

Лещенко В. О., Заболотна Н. І.

**АНАЛІЗ СУЧАСНИХ МЕТОДІВ ТА БІОТЕХНІЧНИХ СИСТЕМ ДЛЯ ДІАГНОСТИКИ  
ОНКОПАТОЛОГІЙ ЗА ЦИТОЛОГІЧНИМИ ЗОБРАЖЕННЯМИ**

У статті проведено комплексний аналіз сучасних методів та біотехнічних систем для автоматизованої діагностики онкопатологій за цитологічними зображеннями. Актуальність дослідження зумовлена зростанням онкологічної захворюваності у світі – за оцінками GLOBOCAN 2022, щорічно реєструється близько 20 мільйонів нових випадків раку, а в Україні понад 38% випадків діагностується на пізніх стадіях, що підкреслює гостру потребу в ефективних засобах раннього скринінгу. Метою роботи є систематичний аналіз поточного стану методів комп'ютерного аналізу цитологічних зображень, виявлення ключових невирішених проблем та обґрунтування напрямків подальших досліджень у галузі біомедичної інженерії. Розглянуто медико-технічні аспекти формування цитологічних зображень, включаючи типи мікроскопії, методики підготовки препаратів та характерні артефакти. Проаналізовано комерційні CAD-системи для цервікальної цитології, зокрема ThinPrep Imaging System, BD FocalPoint GS та Genius Digital Diagnostics System, а також основні відкриті датасети (SIPaKMeD, Herlev, ISBI 2014/2015, CRIC, Cx22). Виконано порівняльний аналіз класичних методів сегментації клітин (порогова обробка Оцу, метод вододілу, активні контури) та архітектур глибокого навчання (U-Net, Mask R-CNN, Cellpose, StarDist, HoVer-Net), який показав суттєву перевагу нейромережових підходів за точністю та стійкістю. Досліджено методи класифікації від класичних алгоритмів машинного навчання (SVM, Random Forest) до сучасних глибоких нейронних мереж (ResNet, EfficientNet, Vision Transformers). Виявлено три ключові невирішені проблеми: недостатня точність інстансної сегментації перекриваючих клітин у реальних препаратах, значний доменний зсув між лабораторіями, що призводить до суттєвого зниження продуктивності навчених моделей, а також обмеженість та незбалансованість наявних відкритих датасетів. Обґрунтовано напрямки подальших досліджень, спрямований на розробку біотехнічної системи, яка поєднувала б модифікований метод інстансної сегментації з автоматичною класифікацією за морфологічними ознаками та забезпечувала стійкість до доменного зсуву для потреб ранньої діагностики онкопатологій.

**Ключові слова:** цитологічні зображення, комп'ютерна діагностика, сегментація клітин, нейронні мережі, CAD-системи, онкопатологія, біотехнічна система.

**Постановка проблеми.** Онкологічні захворювання залишаються однією з провідних причин смертності у світі. За даними Міжнародного агентства з дослідження раку (IARC), у 2022 році зареєстровано близько 20 мільйонів нових випадків онкологічних захворювань та 9,7 мільйона смертей від раку [1]. Прогнозується, що до 2050 року кількість нових випадків зросте до 35 мільйонів, що зумовлено демографічними змінами та старінням населення [1, 2]. В Україні ситуація є не менш тривожною: за прогнозами, у 2022 році очікувалося близько 179 тисяч нових випадків раку, при цьому рак шийки матки залишається четвертим за поширеністю серед жінок із стандартизованим показником захворюваності 17,5 на 100 000 [3]. Додатковим викликом для української системи охорони здоров'я стала повномасштабна військова агресія, яка призвела до руйнування інфраструктури, дефіциту кадрів та зниження охоплення скринінговими програмами [4].

Цитологічний аналіз є одним із ключових методів ранньої діагностики передракових станів та злоякісних новоутворень. Його переваги порівняно з гістологічним дослідженням полягають у малоінвазивності забору матеріалу, відносній швидкості отримання результатів та значно нижчій вартості процедури. Водночас традиційний (ручний) цитологічний аналіз характеризується значною суб'єктивністю: міжекспертна узгодженість при оцінці препаратів Папаніколау становить лише 50–70% навіть серед досвідчених цитопатологів [5]. Крім того, спостерігається глобальний дефіцит кваліфікованих спеціалістів – за даними американських досліджень, кількість цитотехнологів зменшилася на 38% з 2005 до 2022 року [6].

