

2. МЕТОДИ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ РІЗНОЇ ПРИРОДИ ДЛЯ ВИПАДКУ НЕЗБАЛАНСОВАНОГО НАБОРУ

Антон Охріменко, аспірант
Кафедра математичного моделювання і аналізу даних
Навчально-науковий Фізико-технічний інститут
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

antoh-ipt21@iit.kpi.ua

ВСТУП

Незбалансованість навчальних вибірок — це фундаментальна проблема при навчанні будь-якої моделі машинного навчання. Окрім очевидної схильності модель добре вивчати лише ті дані, які найбільш представлені у наборі даних, виникає проблема оцінки якості її роботи на тестовому наборі даних, який так само може бути незбалансованим. Іншою перешкодою є перекриття класів — подібність екземплярів різних класів між собою аж до неможливості розрізнення. У випадку, коли ці дві проблеми наявні одночасно, створення ефективної моделі машинного навчання значно ускладнене.

Розширення набору даних за рахунок найменш представлених класів зведе до нуля проблему незбалансованості класів. В свою чергу, збільшення кількості ознак, які характеризують кожний екземпляр даних, може вирішити проблему перекриття класів. Для задачі сегментації зображень це еквівалентно збільшенню кількості каналів у зображенні. У будь-якому випадку, виникає необхідність отримати нові дані для розширення вже наявного набору. Це не завжди можливо, через фізичну неможливість, через обмеження бюджету або ж з інших причин. Наприклад, для задачі сегментації супутникових знімків, неможливо отримати додаткові зображення через фізичні обмеження руху супутників. Для задачі сегментації гістологічних мікроскопічних знімків, отримання нових даних потребує виконання складних операцій з отримання культури клітин, їх фарбування та дослідження за допомогою мікроскопу.

У наступних підрозділах розглядаються методи компенсації дисбалансу класів у задачі сегментації без використання додаткових навчальних даних.

2.1. ВДОСКОНАЛЕННЯ МЕТОДУ ГЛИБИННОГО НАВЧАННЯ ДЛЯ КЛАСИФІКАЦІЇ ЗЕМНОГО ПОКРИВУ ТА ВИЯВЛЕННЯ ТИПУ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР

Карти класифікації земного покриву, в тому числі типів сільськогосподарських культур, мають важливе значення для сільського господарства, моніторингу довкілля та розвитку сільських громад. Точна класифікація сприяє розв'язанню проблем виробництва сільськогосподарських культур та оцінці урожайності [1], оцінці ризиків та аналізу умов навколишнього середовища [2]. Ці карти є необхідними для оцінки земельного покриття та землекористування, дозволяючи вивчати соціальні та економічні аспекти розвитку сільських територій. Вони також допомагають оцінити важливі екологічні та сільськогосподарські показники для оцінки ступеню опустелювання, стійкого використання землі та досягнення цілей сталого розвитку [2]. Крім того, інвентаризація сільськогосподарських культур за допомогою класифікації супутникових даних стала необхідною для прийняття рішень на рівні країни в багатьох регіонах світу. Послуги, такі як CropScare Сполучених Штатів [3] і система Sen-4-CAP Європейського Союзу [4], підтримують сільськогосподарську політику в Північній Америці і Європі.

2.1.1. ОГЛЯД ЛІТЕРАТУРИ

Стрімкий ріст індустрії та доступності даних спостереження Землі призвів до нових досліджень щодо можливостей впровадження штучного інтелекту для спостереження за земною поверхнею. Сучасні методи глибокого навчання пропонують потенціал покращення якості супутникових продуктів для точної оцінки сільського господарства та економічних аспектів розвитку сільських територій [5]. Глибоке навчання, зокрема згорткові нейронні мережі, стало перспективним підходом до технологій картографування відповідно до мінливих потреб у класифікації [6]. В той же час глибокі моделі навчання стикаються з двома критичними викликами, що ускладнюють масштабованість, універсальність та застосування в реальному світі: необхідність великої кількості розмічених супутникових даних і наявність

2.2. Методи сегментації зображень різної природи для випадку...

дисбалансу даних через розподіл покриття земної поверхні та типів сільськогосподарських культур. Це призводить до існування менш представлених і більш представлених класів. Крім того, складність взаємозв'язків між багатоспектральними та текстурними характеристиками обмежує можливості методів аугментації.

Зазначені вище питання не можуть бути вирішені за допомогою традиційних технік аугментації комп'ютерного зору, таких як обертання, випадкове видалення та змішування [7]. Замість цього, найбільш ефективним рішенням є використання передових технологій, а саме генеративних змагальних нейронних мереж (GANs) [8]. Незважаючи на популярність GANs в різних наукових галузях, їх застосування для обробки даних дистанційного спостереження земної поверхні залишається недостатньо розвиненим і в основному обмежується обробкою зображень і задачами супер-роздільності. Дослідження [9] представляє GAN-підходи для підвищення просторового розрізнення та масштабування супутникових зображень і підтверджує високу якість аналізу текстурних характеристик реального світу такими GAN. В [10] показано, що модель GAN може генерувати реалістичні супутникові зображення за текстурними та багатоспектральними характеристиками. У дослідженні [11] показано, що моделі GAN здатні вивчати закономірності часових змін багатоспектральних характеристик супутникових даних для прогнозу наступного зображення в послідовності і, таким чином, заповнювати прогалини в часових рядах. Схожий висновок був зроблений для складних багатоспектральних характеристик різних типів сільськогосподарських культур у дослідженні [12]. Крім того, такі методи, як «pixel-to-pixel» (pix2pix), вже дозволяють створювати реалістичні супутникові зображення за допомогою існуючих карт покриття земної поверхні [13]. З врахуванням потенціалу використання моделей GAN нижче запропоновано новий метод застосування GAN для вирішення питань дисбалансу даних у класифікації земної поверхні та типів сільськогосподарських культур. Цей підхід можна розглядати як першу методологію аугментації супутникових зображень на основі GAN, використовуючи підхід pix2pix [7], тим самим вперше вводячи цей метод аугментації в галузь використання даних дистанційного зондування Землі.

2.1.2. ПОСТАНОВКА ЗАДАЧІ

При виконанні досліджень в якості пілотної було обрано територію Київської області, яка розташована в центральній-північній частині України та має площу 28 131 км². Область розташована на межі двох природних зон — полісся та лісостепу. Частка землі сільськогосподарського призначення в Київській області становить 44% від загальної площі, основна частина якої розташована в південній частині області (рис. 1).

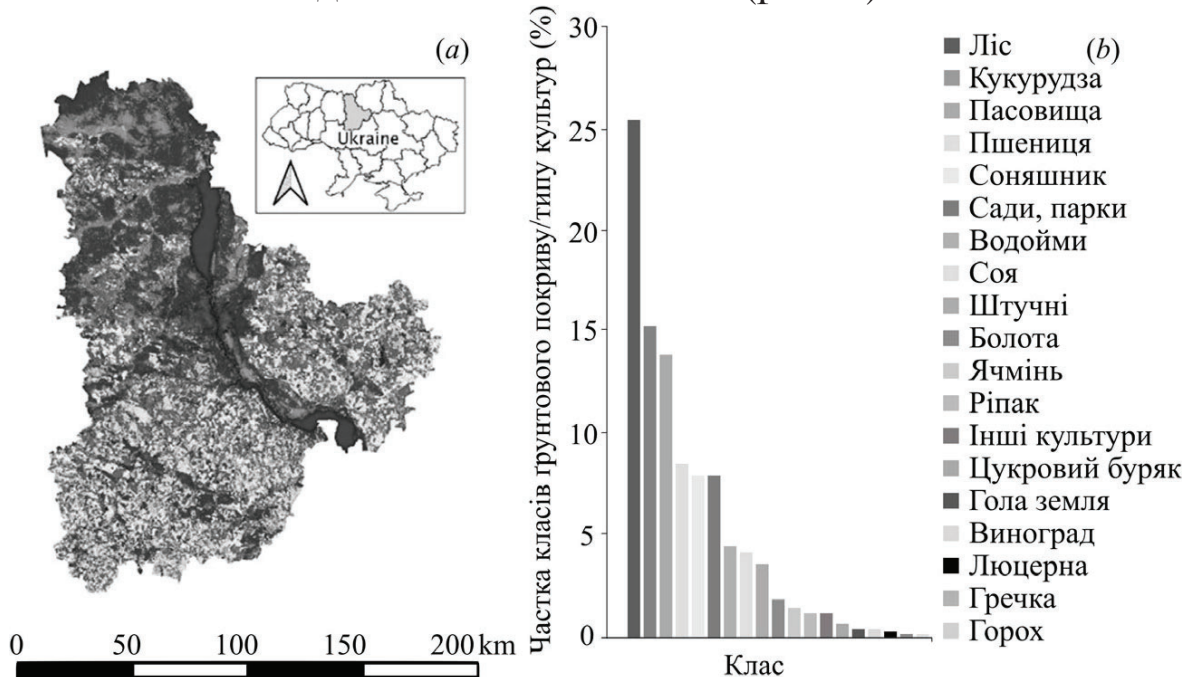


Рис. 1. Ґрунтовий покриття та розподіл культур для території досліджень; а) Карта посівів Київської області 2021; б) Частка рослинного покриття та розподіл культур, представлених у Київській області

При проведенні експерименту була використана карта класифікації на основі супутникових знімків Sentinel-1 і Sentinel-2, надана Інститутом космічних досліджень НАНУ-ДКАУ, отримана в межах проекту Всесвітнього банку по інвентаризації сільськогосподарських земель в Україні з точністю 95% для 2021 р. [14].

