

**АНАЛІЗ СУЧАСНОГО СТАНУ МЕТОДІВ  
КОМП'ЮТЕРНОГО ЗОРУ І ГЛИБИННИХ  
НЕЙРОННИХ МЕРЕЖ ДЛЯ ЕКОЛОГО-  
ЕКОНОМІЧНОГО АНАЛІЗУ**

**1. АНАЛІЗ МЕТОДІВ ТА МОДЕЛЕЙ ГЛИБИННОГО НАВЧАННЯ  
ДЛЯ РОЗВ'ЯЗАННЯ ПРИКЛАДНИХ ЗАДАЧ**

Наталія Куссуль, професор  
Володимир Кузін, магістр  
Андрій Шелестов, професор

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

nataliia.kussul@lil.kpi.ua, volHkuz@gmail.com,  
andrii.shelestov@gmail.com

**ВСТУП**

За останні роки штучний інтелект (ШІ) став однією з головних чинників швидкого розвитку та значного прогресу, революціонізувавши різні галузі та аспекти нашого повсякденного життя. Штучний інтелект (ШІ) став трансформаційною галуззю з величезним потенціалом для революційних змін у різних сферах - від охорони здоров'я та фінансів до транспорту та освіти [1]. ШІ охоплює різноманітний набір технологій і методологій, які дозволяють машинам виконувати завдання, що традиційно вимагали людського інтелекту [2]. Серед багатьох досягнень у галузі штучного інтелекту методи глибинного навчання привернули до себе значну увагу завдяки своїй ефективності у вирішенні складних проблем у різних прикладних сферах [3-5].

Швидкий розвиток ШІ можна пояснити кількома факторами. Перш за все, розвиток обчислювальних потужностей і доступність

величезних наборів даних сприяли навчанню складних моделей ШІ. Крім того, прориви в розробці алгоритмів машинного навчання, зокрема глибинного навчання, дозволили системам ШІ видобувати важливі закономірності та інформацію з величезних масивів даних. Поява хмарних обчислень і спільноти з відкритим вихідним кодом полегшила доступ до інструментів і фреймворків ШІ, сприяючи широкому впровадженню і співпраці в різних дисциплінах. Крім того, збільшення інвестицій як з приватного, так і з державного секторів сприяло розвитку досліджень і розробок у галузі ШІ, що призвело до появи інноваційних застосувань і технологій.

Оскільки ШІ продовжує розвиватися, його вплив на суспільство є глибоким і багатогранним. Геопросторовий інтелект є важливою сферою застосування ШІ. Інтелектуальне дослідження Землі, що базується на ШІ та глибокому навчанні, аналізує геопросторові дані з різних джерел, щоб отримати практичні знання про планету. Глибокі нейронні мережі ефективно обробляють складні набори геопросторових даних для розв'язання таких задач, як класифікація ґрунтового покриву, виявлення об'єктів і моделювання клімату, автоматизуючи завдання, які раніше вимагали значних людських зусиль. Поточні дослідження спрямовані на покращення інтерпретованості, узагальнення та застосовності моделей до нових типів даних, що обіцяє широке впровадження глибинного навчання у застосуваннях інтелектуального дослідження Землі, трансформативне бачення клімату, екології, ресурсів тощо, а також загальне покращення розуміння планетарних процесів та розвитку економіки.

Останні досягнення в галузі ШІ та інструменти для впровадження машинного навчання проаналізовано в багатьох оглядах [6, 7]. Існує кілька досліджень, присвячених моделям штучного інтелекту та глибинного навчання для інтелектуального дослідження та спостереження Землі [8]. Водночас ця сфера розвивається настільки швидко, що вимагає переосмислення загальної картини кожні кілька років. У даному розділі містяться результати аналізу сучасного стану глибинного навчання та його застосування для інтелектуального дослідження Землі, управління та прийняття рішень, акцентуючи увагу на практичному досвіді його застосування в Україні.

### 1.1. ВИЗНАЧЕННЯ ШТУЧНОГО ІНТЕЛЕКТУ

Перш ніж приступити до класифікації методів ШІ, почнемо з визначення ШІ. Штучний інтелект (ШІ) - це здатність машин виконувати когнітивні функції та завдання, які зазвичай вимагають людського інтелекту та сприйняття. Це наука та інженерія розробки комп'ютерних систем, які можуть брати участь у процесах мислення, подібних до людських, таких як навчання, міркування, розв'язання проблем, сприйняття, прогнозування, планування та прийняття рішень [9]. Алгоритми ШІ можуть реалізовувати навчання, розпізнавання, розв'язання проблем, розуміння мови та/або логічне мислення.

На сьогоднішній день моделі ШІ поділяються на сильні та слабкі. Терміни «сильний ШІ» і «слабкий ШІ» представляють різні рівні здатності штучного інтелекту [10]. Слабкий ШІ, також відомий як «вузький ШІ», призначений для вирішення конкретної задачі або обмеженого набору задач, фокусуючись на одній когнітивній здатності. Більшість з того, що ми сьогодні називаємо «слабким ШІ», складається зі спеціалізованих алгоритмів, розроблених для досягнення успіху в розв'язанні конкретних задач, що базуються на використанні великих навчальних даних. Таким алгоритмам бракує справжнього інтелекту, що виходить за межі запрограмованих функцій, і вони зазвичай використовуються в сучасних програмах ШІ, таких як чат-боти, розпізнавання обличчя, медичні програми, віртуальні асистенти тощо [11].

З іншого боку, сильний ШІ, також відомий як «загальний ШІ» або «повний ШІ», має на меті продемонструвати інтелект і пізнання на рівні людини [12]. Він має здатність розуміти, міркувати, планувати, вирішувати проблеми, абстрактно мислити і швидко вчитися на власному досвіді. На відміну від слабого ШІ, сильний ШІ демонструє узагальнений інтелект у різних сферах і може переносити знання з однієї задачі на інші, не пов'язані з нею. Наразі сильний ШІ залишається гіпотетичною концепцією, а його розробка - довгостроковою метою [13].

### 1.2. МАШИННЕ НАВЧАННЯ

ШІ включає в себе різні підходи та моделі, і однією з основних парадигм є машинне навчання. Машинне навчання - це галузь штучного інтелекту (ШІ), яка передбачає використання математичних методів, що дозволяють машинам здобувати знання

на основі досвіду. Воно зосереджене на розробці алгоритмів, заснованих на емпіричних і навчальних даних, для оптимізації результатів і поліпшення прогнозів за допомогою процесів навчання [14].

ML застосовується для вирішення наукових і прикладних задач у таких галузях, як хімія, медицина, астрономія, біологія, сільське господарство, муніципальне господарство, промисловість, будівництво, моделювання навколишнього середовища, обробка природної мови, геопросторовий аналіз тощо. Різні парадигми та сфери застосування ШІ, що розглядаються в [15], показані на рис. 1. У нашому дослідженні основна увага буде зосереджена на машинному навчанні, а точніше, на моделях глибокого навчання та їх застосуванні для спостереження та інтелектуального дослідження Землі та підтримки прийняття рішень.



Рис. 1. Підрозділи штучного інтелекту [15].

Існують наступні основні методи машинного навчання [16]:

*Навчання з вчителем* (supervised learning) передбачає використання маркованих даних для навчання алгоритмів, що дозволяє їм прогнозувати результати. Цей підхід охоплює регресію, яка має справу з безперервними даними, і класифікацію дискретних даних. Успіх навчання з учителем значною мірою залежить від доступу до великих наборів даних. Воно знаходить

## 1.1. Аналіз методів та моделей глибинного навчання для розв'язання ...

---

застосування в різних галузях, включаючи розпізнавання зображень і мови, медичну діагностику і спостереження Землі для розв'язанні таких задач, як класифікація ґрунтового покриву, картографування рослинності і виявлення об'єктів.

Навчання без вчителя (unsupervised learning) використовує немарковані дані для виявлення закономірностей і групування даних у кластери без конкретних прогнозів. Алгоритми, що використовуються в навчанні без вчителя, включають кластеризацію, зменшення розмірності та виявлення аномалій. Застосування навчання без вчителя можна знайти в задачах супутникового моніторингу, таких як виявлення аномалій для ідентифікації нестандартних ситуацій на супутникових знімках, класифікація для групування схожих регіонів при картографуванні ґрунтового покриву і зменшення розмірності для обробки великомасштабних супутникових даних.

*Напівкеруване навчання* (semi-supervised learning) використовує переваги як невеликих маркованих наборів даних, так і великих немаркованих наборів даних. Такий підхід дозволяє поєднати інформацію з маркованих даних із закономірностями, виявленими у немаркованих даних, для покращення навчання моделі. У супутниковому моніторингу цей метод може бути корисним для розв'язання таких задач, як картографування ґрунтового покриву, де обмежені марковані зразки можуть бути доповнені великим обсягом немаркованих супутникових знімків для підвищення точності та узагальнення.

*Навчання з підкріпленням* (reinforcement learning) передбачає, що агент навчається через взаємодію з навколишнім середовищем методом спроб і помилок. Він отримує винагороду за дії, які наближають його до заздалегідь визначеної мети. Агент приймає рішення на основі минулого досвіду (використання) і пробує нові дії (дослідження), щоб оптимізувати процес навчання. У супутниковому моніторингу навчання з підкріпленням може застосовуватися для розв'язання таких задач, як автономна навігація дронів для збору конкретних даних, оптимізація орбіт супутників для покращення збору даних або адаптивне планування роботи датчиків для отримання релевантної інформації.

*Глибинне навчання* - це потужна підгрупа машинного навчання, яка використовує нейронні мережі з декількома шарами для створення ієрархічних представлень і моделювання складних взаємозв'язків у даних. Здатність ефективно обробляти величезні обсяги даних робить його добре придатним для різних застосувань, в тому числі в супутниковому моніторингу. Глибинне навчання

використовується для таких задач, як аналіз зображень і відео, виявлення об'єктів, семантична сегментація аеро- або супутникових знімків, класифікація землекористування і рослинного покриву, виявлення змін і навіть моніторинг змін у навколишньому середовищі, таких як вирубка лісів або зростання міст. Нещодавні досягнення в галузі глибинного навчання значною мірою зумовлені зростанням обчислювальних потужностей і доступністю великих масивів даних супутникового моніторингу.

У наступних розділах розглянуті кілька часто використовуваних архітектур глибинного навчання для розв'язання задач супутникового спостереження Землі, описані їх застосування та корисність у цій галузі. Ми починаємо з керованих згорткових і рекурентних нейронних мереж, за якими слідує генеративні змагальні мережі для навчання без вчителя. Нарешті, ми досліджуємо найбільш популярні моделі трансформерів і їх застосування. Це висвітлення провідних підходів глибокого навчання дає контекст щодо їх ролі в просуванні досліджень і практики супутникового моніторингу. Маючи ці передумови, розглядаються також виклики, що залишилися, і майбутні перспективи глибинного навчання для супутникового моніторингу.

### **1.3. ОГЛЯД МОДЕЛЕЙ ТА МЕТОДІВ ГЛИБИННОГО НАВЧАННЯ ДЛЯ СУПУТНИКОВОГО МОНІТОРИНГУ**

Глибинне навчання — це форма машинного навчання, побудована на штучних нейронних мережах [17] із кількома прихованими шарами, що дозволяє моделям вивчати складні шаблони з великих обсягів даних. Глибина мережі дозволяє вивчати абстрактні представлення та вилучати ієрархічні ознаки безпосередньо з необроблених даних [18], усуваючи необхідність розробки функцій вручну.