**Аналіз останніх досліджень і публікацій.** У сучасних дослідженнях автоматизованої цитодіагностики основну увагу приділено трьом напрямкам: побудові САД-систем для цифрової патології [7-9], удосконаленню методів сегментації клітин і ядер, зокрема для перекриваючих клітин [10-17], а також застосуванню методів глибокого навчання для класифікації цитологічних зображень [18-23]. Попри суттєвий прогрес, у наукових публікаціях залишаються недостатньо вирішеними питання стійкої інстансної сегментації у складних препаратах, обмеженості відкритих анованих датасетів, доменного зсуву між лабораторіями та перенесення високих експериментальних результатів у реальну клінічну практику.

**Мета дослідження.** Мета статті – систематично проаналізувати сучасні методи комп'ютерного аналізу цитологічних зображень для діагностики онкопатологій, визначити основні невирішені проблеми та обґрунтувати напрями подальших досліджень у біомедичній інженерії.

**Основний матеріал дослідження.** Рання діагностика онкологічних захворювань є вирішальним фактором, що визначає ефективність лікування та прогноз виживаності пацієнтів. За оцінками GLOBOCAN 2022, рак легенів (12,4% усіх нових випадків) та рак молочної залози (11,6%) є найпоширенішими формами онкопатологій у світі, тоді як рак шийки матки залишається четвертим за частотою у жінок із щорічною захворюваністю близько 660 тисяч випадків та 350 тисячами смертей [1, 2]. В країнах з низьким та середнім індексом людського розвитку (HDI) спостерігається непропорційно вища смертність від раку, що пояснюється пізньою діагностикою та обмеженим доступом до якісного лікування [2].

В Україні, за даними Національного канцер-реєстру, понад 38% випадків раку у 2022 році діагностовано на III–IV стадіях, а показник однорічної летальності становив 23,9% [4]. Для раку шийки матки частка пізніх стадій при первинній діагностиці досягала 36% у 2023 році, що вдвічі перевищує показник 2013 року (16,1%) [24]. Ці дані свідчать про критичну потребу у впровадженні ефективних скринінгових програм із залученням сучасних діагностичних технологій.

Біотехнічна система формування цитологічних зображень. Формування якісного цитологічного зображення є першим і визначальним етапом автоматизованого аналізу. Основою біотехнічної системи візуалізації є мікроскоп із цифровою системою реєстрації зображень. Залежно від типу мікроскопії, що застосовується, розрізняють кілька підходів до отримання цитологічних зображень.

Світлова (оптична) мікроскопія є найбільш поширеним методом у рутинній цитологічній практиці. Сучасні моделі дослідницьких мікроскопів, зокрема Olympus BX53, Nikon Eclipse Ni-U, Zeiss Axio Imager.Z2 та Leica DM6 B, забезпечують збільшення до 1000× з роздільною здатністю до 0,2 мкм та оснащені цифровими камерами з роздільною здатністю від 5 до 20 мегапікселів. Ці системи дозволяють формувати зображення у яскравому полі, темному полі, фазовому контрасті та диференційному інтерференційному контрасті (DIC).

Флуоресцентна мікроскопія, зокрема конфокальна лазерна скануюча мікроскопія (CLSM), забезпечує значно вищу контрастність та можливість оптичного зрізування зразка,

що дозволяє отримувати тривимірні реконструкції клітинних структур. Однак висока вартість обладнання та необхідність спеціального фарбування (флуорохроми) обмежують широке клінічне використання цього методу [25].

Електронна мікроскопія (сканувальна та трансмісійна) забезпечує надвисоку роздільну здатність (до нанометрового рівня), проте потребує складної підготовки зразків та не придатна для рутинного скринінгу через тривалість процедури та високу вартість.

*Підготовка цитологічних препаратів.* Якість цитологічного зображення безпосередньо залежить від методики підготовки препарату. Основними етапами є забір матеріалу, фіксація, фарбування та нанесення на предметне скло. Серед методів фарбування найбільш широко використовуються метод Папаніколау (PAP), який є стандартом для цервікальної цитології і забезпечує поліхромне забарвлення клітинних компонентів, та метод Романовського-Гімзи, що застосовується для гематологічних та негінекологічних цитологічних досліджень [5]. Сучасна технологія рідинної цитології (Liquid-Based Cytology, LBC), зокрема системи ThinPrep (Hologic) та SurePath (BD), суттєво покращила якість препаратів шляхом рівномірного розподілу клітин у тонкому шарі та зменшення кількості артефактів порівняно з конвенційними мазками [6, 7].

*Артефакти та проблеми якості зображень.* На етапі формування цитологічних зображень виникає низка типових артефактів, які суттєво ускладнюють подальший автоматизований аналіз. До основних проблем належать: нерівномірне освітлення поля зору, що призводить до варіації інтенсивності фону; розфокусування окремих ділянок зображення через неоднорідну товщину препарату; шум цифрової камери; а також перекриття (злипання) клітин, яке є однією з найбільш складних задач для алгоритмів автоматичної сегментації [10, 11]. Додатково спостерігаються варіації кольорних характеристик між різними партіями реагентів та лабораторіями, що створює проблему доменного зсуву (domain shift) при застосуванні навчених моделей [18].