Супутникові дані, які було використано для сегментації сільськогосподарських культур, були отримані за допомогою супутника Sentinel-2 за липень у вигляді безхмарного композиту зі спектральними каналами червоного, зеленого, синього та ближнього інфрачервоного діапазонів. Композит мав просторове розрізнення 10 м із рівнем обробки L2, включаючи радіометричну

2.2. Методи сегментації зображень різної природи для випадку...

корекцію, атмосферну корекцію, маскування хмар та мозаїку. Далі композит був розділений на 4250 зображень розміром 256×256 пікселів і 4 канали, та розподілений на незалежні набори даних для навчання і тестування з 2125 зображень у кожному. З точки зору представленості пікселів малих класів набір даних, який використовувався для навчання як генератора GAN, так і моделі класифікації, для навчання є сильно незбалансованим.

Ефективна архітектура GAN дозволяє генерувати реалістичні синтетичні супутникові зображення для аугментації навчальних даних. Взаємозв'язок між супутниковими зображеннями та сегментаційними масками ґрунтується на використанні текстурних та багатоспектральних ознак в межах зображень. Було розроблено модель, яка здатна генерувати реалістичні супутникові зображення для будь-якої штучно створеної сегментаційної маски. Запропонований метод [15] дозволяє створювати синтетичні пари супутникових зображень і масок, генерувати розподіли сільськогосподарських культур, які не зустрічаються у природному середовищі, та надає можливість керування балансом класів в наборі даних [16]. У даному дослідженні для створення синтетичних супутникових зображень з 4 каналами було використано сегментаційні маски розміром 256×256 пікселів.

Для побудови моделі генератора використано трансляцію зображень в зображення [17], яка спрямована на перетворення зображень з одного домена в інший із відмінною розмірністю простору ознак. Результат аугментаційного алгоритму показано на рис. 2.

Спочатку модель $pix-2-pix$ GAN була налаштована на генерацію реалістичних супутникових зображень на основі сегментаційної карти. Потім реальні сегментаційні карти були модифіковані для отримання штучних масок з більшою часткою менш представлених класів. Після цього початкові і згенеровані супутникові дані були об'єднані в єдину навчальну колекцію даних, яка використовувалася для навчання моделі сегментації з вищою точністю розділення менш представлених класів.

2.1.3. РОЗВ'ЯЗАННЯ ЗАДАЧІ

В якості базової моделі для задачі семантичної сегментації була використана класична модель U-Net із вхідним розміром $256 \times 256 \times 4$ пікселів. U-Net — це широко використовувана архітектура глибокого навчання при розв'язанні задач моніторингу земель [18]. Її ключовою властивістю є включення додаткових

з'єднань від шарів, які зменшують розмір простору ознак до відповідних шарів, які розширюють його.

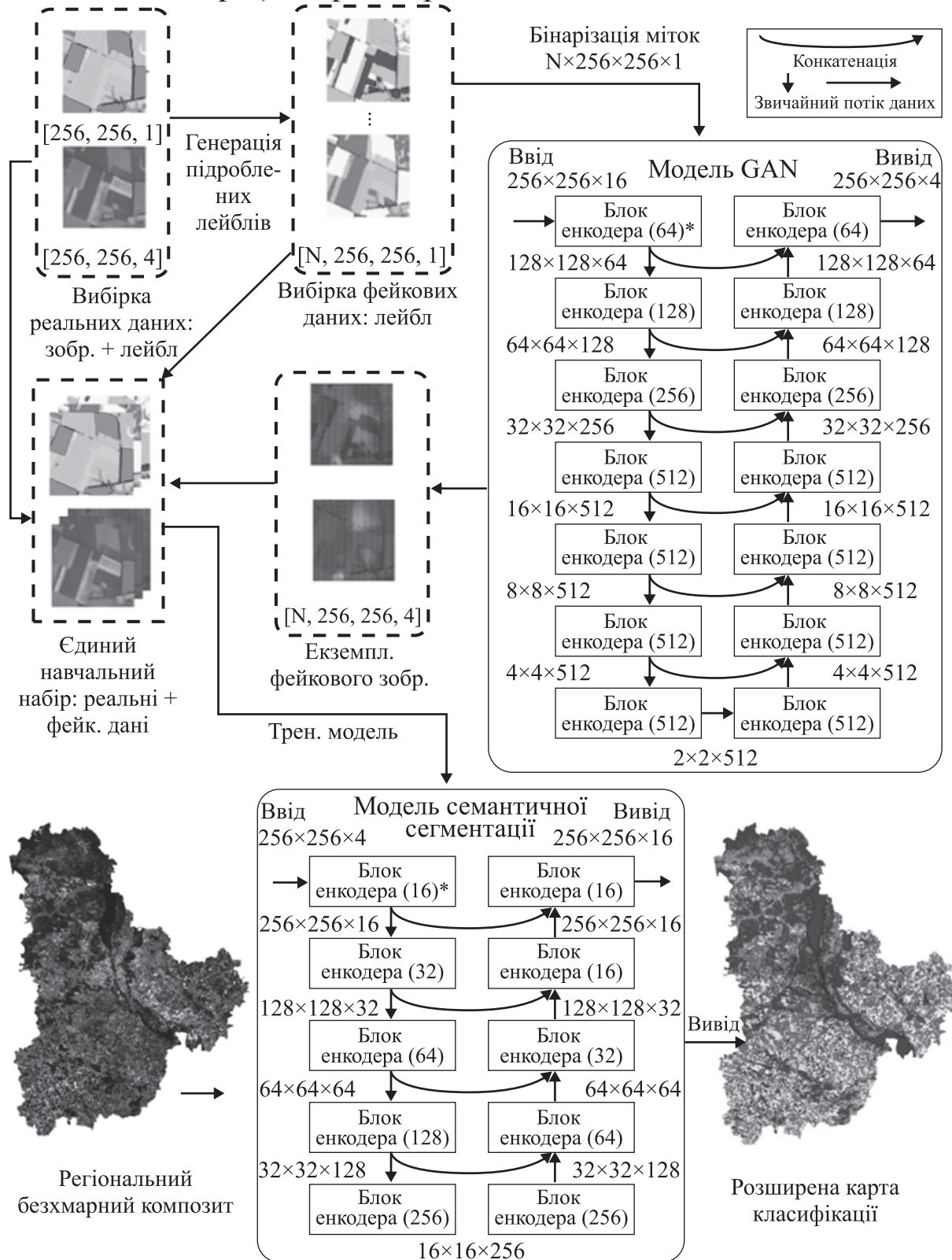


Рис. 2. Пропонована схема аугментації GAN для сегментування типів сільськогосподарських культур з використанням моделі сегментації глибокого навчання.

2.2. Методи сегментації зображень різної природи для випадку...

Така архітектура дозволяє виділяти високорівневі ознаки на кожному кроці, зберігаючи при цьому просторову інформацію в зображенні. Крім того, це допомагає вирішувати проблему зникнення градієнтів при певних умовах.

У класичній моделі U-Net [18] кожний блок енкодера або декодера містить два 3×3 згорткових шари, за якими застосовується функція активації ReLU. Водночас кількість каналів подвоюється на кожному блоку енкодера (крім першого, де є 64 канали). На кожному блоку декодера вона зменшується вдвічі. Max-pooling із кроком 2 в блоках енкодера використовується для зменшення розміру карт ознак, тоді як в декодері для масштабування використовуються зворотні згортки. Вихідний шар мережі представляє собою згортку 1×1 з сигмоїдною функцією активації.

Для навчання моделі сегментації U-Net для класифікації сільськогосподарських культур був використаний алгоритм оптимізації Adam із швидкістю навчання 2×10^{-4} та функція втрати focal loss з параметром гамма, який дорівнював 2. Було обрано саме цю функцію втрати, оскільки вона є дуже популярною і, у більшості випадків, найбільш ефективною функцією втрат при розв'язанні задач з використанням незбалансованих наборів даних.

2.1.3.1. ГЕНЕРАТИВНА ЗМАГАЛЬНА МЕРЕЖА

Генеративна змагальна мережа (GAN) для генерації супутникових даних складається з двох моделей нейронних мереж: генератора для побудови зображення на основі сегментаційних масок, і дискримінатора, який дозволяє відрізнити реальні та штучно згенеровані зображення (рис. 3).

Генератор використовується для створення реалістичних фрагментів супутникових зображень, які не може розрізнити дискримінатор. В свою чергу, дискримінатор навчається розрізняти реальні та штучні зображення. Обидві моделі працюють з вхідними даними розміром $256 \times 256 \times 4$ пікселів.

Для побудови генератора використано модель pix-2-pix на основі U-Net з 7 блоками енкодера і декодера, в яких присутні 6 обхідних з'єднань [19] та гіперболічний тангенс в якості активаційної функції на вихідному шарі. Для дискримінатора використано глибоку згорткову нейронну мережу з 5 шарами 2D-згортки, включаючи нормалізацію пакету та активаційні функції Leaky ReLU, а також вихідний згортковий шар із

сигмоїдною активаційною функцією. Функцію втрат дискримінатора оптимізовано за допомогою бінарної крос-ентропії:

$$BCE = \frac{1}{n} \sum_{i=1}^n y_i \times \log_2 \hat{y}_i + (1 - y_i) \times \log_2 (1 - \hat{y}_i), \quad (1)$$

де n — кількість вибірок у навчальному датасеті, y_i — мітка екземпляру i та \hat{y}_i — вихід нейронної мережі для цієї вибірки. Функція втрати генератора комбінує бінарну крос-ентропію дискримінатора та середньоквадратичну помилку між створеними зображеннями та реальними зображеннями за формулою:

$$Loss_{gen} = MSE(y_i, \hat{y}_i) + 0.01 \times BCE(y_i, \hat{y}_i). \quad (2)$$

Для оптимізації параметрів під час навчання використано алгоритм Adam з швидкістю навчання 2×10^{-4} та коефіцієнтом $\beta=0.5$.