Поширені архітектури глибокого навчання показані на рис. 2 і включають згорткові нейронні мережі (CNN) для додатків комп'ютерного зору [19], таких як класифікація зображень, рекурентні нейронні мережі (RNN) для обробки послідовних даних, таких як текст і мова [20], і генеративні змагальні мережі (GAN) [21] для генерування нових даних. Однією з популярних архітектур, яка часто використовується для навчання за подібністю, є сіамська нейронна мережа [22]. Зовсім недавно

## 1.1. Аналіз методів та моделей глибокого навчання для розв'язання ...

моделі-трансформери, що використовують механізми уваги, досягли найсучасніших відомих результатів у задачах обробки природної мови [23].

Глибине навчання вимагає великих наборів маркованих даних і значних обчислювальних потужностей, до складу яких часто входять графічні процесори. Воно дозволило просунутись вперед у таких сферах, як комп'ютерний зір [24], обробка мовлення [25] та обробка природної мови [26]. Однак моделі глибокого навчання можуть бути складними «чорними скриньками», яким бракує інтерпретованості та прозорості [27].

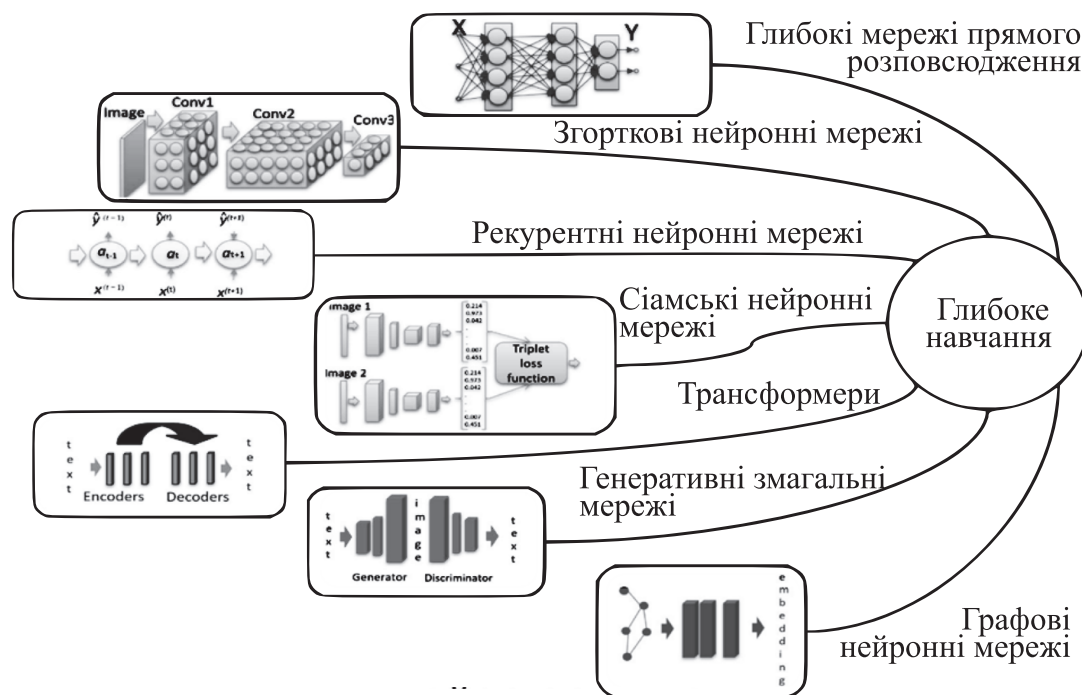


Рис. 2. Глибинні мережі [15]

Наступні розділи зосереджені на представленні декількох архітектур, що часто використовуються у супутниковому моніторингу. До них відносяться керовані моделі згорткових нейронних мереж (CNN) і рекурентних нейронних мереж (RNN), а також моделі генеративних змагальних мереж (GAN) і все більш популярні трансформери.

### 1.4. ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ (CNN)

Згорткові нейронні мережі (CNN) - це спеціалізований тип глибоких нейронних мереж, який широко використовується для аналізу зображень. Ключовою властивістю CNN є використання

згорткових шарів для застосування фільтрів до вхідних даних, які виявляють просторові шаблони, такі як краї, і створюють карти ознак. Кілька згорткових шарів об'єднуються в стек для вивчення ієрархічних представлень ознак. Шари об'єднання використовуються для зменшення вибірки карт ознак, зменшуючи обчислювальне навантаження та попереджуючи перенавчання, причому найчастіше використовується максимальне об'єднання шарів. Повнозв'язні шари слідує за згортковими шарами для отримання остаточних оцінок класів або прогнозування. CNN повторно вимагають великі марковані навчальні дані для налаштування фільтрів і параметрів моделі (рис. 3). Вони використовують інваріантність перетворення та композиційність візуального сприйняття для ефективного навчання ознак з пікселів. CNN знаходять різні застосування, такі як класифікація зображень, виявлення об'єктів, семантична сегментація, генерація зображень і навігація роботів. Серед популярних архітектур CNN - LeNet [28], AlexNet [29], VGG [30], ResNet [31], Inception [32], YOLO [33] та інші.

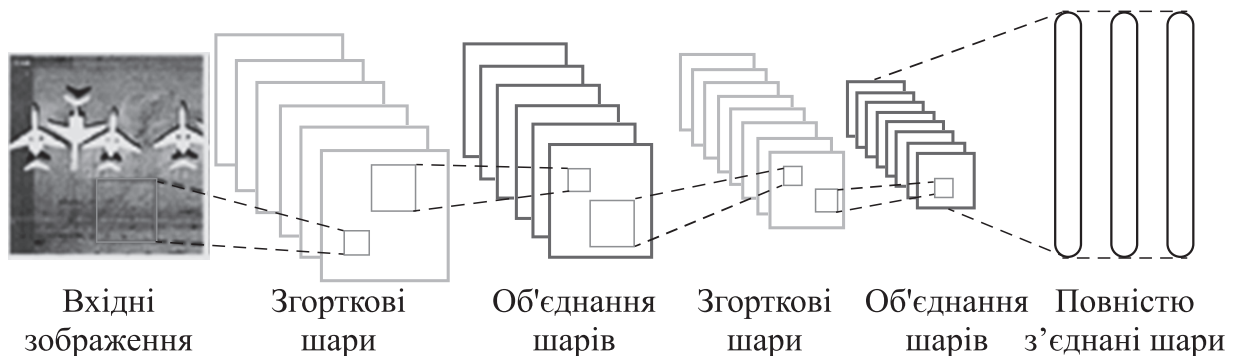


Рис. 3. Структура CNN, що складається із згорткових, об'єднаних і повнозв'язаних шарів [34]

Згорткові нейронні мережі (CNN) стали домінуючою технологією для розв'язання різних задач супутниковому моніторингу завдяки їх ефективності для аналізу супутникових даних. CNN знаходять застосування у супутникових спостереженнях завдяки наступним можливостям:

- Класифікація зображень: CNN відмінно справляються з класифікацією супутникових зображень за різними категоріями земного покриття або землекористування, а також з ідентифікацією конкретних об'єктів. Відомі архітектури, такі як ResNet і VGGNet, дають змогу отримувати високоякісні результати на основі наборів даних аеро- і супутникових знімків [35].



## 1.1. Аналіз методів та моделей глибокого навчання для розв'язання ...

- Виявлення об'єктів: CNN у поєднанні з такими методами, як регресія з обмежувальною рамкою і блоків прив'язок, дозволяють виявляти такі об'єкти, як транспортні засоби, кораблі і будівлі на аерознімках. Швидша R-CNN [36] є популярною моделлю CNN для виявлення об'єктів.

- Семантична сегментація: Використовуючи повні згорткові мережі [37], зображення супутникового моніторингу можна сегментувати на рівні пікселів, присвоюючи кожному пікселю мітку класу. U-Net і SegNet є загальнозживаними архітектурами CNN для розв'язання цієї задачі [38] (рис. 4).

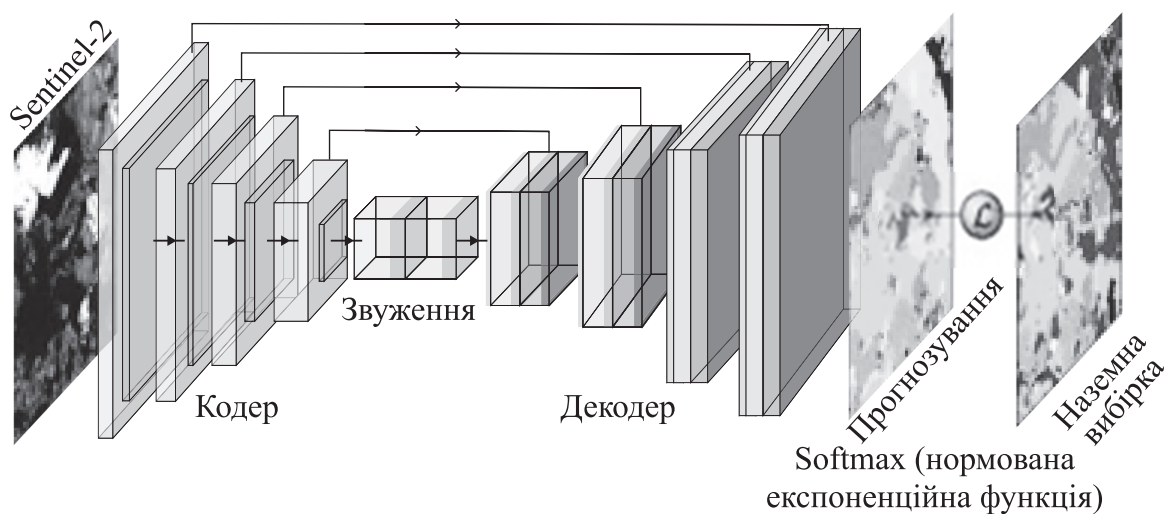


Рис. 4. Архітектура U-Net [39]

- Виявлення змін: Сіамські CNN (рис. 5), що включають дві вхідні гілки, дозволяють ефективно виявляти зміни між різночасовими зображеннями, що робить їх цінними для ідентифікації районів, які постраждали від стихійних лих [40-42].

- Злиття зображень: CNN використовуються для злиття мультиспектральних і панхроматичних зображень, створюючи кольорові зображення високого просторового розрізнення [44].

- Надвисоке просторове розрізнення: CNN, такі як SRCNN, дозволяють підвищувати просторове розрізнення, покращуючи якість зображення. SRCNN, або згорткова нейронна мережа для підвищення просторового розрізнення, - це новаторська архітектура згорткової нейронної мережі (CNN), що використовується для отримання надвисокого просторового розрізнення для одного зображення [45, 46]. Вона була запропонована в 2014 році і стала однією з перших CNN, застосованих до розв'язання задач відновлення зображень з

надвисоким просторовим розрізненням Архітектура SRCNN складається з трьох згорткових шарів - виділення фрагментів, нелінійного відображення та реконструкції (рис. 6).