Системи автоматизованої комп'ютерної діагностики (Computer-Aided Diagnosis, CAD) у цитопатології призначені для допомоги цитотехнологам та патологам у виявленні аномальних клітин, підвищення чутливості діагностики та оптимізації робочого процесу лабораторії. Історично перші CAD-системи для цервікальної цитології з'явилися у 2000-х роках і базувалися на алгоритмах класичного комп'ютерного зору [6, 7].

*Класифікація CAD-систем.* За типом задачі, що вирішується, CAD-системи поділяють на: (1) скринінгові – призначені для первинного сортування препаратів за ступенем підозрілості; (2) діагностичні – спрямовані на класифікацію клітин або ділянок зображення за конкретними діагностичними категоріями (наприклад, за системою Бетесда); (3) прогностичні – оцінюють імовірність прогресії захворювання. За архітектурою розрізняють класичні pipeline-системи, де кожен етап (сегментація, виділення ознак, класифікація) реалізований окремим модулем, та end-to-end системи на основі глибокого навчання, де нейронна мережа безпосередньо пов'язує вхідне зображення з діагностичним висновком [18, 19].

Серед комерційних систем, що отримали регуляторне схвалення, найбільш відомими є ThinPrep Imaging System (TIS) від Hologic (FDA, 2003), яка аналізує 22 поля зору з найбільш підозрілими клітинами; BD FocalPoint GS (FDA, 2008), що додатково категоризує слайди за квантилями імовірності аномалій; та нове покоління – Genius Digital Diagnostics System (Hologic), який отримав FDA-дозвіл у лютому 2024 року і є першою та єдиною схваленою системою цифрової цитології, що поєднує об'ємне сканування (14 Z-площин) з алгоритмом глибокого навчання [6, 8]. За даними ретроспективних досліджень, система Genius продемонструвала чутливість 98,2–100% для виявлення ASC-US+ за 95% довірчого інтервалу, забезпечуючи при цьому скорочення часу перегляду препарату приблизно на 50% [6, 8, 9].

Серед дослідницьких платформ варто відзначити систему BestCyte (CellSolutions, США) для автоматизованого сортування клітин, CytoProcessor (DATEXIM, Франція), а також ряд рішень на базі хмарних обчислень, зокрема від PathAI та Google Health, які активно розвивають алгоритми аналізу цитологічних та гістологічних зображень [18, 19]. Узагальнену порівняльну характеристику основних комерційних CAD-систем наведено в табл. 1.

Таблиця 1 – Порівняльна характеристика комерційних CAD-систем для цервікальної цитології

Система	Виробник	Рік FDA	Метод аналізу	Тип	Ключова особливість	Джерело
ThinPrep Imaging System	Hologic	2003	Класичний CV	Скринінг	22 FoV з підозрілими клітинами	[7], [8]
BD FocalPoint GS	Becton Dickinson	2008	Класичний CV	Скринінг	Квантильна категоризація слайдів	[7]
Genius Digital Diagnostics	Hologic	2024	Deep Learning + Vol. imaging	Скринінг/Діагностика	14 Z-площин, AI-галерея, ~50% менше часу	[6], [8], [9]
BestCyte	CellSolutions	CE (досл.)	Класичний CV + ML	Скринінг	Автоматичне сортування клітин	[18], [19]
CytoProcessor	DATEXIM	CE (досл.)	CNN	Діагностика	Хмарна платформа аналізу	[19]

*Відкриті датасети.* Розвиток алгоритмів комп'ютерного аналізу цитологічних зображень значною мірою залежить від наявності якісних анотованих наборів даних. Серед найбільш відомих відкритих датасетів слід виділити: SIPaKMeD (4049 клітин, 5 класів, Університет Патр), який містить зображення окремих клітин, вирізаних з мазків Папаніколау [26]; Herlev (917 клітин, 7 класів), класичний датасет з ручною анотацією ядра та цитоплазми [27]; ISBI 2014/2015, створені для змагань із сегментації перекриваючих цервікальних клітин, що включають як синтетичні, так і реальні зображення [10]; набір CRIC (400 зображень із різними діагностичними категоріями) [28]; та Cx22 (1320 зображень, із сегментацією ядер та цитоплазми) [29]. Суттєвими обмеженнями існуючих датасетів є їх відносно невеликий розмір, обмежена кількість діагностичних категорій, відсутність стандартизованого протоколу анотації та зміщення, пов'язане з конкретним обладнанням та лабораторією, де зразки були отримані.