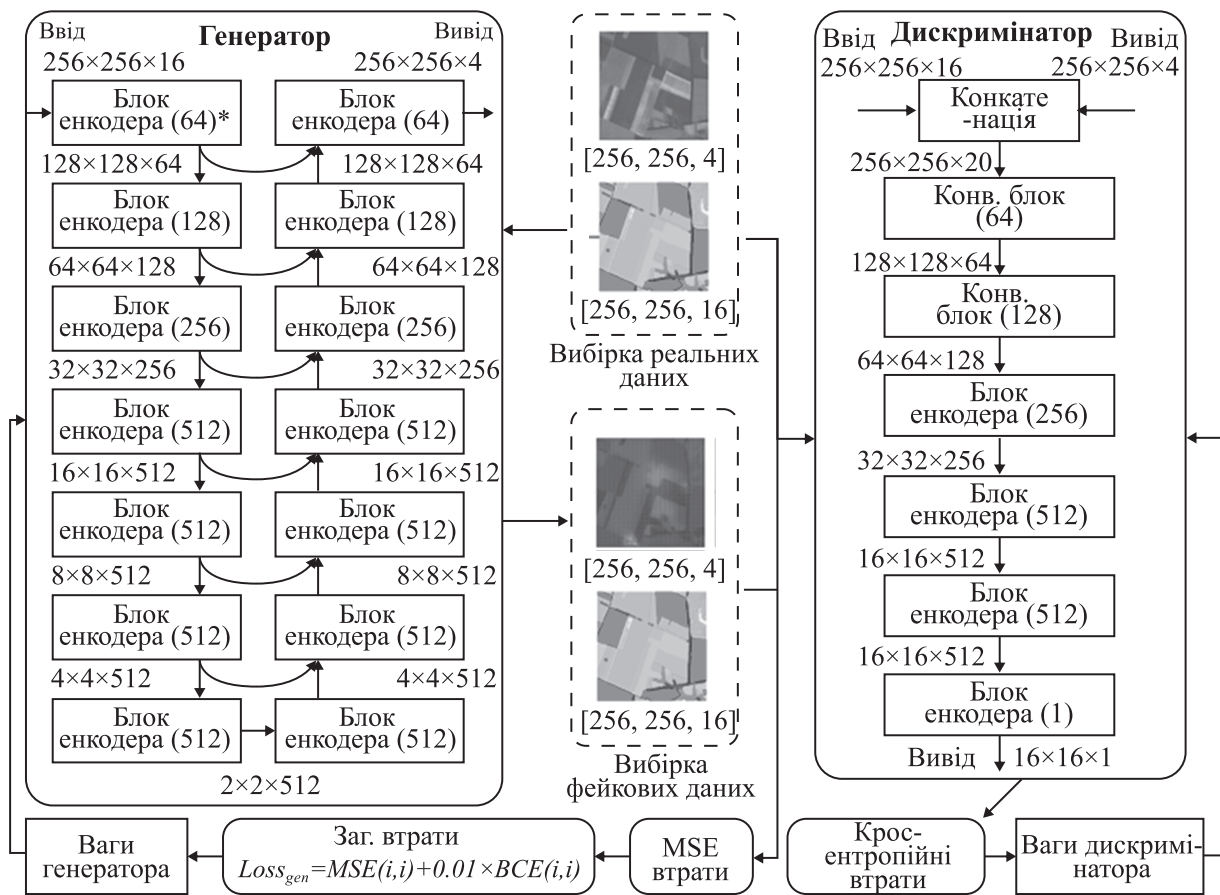


Рис. 3. Схема навчання моделі GAN

2.1.3.2. СХЕМА АУГМЕНТАЦІЇ

Для навчання генерації зображень за допомогою GAN шляхом модифікації початкових масок із навчального набору даних були створені штучні сегментаційні маски. Цей підхід дозволяє зберегти просторові шаблони сільськогосподарських полів, подібних до реального середовища. Спочатку підраховується кількість пікселів для кожного класу, здійснюється їх сортування за частотою та створюється словник заміन. Цей словник перетворює найбільш поширений клас на найменш поширений, другий найбільш поширений на другий найменш поширений і т.д. Нарешті, всі пікселі класу сільськогосподарських культур замінюються на нові; кількість пікселів, які займають найменш представлені класи, збільшується, і стає одного порядку з кількістю пікселів найбільш представлених класів. Таким чином, вирішується проблема дисбалансу навчальних даних.

2.1.3.3. КЛАСИЧНИЙ ПІДХІД ДО АУГМЕНТАЦІЇ

Запропонований метод було порівняно з класичними методами генерації штучних зображень [7]. В першому методі для оцінки статистичних характеристик кожного класу та генерації синтетичних сегментаційних масок використовується метод Монте-Карло. Потім заповнюються синтетичні супутникові зображення шляхом вибору по черзі пікселів на основі ймовірнісного розподілу вибраного класу сільськогосподарських культур у багатоспектральному просторі ознак. Цей підхід, який часто використовується при розв'язанні задач глибокого навчання, допомагає контролювати розподіл даних та запобігає перенавчанню, надаючи набір сегментаційних масок та синтетичних супутникових зображень для навчання поряд із реальними даними.

Другий метод, запропонований в [20], включає в себе вибір пікселів безпосередньо з реальних зображень за допомогою тієї ж штучної сегментаційної маски, що і попередні методи. Випадковий реальний набір даних вибирається для кожної маски та класу, забезпечуючи достатню кількість відповідних пікселів сільськогосподарських культур. Потім шляхом вибору пікселів з реального зображення заповнюється призначений клас на синтетичному зображенні. Цей підхід дозволяє зберегти стабільність значень пікселів шляхом вибору з одного класу в малій області, враховуючи незначні відмінності у властивостях пікселів на різних регіонах.

2.1.4. ОБГОВОРЕННЯ І ВИСНОВКИ

В експерименті було створено 2384 синтетичних супутникових зображень з штучними масками як з використанням методів GAN, так і статистичних методів. Потім було навчено чотири моделі: (I) з використанням реальних супутникових даних; (II) на основі комбінування реальних супутникових даних із синтетичними зображеннями на основі статистичного методу; (III) з використанням синтетичних зображень на основі методу вибірки пікселів; (IV) на основі синтетичних зображень, отриманих методом GAN. Було застосовано стандартні аугментації, такі як обертання і відображення, і використано функцію втрати focal loss для вирішення проблеми дисбалансу класів. Далі результати цих чотирьох методів було перевірено на незалежному наборі даних для тестування, який складався із 2 125 реальних зображень. Отримані результати представлено в табл. 1.

Модель (I) досягла 77.3% загальної точності (Overall accuracy, OA) і 64.1% перетину над об'єднанням (Intersection over Union, IoU), але середні точності для класу сільськогосподарських угідь (Average accuracies of cropland, AAC) для точності користувача (User Accuracy, UA), точності виробника (Producer Accuracy, PA) та IoU були дуже низькими через великий дисбаланс представлення класів сільськогосподарських культур у реальних даних. Генерація нових зразків із використанням статистичної моделі покращила представленість малих класів у наборі даних і не спричинила перенавчання. В результаті точність для більшості класів моделі (II) не змінилася, але точність малих класів значно зросла, що призвело до підвищення AAC на 8.5% для UA, 5.3% для PA і 9.0% для IoU. Метод вибірки пікселів в цілому показав схожі результати, і AAC для малих класів моделі (III) збільшилося на 6.7% для UA, 6.2% для PA і 8.0% для IoU.

Модель (IV), навчена за допомогою запропонованої методології аугментації GAN, випередила модель (I), модель (II) та модель (III). В порівнянні з моделлю (II) середні значення UA зросли на 2.7%, PA на 1.1% і IoU на 1.2%. Також OA та загальний IoU збільшилися на 1.4% та 1.6% відповідно.

Точність UA та PA для 4 з 8 класів сільськогосподарських культур та IoU для 7 із 8 класів у моделі (IV) є найвищими у експерименті. Є лише 2 винятки, де показники інших моделей значно вищі, ніж у моделі (IV): це кукурудза, де модель (IV) має найкращий результат лише за PA та IoU, та цукровий буряк, де модель (IV) випередила модель (III) лише за UA.

2.2. Методи сегментації зображень різної природи для випадку...