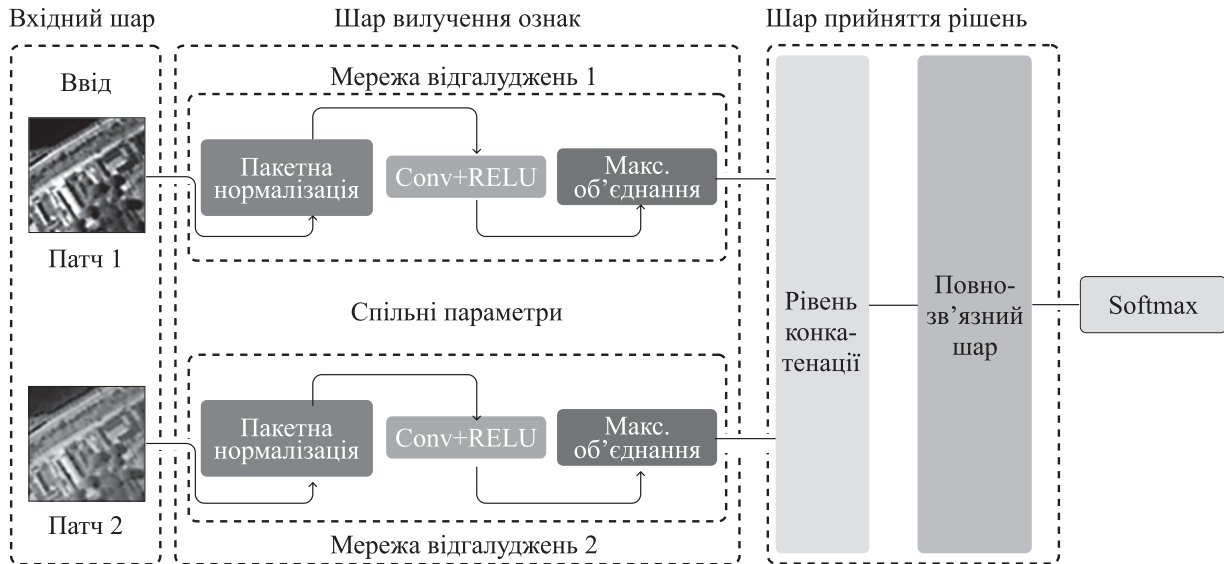


Рис. 5. Структура сіамської мережі [43]

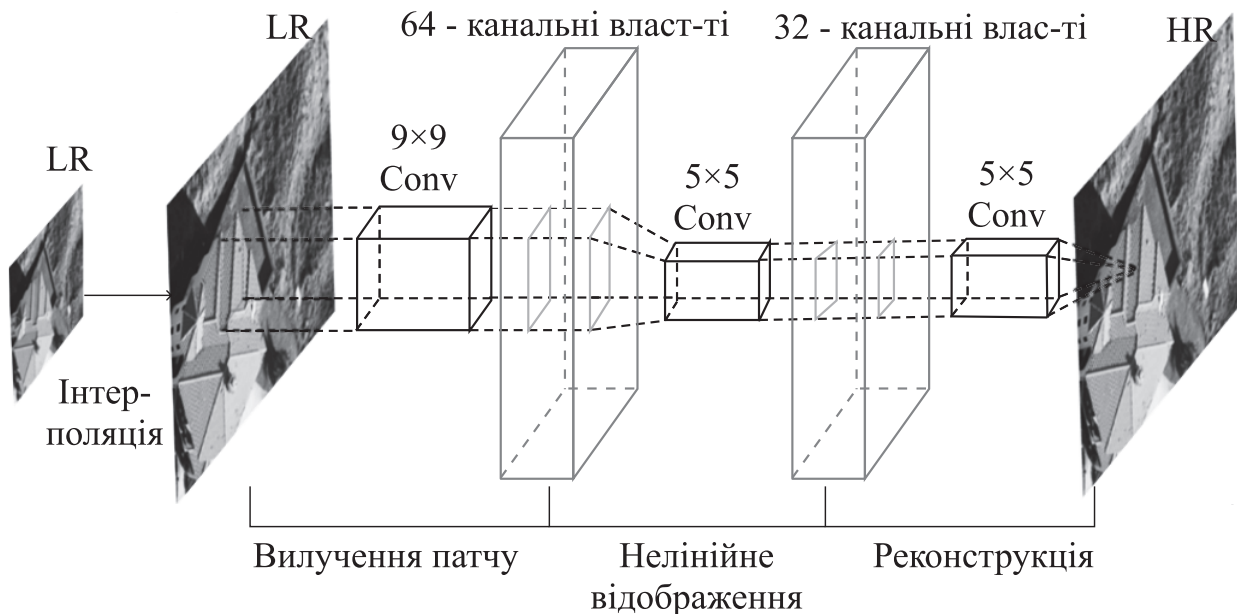


Рис. 6. Ескіз архітектури SRCNN [48]

Вона продемонструвала найвищу продуктивність порівняно з попередніми методами підвищення просторового розрізнення, заснованими на прикладах та розрідженому кодуванні. SRCNN поступово масштабує зображення низького просторового розрізнення, навчаючись наскрізному відображенню від низького

до високого просторового розрізнення. Це продемонструвало, що CNN можуть ефективно вирішувати проблеми низькорівневої обробки зображень, розширюючи сферу їх застосування за межі задач технічного зору високого рівня. Архітектура SRCNN зумовила підвищення інтересу до використання глибинного навчання для інших задач відновлення зображень, таких як фільтрація шуму та зменшення розмитості. Згодом на основі архітектури та методології навчання SRCNN були розроблені вдосконалені CNN з надвисоким просторовим розрізненням [47].

- Виявлення аномалій: CNN можуть вивчати закономірності супутникових знімків та ідентифікувати області, що відхиляються від норми, вказуючи на потенційну аномалію [49, 50].

Згорткові нейронні мережі (CNN) були адаптовані для обробки даних з тривимірних хмар точок, отриманих з LiDAR-сенсорів, у таких задачах, як класифікація, сегментація та ідентифікація об'єктів [51, 52]. Таку адаптацію можна здійснити кількома методами. Об'ємні CNN [53] використовують представлення хмари точок у вигляді тривимірної воксельної сітки, яка обробляється стандартними 3D CNN. Однак обмеження пам'яті може призвести до низької роздільної здатності. Багаторакурсні CNN [54] генерують декілька 2D-проекцій з хмари точок, які потім подаються в звичайні 2D CNN. Агрегування ознак дає прогноз для кожної точки.

PointNet [55] безпосередньо споживає сирі хмари точок, використовуючи MLP з максимальним об'єднанням для вивчення точкових ознак, інваріантних до перестановки вхідних даних, що робить його новаторським підходом. Локальне об'єднання ознак застосовує стандартні CNN на локальних фрагментах області навколо кожної 3D-точки з подальшим об'єднанням ознак. Графові CNNs [56] моделюють хмару точок як граф, з точками як вузлами, і використовують згортки графа над локальними областями для вивчення ознак.

Неперервні згортки визначають згортки на неперервній області з використанням радіальних базисних функцій з центром у точках. Кореляція ядра та деформація вивчають ваги ядра та поля деформації, щоб відповідати різним локальним геометріям. Вісімкове розбиття ієрархічно ділить простір застосовує 3D CNN до вісімкових вокселів, що містять точки [57].

## 1.5. РЕКУРЕНТНІ НЕЙРОННІ МЕРЕЖІ (RNN)

Рекурентні нейронні мережі (RNN) - це спеціалізовані штучні нейронні мережі, призначені для обробки послідовних даних, таких як текст, аудіо, часові ряди та відео.

Ключовими характеристиками RNN є циклічні зв'язки, які зберігають інформацію в послідовності, забезпечуючи динамічну поведінку в часі. RNN, такі як LSTM [58] та GRU [59], дозволяють фіксувати довгострокові залежності та обробляти послідовності різної довжини, рекурсивно застосовуючи рекурентну функцію на кожному часовому кроці (рис. 7).

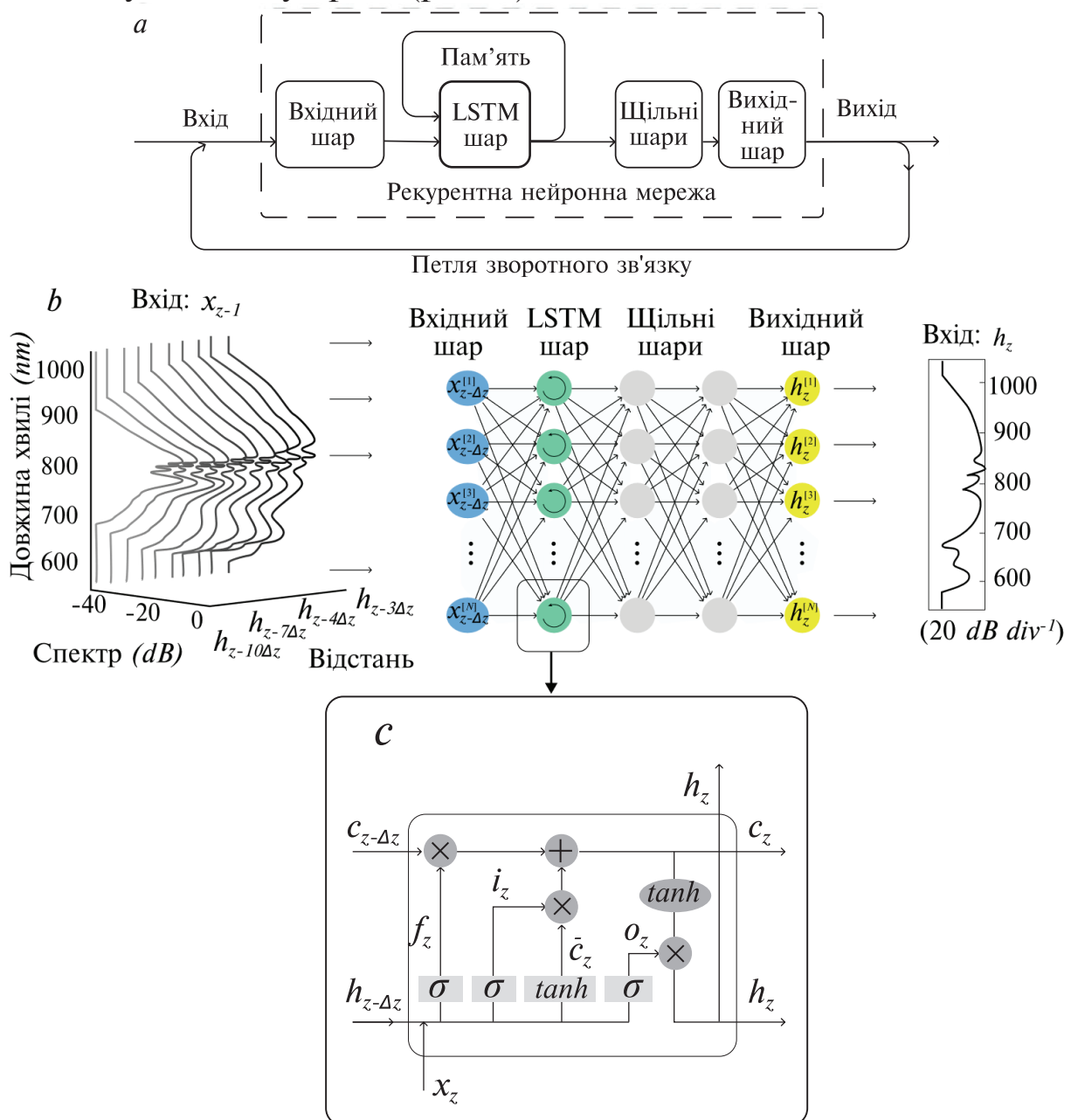


Рис. 7. Рекурентна нейронна мережа [67]

## 1.1. Аналіз методів та моделей глибокого навчання для розв'язання ...

---

RNN знаходять широке застосування у моделюванні мови [60], розпізнаванні мовлення [61], генерації тексту [62], машинному перекладі [63], субтитруванні зображень/відео [64] та прогнозуванні часових рядів [65]. Однак під час навчання на довгих послідовностях вони можуть зіткнутися з такими проблемами, як зникаючі та вибухові градієнти. Для покращення фокусування на релевантному вхідному контексті до RNN часто додають механізми уваги [66].

Це питання буде розглянуто в підрозділі 1.8. RNN є цінним інструментом для моделювання послідовних даних, використовуючи їх короткочасну пам'ять і динамічну часову поведінку, отриману завдяки циклічним зв'язкам.

Рекурентні нейронні мережі (RNN) знаходять застосування в різних задачах супутникового моніторингу, які вимагають моделювання послідовних даних. До таких застосувань відносяться:

- Аналіз часових рядів: RNN, особливо варіанти LSTM, використовуються для прогнозування врожайності сільськогосподарських культур, тенденцій вирубки лісів, кліматичних змін тощо на основі часових рядів супутникових знімків [68, 69].