Сегментація клітин та їх субклітинних структур (ядер, цитоплазми) є фундаментальним етапом автоматизованого цитологічного аналізу. Точність сегментації безпосередньо впливає на якість подальшої класифікації, оскільки морфометричні параметри клітин (площа ядра, площа цитоплазми, співвідношення ядро/цитоплазма, периметр, еліптичність) є ключовими діагностичними ознаками в цитопатології [5, 18].

*Класичні методи сегментації.* Порогова обробка методом Оцу (Otsu) забезпечує автоматичне визначення оптимального порогу бінаризації на основі максимізації міжкласової дисперсії. Цей метод ефективний при достатньому контрасті між об'єктом та фоном, однак значно втрачає точність при нерівномірному освітленні та низькому контрасті, характерному для цитологічних зображень. Метод вододілу (watershed) базується на топографічній інтерпретації градієнтного зображення і дозволяє розділяти дотичні об'єкти, проте схильний до надлишкової сегментації. Активні контури використовують ітеративну деформацію кривої під впливом внутрішніх та зовнішніх енергій, забезпечуючи плавні межі об'єктів, але потребують вдалої ініціалізації та чутливі до локальних мінімумів [10, 12, 13].

*Методи на основі глибокого навчання.* Архітектура U-Net, запропонована Ronneberger et al. у 2015 році, стала фактичним стандартом для семантичної сегментації медичних зображень

завдяки ефективному енкодер-декодерному дизайну з пропусковими з'єднаннями, що дозволяє зберігати просторову інформацію на різних масштабах. Варіанти U-Net (U-Net++, Attention U-Net, U<sup>2</sup>-Net) демонструють стабільну продуктивність при роботі з невеликими наборами даних, типовими для медичної візуалізації, досягаючи коефіцієнта Dice до 0,91–0,95 для сегментації ядер при достатній кількості навчальних зразків [12–14].

Mask R-CNN, розроблена на базі Faster R-CNN, є архітектурою інстансної сегментації, яка одночасно виконує детекцію, класифікацію та попиксельну сегментацію кожного об'єкта окремо. Дослідження Rettenberger et al. (2023) показало, що Mask R-CNN перевершує U-Net саме при сегментації перекриваючих клітин, що робить її особливо перспективною для цитологічних застосувань [15]. Спеціалізовані архітектури, такі як HoVer-Net (визначення меж через передбачення горизонтальних та вертикальних карт відстаней), StarDist (апроксимація об'єктів зірчастопуклими полігонами) та Cellpose (градієнтне поле потоку для визначення належності пікселів), розроблені саме для задач клітинної сегментації та демонструють конкурентні результати на різноманітних типах мікроскопічних зображень [12, 16].

*Проблема перекриття (злипання) клітин.* Перекриття клітин є однією з найбільш складних задач у цитологічній сегментації, особливо актуальною для конвенційних мазків. За результатами змагань ISBI 2014/2015, навіть найкращі методи того часу демонстрували задовільну точність лише для кластерів з не більше ніж трьома клітинами та коефіцієнтом перекриття до 0,3 [10]. Сучасні підходи до розв'язання цієї задачі включають: використання інстансної сегментації (Mask R-CNN, що досягає DSC 0,92 на даних ISBI); мережі IRNet (Instance Relation Network), яка моделює відношення між перекриваючими екземплярами [17]; методи на основі shape priors, що використовують апріорні знання про форму клітин [18]; а також комбіновані підходи, що поєднують класичну обробку (маркерний watershed) з нейромережевою детекцією [12]. Порівняння класичних і сучасних методів сегментації цитологічних зображень наведено в табл. 2.

Таблиця 2 – Порівняння методів сегментації цитологічних зображень

Метод	Тип	Dice/DSC	IoU	Задача	Обмеження	Джерело
Otsu thresholding	Класичний	0,70–0,80	0,55–0,68	Ядра	Чутливість до освітлення, не розділяє кластери	[10, 13]
Watershed	Класичний	0,75–0,85	0,60–0,74	Ядра, клітини	Надлишкова сегментація	[10, 12]
Active contours	Класичний	0,78–0,87	0,65–0,77	Цитоплазма	Потребує ініціалізації, повільний	[10, 12]
U-Net	DL (семант.)	0,91–0,95	0,84–0,91	Ядра, клітини	Не розділяє перекриваючі екземпляри	[12–14]
U-Net++	DL (семант.)	0,92–0,96	0,85–0,92	Ядра	Аналогічно U-Net щодо інстансів	[12–14]
Mask R-CNN	DL (інстансна)	0,88–0,92	0,79–0,86	Клітини, ядра	Потребує більше даних, повільніший	[15]
Cellpose	DL (інстансна)	0,85–0,91	0,75–0,84	Клітини (різні типи)	Узагальнена модель, не спеціалізована	[16]
StarDist	DL (інстансна)	0,83–0,90	0,72–0,82	Ядра (зірчasto-опуклі)	Обмеження на форму об'єктів	[12]
HoVer-Net	DL (інстансна)	0,86–0,93	0,77–0,87	Ядра + класифікація	Складна архітектура, ресурсоємний	[12]