Таблиця 1 Порівняння метрик моделей класифікації с/г культур (найвищі метрики виділено жирним шрифтом)

с/г культура	Модель (I)			Модель (II)			Модель (III)			Модель (IV)		
	UA (%)	PA (%)	IoU	UA (%)	PA (%)	IoU	UA (%)	PA (%)	IoU	UA (%)	PA (%)	IoU
соняшник	80.4	92.7	75.7	83.5	91.0	77.1	82.8	89.6	75.5	82.8	92.9	77.9
кукурудза	83.8	90.0	76.6	83.9	89.4	76.3	82.6	89.6	75.4	87.6	88.2	78.4
пшениця	71.5	83.5	62.7	72.2	83.2	63.1	72.7	82.4	63.0	73.2	85.8	65.3
соя	80.8	54.9	48.6	73.4	60.7	49.8	76.7	55.1	47.1	73.0	68.2	54.5
ріпак	37.7	29.3	19.7	42.7	32.1	22.4	44.4	31.9	22.8	52.0	37.1	27.6
ячмінь	45.8	22.4	17.7	44.4	22.9	17.8	40.1	22.3	16.7	48.2	30.0	22.7
цукровий бурак	0.0	0.0	0.0	61.2	33.8	27.8	50.2	47.9	32.4	65.8	15.5	14.3
інші	26.6	3.2	2.7	32.7	4.9	41.0	30.1	6.9	34.7	33.2	9.7	43.8
ААС (%)	53.3	47.0	37.9	61.8	52.3	46.9	59.9	53.2	45.9	64.5	53.4	48.1
ОА (%)	77.3		64.1	77.4		64.6	77.0		64.0	78.8		66.2

Загалом аугментація GAN покращила продуктивність класифікації сільськогосподарських культур на 11.2%, 6.4% і 10.2% для середніх значень UA, PA та IoU відповідно, при цьому ОА та загальний IoU покращилися на 1.5% і 2.1%. Отримані карти класифікації на основі моделей (I)–(IV) представлені на рис. 4.

Слід відзначити, що для основних класів, таких як кукурудза, пшениця та соняшник, які є більш представленими, якість в обох картах сегментації є високою. В той же час для менш представленого класу ріпаку (темно-червоного кольору) результат, отриманий на основі моделі (I), має дефекти та артефакти та набагато меншу якість порівняно з моделлю (IV). Це можна пояснити тим, що GAN, на відміну від класичних статистичних методів генерації, дозволяє створити штучні приклади, які збережуть подібність не лише в точкових статистичних характеристиках класів, але й у відповідних текстурах.

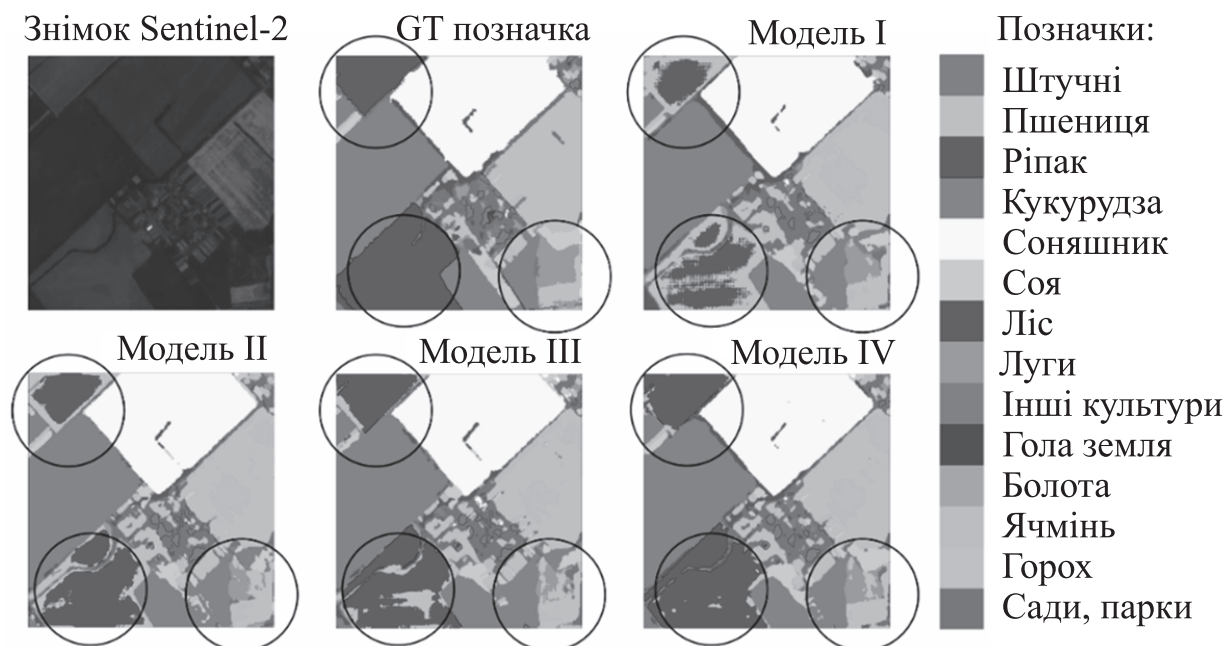


Рис. 4. Порівняння реальної карти сегментації (GT позначка) і результатів класифікації, отриманих лише на основі реальних зображень Sentinel-2 (Модель I), із даними, створеними за допомогою статистичного методу (Модель II), із даними, створеними за допомогою методу вибірки пікселів (Модель III), та із даними, створеними за допомогою методу GAN (Модель IV)

2.2. ВДОСКОНАЛЕННЯ МЕТОДУ ГЛИБИННОГО НАВЧАННЯ ДЛЯ СЕГМЕНТАЦІЇ ГІСТОЛОГІЧНИХ ЗОБРАЖЕНЬ

Датасети гістологічних зображень, як і датасети супутникових знімків, характеризуються сильним дисбалансом класів, а отримання додаткових даних значно ускладнене або ж зовсім неможливе. Гістологічні зображення тканин широко використовуються при виявленні патологій, зокрема злоякісних пухлин. Використання методів комп'ютерного зору дозволяє частково автоматизувати обробку даних, що може використовуватися як допоміжний інструмент під час роботи лікарів при встановленні клінічного діагнозу [22]. У порівнянні з традиційними алгоритмами машинного зору, де ознаки виділяються вручну, глибокі нейронні мережі показують кращі результати у задачах сегментації тканин, окремих клітин чи виділенні границь (меж) [25].

2.2.1. ОГЛЯД ЛІТЕРАТУРИ

Подальше покращення результатів може відбуватися за рахунок зміни архітектури моделі або її окремих шарів. У роботі [21] розглядається модифікації згорткового шару, де вхідні дані ущільнюються безпосередньо перед операцією згортки. Це дозволяє виділяти більше ознак на кожному послідовному шарі нейронної мережі, зберігаючи при цьому просторові властивості. Як наслідок, загальна кількість шарів зменшується, а також зменшується час обробки зображення.

Для боротьби з незбалансованістю навчальних вибірок, присутньою у більшості публічно доступних наборах даних, авторами [23] запропоновано використання зваженої функції втрат відповідно до частоти класів у вибірці. Менш представленим класам відповідають більші вагові коефіцієнти, а більш представленим класам — відповідно менші. Таким чином, вплив кожного класу на зміни параметрів моделі стає приблизно рівним. Також у даному розділі представлено алгоритм післяопераційної обробки результатів роботи моделі, який враховує впевненість моделі та передбачену відстань до межі клітини для кожного пікселя. Таким чином отримується повна оцінка форми клітини та фільтруються хибні передбачення пікселів фону.

Подібний підхід продемонстровано у роботі [24] присвяченій сегментації уражень шкіри, де в ролі функції втрат використовується зважена контурна крос-ентропія (WCCE). Вона змушує модель приділяти більше уваги контурам областей інтересу, що загалом покращує якість сегментації та точного розпізнавання границь уражених зон.

У роботі [26] для подолання проблеми незбалансованості класів використовується узагальнена функція втрат Дайса (Dice Loss). Вклад кожного класу в остаточне значення функції регулюються ваговими коефіцієнтами, які обернено пропорційні площі, що даний клас займає на конкретному зображенні. Окрім того, наведено кроки попередньої обробки даних, такі як зміна розмірів та розбиття зображень, нормалізація та аугментація.

Більшість публічно доступних набірів даних складаються з невеликої кількості зображень, а деякі класи зустрічаються набагато частіше за інші [23]. Можливо розширити навчальну вибірку за рахунок штучно згенерованих зображень. Генеративні змагальні мережі здатні створювати реалістичні гістологічні зображення за заданою сегментаційною маскою [28], що дозволяє згенерувати багато зображень з найменш представленими класами

та вирівняти розподіл класів у об'єднаному датасеті з реальний та штучних даних.

Для покращення розпізнавання границь окремих клітин, автори [27] вводять додатковий клас, який відповідає цій межі, а його пікселі знаходяться поруч зі границею на невеликій відстані. Загальне число пікселів такого класу дуже мале, що посилює незбалансованість класів у датасеті. Тому при навчанні використовується зважена крос-ентропія, де нововведеному класу відповідає більший ваговий коефіцієнт у порівнянні з рештою класів. Також більші вагові коефіцієнти отримали пікселі, як клітин так і фону, які лежать у безпосередній близькості від нового класу, тобто поруч з границею — чим далі від границі, тим менший коефіцієнт.

2.2.2. ПОСТАНОВКА ЗАДАЧІ

Для подальших експериментів у цьому розділі використано датасет MoNuSAC2020 [29], який був вперше представлений на змаганні з сегментації гістологічних знімків MoNuSAC (A Multi-organ Nuclei Segmentation and Classification Challenge) у 2020 році. Він представляє собою набір отриманих з мікроскопу зображень біологічних тканин, пофарбованих гематоксилином-еозином. У анотації датасету вказано клас та границі кожної клітини.

Загалом датасет містить 310 зображень різного розміру, з них 209 для навчання та 101 для тестування. На зображеннях присутні чотири типи клітин з чотирьох органів. Загальне число розмічених клітин перевищує 46 000, у зборі даних брали участь 37 різних лікувальних закладів та 71 пацієнт. Датасет має 5 класів: епітеліальні клітини, лімфоцити, нейтрофіли, макрофаги та один додатковий клас «Фон», що позначає відсутність перелічених клітин. Приклад розміченого зображення з датасету зображено на рис. 5.