- Відстеження об'єктів: RNN можуть моделювати рухи та появу об'єктів на послідовних відеоданих супутникового моніторингу [70], що дозволяє відстежувати автомобілі, кораблі, літаки тощо.

- Виявлення змін: RNN допомагають виявляти зміни в ландшафтах, міській інфраструктурі тощо з плином часу, використовуючи послідовні супутникові зображення [71].

- Підписи до зображень: RNN дозволяють генерувати текстові описи вмісту та активностей на супутникових знімках [72], наприклад, щодо змін берегової лінії.

- Аналіз хмари точок: Застосовується до послідовного сканування LiDAR-даних. RNN аналізують динамічні хмари точок для таких задач, як 4D-сегментація [73].

- Генерація зображень: умовні RNN, такі як LSTM-GAN, можуть створювати реалістичні супутникові зображення, вивчаючи просторові кореляції [74]. Архітектура GAN буде розглянута в підрозділі 1.6.

- Виявлення аномалій: RNN вивчають нормальні часові закономірності в послідовних даних повторного супутникового моніторингу, що дозволяє виявляти аномальні події [75].

- Класифікація послідовності зображень: RNN класифікують відеокадри спостереження землі або фрагменти супутникових знімків за категоріями землекористування [76].

- Видалення хмар: RNN вивчають хмарні патерни на часових рядах супутникових знімків для реконструкції чистих композитних зображень [77].

Підсумовуючи, архітектури RNN, зокрема LSTM, є цінними для розв'язання задач спостереження Землі, пов'язаних з динамікою в часі, завдяки їх здатності моделювати довготривалий контекст і пам'ять, що робить їх придатними для послідовного моделювання.

## 1.6. ГЕНЕРАТИВНІ ЗМАГАЛЬНІ МЕРЕЖІ

Генеративні змагальні мережі (GAN) [78] - це тип алгоритмів ШІ, які використовуються в машинному навчанні без вчителя (рис. 8). GAN складаються з генератора та дискримінатора, які змагаються в грі з нульовою сумою [79]. Генератор виробляє синтетичні дані, тоді як дискримінатор відрізняє справжні зразки від фальшивих. Завдяки цьому змагання генератор набуває досвіду у створенні все більш реалістичних синтетичних даних. Застосування GAN включають створення фотореалістичних зображень [80], синтетичного звуку [81], перетворення зображення у зображення [82], генерування зображень з тексту [83], виявлення аномалій [84] та доповнення даних [85]. GAN пропонують потужний підхід для реалізації навчання без вчителя, що дозволяє генерувати реалістичні синтетичні дані в різних прикладних областях.

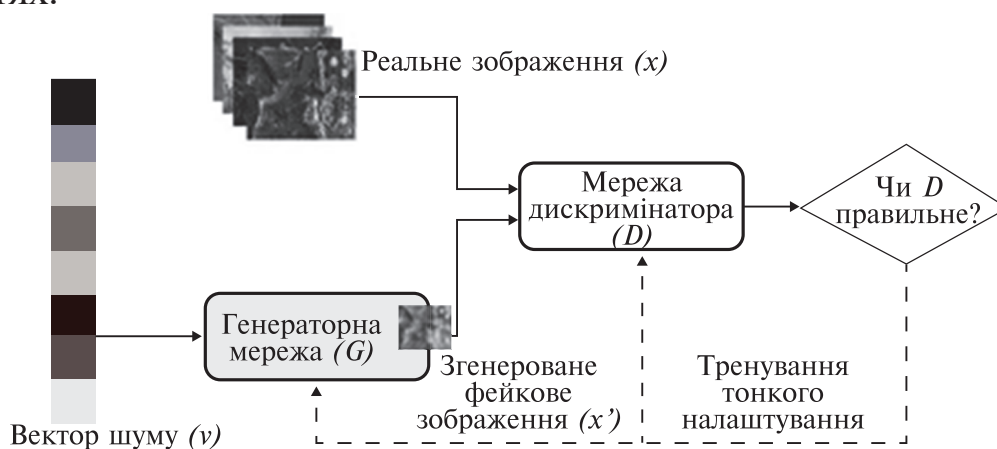


Рис. 8. Загальна ілюстрація структури генеративної змагальної мережі (GAN) [86]

## 1.7. ВИКОРИСТАННЯ ГЕНЕРАТИВНИХ ЗМАГАЛЬНИХ МЕРЕЖ У СУПУТНИКОВОМУ МОНІТОРИНГУ

Мережі GAN знайшли різноманітні застосування в супутниковому моніторингу, надаючи інноваційні рішення для розв'язання задач аналізу та інтерпретації зображень. Нижче наведено деякі задачі в цій галузі, де GAN ефективно використовуються.

**Покращення зображення.** GAN відіграють вирішальну роль у підвищенні якості супутникових зображень із низьким просторовим розрізненням. Генеруючи додаткові реалістичні деталі та покращуючи просторове розрізнення, GAN дозволяють аналітикам отримувати чіткішу інформацію із супутникових зображень, допомагаючи точніше тлумачити та приймати рішення. Розкриваючи дрібніші деталі, технології суперроздільності на основі GAN дозволяють покращити візуалізацію та ідентифікацію таких об'єктів, як транспортні засоби, дерева та дороги. Зокрема, проблема була ефективно вирішена в [87]. У дослідженні пропонується методологія підвищення просторового розрізнення (супер-розрізнення) існуючих зображень із низьким просторовим розрізненням до зображень із вищим просторовим розрізненням. Основна ідея полягає в тому, щоб використовувати передові методи глибокого навчання - Generative Adversarial Networks (GAN) і навчити їх збільшувати просторове розрізнення супутникових зображень. Експериментальні результати для даних Sentinel-2 показали, що цей підхід ефективний і може бути використаний для створення продуктів надвисокого просторового розрізнення.

**Видалення хмар і заповнення прогалін.** Хмари часто закривають важливі ділянки супутникових зображень, перешкоджаючи їх аналізу. GAN можна навчити на безхмарних прикладах для створення безхмарних версій зображень, ефективно видаляючи хмарний покрив і відкриваючи приховану інформацію. У [88] автори пропонують метод видалення хмар для одного супутникового зображення. Складність цієї задачі полягає в труднощах відрізнити хмари від білих об'єктів на земній поверхні та коректно заповнити нерівні відсутні області. Автори пропонують двоетапний алгоритм. Перший етап – це сегментація хмар, тобто виділення хмар і видалення тонких хмар (димки) безпосередньо за допомогою U-Net. Другий етап – відновлення зображення, тобто видалення густої хмари та відновлення відповідних нерегулярних відсутніх областей за допомогою

генеративної змагальної мережі GAN. Подібним чином GAN є ефективними для заповнення відсутніх частин супутникових зображень, які є результатом несправності датчика або поганого покриття. Зокрема, у [89] представлено нову нейронну систему, засновану на умовних глибоких генеративних змагальних мережах (сGAN), оптимізованих для заповнення прогалів супутникових зображень, використовуючи сусідні значення. Достовірно доповнюючи відсутній контент, GAN дозволяють здійснювати більш повний і безперервний моніторинг областей, що становлять інтерес.

**Виявлення аномалії.** Використовуючи GAN, можна ефективно виявляти аномалії в супутникових даних. GAN можуть вивчати шаблони нормальності із супутникових даних, і відхилення від цих шаблонів можуть вказувати на аномалії, такі як незаконне вирубування лісів, будівництво чи забруднення, допомагаючи своєчасно та точно ідентифікувати потенційні проблеми. Запропонований метод на основі GAN [90] було апробовано на мультитимових зображеннях Landsat за 2000-2019 роки для посушливого регіону в Алжирі. Він демонструє кращу продуктивність у виявленні змін опустелювання порівняно з DBN, DBM, CNN та іншими методами машинного навчання.

**Симуляція.** GAN дозволяють генерувати реалістичні змодельовані супутникові зображення для сценаріїв, коли реальних даних бракує або їх недостатньо. Ці змодельовані зображення можуть бути цінними для розширення наборів даних, дозволяючи дослідникам виконувати тести та аналізи в різних гіпотетичних ситуаціях. Автори [91] пропонують використовувати GAN для моделювання/симуляції географічних процесів, оцінки ефективності наборів даних для задачі Game of Life і розробки більш складних її варіантів як орієнтирів для майбутніх досліджень.

**Реєстрація зображень.** GAN використовуються для вирівнювання зображень, отриманих у різний час або з різних датчиків/кутів, забезпечуючи узгодженість даних. Таке вирівнювання підвищує точність виявлення змін і дозволяє проводити надійний аналіз часових рядів. У роботі [92] автори запропонували підхід на основі GAN для вирішення проблеми кореєстрації оптичних і SAR зображень. Спочатку вони використовували мережу на основі GAN для генерації псевдо-радарного зображення з оптичного зображення, щоб забезпечити порівнянну інформацію про інтенсивність між двома зображеннями. Після цього застосувались методи, засновані на



площі та ознаках, для вирівнювання псевдо-радарного зображення з автентичним зображенням. Результати експерименту продемонстрували незамінність цього методу, особливо підкресливши ефективність генератора GAN. Зокрема, генератор GAN точно і надійно відтворює геометричну структуру оптичного зображення, відкриваючи тим самим нові можливості для MMRS (мультимодального супутникового моніторингу).

**Виявлення об'єктів.** GAN сприяють покращенню виявлення і класифікації об'єктів на супутникових знімках, таких як будівлі, дороги, літаки і кораблі. Використовуючи методи виявлення об'єктів на основі GAN, аналітики можуть отримати більш детальну і точну інформацію про об'єкти, присутні на знімках. У [93] було представлено SERGAN, генеративну змагальну мережу GAN з покращенням країв зображення. В цю архітектуру було включено модуль покращення країв зображень, щоб підкреслити краї об'єктів на зображеннях з надвисоким просторовим розрізненням (SR), тим самим сприяючи високоточному виявленню об'єктів.

**Доповнення (аугментація) навчальної вибірки.** При розв'язанні задач машинного навчання на супутникових даних часто виникає проблема незбалансованості навчальної вибірки, яку неможливо подолати традиційними підходами, оскільки вони не враховують фізичну природу багатоспектральних зображень а також їх семантичний контекст. В роботі [94] запропоновано метод аугментації навчальної вибірки для задач семантичної сегментації супутникових даних на основі GAN. Запропонований метод враховує спектральні і текстурні характеристики різних типів земного покриття і дозволяє підвищити точність класифікації міноритарних класів на 10% відсотків

## 1.8. ТРАНСФОРМЕРИ

Трансформери - це передова неймережева архітектура, яка набула значної популярності в глибокому навчанні, зокрема для задач обробки природної мови (NLP). Трансформери були представлені у 2017 році [5], авторами якої були дослідники з проєкту Google Brain. На відміну від рекурентних нейронних мереж (RNN), трансформери значною мірою покладаються на механізм уваги для встановлення глобальних залежностей між входом і виходом. Архітектура трансформерів показана на рис. 9.