*Комп'ютерна морфометрія.* Після сегментації клітин виконується кількісне визначення їх геометричних параметрів, що є основою для подальшої класифікації. Основними морфометричними ознаками є: площа ядра та цитоплазми; периметр ядра; співвідношення

ядро/цитоплазма (N/C ratio), яке є одним із ключових критеріїв малігнізації; еліптичність (відхилення форми ядра від кола); текстурні характеристики хроматину (ентропія, однорідність за Haralick); а також інтенсивність та розподіл фарбування [5, 18]. Точність визначення цих параметрів безпосередньо залежить від якості попередньої сегментації, що підкреслює критичну роль цього етапу в загальному аналізі.

*Класичні методи машинного навчання.* До появи глибокого навчання класифікація цитологічних зображень базувалася на попередньому виділенні ручних ознак з подальшою класифікацією за допомогою алгоритмів машинного навчання. Серед найбільш ефективних класифікаторів цього покоління – метод опорних векторів (SVM), який демонстрував точність 85–95% на різних датасетах завдяки здатності працювати у високорозмірних просторах ознак; випадковий ліс (Random Forest), стійкий до перенавчання та здатний оцінювати важливість ознак; а також метод k-найближчих сусідів (k-NN), простий в реалізації, але чутливий до вибору метрики та розмірності [18, 19]. Типові набори ознак включали текстурні дескриптори (матриця суміжності Haralick, Local Binary Patterns), геометричні параметри клітин та ядер, а також колірні гістограми у просторах RGB та HSV.

*Глибоке навчання для класифікації.* Згорткові нейронні мережі (CNN) суттєво змінили підхід до класифікації цитологічних зображень, усунувши необхідність ручного виділення ознак. Глибокі архітектури автоматично навчаються ієрархічним представленням зображень – від низькорівневих (краї, текстури) до високорівневих (морфологічні патерни клітин) [18, 19]. Серед найбільш ефективних архітектур для класифікації цитологічних зображень:

ResNet (Residual Network) – використовує залишкові з'єднання (skip connections), що дозволяє навчати дуже глибокі мережі (50–152 шари) без деградації градієнтів. На датасеті SIPaKMeD ResNet-50 досягає точності 95–98% при класифікації 5 класів цервікальних клітин [18, 19, 26].

VGG-16/VGG-19 – класичні глибокі мережі з послідовною архітектурою, що часто використовуються як попередньо навчені екстрактори ознак (transfer learning) для цитологічних задач [18].

EfficientNet – сімейство мереж, оптимізованих за балансом глибини, ширини та роздільної здатності, що демонструє конкурентну точність при значно меншій кількості параметрів [19].

Vision Transformers (ViT) – архітектура, що базується на механізмі самоуваги (self-attention) і дозволяє моделювати глобальні залежності в зображенні. Однак ViT потребує значно більшого обсягу навчальних даних порівняно з CNN і, за результатами досліджень на цитологічних даних, демонструє нижчу точність при обмежених наборах (Dice 0,58 проти 0,67 для U-Net на датасеті HRASPP) [13].

*Класифікація на рівні клітини та на рівні слайда.* Задачу класифікації цитологічних зображень можна розглядати на двох рівнях. Класифікація на рівні клітини (cell-level) передбачає попередню сегментацію та класифікацію кожної клітини окремо. Класифікація на рівні слайда (slide-level або whole-slide image classification) присвоює діагностичну категорію всьому препарату, що більш наближено до реального клінічного робочого процесу [18, 19]. Сучасні end-to-end підходи, зокрема на базі множинного навчання (Multiple Instance Learning, MIL), дозволяють навчати моделі slide-level класифікації без детальної поклітинної анотації, використовуючи лише слабкі мітки на рівні слайда [21]. Узагальнене порівняння методів класифікації цитологічних зображень наведено в табл. 3.

*Ключові проблеми класифікації.* Незважаючи на високі показники точності, продемонстровані на стандартних датасетах, впровадження алгоритмів класифікації в реальну клінічну практику стикається з низкою серйозних проблем.