Зображення з датасету поділено на фрагменти розміром по 256×256 кожне. Більшість оригінальних зображень мають значно більший розмір, аніж вказаний, тому у результаті загальне число фрагментів становило 1 131 у наборі для навчання, та 439 у наборі для тестування.

Для навчального набору даних обраховано відсоток ненадійних точок по кожному класу та загалом. Дані наведено у табл. 2. Більша половина представників усіх класів була визнана ненадійними екземплярами. Класи «Нейтрофіл» та «Макрофаг» є найменш чисельними та одночасно найбільш ненадійними

2.2. Методи сегментації зображень різної природи для випадку...

класами. Незважаючи на це, явна тенденція щодо залежності відсотку надійності від числа представників класу відсутня. Наприклад, більш чисельний клас «Епітелій» має меншу частку надійних екземплярів ніж клас «Лімфоцит».

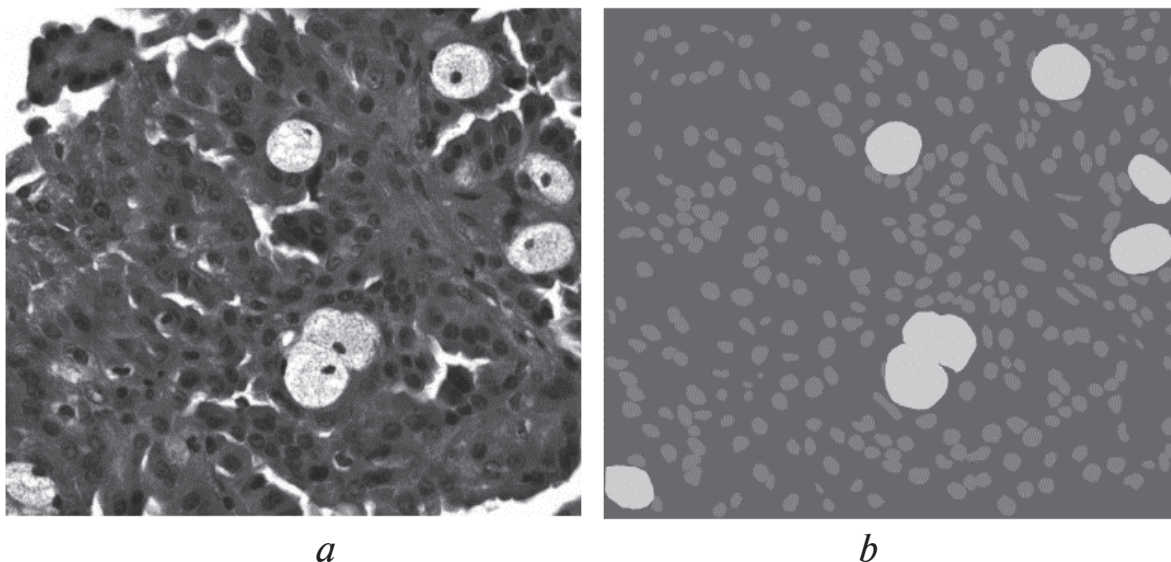


Рис. 5. Приклад зображення (а) та істинної маски сегментації (б) з датасету MoNuSAC2020

Таблиця 2 Загальне число пікселів та частка ненадійних даних по кожному класу

Назва класу	Кількість пікселів, тис.	Відсоток надійних точок, %
Фон	60 719	44.68
Епітелій	8 990	34.66
Лімфоцит	3 138	42.28
Нейтрофіл	136	24.12
Макрофаг	1 136	24.06
Разом	74 121	43.01

2.2.3. РОЗВ'ЯЗАННЯ ЗАДАЧІ

Для розв'язання поставленої задачі пропонується використовувати просторові вагові коефіцієнти при обчисленні функції втрат. Для кожного фрагменту зображення на основі значень його пікселів генерується відповідна йому вагова маска.

Отримані маски використовуються для модифікації процесу тренування моделей глибокого навчання, шляхом зміни функції

втрат. Звичайну функцію втрат можна описати наступною залежністю:

$$Loss = F(\hat{Y}, \hat{Y}'), \quad (3)$$

де \hat{Y}, \hat{Y}' — істинна маска сегментації та маска, яку вивела модель, відповідно.

Запропонований метод ускладнює цю функцію, додаючи залежність від маски, яка виконує роль вагових коефіцієнтів для кожного пікселю:

$$Loss = F(\hat{Y}, \hat{Y}', \hat{M}), \quad (4)$$

де \hat{M} — вагова маска такої ж розмірності як і \hat{Y} .

Слід зазначити відмінності даного способу від інших методів введення вагових коефіцієнтів у функцію втрат. Відмінності від застосування вагових коефіцієнтів для кожного окремого класу у даному випадку очевидні, оскільки значення вагової маски у запропонованому підході не залежить прямо від істинного класу пікселів на фрагментів. Інший же підхід, який характеризується застосуванням вагових коефіцієнтів для кожного фрагменту з набору даних, є більш вузьким у порівнянні з представленим. Він має лише один коефіцієнт на індивідуальний фрагмент, який входить до набору даних, а запропонований метод має по одному коефіцієнту для кожного пікселю у кожному фрагменті. Відповідно, сукупність таких коефіцієнтів, впорядкована за координатами відповідних пікселів, може розглядатися як окреме зображення або ж маска, звідки і пішла її назва.

2.2.3.1. ПОКРАЩЕННЯ АЛГОРИТМУ ВИЯВЛЕННЯ НЕОДНОЗНАЧНИХ ДАНИХ

Перші експерименти на датасеті MoNuSAC2020 з використанням вагових масок показали необхідність вдосконалення даного методу, а саме в частині алгоритму виявлення неоднозначних екземплярів даних.

Оригінальна версія алгоритму розглядає кожен піксель як окремий елемент, жодним чином не пов'язаний з сусідніми пікселями, замінюючи задачу сегментації на задачу по-піксельної класифікації. Незважаючи на таке спрощення, це показувало хороший результат в області обробки супутникових знімків, де вирішальним фактором є колір, а не форма, розмір чи інші ознаки. У даній же області, колір хоч і є важливою ознакою, але й інші властивості, зокрема форма, мають більше значення ніж раніше.

2.2. Методи сегментації зображень різної природи для випадку...

Тому виникла необхідність врахувати просторові властивості при визначенні неоднозначних пікселів.

При визначенні, які пікселі є сусідніми у просторі ознак, у першій версії алгоритму обчислювалася евклідова відстань між ними у просторі ознак:

$$\|\vec{X}_{i,j} - \vec{X}_{p,q}\|_{L_2}, \quad (5)$$

де $\hat{Y}: \{\vec{X}_{i,j}\}$ — зображення \hat{X} , що складається з пікселів $\vec{X}_{i,j}$, а i, j та p, q — координати пікселів на зображенні. Довжина вектору $\vec{X}_{i,j}$ дорівнює числу каналів у зображенні, в даному випадку це 3. Значення каналу k в індивідуальному пікселі позначимо як $X_{i,j,k}$.

Основна ідея вдосконалення алгоритму полягає у тому, що при визначенні відстані у просторі ознак між двома пікселями, враховувати значення їхніх сусідніх пікселів в межах деякого радіусу R . Очевидно, що такі сусідні пікселі мають впливати на кінцеву величину відстані менше, ніж цільові два пікселі. Тому доцільно враховувати відстані між сусідніми пікселями з певним, нижчими за одиницю коефіцієнтами. Нехай K — така матриця коефіцієнтів, розмірами $[2R+1 \times 2R+1]$. Центральний її елемент дорівнює одиниці, решта знаходяться в межах $[0,1)$:

$$K = \begin{pmatrix} K_{1,1} & \cdots & K_{1,R+1} & \cdots & K_{1,2R+1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ K_{R+1,1} & \cdots & 1 & \cdots & K_{R+1,2R+1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ K_{2R+1,1} & \cdots & K_{2R+1,R+1} & \cdots & K_{2R+1,2R+1} \end{pmatrix}. \quad (6)$$

$$0 \leq K_{i,j} < 1, i \neq R+1, j \neq R+1$$

Різниця між каналом k двох цільових пікселів $\vec{X}_{i,j}$ та $\vec{X}_{p,q}$, з урахуванням сусідніх пікселів обчислюється за наступною формулою:

$$\|X_{i,j,k} - X_{p,q,k}\|_{ch} = \sum_{f=-R}^{+R} \sum_{g=-R}^{+R} K_{f+R+1,g+R+1} (X_{i+f,j+g,k} - X_{p+f,q+g,k}). \quad (7)$$

Це зважена середня різниця значень певного каналу пікселів. З урахуванням дистрибутивності множення, та комутативності додавання перепишемо ще в наступному вигляді:

$$\begin{aligned} \|X_{i,j,k} - X_{p,q,k}\|_{ch} = & \sum_{f=-R}^{+R} \sum_{g=-R}^{+R} K_{f+R+1,g+R+1} X_{i+f,j+g,k} - \\ & - \sum_{f=-R}^{+R} \sum_{g=-R}^{+R} K_{f+R+1,g+R+1} X_{p+f,q+g,k}. \end{aligned} \quad (8)$$

Отримані два доданки є нічим іншим як формулою згортки з ядром K над k каналом пікселів $\vec{X}_{i,j}$ та $\vec{X}_{p,q}$ відповідно. Для того, щоб уникнути обчислення згортки кожного разу при визначенні різниці між двома пікселями, є сенс обрахувати згортку заздалегідь для всього зображення. Нехай \hat{X}' — зображення, отримане шляхом по-канальної згортки ядра K з оригінальним зображенням \hat{X} :

$$\begin{aligned} X_{i,j,k}' = & \sum_{f=-R}^{+R} \sum_{g=-R}^{+R} K_{f+R+1,g+R+1} X_{i+f,j+g,k}, \forall i,j,k \\ \hat{X}' = & K * \hat{X}. \end{aligned} \quad (9)$$

У результаті, відстань між двома пікселями у просторі ознак з урахуванням їх сусідів ($L2_{nn}$) обчислюється як:

$$\begin{aligned} \|\vec{X}_{i,j} - \vec{X}_{p,q}\|_{L2_{nn}} = & \sqrt{\sum_k (X_{i,j,k}' - X_{p,q,k}')^2} = \\ & \|\vec{X}_{i,j}' - \vec{X}_{p,q}'\|_{L2}. \end{aligned} \quad (10)$$

Як висновок, визначення неоднозначних пікселів у наборі зображень з урахуванням просторовості, математично еквівалентне застосуванню згортки з ядром K до всього набору зображень, та запуску на ньому оригінального, по-піксельного алгоритму виявлення неоднозначних даних. Застосування згортки зводить дану задачу до попередньої.