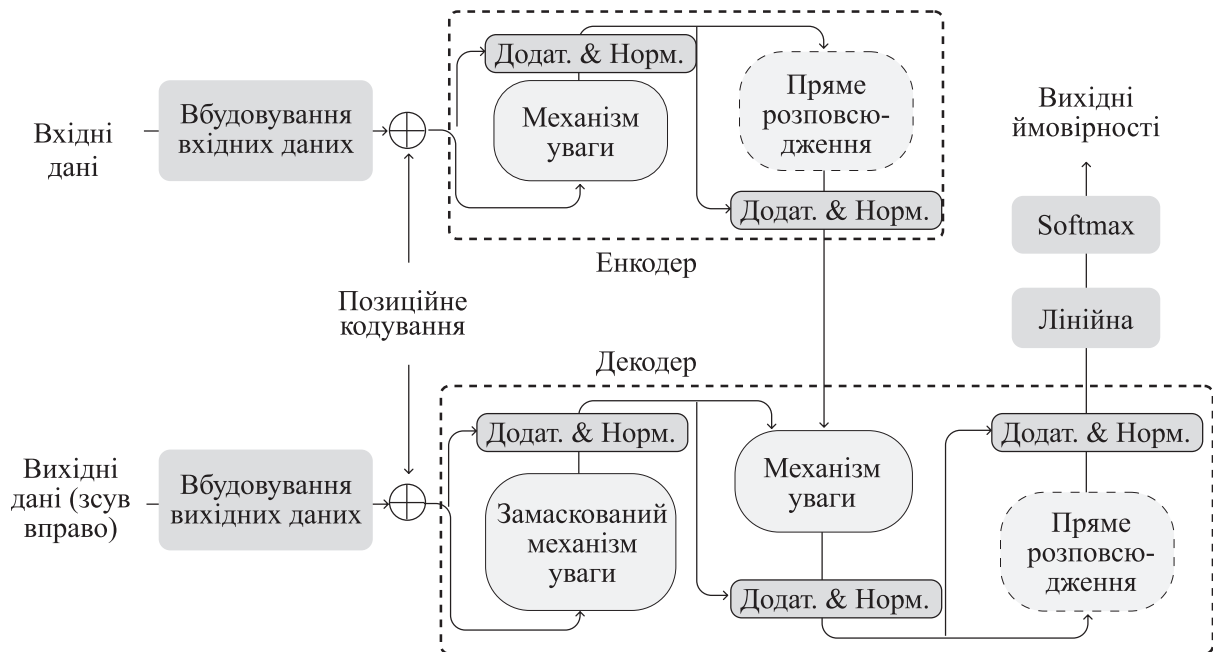


Рис. 9. Архітектура трансформерів [23]

Трансформери не використовують згортку чи рекурентність, працюючи лише з рівнем уваги. Вони розглядають кожну точку даних як незалежну від інших. Таким чином, позиційна інформація додається до моделі явно, щоб зберегти інформацію про порядок або розташування. Енкодер трансформера обробляє всю вхідну послідовність одразу, генеруючи вхідні представлення. У завданнях NLP декодер генерує вихідну послідовність токен за токеном. Важливою особливістю є здатність до багатосторонньої уваги, що дозволяє моделі фокусуватися на різних позиціях у вхідній послідовності. Щоб включити інформацію про положення токенів у послідовності, трансформери використовують позиційне кодування. Відсутність рекурентності дозволяє трансформерам навчатися паралельно, що робить їх більш ефективними, ніж RNN. Трансформери досягли найсучасніших результатів у задачах перекладу, перевершивши показники RNN і CNN, що значно сприяло їхній популярності. Серед відомих моделей трансформерів - BERT [95], GPT-2/3 [7, 8], RoB-ERTa [96] і T5 [97], які сприяли просуванню вперед у галузі NLP.

Трансформери набувають все більшої популярності в різних застосуваннях супутникового моніторингу завдяки їх здатності моделювати довгострокові залежності в послідовних даних. У контексті супутникового моніторингу трансформери довели свою ефективність у кількох сферах:

Задачі класифікації зображень виграють від використання трансформерів, оскільки вони розглядають вхідні зображення як

## 1.1. Аналіз методів та моделей глибокого навчання для розв'язання ...

---

послідовності ділянок. Ця адаптація, відома як архітектура Vision Transformer (ViT) [98], продемонструвала високу продуктивність у тестах класифікації зображень.

Розв'язки задачі детекції об'єктів можна покращити за допомогою використання трансформерів детекції (Detection Transformers, DETR) [99], які замінюють традиційну архітектуру CNN і ручні етапи постобробки на архітектуру трансформерного кодера-декодера. Це дозволяє використовувати наскрізне навчання для ідентифікації об'єктів. Для семантичної сегментації зображень трансформери можуть бути адаптовані за допомогою підходу Semantic Segmentation Transformer (SETR) [100]. Тут трансформерний декодер передбачає попиксельні мітки, підвищуючи точність семантичної сегментації. Трансформери також цінні в задачах підвищення роздільної здатності [101].

Механізм уваги, притаманний трансформерам, виявляється корисним для виявлення аномалій на супутникових зображеннях [102], оскільки він дозволяє ідентифікувати патерни аномалій. У сфері підписів до супутникових зображень архітектури на основі трансформерів, такі як Multimodal Attention Transformer (MAT) [103], здатні генерувати контекстні підписи. Аналіз хмари точок на основі даних LiDAR також може отримати вигоду від трансформерів [104], оскільки вони можуть безпосередньо обробляти тривимірні хмари точок для задач класифікації та сегментації.

Загалом, трансформери є ефективним архітектурним рішенням для моделювання просторово-часових зв'язків у супутникових даних. Їх механізм уваги добре підходить для обробки різномасштабних і мультимодальних даних спостереження Землі, що сприяє їх зростаючому використанню в цій галузі.

## 1.9. ГРАФОВА НЕЙРОННА МЕРЕЖА (GNN)

Графова нейронна мережа (ГНМ, GNN) [104] - це модель глибокого навчання, призначена для аналізу даних, структурованих у вигляді графів. Графи складаються з вузлів (вершин) і ребер (зв'язків між вузлами), а ГНМ використовують властивості вузлів і зв'язки між ребрами для вилучення важливої інформації. Ці моделі безпосередньо оперують даними графа,

поширюючи інформацію про властивості вузлів по ребрах для аналізу і навчання.

Існують різні архітектури GNN (рис. 10), включаючи графові згорткові мережі (GCN) [105], графові мережі уваги (GAT) [106] та графові рекурентні нейронні мережі (GRNN) [107], [108]. GNN запам'ятовують представлення вузлів, які кодуєть як ознаки вузлів, а також топологію графа, що лежить в основі, фіксує контекст мережі у вкладених вузлах.

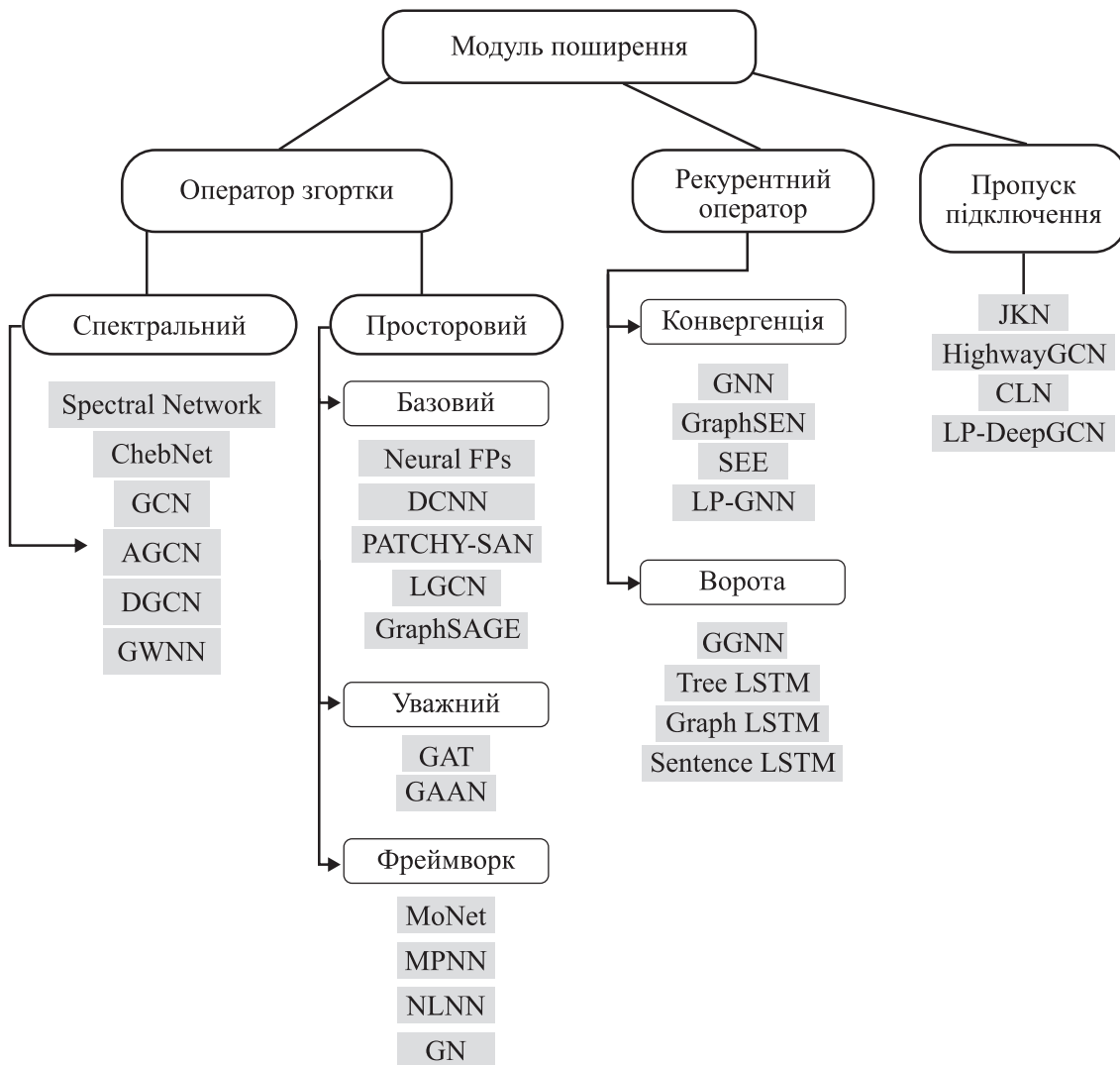


Рис. 10. Огляд обчислювальних модулів [109]

Для задач класифікації вузлів, GNN розглядають сусідні вузли та ребра, щоб зробити прогнозування для даного вузла. У задачах класифікації графів GNN вивчають репрезентації для цілих графів, щоб класифікувати їх.

GNN мають можливість динамічно оновлювати представлення вершин, пропонуючи та трансформуючи ознаки на ребрах, що дозволяє їм ефективно моделювати залежності.

Ці моделі знаходять застосування в різних галузях, таких як аналіз соціальних мереж [110], прогнозування властивостей молекул [111], рекомендаційні системи [112], графи знань [113] та прогнозування трафіку [114].

Графові нейронні мережі (ГНМ) знайшли різноманітне застосування в супутниковому моніторингу та геопросторовому аналізі завдяки своїй здатності моделювати дані у вигляді графів і ефективно фіксувати взаємозв'язки. Ось деякі застосування ГНМ у спостереженні землі.

**1. Сценарні графи.** GNN моделюють супутникові зображення у вигляді сценарних графів, представляючи об'єкти як вузли, а зв'язки між ними - як ребра. Такий підхід дозволяє отримувати контекстну інформацію, покращуючи загальне розуміння та аналіз. Наприклад, в [115] розглядається проблема класифікації багатомісних сцен на основі супутникових зображень з дуже високим просторовим розрізненням (VHR) шляхом дослідження глибокої графової згорткової мережі (ГЗМ). Оскільки задана сцена надвисокого розрізнення містить декілька локальних ознак, традиційні системи класифікації з однією міткою не передають справжньої семантики сцени.

З іншого боку, очікується, що підходи до класифікації за кількома мітками допоможуть краще охарактеризувати область, що розглядається. В межах багатокласової класифікації дане зображення представляється у вигляді графа суміжності регіонів (RAG) відповідних локальних областей.

