.
- [115] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. S. Marcal and J. Rozeira, “PH2 – A dermoscopic image database for research and benchmarking”, в *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, c. 5437—5440. DOI: 10.1109/EMBC.2013.6610779.
- [116] J. Kawahara, S. Daneshvar, G. Argenziano and G. Hamarneh, “Seven–point checklist and skin lesion classification using multitask multimodal neural nets”, *IEEE Journal of Biomedical and Health Informatics*, т. 23, № 2, c. 538–546, 6ep. 2019, ISSN: 2168–2194. DOI: 10.1109/JBHI.2018.2824327.
- [117] M. F. J. Acosta, L. Y. C. Tovar, M. B. Garcia–Zapirain and W. S. Percybrooks, “Melanoma diagnosis using deep learning techniques on dermoscopic images”, *BMC Medical Imaging*, т. 21, № 1, c. 6, 2021. DOI: 10.1186/s12880-

020-00534-8. url: <https://bmcmmedimaging.biomedcentral.com/articles/10.1186/s12880-020-00534-8>.

- [118] N. Kurtansky et al., “The SLICE–3D dataset: 400,000 skin lesion image crops extracted from 3D TBP for skin cancer detection”, *Scientific Data*, т. 11, № 1, с. 884, 2024. DOI: 10.1038/s41597-024-03743-w.

ДОДАТОК А. СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Sydorskyi, V. (2024). Interactive decision support system for lung cancer segmentation. *System Research and Information Technologies*, 2, 68–81. DOI: <https://doi.org/10.20535/SRIT.2308-8893.2024.2.05> ISSN: 1681–6048 (Scopus Q4)

2. Sydorskyi, V. (2025). Evaluating loss functions for 3D lung nodules segmentation. In *System analysis and data mining* (pp. 279–293). *Studies in Systems, Decision and Control* (Vol. 609) Cham: Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-97529-5_17 ISSN: 2198–4182 (Scopus)

3. Сидорський, В. С. (2025). Класифікація легеневих утворів на фрагментах КТ-зображень із використанням тривимірних згорткових нейронних мереж. *Таврійський науковий вісник. Серія: Технічні науки*, 1(5), 399–412. DOI: <https://doi.org/10.32782/tnv-tech.2025.5.1.42> ISSN: 2786–4596 (фахове видання категорії Б)

У матеріалах конференцій:

4. Sydorskyi, V., Krashenyi, I., Savka, D., & Zarichkovyi, O. (2022). Semi-supervised segmentation of functional tissue units at the cellular level. *CEUR Workshop Proceedings*, 3384, 11–23. https://ceur-ws.org/Vol-3384/Paper_2.pdf ISSN: 1613–0073 (Conference Paper). Внесок: Сидорський В.С. – розробка основних методів і фінальної системи; Крашений І.Е. – наукова редакція, дослідження архітектур нейронних мереж; Саква Д.Ю. – дослідження даних і різних методів забарвлення, часткова розробка стратегії аугментації; Зарічковий О.А. – огляд літератури та існуючих методів.

5. Sydorskyi, V. S. (2023). Interactive intelligent decision support system for lung cancer diagnosis (E. M. Kiseleva, Ed.). In *Mathematical support and software for intelligent systems (MSSIS–2023): Abstracts of the XXI International scientific and practical conference* (pp. 45–46). Dnipro: DNU. <http://mpzis.dnu.dp.ua/wp-content/uploads/2023/11/mpzis-2023.pdf>

ДОДАТОК Б. АКТ ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОГО ДОСЛІДЖЕННЯ



Act of the implementation of the dissertation research results

of
Sydorskyi Volodymyr Sergiyovych

In his PhD research, Volodymyr Sydorskyi focused on building and exploring Intelligent Decision Support Systems (IDSS) for analyzing medical images, with particular emphasis on CT modality and the task of lung cancer detection. During his research, several IDSS solutions were developed, including systems for lung cancer and melanoma detection. Alongside these systems, a number of methods and models were proposed to improve diagnostic accuracy and to enable deeper interaction between users (doctors and other medical specialists) and the systems. This research contributes to improving diagnostic accuracy, increasing doctors' efficiency, and reducing routine workload.

A series of approaches and results—such as systematic research on 3D semantic segmentation loss functions and their modifications, the development of a lung nodule classification system, and several architectural components of the proposed framework—are being applied in the Better Medicine AI product, specifically in the Kidney Cancer and Lung Cancer detection systems. The Kidney Cancer detection system is currently in the process of deployment into clinical IT systems, while the Lung Cancer detection system is expected to be finalized in the near future.

Signed on behalf of Better Medicine OÜ

Priit Salumaa
CEO and member of the board
Better Medicine

Better Medicine OÜ | Reg. No. 16093156 | VAT EE102413320
Narva mnt 9, Tartu linn, Tartu linn 51009, Estonia
+372 5334 0822 | info@bettermedicine.ai | bettermedicine.ai

ДОДАТОК В. АКТ ВПРОВАДЖЕННЯ В НАВЧАЛЬНИЙ ПРОЦЕС



УКРАЇНА

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
(КПІ ім. Ігоря Сікорського)

НАВЧАЛЬНО-НАУКОВИЙ
ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ

пр-т Берестейський, 37, м. Київ, 03056, тел. (044) 204 81 46 тел. (044) 204 80 14
<https://www.iasa.kpi.ua> e-mail: ipsa@kpi.ua ЄДРПОУ 02070921

«ЗАТВЕРДЖУЮ»

В. о. директора ІН ІПСА
КПІ ім. Ігоря Сікорського

Інститут прикладного системного аналізу
Костянтин ЄФРЕМОВ
«02» березня 2026 року

АКТ

про впровадження результатів дисертаційної роботи
Сидорського Володимира Сергійовича

на тему «Інтелектуальна система підтримки прийняття рішень на базі нейронних мереж
глибокого навчання для задач медичної діагностики»
у навчальний процес ІН ІПСА КПІ ім. Ігоря Сікорського

Цим актом підтверджується, що у процесі виконання дисертаційного дослідження Сидорський В.С. отримав низку наукових результатів стосовно використання глибоких згорткових нейронних мереж, методів доменної адаптації, а також гібридних нейронних мереж. Ці результати впроваджено у навчальний процес кафедри штучного інтелекту навчально-наукового інституту прикладного системного аналізу КПІ ім. Ігоря Сікорського в межах вибіркового освітнього компонента «Інтелектуальна обробка голосової інформації» освітньо-професійної програми «Системи і методи штучного інтелекту» другого (магістерського) рівня вищої освіти, який забезпечує Сидорський В.С.

Автором дисертаційної роботи оновлено силабус дисципліни та внесено відповідні зміни до лекційних і практичних занять з метою врахування отриманих результатів, зокрема додано розділ «Класифікація сигналів». У межах цього розділу студенти вивчають методи доменної адаптації для класифікації акустичних сигналів, а також гібридні нейронні мережі, розроблені Сидорським В.С. в межах дисертаційної роботи, та застосовують їх у своїх індивідуальних завданнях.

В. о. завідувачки кафедри
штучного інтелекту

Ірина ДЖИГИРЕЙ