Нову версію алгоритму використано для визначення неоднозначних пікселів у датасеті MoNuSAC2020. У якості ядра

2.2. Методи сегментації зображень різної природи для випадку...

згортки K , вибрано гауссівське ядро розмірами 5×5 . Застосування такого ядра до зображення є нічим іншим як гауссівським розмиттям цього зображення. Далі застосовано звичайну, по-піксельну версію алгоритму на невеликій підмножині датасету. На її основі визначено приналежність до неоднозначних даних для повного датасету та згенеровано відповідні вагові маски. Статистика щодо частки неоднозначним даних по кожному класу та загалом представлено у табл. 3.

Для класів «Фон», «Епітелій» та «Лімфоцит» значно виросла частка надійних екземплярів. Це можна спробувати пояснити тим, що представники даних класів є подібними до інших, якщо розглядати задачу по-піксельної класифікації. Натомість, врахування просторових властивостей зображення, а саме значень у сусідніх пікселях, дає змогу краще розрізняти дані класи. Натомість, для класів «Нейтрофіл» та «Макрофаг» ситуація протилежна – збільшилася частка ненадійних екземплярів, для цих класів врахування просторовості зменшило їх відмінності від інших класів. Це може бути пов'язаним з наявністю викидів та екстремальних значень серед представників цих класу, вплив яких зменшився через врахування просторовості. Іншим поясненням може бути малі розміри цих клітин: розглядаючи окремі екземпляри класів, можна зауважити що значна частка їхніх сусідів будуть вже представниками інших класів.

Таблиця 3 Загальне число пікселів та частка ненадійних даних по кожному класу з урахуванням сусідніх пікселів

Назва класу	Кількість пікселів, тис.	Відсоток надійних точок, %
Фон	60 719	78.43
Епітелій	8 990	53.45
Лімфоцит	3 138	52.74
Нейтрофіл	136	1.23
Макрофаг	1 136	4.49
Разом	74 121	73.04

2.2.3.2. НАБОРИ ВАГОВИХ МАСОК

Згідно з даними табл. 3, майже всі пікселі у найбільшому класі «Фон» є надійними, а більшість пікселів з інших класів є ненадійними. У попередніх експериментах з супутниковими

знімками перевірялося, даним якого саме типу слід надати перевагу через вагові маски — чи надійним, чи ненадійним. У даному випадку, з огляду на статистику датасету, очевидним є надати перевагу неоднозначним даним. Надати перевагу нормальним пікселям — означає зробити фокус моделі на найбільш представленому класі, який відповідає фону, та значно зменшити акцент моделі на найменш представлені класи, які відповідають різним типам клітин.

Згенеровано 2 набори вагових масок, де неоднозначні пікселі мали більші значення відповідних коефіцієнтів у вагових масках. Для повноти експерименту, також було згенеровано ще один набір вагових масок, де фокус надавався нормальним даним. Морфологічні перетворення було вирішено не застосовувати через неоднорідну структуру вагових масок. Повна інформація про набори представлена у табл. 4.

Приклади вагових масок з різних наборів зображені на рис. 6.

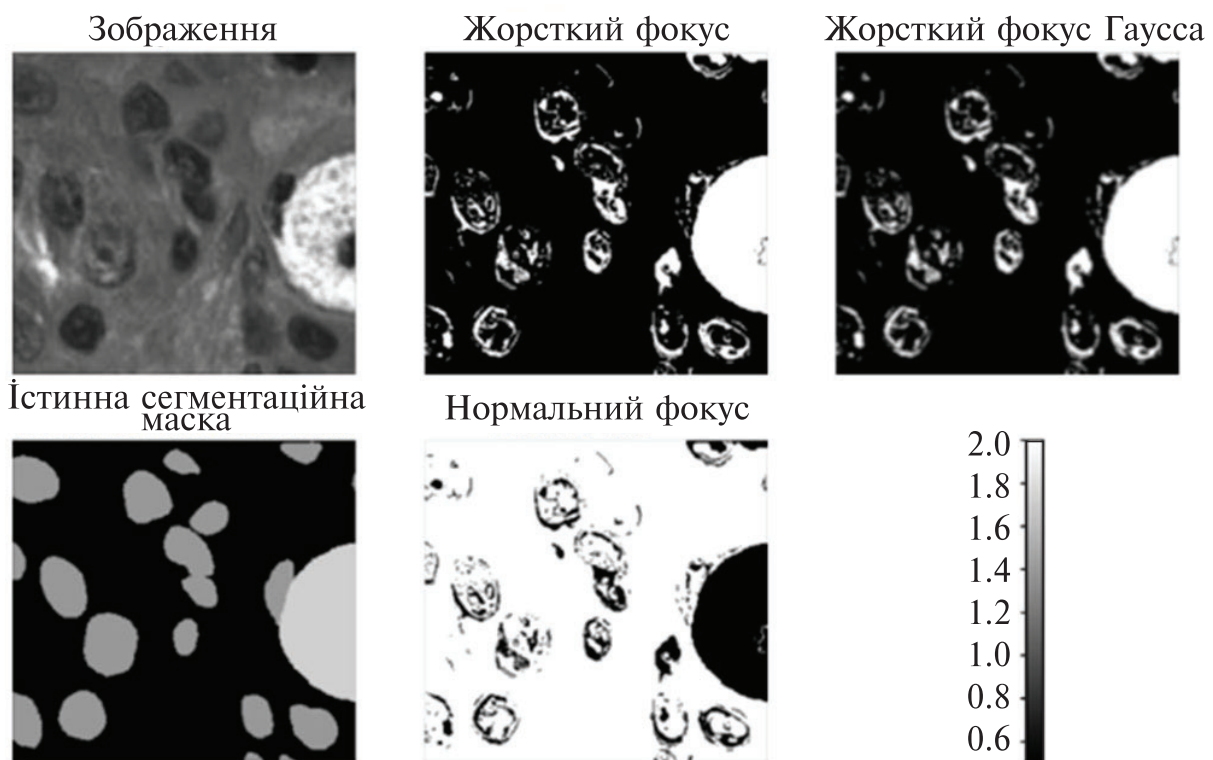


Рис. 6. Візуалізація гістологічного зображення та відповідної йому маски з набору

З цих зображень можна зробити висновок, що деякі клітини повністю віднесені до неоднозначних даних, а інші — лише частково. У таких клітинах до неоднозначних даних віднесені

2.2. Методи сегментації зображень різної природи для випадку...

лише їх периферичні області, що вказує на особливу складність розпізнавання межі клітини та фону. Більшість пікселів, що відповідають фону, віднесені до надійних даних, окрім невеликої частки структур, що схожі за своїми спектральними властивостями на клітини.

Таблиця 4 Список перетворень, застосованих до різних наборів вагових масок)

№	Назва набору	Фокус вагових коефіцієнтів	Гауссове розмиття
1	Normal focus	нормальні пікселі (4:1)	-
2	Hardcase focus	неоднозначні пікселі (4:1)	-
3	Hardcase focus Gauss	неоднозначні пікселі (4:1)	5×5

2.2.3.3. АРХІТЕКТУРА МОДЕЛІ ТА СХЕМА ЕКСПЕРИМЕНТУ

У зв'язку з набагато складнішою структурою даних на зображенні та великою варіабельністю текстур, прийнято рішення використати іншу архітектуру моделі, а саме SegNet [30], який має вдвічі більше параметрів за попередню модель, засновану на U-Net. На рис. 7 представлена архітектура моделі. Вона приймає на вхід дані розмірністю $[N \times 256 \times 256 \times 3]$, та видає результат розмірністю $[N \times 256 \times 256 \times 5]$, де N — кількість зображень у навчальному пакеті.

Оптимізація параметрів моделі виконувалася з використанням оптимізатора Adam за параметрами $\beta_1=0.9$ та $\beta_2=0.999$. Швидкість навчання змінювалася динамічно в межах $[1 \cdot 10^{-1}, 1 \cdot 10^{-5}]$.

Для навчання базової моделі у якості функції втрат використовувалося стандартна крос-ентропія. Під час навчання моделі з використанням методу вагових масок, використовувалася модифікована крос-ентропія:

$$BCE(\hat{Y}, \hat{Y}', \hat{M}) = \frac{-1}{n} \sum_{i,j} M_{ij} \sum_{k=1}^c Y_{ijk} \log(Y_{ijk}'), \quad (11)$$

де \hat{Y}, \hat{Y}' — сегментаційні маски, істинна і передбачена відповідно, \hat{M} — вагова маска, n — кількість пікселів в одному фрагменті, C — кількість класів.