**2. Сегментація.** GNN покращують сегментацію зображень, розглядаючи пікселі або суперпікселі як вузли графа і поширюючи інформацію між ними на основі близькості та схожості. Це допомагає досягти більш точних і послідовних результатів сегментації. В [116] автори пропонують підхід з використанням графових згорткових мереж (GCN) на основі суперпікселів для семантичної сегментації SAR-зображень. Суперпікселі та їх характеристики використовуються для побудови графа, де суперпікселі є вузлами, а ребра відображають просторові відношення. Поєднуючи суперпікселі, глибоке вилучення ознак і GCN, запропонований підхід спрямований на використання локальних ознак зображення, а також нелокальних просторових зв'язків для точної сегментації SAR-зображень. Експерименти на

даних Sentinel-1 демонструють покращену продуктивність порівняно з іншими методами.

**3. Класифікація.** Інтеграція шарів GNN у згорткові нейронні мережі (CNN) покращує класифікацію ґрунтового покриву та типів культур. Доповнюючи спектральні дані просторовими зв'язками, GNN сприяють підвищенню ефективності класифікації. Автори статті [117] пропонують новий метод класифікації гіперспектральних зображень під назвою FCGN (feature fusion of convolutional neural network and graph convolutional network). Вони об'єднують ознаки з двох гілок, щоб отримати багатший набір ознак, які включають як локальні деталі, так і глобальну контекстну інформацію. Вони демонструють покращену точність класифікації порівняно з використанням лише GCN або CNN на стандартних гіперспектральних наборах даних, таких як Indian Pines, Pavia University та Salinas.

**4. Виявлення об'єктів.** GNN можна застосовувати для виявлення об'єктів на аерознімках, формулюючи задачу як класифікацію вершин графа, враховуючи навколишній контекст. Ребра графа з'єднують об'єкти і фонові області, допомагаючи в локалізації об'єктів. В [118] запропоновано нову гібридну багатопотокову ієрархічну мережу вбудовування графів, керовану увагою (HA-MHGEN) для виявлення багатокласових геопросторових об'єктів на супутникових зображеннях з високим просторовим розрізненням. Метод будує ієрархічні просторові графи для представлення різномасштабних просторових відношень між об'єктами та семантичні графи, що інтегрують словесні вставки міток об'єктів для відображення контекстних семантичних відношень. Багатомасштабна мережа згорткових графів використовує ці графи для вивчення внутрішньо- та міжоб'єктних просторових і семантичних зв'язків. Модуль перехресної уваги об'єднує просторові та семантичні ознаки, дозволяючи моделі вловлювати явні та неявні зв'язки з обох доменів. Комплексні експерименти на наборах даних DOTA, DIOR та NWPU VHR-10 демонструють ефективність запропонованого методу порівняно з іншими сучасними підходами до виявлення об'єктів супутникового моніторингу. Модель досягає вищої продуктивності завдяки спільному навчанню взаємодоповнюючих просторових та семантичних патернів.

**5. Виявлення змін.** Пари часових зображень можуть бути представлені у вигляді графів, що дозволяє GNN виявляти зміни між часовими кроками на основі модифікацій структури графа та

вузлових ознак. Це полегшує моніторинг та аналіз динамічних середовищ. Автори [119] пропонують наскрізну динамічну нейронну мережу на рівні графа (DGLNN) для виявлення змін на радарних зображеннях із синтезованою апертурою (SAR). DGLNN використовує локальну структуру кожного сусіднього пікселя на рівні графа для навчання дискримінативного графового представлення. Він використовує підхід  $k$ -найближчих сусідів для реконструкції ребер між вузлами замість фіксованих ребер, що дозволяє кожному вузлу використовувати ознаки різних сусідів. Експерименти на міждомених зображеннях SAR показують, що DGLNN досягає найвищої продуктивності порівняно з іншими методами глибинного навчання для виявлення змін SAR. Графовий підхід ефективно справляється з шумом і відсутністю явної структури на SAR-зображеннях.

**6. Аналіз хмар точок.** GNN особливо добре підходять для аналізу хмар точок LiDAR, де вузли представляють окремі точки і ребра, визначені за допомогою тривимірної близькості. Такий підхід дозволяє розв'язувати такі задачі, як класифікація та сегментація хмари точок. Автори [120] пропонують нову нейронну мережу згортки геометричних моментів графа (GGM-Net) для класифікації форми хмари точок. Запропоновано новий оператор згортки, який називається згортка GGM, що явно кодує локальну геометричну структуру множини точок, вивчаючи особливості з геометричних моментів точок та їхніх локальних сусідів. Це призводить до кращого розуміння геометрії поверхні та підвищує стійкість.

**7. Злиття даних.** GNN дозволяють об'єднувати інформацію з різних джерел, таких як гіперспектральні, LiDAR та оптичні дані, представляючи їх у вигляді вузлів графа та встановлюючи ребра для зв'язку між ними. Це полегшує комплексну інтеграцію даних. Автори пропонують підхід на основі графів [121] для об'єднання спектральної, просторової та висотної інформації з гіперспектральних зображень і даних LiDAR. Вони виділяють морфологічні особливості з гіперспектрального зображення і даних LiDAR за допомогою морфологічних відкриттів і закриттів. Ці особливості фіксують просторову інформацію та інформацію про висоту. Об'єкти нормалізуються до однієї розмірності і будується граф злиття, де вузли з'єднуються лише тоді, коли вони мають схожі спектральні, просторові та висотні характеристики. Потім характеристики проєктуються в низьковимірний підпростір, де зберігаються сусідські зв'язки. Це дозволяє зменшити розмірність зв'язку та об'єднати різномодальні дані. За словами

авторів, цей підхід на основі графів підвищує точність класифікації порівняно з використанням окремих джерел даних або простим нагромадженням ознак. В роботі [122] досліджено способи представлення багаточасткових графів для задач машинного навчання на графах із використанням топологічних методів, зокрема персистентних гомологій (ПГ) хмар точок. Також розглянуто векторні представлення графів, отримані за допомогою білінійних моделей і моделей трансляції, таких як TuskeR, MurE та PairRE. Як приклад використано багаточастковий граф із 271 типом вершин і двома типами ребер. Обчислено ПГ до другого виміру, що охоплює 1-, 2- та 3-вимірні топологічні особливості графів. Персистентні діаграми використано для представлення ПГ, а їх точки проаналізовано статистично за допомогою коефіцієнтів ексцесу, асиметрії, середнього та відхилення. Аналіз показав, що різні моделі представлень мають різні характеристики з точки зору ПГ, що вказує на їхню топологічну неоднаковість. Це підтверджує, що вибір моделі суттєво впливає на якість і точність вивчення багаточасткових графів.

**8. Напівкероване навчання.** GNN використовують зв'язність графів для напівкерованого навчання, коли кількість маркованих супутникових даних обмежена. Вони можуть ефективно використовувати велику кількість немаркованих даних для покращення продуктивності моделі. Автори пропонують нову графову згорткову мережеву модель [123] під назвою DGCGAN, яка поєднує GCN і DCGAN для напівкерованої класифікації вузлів. DGCGAN усуває два основні обмеження існуючих GCN - низьку точність, коли позначених вузлів мало, та недостатню робастність. Спочатку він відображає дані графа у високонелінійний простір, використовуючи симетричне нормалізоване перетворення Лапласа для отримання характеристик вузлів. Потім він перетворює ознаки вузлів у звичайні структуровані дані, такі як зображення, використовуючи бікубічну інтерполяцію, і подає їх як позитивні вибірки на дискримінатор. Модуль з розширеними можливостями вибірково покращує характеристики вузлів для кращої класифікації. Завдяки включенню DCGAN модель розширює розмір вибірки і додає обмеження для підвищення робастності. Експерименти на мережах цитування та наборі супутникових даних показують, що DGCGAN перевершує відомі методи напівкерованого навчання, особливо коли маркованих даних дуже мало. Запропоновані методи допомагають подолати обмеження GCN і підвищити продуктивність у задачах класифікації напівкерованих вузлів.



## 1.1. Аналіз методів та моделей глибинного навчання для розв'язання ...

---

З розвитком досліджень у галузі GNN можна очікувати все більшого застосування і впливу цих потужних моделей у супутниковому моніторингу, що дасть змогу проводити більш точний і всебічний аналіз у різних прикладних галузях.

Інноваційні моделі та технології надають нові можливості, але водночас створюють нові виклики. Нещодавно [124] було визначено деякі ключові виклики у використанні ШІ та машинного навчання для супутникового моніторингу. Ці виклики включають обмеженість навчальних наборів даних, складність моделей супутникового моніторингу, обробку великих даних, адаптацію до нетрадиційних джерел даних, труднощі трансферного навчання, недостатню інтерпретованість, високий бар'єр для входу, проблеми навчання та архітектурні обмеження. На їх подолання спрямовано сучасні дослідження в сфері супутникового інтелекту.

### ВИСНОВКИ

Аналіз поточного стану досліджень в сфері глибинного навчання [125, 126, 127] свідчить про прогрес у вирішенні цих проблем. Такі стратегії, як трансферне навчання, доповнення даних і нові мережеві архітектури покращили продуктивність моделей і розв'язання складних задач. Розподілене навчання, хмарні обчислення і спеціалізовані мікросхеми зробили обробку великих обсягів супутникових даних більш можливою. Такі методи, як адаптація до предметної області, карти уваги і зрозумілий ШІ, були прийняті для покращення можливості передачі і інтерпретації даних. Фреймворки з відкритим вихідним кодом зробили глибинне навчання більш доступним. Такі стратегії, як вибір спектральних каналів і напівконтрольоване навчання, частково вирішили проблеми навчання. Інновації в CNN, трансформерах і GNN розширили можливості моделей для задач супутникового моніторингу.

Це означає, що машинне навчання інтенсивно розвивається в сфері спостереження Землі. Незважаючи на прогрес, все ще залишаються проблеми, зокрема, різноманітність наборів даних, проблеми узагальнення і потреба в моделях, результати яких краще інтерпретуються. Поточні зусилля спрямовані на подолання цих перешкод і розвиток застосувань ШІ і ML у галузі супутникових спостережень. У табл. 1 наведено порівняння

## Частина 1. Аналіз сучасного стану методів комп'ютерного зору ...

специфіки проблем застосування AI/ML в сфері спостереження Землі з традиційною обробкою зображень.

Як видно з табл. 1, супутниковий моніторинг часто є проблематичним через великі, різноманітні мультимодальні дані, де мітки, домени і завдання відрізняються від традиційної обробки зображень. Рішення включають використання немаркованих даних, навчання на основі перенесення моделей (transfer learning) та інші підходи. На сьогоднішній день рішення на основі ML активно розвиваються. У наступних розділах описано підходи та їх застосування, спрямовані на подолання описаних вище викликів у впровадженні рішень ML Earth Intelligence для прийняття рішень та сталого розвитку.