**Таблиця 3 – Порівняння методів класифікації цитологічних зображень**

Метод/Архітектура	Датасет	Патологія	Accuracy	Sensitivity	AUC	Рік	Джерело
SVM + Haralick	Herlev	Церв. цитологія	92,4%	89,1%	0,94	2019	[18]
Random Forest + LBP	SIPaKMeD	Церв. цитологія	90,5%	87,3%	0,93	2020	[18]
ResNet-50	SIPaKMeD	Церв. цитологія	97,5%	96,8%	0,99	2021	[18, 19]
VGG-16 (transfer)	Herlev	Церв. цитологія	94,8%	92,1%	0,97	2020	[18]
EfficientNet-B3	SIPaKMeD	Церв. цитологія	98,1%	97,4%	0,99	2022	[18, 19]
DeepCervix (гібрид)	SIPaKMeD + Herlev	Церв. цитологія	99,0%	98,7%	0,99	2021	[19]
ViT (Vision Transf.)	SIPaKMeD	Церв. цитологія	95,2%	93,8%	0,98	2023	[13]
Cytologist-in-Loop AI	Клінічний (8000 WSI)	Церв. скринінг	–	81,6%	–	2024	[23]

По-перше, відкриті датасети є відносно невеликими та незбалансованими: патологічні класи зазвичай представлені значно меншою кількістю зразків порівняно з нормальними, що призводить до зміщення класифікатора [18, 19]. По-друге, відсутність стандартизованої розмітки – різні експерти можуть по-різному класифікувати одну й ту ж клітину, а міжекспертна узгодженість при оцінці помірних дисплазій (LSIL/ASC-US) є особливо низькою [5]. По-третє, доменний зсув (domain shift) між лабораторіями, зумовлений відмінностями в обладнанні, протоколах фарбування та підготовки препаратів, призводить до значного зниження точності моделей при застосуванні до даних з нових джерел. Дослідження показують, що метрики на приватних (інституціональних) датасетах часто суттєво нижчі, ніж на публічних: наприклад, DSC найкращої моделі може суттєво знижуватися при застосуванні до даних з нових джерел [18, 22].

## ВИСНОВКИ

Аналіз сучасних методів і біотехнічних систем для діагностики онкопатологій за цитологічними зображеннями свідчить про значний прогрес у цій галузі, особливо завдяки впровадженню технологій глибокого навчання. Комерційні CAD-системи нового покоління, зокрема Genius Digital Diagnostics System, демонструють клінічно значущу ефективність і поступово впроваджуються у лабораторну практику. Архітектури глибокого навчання (U-Net та її модифікації, Mask R-CNN, спеціалізовані мережі для сегментації клітин) значно перевершують класичні методи обробки зображень за точністю та стійкістю.

Водночас аналіз виявив низку суттєвих невирішених проблем, що обмежують впровадження автоматизованих систем у клінічну практику. Задача інстансної сегментації перекриваючих клітин залишається складною: сучасні методи забезпечують прийнятну точність лише при помірному перекритті та невеликих кластерах, тоді як реальні препарати часто мають складніші конфігурації. Доменний зсув між лабораторіями є серйозним бар'єром для впровадження, оскільки моделі, навчені на даних однієї установи, значно втрачають точність на зразках з інших джерел. Також розвиток стримують обмеженість і незбалансованість відкритих датасетів, відсутність стандартизованого протоколу анотації та дефіцит анотованих зразків для рідкісних патологій.

Особливо актуальною є проблема автоматизованого аналізу для окремих нозологій, зокрема гінекологічної цитології. Потреба в ефективному скринінгу раку шийки матки є критичною для України через високу частку пізніх стадій діагностики та обмежене охоплення скринінговими програмами. Розробка біотехнічної системи, що поєднує модифікований метод інстансної сегментації клітин, автоматичну класифікацію за морфологічними ознаками та стійкість до доменного зсуву, є актуальним науковим і практичним завданням біомедицинської інженерії.

Отже, подальші дослідження слід зосередити на розробці удосконаленого методу інстансної сегментації для ефективного розділення кластерів перекриваючих клітин, створенні класифікатора для конкретних нозологій на реальних клінічних зразках, а також розробці інтегрованої біотехнічної системи візуалізації та аналізу для цитологічних лабораторій.