Всього було навчено 4 моделі — одна базова і 3 — з використанням різних наборів вагових масок.

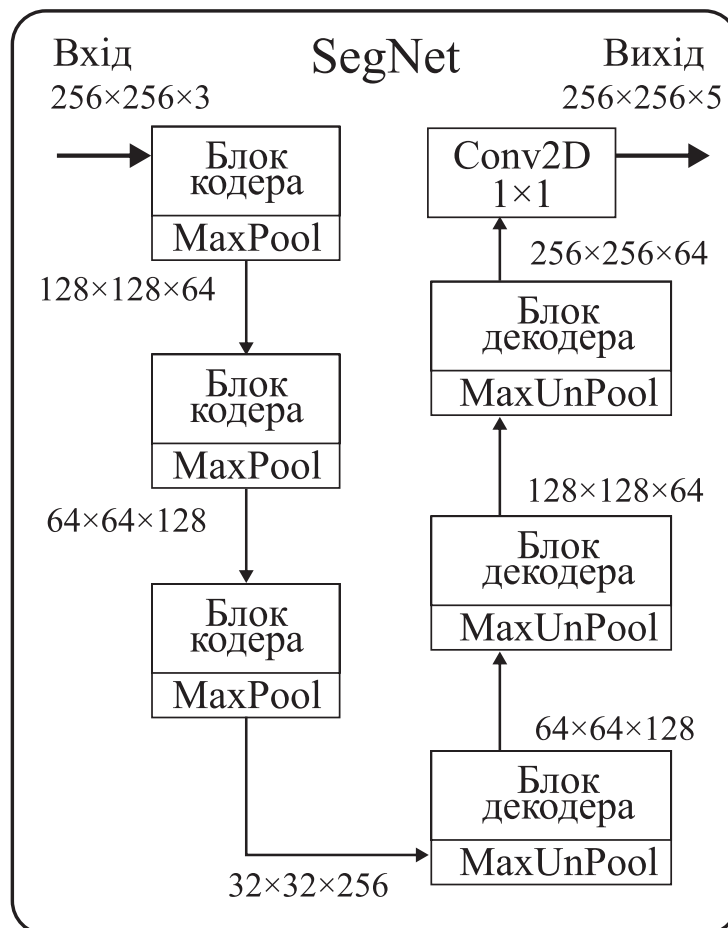


Рис. 7. Архітектура сегментаційної моделі

2.2.4. ОБГОВОРЕННЯ І ВИСНОВКИ

Навчені моделі оцінено на незалежному наборі даних, який не використовувався при навчанні. Обраховано метрики, такі як точність, влучність (precision), повнота (recall) та IoU. Детальні результати наведені у табл. 5.

Як і очікувалось, модель, що тренувалася з акцентом на нормальні дані, не продемонструвала покращення метрик малих класів у порівнянні з базовою моделлю.

2.2. Методи сегментації зображень різної природи для випадку...

Таблиця 5 Порівняння метрик моделей сегментації гістологічних зображень (метрики, вищі за базову модель, виділено жирним шрифтом)

Назва класу	Model (I)			Model (II)			Model (III)			Model (IV)		
	UA (%)	PA (%)	IoU	UA (%)	PA (%)	IoU	UA (%)	PA (%)	IoU	UA (%)	PA (%)	IoU
Фон	93.0	96.1	89.6	93.6	96.0	90.1	93.0	95.7	89.3	92.9	94.5	88.1
Епітелій	81.9	71.9	62.1	76.9	74.3	60.8	82.7	73.5	63.7	83.7	69,7	61.4
Лімфоцит	81.0	71.3	61.1	77.5	72.5	59.9	84.8	68.2	60.7	79,0	74,4	62.1
Нейтрофіл	0.0	0.0	0.0	0.0	0.0	0.0	48.3	42.9	29.4	23,0	10.1	7.6
Макрофаг	4.9	3.0	1.9	24.8	2.3	2.1	11.4	11.9	6.2	5.4	9.6	3.6
ААС (%)	52.1	48.5	42.9	54.6	49.0	42.6	64.0	58.4	49.9	56.8	51.7	44.6
ОА (%)	90.5	83.4	83.4	90.7	83.6	83.6	90.3	83.4	83.4	89.1	89.1	82.6

Інші дві моделі розрізняють малі класи суттєво краще за базову модель. Вагові маски цих моделей мали більше значення коефіцієнтів для неоднозначних пікселів, ніж для нормальних, а до набору масок останньої моделі також було застосовано Гауссове розмиття з ядром 5×5 . Модель, що вчилася з використанням набору масок *Hardcase focus*, перевищує базову модель майже в усіх метриках, які стосуються малих класів, а загальні метрики змінилися незначно. Особливо слід відзначити те, що базова модель нездатна розпізнати клас «Макрофаг», а використання вагових масок дало можливість правильно розпізнавати цей клас майже в половині випадків (рис. 8).

У попередніх дослідах, для задачі сегментації найкраще працювали вагові маски, де більші значення коефіцієнтів відповідали неоднозначним пікселям. У даному експерименті ситуація протилежна. Це може бути пов'язано з наявністю домінантного класу «Фон», число екземплярів якого набагато більше за усі інші класи разом узяті. Алгоритм визначення неоднозначним даних відніс більшість екземплярів цього класу до однієї категорії. Як наслідок, модель мала шанси вивчити малі

класи лише якщо цій категорії буде присвоєний низький пріоритет.

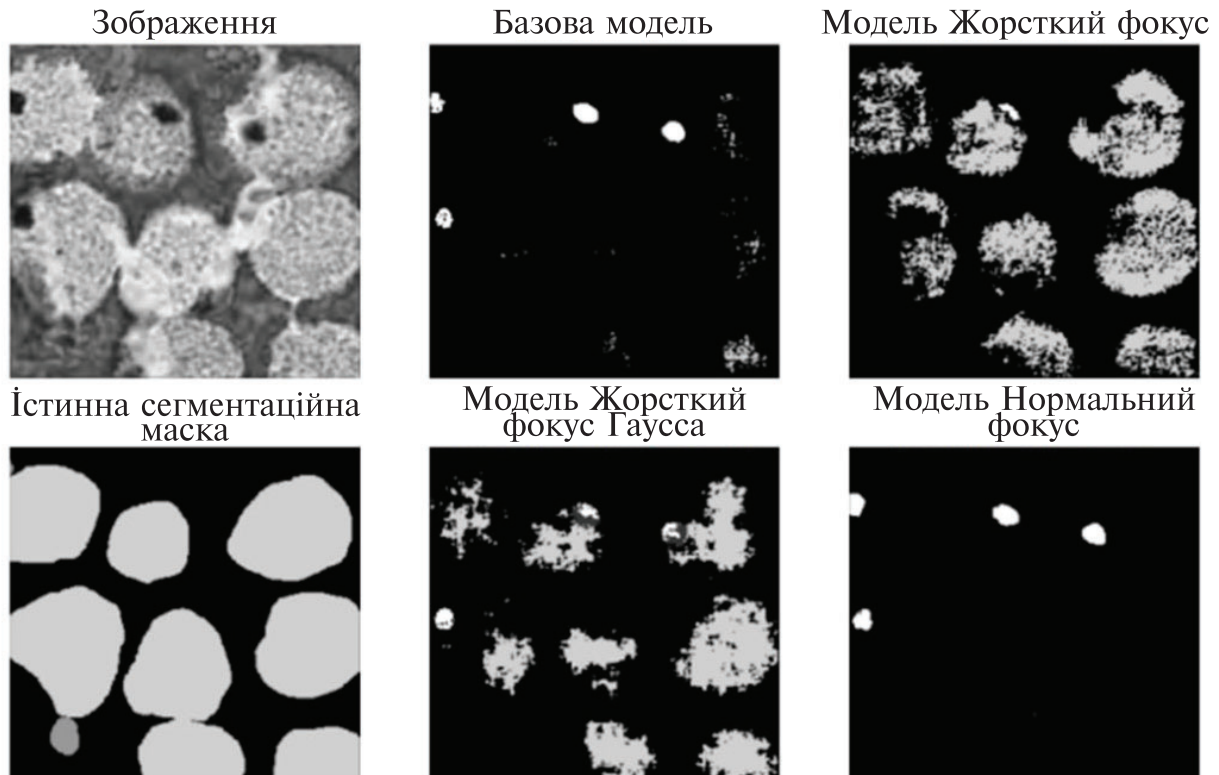


Рис. 8. Візуалізація гістологічних зображень, істинних масок сегментації (істинні класи), результатів роботи базової моделі та результати роботи моделей з використанням вагових масок

Представники доміантного класу знаходяться у просторі ознак у великих кількостях, і у разі наявності проблеми перекриття класів, з більшою імовірністю будуть вибрані як сусіди під час роботи алгоритму виявлення неоднозначних даних. Це узгоджується з реальними результатами, коли модель має тенденцію вивчати найбільш представлений клас, ігноруючи малі класи. У даному датасеті, приблизно половина пікселів класів «Епітелій» та «Лімфоцит» знаходилися у просторі ознак достатньо далеко, щоб бути чітко розрізненими від інших класів та не отримати такого суттєвого фокусу під час навчання, як класи «Нейтрофіл» чи «Макрофаг». Тим не менш, метрики цих класів також збільшилися у порівнянні з базовою моделлю.

ВИСНОВКИ

На основі результатів проведених досліджень та отриманих результатів можна зробити наступні висновки. Запропоновано

2.2. Методи сегментації зображень різної природи для випадку...