Таблиця 1. Порівняння проблем ШІ/ML у галузі супутникових спостережень з традиційною обробкою зображень

Виклик	Супутниковий моніторинг	Традиційна обробка зображень	Потенційні рішення
Розмір даних	Дуже великі (ГБ/ТБ)	Менші (МБ/ГБ)	Хмарне сховище, розподілене навчання
Різнманітність даних	Багатоспектральні, гіперспектральні, SAR, LiDAR	RGB, градації сірого	Мультимодальні моделі, злиття даних
Дефіцит міток	Дорогий збір базових даних	Комп'ютерне маркування	Трансферне навчання, напівконтрольоване навчання
Просторова роздільна здатність	Змішані пікселі з низькою роздільною здатністю	Чітке розділення об'єктів	Контрольоване попереднє навчання, синтез зображень
Часовий аналіз	Необхідне моделювання часових рядів	Окремих зображень достатньо	Рекурентні мережі, моделі виявлення змін
Зміщення домену	Узагальнення через датчики, географія	Послідовні домени	Адаптація предметної області, налагодження моделі
Інтерпретованість	Важлива для оперативного використання	Не дуже важлива для	Зрозумілі методи ШІ

## 1.1. Аналіз методів та моделей глибинного навчання для розв'язання ...

		оперативного використання	
Оцінювання	Необхідні спеціальні метрики	Достатньо стандартних показників	Схеми перевірки для конкретного домену

### ПЕРЕЛІК ПОСИЛАНЬ

1. Bali O., Bali J. Artificial intelligence in ophthalmology and healthcare: An updated review of the techniques in use. *Indian Journal of Ophthalmology*. 2021. Vol. 69, no. 1. pp. 8-13. DOI: 10.4103/ijo.ijo\_1848\_19.
2. Makridakis S. The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*. 2017. Vol. 90. pp. 46–60. DOI: 10.1016/j.futures.2017.03.006.
3. Khan M. *Advancements in Artificial Intelligence: Deep Learning and Meta-Analysis*. Center for Open Science. 2023. DOI: 10.5281/zenodo.8106837.
4. Rajesh, S., C., K.: *Advancements in Deep Learning: A Comprehensive Review*. *REST Journal on Data Analytics and Artificial Intelligence*. 2023. Vol. 2, no. 2. pp. 65-69. DOI: <https://doi.org/10.46632/jdaai/2/2/7>.
5. Vaswani, A., Shazeer, N., Parmar, N., et al. Attention is All You Need. In *NIPS*. Presented at the (2017).
6. Nguyen, G., Dlugolinsky, S., Bobok, M. et al. Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. *Artificial Intelligence Review*. 2019. Vol. 52. pp. 77–124. DOI: <https://doi.org/10.1007/s10462-018-09679-z>.
7. Sarker, I.H. *Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions*. *SN Computer Science*. 2021. Vol. 2, no. 420. pp. DOI: <https://doi.org/10.1007/s42979-021-00815-1>.
8. Brown B., Mann T., Ryder B., et al. Language Models are Few-Shot Learners. *Advances in neural information processing systems*. 2020. Vol. 33. pp. 1877-1901.
9. Saleh Z. *Artificial Intelligence Definition, Ethics and Standards*. The British University in Egypt. 2019.
10. Goertzel B., Pennachin C. *Artificial general intelligence*. New York: Springer. 2007. P. 518.
11. Kim S.-K., Huh J.-H. Consistency of Medical Data Using Intelligent Neuron Faster R-CNN Algorithm for Smart Health Care

Application. Healthcare. 2020. Vol. 8, no. 185. pp. 1-24. DOI: <https://doi.org/10.3390/healthcare8020185>.

12. J. Borana: Applications of Artificial Intelligence & Associated Technologies. Science [ETEBMS-2016], 2016. Vol. 5, no 6. pp. 64-67.

13. Reza Tizhoosh H., Pantanowitz L. Artificial Intelligence and Digital Pathology: Challenges and Opportunities. Journal of Pathology Informatics. 2018. Vol. 9, no. 38. pp. 1-6. DOI: [https://doi.org/https://doi.org/10.4103/jpi.jpi\\_53\\_18](https://doi.org/https://doi.org/10.4103/jpi.jpi_53_18).

14. Grilli E., Remondino F. Classification of 3D Digital Heritage. Remote Sensing. 2019. Vol. 11, no. 7. DOI: <https://doi.org/10.3390/rs11070847>.

15. Mukhamediev R.I., Popova Y., et al. Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges. Mathematics. 2022. Vol. 10, no. 2552. DOI: <https://doi.org/10.3390/math10152552>.

16. Nassif A.B., Shahin I., Attili I., et al. Speech Recognition Using Deep Neural Networks: A Systematic Review. IEEE Access. 2019, Vol. 7, pp. 19143–19165 (2019). DOI: <https://doi.org/10.1109/access.2019.2896880>.

17. Schmidhuber J. Deep learning in neural networks: An overview. Neural Networks. 2015. Vol. 61, pp. 85–117. DOI: <https://doi.org/10.1016/j.neunet.2014.09.003>.

18. Litjens G., Kooi T., Bejnordi B.E., et al. A survey on deep learning in medical image analysis. Medical Image Analysis. 2017. Vol. 42. pp. 60–88. DOI: <https://doi.org/10.1016/j.media.2017.07.005>.

19. LeCun Y., Kavukcuoglu K., Farabet C. Convolutional networks and applications in vision. Proceedings of 2010 IEEE International Symposium on Circuits and Systems. 2010. pp. 253-256. DOI: [10.1109/ISCAS.2010.5537907](https://doi.org/10.1109/ISCAS.2010.5537907).

20. Zhang, L., Wang, S., Liu, B. Deep learning for sentiment analysis: A survey. WIREs Data Mining and Knowledge Discovery. 2018. Vol. 8. DOI: <https://doi.org/10.1002/widm.1253>.

21. Goodfellow I., Pouget-Abadie J., Mirza M., et al. Generative adversarial networks. Communications of the ACM. 2020. Vol. 63. pp. 139–144. DOI: <https://doi.org/10.1145/3422622>.

22. Li Y., Chen C.L.P. Zhang T. A Survey on Siamese Network: Methodologies, Applications, and Opportunities. IEEE Transactions on Artificial Intelligence. 2022. Vol. 3. pp. 994–1014. DOI: <https://doi.org/10.1109/tai.2022.3207112>.

23. Lin T., Wang Y., Liu X., et al. A survey of transformers. AI Open. 2022. Vol. 3. pp. 111–132. DOI: [10.1016/j.aiopen.2022.10.001](https://doi.org/10.1016/j.aiopen.2022.10.001).

24. Chai J., Zeng H., Li A., et al. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*. 2021. Vol. 6, no. 100134. DOI: 10.1016/j.mlwa.2021.100134.
25. Mehrish A., Majumder N., Bharadwaj R., et al. A review of deep learning techniques for speech processing. *Information Fusion*. 2023. Vol. 99, no. 101869 DOI: 10.1016/j.inffus.2023.101869.
26. Otter D.W., Medina J.R., Kalita J.K. A Survey of the Usages of Deep Learning for Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*. 2021. Vol. 32. pp. 604–624. DOI: 10.1109/tnnls.2020.2979670.
27. Castelvechi D. Can we open the black box of AI? *Nature*. 2016. Vol. 538. pp. 20–23. DOI: 10.1038/538020a.
28. Artiemjew P., Chojka A., Rapiński J. Deep Learning for RFI Artifact Recognition in Sentinel-1 Data. *Remote Sensing*. 2020. Vol. 13, no.7. DOI: 10.3390/rs13010007.
29. Han X., Zhong Y., Cao L., et al. Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification. *Remote Sensing*. 2017. Vol. 9, no. 848. DOI: 10.3390/rs9080848.
30. Muhammad U., Wang W., Chattha S.P., et al. Pre-trained VGGNet Architecture for Remote-Sensing Image Scene Classification. *International Conference on Pattern Recognition (ICPR)*. IEEE. 2018. pp. 1622-1627. DOI: 10.1109/ICPR.2018.8545591.
31. He K., Zhang X., Ren S., et al. Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016. pp. 770-778.
32. Ma H., Liu Y., Ren Y., Wang D., Yu L., Yu J. Improved CNN Classification Method for Groups of Buildings Damaged by Earthquake, Based on High Resolution Remote Sensing Images. *Remote Sensing*. 2020. Vol. 12, no. 260 DOI: 10.3390/rs12020260.
33. Zakria Z., Deng J., Kumar R., Khokhar M.S., Cai J., Kumar J. Multiscale and Direction Target Detecting in Remote Sensing Images via Modified YOLO-v4. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022. Vol. 15, pp. 1039–1048. DOI: 10.1109/jstars.2022.3140776.
34. Zhou W., Newsam S., Li C., Shao Z. Learning Low Dimensional Convolutional Neural Networks for High-Resolution Remote Sensing Image Retrieval. *Remote Sensing*. 2017. Vol. 9, 489. DOI: 10.3390/rs9050489.
35. Alsabhan W., Alotaiby T., Dudin B. Detecting Buildings and Nonbuildings from Satellite Images Using U-Net. *Computational*

Intelligence and Neuroscience. 2022, pp. 1–13. DOI: 10.1155/2022/4831223.

36. Girshick R., Donahue J., Darrell T., Malik J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition. 2014. pp. 580–587.

37. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. pp. 3431–3440

38. Wang X., Jing S., Dai H., Shi A. High-resolution remote sensing images semantic segmentation using improved UNet and SegNet. Computers and Electrical Engineering. 2023. Vol. 108, no. 108734. DOI: 10.1016/j.compeleceng.2023.108734.

39. Garcha Rodriguez C., Vitria J., Mora O. Uncertainty-Based Human-in-the-Loop Deep Learning for Land Cover Segmentation. Remote Sensing. 2020. Vol. 12, no. 3836. DOI: 10.3390/rs12223836.

40. Yang L., Chen Y., Song S., Li F., Huang G. Deep Siamese Networks Based Change Detection with Remote Sensing Images. Remote Sensing. 2021. Vol. 13, no. 3394. DOI: 10.3390/rs13173394.

41. Chen T., Lu Z., Yang Y., Zhang Y., Du B., Plaza A. A Siamese Network Based U-Net for Change Detection in High Resolution Remote Sensing Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2022. Vol. 15, pp. 2357–2369.

42. Zhang R., Zhang H., Ning X., Huang X., Wang J., Cui, W. Global-aware siamese network for change detection on remote sensing images. ISPRS Journal of Photogrammetry and Remote Sensing. 2023. Vol. 199, pp. 61–72.

43. Dong Y., Jiao W., Long T., Liu L., He G., Gong C., Guo Y. Local Deep Descriptor for Remote Sensing Image Feature Matching. Remote Sensing. 2019. Vol. 11, no. 430. DOI: 10.3390/rs11040430.

44. Lu Y., Yang J., Xiao L. Multi-Supervised Recursive-CNN for Hyperspectral and Multispectral Image Fusion. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. 2021. pp. 2440–2443. DOI: 10.1109/IGARSS47720.2021.9553598.

45. Dong C., Loy C.C., He K., Tang X. Learning a Deep Convolutional Network for Image Super-Resolution. Computer Vision – ECCV. 2014. Springer International Publishing, Cham. 2014. pp. 184–199.

46. Dong C., Loy C.C., He K., Tang X. Image Super-Resolution Using Deep Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2016. Vol. 38, pp. 295–307.

47. Lin G., Wu Q., Huang X., Qiu L., Chen X. Deep Convolutional Networks-Based Image Super-Resolution. In: *Intelligent Computing Theories and Application*. Springer International Publishing, Cham (2017). pp. 338–344.

48. Pashaei M., Starek M.J., Kamangir H., Berryhill J. Deep Learning-Based Single Image Super-Resolution: An Investigation for Dense Scene Reconstruction with UAS Photogrammetry. *Remote Sensing*. 2020. Vol. 12, no. 1757. DOI: 10.3390/rs12111757.

49. Zhang L., Cheng B. Transferred CNN Based on Tensor for Hyperspectral Anomaly Detection. *IEEE Geoscience and Remote Sensing Letters*. 2020. Vol. 17, pp. 2115–2119. DOI: 10.1109/lgrs.2019.2962582.

50. Zhang L., Cheng B. Fractional Fourier Transform and Transferred CNN Based on Tensor for Hyperspectral Anomaly Detection. *IEEE Geoscience and Remote Sensing Letters*. 2022. Vol. 19, pp.1–5. DOI: 10.1109/lgrs.2021.3072249.

51. Maturana D., Scherer S. 3D Convolutional Neural Networks for landing zone detection from LiDAR. 2015 IEEE International Conference on Robotics and Automation (ICRA). 2015. pp. 3471–3478.