### Список використаних джерел

1. Bray F., Laversanne M., Sung H. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2024. Vol. 74, no. 3. P. 229-263. DOI: <https://doi.org/10.3322/caac.21834>.
2. Filho A. M., Laversanne M., Ferlay J. et al. The GLOBOCAN 2022 cancer estimates: Data sources, methods, and a snapshot of the cancer burden worldwide. *International Journal of Cancer*. 2025. Vol. 156, no. 7. P. 1336-1346. DOI: <https://doi.org/10.1002/ijc.35278>.
3. Ryzhov A., Bray F., Ferlay J. et al. Recent cancer incidence trends in Ukraine and short-term predictions to 2022. *Cancer Epidemiology*. 2020. Vol. 65. Article 101663. DOI: <https://doi.org/10.1016/j.canep.2019.101663>.
4. National Cancer Registry of Ukraine. The oncoepidemiological situation and the state of cancer care organization in Ukraine in 2022-2023. *Cancer in Ukraine, 2022-2023: Bulletin of National Cancer Registry of Ukraine*. Kyiv, 2023. Vol. 25. P. 7-29.
5. Nayar R., Wilbur D. C., eds. *The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes*. 3rd ed. Cham: Springer, 2015. DOI: <https://doi.org/10.1007/978-3-319-11074-5>.
6. Murphy K. M., Weatherhead K., Chenault C. et al. Impact of the Genius Digital Diagnostics System on workflow and accuracy compared with the ThinPrep Imaging System for review of ThinPrep Papanicolaou tests. *American Journal of Clinical Pathology*. 2025. Vol. 164, no. 5. P. 746-751. DOI: <https://doi.org/10.1093/ajcp/aqaf099>.
7. Thrall M. J. Automated screening of Papanicolaou tests: a review of the literature. *Diagnostic Cytopathology*. 2019. Vol. 47, no. 1. P. 20-27. DOI: <https://doi.org/10.1002/dc.23931>.
8. Ikenberg H., Lieder S., Ahr A. et al. Comparison of the Hologic Genius Digital Diagnostics System with the ThinPrep Imaging System - a retrospective assessment. *Cancer Cytopathology*. 2023. Vol. 131, no. 7. P. 424-432. DOI: <https://doi.org/10.1002/cncy.22695>.
9. Cantley R. L. et al. Validation of AI-assisted ThinPrep Pap test screening using the Genius Digital Diagnostics System. *Journal of Pathology Informatics*. 2024. Vol. 15. Article 100391. DOI: <https://doi.org/10.1016/j.jpi.2024.100391>.
10. Lu Z., Carneiro G., Bradley A. P. et al. Evaluation of three algorithms for the segmentation of overlapping cervical cells. *IEEE Journal of Biomedical and Health Informatics*. 2017. Vol. 21, no. 2. P. 441-450. DOI: <https://doi.org/10.1109/JBHI.2016.2519686>.
11. Wan T., Xu S., Sang C. et al. Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. *Neurocomputing*. 2019. Vol. 365. P. 157-170. DOI: <https://doi.org/10.1016/j.neucom.2019.06.086>.

12. Krikid F., Rositi H., Vacavant A. State-of-the-art deep learning methods for microscopic image segmentation: applications to cells, nuclei, and tissues. *Journal of Imaging*. 2024. Vol. 10, no. 12. Art. 311. DOI: <https://doi.org/10.3390/jimaging10120311>.
13. Rasheed A. et al. Techniques and challenges for nuclei segmentation in cervical smear images: a review. *Artificial Intelligence Review*. 2025. Vol. 58. Art. 299. DOI: <https://doi.org/10.1007/s10462-025-11207-9>.
14. Ji J., Zhang W., Dong Y. et al. Automated cervical cell segmentation using deep ensemble learning. *BMC Medical Imaging*. 2023. Vol. 23. Art. 137. DOI: <https://doi.org/10.1186/s12880-023-01096-1>.
15. Rettenberger L., Münke F. R., Bruch R., Reischl M. Mask R-CNN outperforms U-Net in instance segmentation for overlapping cells. *Current Directions in Biomedical Engineering*. 2023. Vol. 9, no. 1. P. 335-338. DOI: <https://doi.org/10.1515/cdbme-2023-1084>.
16. Stringer C. et al. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*. 2021. Vol. 18. P. 100-106. DOI: <https://doi.org/10.1038/s41592-020-01018-x>.
17. Zhou Y. et al. IRNet: Instance Relation Network for Overlapping Cervical Cell Segmentation. In: *Medical Image Computing and Computer Assisted Intervention - MICCAI 2019*. 2019. P. 640-648. DOI: [https://doi.org/10.1007/978-3-030-32239-7\\_71](https://doi.org/10.1007/978-3-030-32239-7_71).
18. Jiang P., Li X., Shen H. A systematic review of deep learning-based cervical cytology screening: from cell identification to whole slide image analysis. *Artificial Intelligence Review*. 2023. Vol. 56. P. 2687-2758. DOI: <https://doi.org/10.1007/s10462-023-10588-z>.
19. Fang M., Liao B., Lei X. et al. A systematic review on deep learning based methods for cervical cell image analysis. *Neurocomputing*. 2024. Vol. 610. Article 128630. DOI: <https://doi.org/10.1016/j.neucom.2024.128630>.
20. Dosovitskiy A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. In: *International Conference on Learning Representations (ICLR)*. 2021.
21. Campanella G. et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*. 2019. Vol. 25. P. 1301-1309. DOI: <https://doi.org/10.1038/s41591-019-0508-1>.
22. Mosquera-Zamudio A., Cancino S., Cárdenas-Montoya G. et al. Deep-learning approaches for cervical cytology nuclei segmentation in whole slide images. *Journal of Imaging*. 2025. Vol. 11, no. 5. Art. 137. DOI: <https://doi.org/10.3390/jimaging11050137>.
23. Wang J. et al. Artificial intelligence enables precision diagnosis of cervical cytology grades and cervical cancer. *Nature Communications*. 2024. Vol. 15. Art. 4369. DOI: <https://doi.org/10.1038/s41467-024-48705-3>.
24. Volodko N., Chopyak V., Mazur Y. Barriers to implementing cervical cancer screening in Ukraine: the path forward. *Proceeding of the Shevchenko Scientific Society. Medical Sciences*. 2025. Vol. 77, no. 1. DOI: <https://doi.org/10.25040/ntsh2025.01.01>.
25. Murphy D. B., Davidson M. W. *Fundamentals of Light Microscopy and Electronic Imaging*. 2nd ed. Hoboken, NJ: Wiley-Blackwell, 2013. DOI: <https://doi.org/10.1002/9781118382905>.
26. Plissiti M. E. et al. SIPaKMeD: A new dataset for feature and image based classification of normal and pathological cervical cells in Pap smear images. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*. 2018. P. 3144-3148. DOI: <https://doi.org/10.1109/ICIP.2018.8451588>.
27. Jantzen J., Norup J., Dounias G., Bjerregaard B. Pap-smear Benchmark Data for Pattern Classification. In: *Proc. NiSIS 2005: Nature Inspired Smart Information Systems (NiSIS), EU Co-ordination Action*. Albufeira, Portugal: NiSIS, 2005. P. 1-9.