новий підхід до аугментації даних для розв'язання задачі класифікації сільськогосподарських культур на основі моделі GAN рix2рix для генерації реалістичних зображень, які відповідають багатоспектральним та текстурним характеристикам. Це дозволило усунути проблему дисбалансу набору даних для глибокого навчання методами семантичної сегментації. Запропонований метод було порівняно з класичними підходами до генерації зображень на основі статистичних характеристик багатоспектральних ознак класів сільськогосподарських культур. Запропонований метод був протестований на основних аугментаціях та функціях втрат, які застосовуються в випадку дисбалансу класів. В результаті запропонований метод виявився кращим порівняно з моделями, навченими на основі лише реальних даних та класичних підходів. Для більшості класів сільськогосподарських культур спостерігалось значне підвищення точності для всіх мало представлених класів. Загальна точність моделі, навченої за запропонованим підходом, покращилася на 1.5% і 2.1% для OA та IoU, в той час як середні значення UA, PA та IoU для класів сільськогосподарських культур зросли на 11.2%, 6.4% та 10.2% порівняно з моделлю, навченою без синтетичних зображень.

Також розглянуто застосування методу вагових масок для сегментації гістологічних знімків. Дані з обраного датасету мали сильний дисбаланс у розподілі класів, а подальший аналіз датасету вказав на наявність проблеми перекриття класів. Запропоновано вдосконалену версію алгоритму виявлення неоднозначних екземплярів даних, яка враховує просторові властивості зображення, а саме значення сусідніх пікселів при обрахунку відстані у просторі ознак. Доведено, що дану версію алгоритму можна звести до попередньої шляхом застосування до зображення операції згортки з ядром коефіцієнтів. Проведено порівняння розробленого методу зі стандартною процедурою навчання моделі на тих самих даних і тією самою архітектурою моделі. За результатами порівняння, використання розробленого методу збільшує точність розпізнавання усіх малих класів, які відповідають різним типам клітин, однак значення загальних метрик залишається незмінними. Для даної задачі найкращі результати давали маски, де ненадійним пікселям присвоювалися більші значення вагових коефіцієнтів. Це пояснюється наявністю домінантного класу, більшість представників якого мали статус надійних.

ПЕРЕЛІК ПОСИЛАНЬ

1. Franch B., Vermote E., Skakun S., Santamaria-Artigas A., Kalecinski N., Roger J.-C., et al. The ARYA crop yield forecasting algorithm: Application to the main wheat exporting countries. *International Journal of Applied Earth Observation and Geoinformation*. 2021. Vol. 104. pp. 102552. DOI: 10.1016/j.jag.2021.102552.
2. Kussul N., Lavreniuk M., Kolotii A., Skakun S., Rakoid O., Shumilo L. A workflow for Sustainable Development Goals indicators assessment based on high-resolution satellite data. *International Journal of Digital Earth*. 2019. Vol. 13, no. 2. pp. 309-321. DOI: 10.1080/17538947.2019.1610807.
3. Boryan C., Yang Z., Mueller R., Craig M. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto International*. 2011. Vol. 26, no. 2. pp. 341-358. DOI: 10.1080/10106049.2011.562309.
4. Defourny P., Bontemps S., Bellemans N., Cara C., Dedieu G., Guzzonato E., et al. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote sensing of environment*. 2019. Vol. 221. pp. 551-568. DOI: 10.1016/j.rse.2018.11.007.
5. Moazzam S.I., Khan U.S., Tiwana M.I., Iqbal J., Qureshi W.S., Shah S.I. A review of application of deep learning for weeds and crops classification in agriculture. *2019 International Conference on Robotics and Automation in Industry (ICRAI)*. 2019. pp. 1-6. DOI: 10.1109/ICRAI47710.2019.8967350.
6. Xu W., Deng X., Guo S., Chen J., Sun L., Zheng X., et al. High-resolution u-net: Preserving image details for cultivated land extraction. *Sensors*. 2020. Vol. 20, no. 15. pp. 4064. DOI: 10.3390/s20154064.
7. Shorten C., Khoshgoftaar T.M. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. 2019. Vol. 6, no. 1. pp. 60. DOI: 10.1186/s40537-019-0197-0.
8. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., et al. Generative adversarial networks.

2.2. Методи сегментації зображень різної природи для випадку...

Communications of the ACM. 2020. Vol. 63, no. 11. pp. 139-144. DOI: 10.1145/3422622.

9. Pham V.D., Bui Q.T. Spatial resolution enhancement method for Landsat imagery using a Generative Adversarial Network. *Remote Sensing Letters*. 2021. Vol. 12, no. 7. pp. 654-665. DOI: 10.1080/2150704x.2021.1918789.

10. Abady L., Barni M., Garzelli A., Tondi B. GAN generation of synthetic multispectral satellite images. *Image and Signal Processing for Remote Sensing XXVI*. 2020. Vol. 11533. pp. 122-133. DOI: 10.1117/12.2575765.

11. Xu Z., Du J., Wang J., Jiang C., Ren Y. Satellite image prediction relying on GAN and LSTM neural networks. *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*. 2019. pp. 1-6. DOI: 10.1109/icc.2019.8761462.

12. Christovam L.E., Shimabukuro M.H., Galo M.D.L.B., Honkavaara E. Pix2pix conditional generative adversarial network with MLP loss function for cloud removal in a cropland time series. *Remote Sensing*. 2021. Vol. 14, no. 1. pp. 144. DOI: 10.3390/rs14010144.

13. Andrade H.J., Fernandes B.J. Synthesis of satellite-like urban images from historical maps using conditional GAN. *IEEE Geoscience and Remote Sensing Letters*. 2022. Vol. 19. pp. 1-4. DOI: 10.1109/lgrs.2020.3023170.

14. Kussul N., Shelestov A., Yailymova H., Yailymov B., Lavreniuk M., Ilyashenko M. Satellite agricultural monitoring in Ukraine at country level: World Bank Project. *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. 2020. pp. 1050-1053. DOI: 10.1109/igarss39084.2020.9324573.

15. Shumilo L., Okhrimenko A., Kussul N., Drozd S., Shkalikov O. Generative adversarial network augmentation for solving the training data imbalance problem in crop classification. *Remote Sensing Letters*. 2023. Vol. 14, no. 11. pp. 1131-1140. DOI: 10.1080/2150704X.2023.2275551.

16. Bowles C., Chen L., Guerrero R., Bentley P., Gunn R., Hammers A., et al. Gan augmentation: Augmenting training data using generative adversarial networks. *arXiv preprint arXiv*. 2018. DOI: 10.48550/ARXIV.1810.10863.

17. Isola P., Zhu J.Y., Zhou T., Efros A.A. Image-to-image translation with conditional adversarial networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. pp. 1125-1134. DOI: 10.1109/cvpr.2017.632.

18. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 2015. pp. 234-241. DOI: 10.1007/978-3-319-24574-4_28.

19. Wang P., Bai X. Thermal infrared pedestrian segmentation based on conditional GAN. *IEEE transactions on image processing*. 2019. Vol. 28, no. 12. pp. 6007-6021. DOI: 10.1109/tip.2019.2924171.

20. Rawat W., Wang Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*. 2017. Vol. 29, no. 9. pp. 2352-2449. DOI: 10.1162/neco_a_00990.

21. Ahmed N. Real-time and accurate deep learning-based multi-organ nucleus segmentation in histology images. *Journal of Real-Time Image Processing*. 2024. Vol. 21. DOI: 10.1007/s11554-024-01420-0.

22. Wu Y., Cheng M., Huang. Sh. et al. Recent Advances of Deep Learning for Computational Histopathology: Principles and Applications / Yawen Wu, Michael Cheng, Shuo Huang et al. // *Cancers*. 2022. Vol. 14. P. 1199. DOI: 10.3390/cancers14051199.

23. Nasir E., Fraz M. NuRISC: Nuclei Radial Instance Segmentation and Classification. *Lecture Notes in Electrical Engineering*. Vol. 810 LNEE. 2023. DOI: 10.1007/978-981-16-6775-6_4.

24. Liang S., Tian Sh., Yu L., Kang X. Improved U-Net based on contour attention for efficient segmentation of skin lesion. *Multimedia Tools and Applications*. 2024. Vol. 83. DOI: 10.1007/s11042-023-16759-y.

25. Kadaskar M., Patil N. Image Analysis of Nuclei Histopathology Using Deep Learning: A Review of Segmentation, Detection, and Classification. *SN Computer Science*. 2023. Vol 4. DOI: 10.1007/s42979-023-02115-2.

26. Hancer E., Traory M., Samet R. et al. An imbalance-aware nuclei segmentation methodology for H&E stained histopathology images. *Biomedical Signal Processing and Control*. 2023. Vol. 83. DOI: 10.1016/j.bspc.2023.104720.

2.2. Методи сегментації зображень різної природи для випадку...

27. Li H., Zhao X. Su A. et al. Color Space Transformation and Multi-Class Weighted Loss for Adhesive White Blood Cell Segmentation. *IEEE Access*. 2020. Vol. 8. DOI: 10.1109/ACCESS.2020.2970485.

28. Li W., Li J., Polson J. et al. High resolution histopathology image generation and segmentation through adversarial training. *Medical Image Analysis*. 2022. Vol. 75. DOI: 10.1016/j.media.2021.102251.

29. Verma R., Kumar N., Patil A. et al. MoNuSAC2020: A Multi-Organ Nuclei Segmentation and Classification Challenge. *IEEE Transactions on Medical Imaging*. 2021. Vol. 40. DOI: 10.1109/TMI.2021.3085712.

30. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. — 2017. — Vol. 39. DOI: 10.1109/TPAMI.2016.2644615.