28. Rezende M. T. et al. CRIC searchable image database as a public platform for conventional Pap smear cytology data. *Scientific Data*. 2021. Vol. 8, no. 1. Art. 151. DOI: <https://doi.org/10.1038/s41597-021-00933-8>.

29. Liu G. et al. Cx22: A new publicly available dataset for deep learning-based segmentation of cervical cytology images. *Computers in Biology and Medicine*. 2022. Vol. 150. Article 106194. DOI: <https://doi.org/10.1016/j.compbiomed.2022.106194>.

Leshchenko V. O., Zabolotna N. I.

## ANALYSIS OF MODERN METHODS AND BIOTECHNICAL SYSTEMS FOR DIAGNOSING ONCOPATHOLOGIES USING CYTOLOGICAL IMAGES

*The article presents a comprehensive analysis of modern methods and biotechnical systems for automated diagnostics of oncopathologies based on cytological images. The relevance of the study is driven by the growing global cancer burden – according to GLOBOCAN 2022 estimates, approximately 20 million new cancer cases are registered annually, while in Ukraine over 38% of cases are diagnosed at advanced stages, highlighting the urgent need for effective early screening tools. The aim of this work is to systematically analyze the current state of computer-aided cytological image analysis methods, identify key unresolved challenges, and substantiate directions for further research in biomedical engineering. The medical and technical aspects of cytological image acquisition are examined, including microscopy types, specimen preparation techniques, and characteristic artifacts. Commercial CAD systems for cervical cytology are analyzed, namely ThinPrep Imaging System, BD FocalPoint GS, and Genius Digital Diagnostics System, along with major publicly available datasets (SIPaKMeD, Herlev, ISBI 2014/2015, CRIC, Cx22). A comparative analysis of classical cell segmentation methods (Otsu thresholding, watershed transform, active contours) and deep learning architectures (U-Net, Mask R-CNN, Cellpose, StarDist, HoVer-Net) is performed, demonstrating a substantial advantage of neural network approaches in terms of accuracy and robustness. Classification methods are investigated, ranging from classical machine learning algorithms (SVM, Random Forest) to modern deep neural networks (ResNet, EfficientNet, Vision Transformers). Three key unresolved problems are identified: insufficient accuracy of instance segmentation for overlapping cells in real clinical specimens, significant domain shift between laboratories leading to substantial degradation of trained model performance, and the limited size and class imbalance of available open datasets. The direction for further research is substantiated, aimed at developing a biotechnical system that would combine a modified instance segmentation method with automatic morphological feature-based classification and ensure robustness to domain shift for the purposes of early oncopathology diagnostics.*

**Keywords:** cytological images, computer-aided diagnosis, cell segmentation, neural networks, CAD systems, oncopathology, biotechnical system.

Стаття надійшла 09.04.2026

Стаття прийнята 11.04.2026р.

Стаття опублікована 30.04.2026р.