52. Jayakumari R., Nidamanuri R.R., Ramiya A.M.: Object-level classification of vegetable crops in 3D LiDAR point cloud using deep learning convolutional neural networks. *Precision Agriculture*. 2021. Vol. 22, pp. 1617–1633. DOI: 10.1007/s11119-021-09803-0.

53. Maturana D., Scherer S. VoxNet: A 3D Convolutional Neural Network for real-time object recognition. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2015. pp. 922–928.

54. Su H., Maji S., Kalogerakis E., Learned-Miller E. Multi-view Convolutional Neural Networks for 3D Shape Recognition. 2015 IEEE International Conference on Computer Vision (ICCV). 2015. pp. 945–953.

55. Charles R.Q., Su H., Kaichun M., Guibas L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. pp. 652–660.

56. Xie Z., Chen J., Peng B. Point clouds learning with attention-based graph convolution networks. *Neurocomputing*. 2020. Vol. 402, pp. 245–255. DOI: 10.1016/j.neucom.2020.03.086.

57. Miller A., Jain V., Mundy J.L. Real-time rendering and dynamic updating of 3-d volumetric data. *Proceedings of the Fourth Workshop on General Purpose Processing on Graphics Processing Units*. ACM, New York, NY, USA. 2011. pp. 1–8.

58. Hochreiter S., Schmidhuber J. Long Short-Term Memory. *Neural Computation*. 1997. Vol. 9, pp. 1735–1780.
59. Chung J., Gulcehre C., Cho K., Bengio Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. arXiv preprint 2014. arXiv:1412.3555.
60. Ganai A.F., Khursheed F. Predicting next Word using RNN and LSTM cells: Stastical Language Modeling. 2019 Fifth International Conference on Image Information Processing (ICIIP). 2019. pp. 469-474.
61. Miao Y., Gowayyed M., Metze F. EESSEN: End-to-end speech recognition using deep RNN models and WFST-based decoding. 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU). 2015. pp. 167-174
62. Lodhwal V., Choudhary G. Survey Paper: Automatic Title Generation for Text with RNN and Pre-trained Transformer Language Model. *International Journal for Research in Applied Science and Engineering Technology*. 2023. Vol. 11, pp. 2117–2124. DOI: 10.22214/ijraset.2023.49713.
63. Johnson M., Schuster M., Le Q.V., et all. Googles Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation. *Transactions of the Association for Computational Linguistics*. 2017. Vol. 5, pp. 339–351.
64. Zhao B., Li X., Lu X. CAM-RNN: Co-Attention Model Based RNN for Video Captioning. *IEEE Transactions on Image Processing*. 2019. Vol. 28, pp. 5552–5565. DOI: 10.1109/tip.2019.2916757.
65. Dey M.K., Chanana S. Short-term forecasting of a univariate time series load using LSTM based RNN model. 2022 International Conference on Futuristic Technologies (INCOFT). 2022. pp. 1-6.
66. Kardakis S., Perikos I., Grivokostopoulou F., Hatzilygeroudis I. Examining Attention Mechanisms in Deep Learning Models for Sentiment Analysis. *Applied Sciences*. 2021. Vol. 11, no. 3883. DOI: 10.3390/app11093883.
67. Salmela L., Tsipinakis N., Foi A., Billet C., Dudley J.M., Genty G. Predicting ultrafast nonlinear dynamics in fibre optics with a recurrent neural network. *Nature Machine Intelligence*. 2021. Vol. 3, pp. 344–354. DOI: 10.1038/s42256-021-00297-z.
68. Ho Tong Minh D., Ienco D., Gaetano R., Lalande N., Ndikumana E., Osman F., Maurel P. Deep Recurrent Neural Networks for Winter Vegetation Quality Mapping via Multitemporal SAR Sentinel-1. *IEEE Geoscience and Remote Sensing Letters*. 2018. Vol. 15, pp. 464–468.



69. Ienco D., Gaetano R., Dupaquier C., Maurel P. Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks. *IEEE Geoscience and Remote Sensing Letters*. 2017. Vol. 14, pp. 1685–1689.

70. Milan A., Rezatofighi S.H., Dick A., Reid I., Schindler K. Online Multi-Target Tracking Using Recurrent Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2017. Vol. 31, no. 1, DOI: 10.1609/aaai.v31i1.11194.

71. Lyu H., Lu H., Mou L. Learning a Transferable Change Rule from a Recurrent Neural Network for Land Cover Change Detection. *Remote Sensing*. 2016. Vol. 8, no. 506.

72. Zhao B. A Systematic Survey of Remote Sensing Image Captioning. *IEEE Access*. 2021. Vol. 9, pp. 154086–154111. DOI: 10.1109/access.2021.3128140.

73. Anders K., Winiwarter L., Lindenbergh R., Williams J.G., Vos S.E., Hufle B. 4D objects-by-change: Spatiotemporal segmentation of geomorphic surface change from LiDAR time series. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2020. Vol. 159, pp. 352–363. DOI: 10.1016/j.isprsjprs.2019.11.025.

74. Wu C., Chen L., Wang G., Chai S., Jiang H., Peng J., Hong Z. Spatiotemporal Scenario Generation of Traffic Flow Based on LSTM-GAN. *IEEE Access*. 2020. Vol. 8, pp. 186191–186198. DOI: 10.1109/access.2020.3029230.

75. Kulshrestha A., Chang L., Stein A. Use of LSTM for Sinkhole-Related Anomaly Detection and Classification of InSAR Deformation Time Series. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022. Vol. 15, pp. 4559–4570. DOI: 10.1109/jstars.2022.3180994.

76. Mou L., Ghamisi P., Zhu X.X. Deep Recurrent Neural Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*. 2017. Vol. 55, pp. 3639–3655.

77. Zhou Y., Wang S., Wu T., et al. For-backward LSTM-based missing data reconstruction for time-series Landsat images. *GIScience Remote Sensing*. 2022. Vol. 59, pp. 410–430. DOI: 10.1080/15481603.2022.2031549.

78. Goodfellow I., Pouget-Abadie J., Mirza M., et al. Generative adversarial networks. *Communications of the ACM*. 2020. Vol. 63, pp. 139–144.

79. Oliehoek F.A., Savani R., Gallego J., van der Pol E., Groÿ R. Beyond Local Nash Equilibria for Adversarial Networks. *Communications in Computer and Information Science*. Springer International Publishing, Cham. 2019. pp. 73–89.

80. Santos J., Frango I. Generating photorealistic images of people's eyes with strabismus using Deep Convolutional Generative Adversarial Networks. 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE). 2020. pp. 1-4.

81. Donahue C., McAuley J., Puckette M. Synthesizing Audio with Generative Adversarial Networks. 2018. arXiv preprint arXiv:1802.04208.

82. 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). IEEE 2017. pp. 1125-1134.

83. Zhang Z. Generative Adversarial Networks for Diverse and Explainable Text-to-Image Generation. 2023. P. 1-25. DOI: 10.33612/diss.507581028.

84. Bashar M.A., Nayak R. TAnoGAN: Time Series Anomaly Detection with Generative Adversarial Networks. 2020 IEEE Symposium Series on Computational Intelligence (SSCI). 2020. pp. 33-43.

85. Sandfort V., Yan K., Pickhardt P.J., Summers R.M. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. Scientific Reports. Vol. 9. 2019. DOI: 10.1038/s41598-019-52737-x.

86. Khelifi L., Mignotte M. Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis. IEEE Access. 2020. Vol. 8, pp. 126385–126400. DOI: 10.1109/access.2020.3008036.

87. Lavreniuk M., Kussul N., Shelestov A., Lavrenyuk A., Shumilo L. Super resolution approach for the satellite data based on the generative adversarial networks. 2022 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 17 – 22 July, 2022, Kuala Lumpur Convention Centre (KLCC), Kuala Lumpur, Malaysia. pp. 1095 – 1098. DOI: 10.1109/IGARSS46834.2022.9884460

88. Jiahao Zheng, Xiao-Yang Liu and Xiaodong Wang Single Image Cloud Removal Using U-Net and Generative Adversarial Networks. IEEE Transactions on Geoscience and Remote Sensing. 2020. Vol. 59. pp. 6371-6385. DOI: 10.1109/TGRS.2020.3027819.

89. Akram Z., Chaker El A. Satellite image inpainting with deep generative adversarial neural networks. IAES International Journal of Artificial Intelligence. Vol. 10, pp.121-130. DOI: 10.11591/IJAI.V10.I1.PP121-130.

90. Zerrouki N., Dairi A., Harrou F., Zerrouki Y., Sun Y. Efficient land desertification detection using a deep learning-driven generative

adversarial network approach: A case study. *Concurrency and Computation: Practice and Experience*. 2021. Vol. 34. DOI: 10.1002/cpe.6604.

91. David J., Michael K. Towards Modeling Geographical Processes with Generative Adversarial Networks (GANs) (Short Paper). 14th International Conference on Spatial Information Theory (COSIT 2019). *Leibniz International Proceedings in Informatics (LIPIcs)*, 2019. Vol. 142, pp. 27:1-27:9, DOI: 10.4230/LIPIcs.COSIT.2019.27.

92. Merkle N., Auer S., Muller R., Reinartz P. Exploring the Potential of Conditional Adversarial Networks for Optical and SAR Image Matching. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2019. Vol. 11, pp. 1811–1820. DOI: 10.1109/jstars.2018.2803212.

93. Rabbi J., Ray N., Schubert M., Chowdhury S., Chao D. Small-Object Detection in Remote Sensing Images with End-to-End Edge-Enhanced GAN and Object Detector Network. *Remote Sensing*. 2020. Vol. 12, no. 1432. DOI: 10.3390/rs12091432.

94. Shumilo, L., Okhrimenko, A., Kussul, N., Drozd, S., & Shkalikov, O. (2023). Generative adversarial network augmentation for solving the training data imbalance problem in crop classification. *Remote Sensing Letters*, 14(11), 1129-1138.

95. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv*. 2019. abs/1810.04805.

96. Liu Y., Ott M., Goyal N., et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. 2019. arXiv preprint arXiv:1907.11692.

97. Raffel C., Shazeer N., Roberts A., et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *The Journal of Machine Learning Research*. 2020. Vol. 21, no. 1, pp. 5485-5551.

98. Bazi Y., Bashmal L., Rahhal M.M.A., Dayil R.A., Ajlan N.A. Vision Transformers for Remote Sensing Image Classification. *Remote Sensing*. 2021. Vol. 13, no. 516. DOI: 10.3390/rs1303051.

99. Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S. End-to-End Object Detection with Transformers. *Computer Vision – ECCV 2020*. Springer International Publishing, Cham. 2020. pp. 213–229.

100. Zheng S., Lu J., Zhao H., et al. Rethinking Semantic Segmentation from a Sequence-to-Sequence Perspective with Transformers. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 6881-6890.

101. Lavreniuk M., Shumilo L., Lavrenyuk A. Generative Adversarial Networks for the Satellite Data Super Resolution Based on the Transformers with Attention. IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, California, USA. 2023. pp. 6294-6297. DOI: 979-8-3503-2010-7/23/.

102. Jin P., Mou L., Xia G.-S et al. Anomaly Detection in Aerial Videos With Transformers. IEEE Transactions on Geoscience and Remote Sensing. 2022. Vol. 60. pp. 1-13. DOI: <https://doi.org/10.1109/tgrs.2022.3198130>.

103. Yu J., Li J., Yu Z., et al. Multimodal Transformer With Multi-View Visual Representation for Image Captioning. IEEE Transactions on Circuits and Systems for Video Technology. 2020. Vol. 30. pp. 4467-4480. DOI: <https://doi.org/10.1109/tcsvt.2019.2947482>.

104. Lu D., Xie Q., Wei M., et al. Transformers in 3D Point Clouds: A Survey.. arXiv:2205.07417. 2022. DOI: <https://doi.org/10.48550/arXiv.2205.07417>.

105. Scarselli F., Gori M., Ah Chung Tsoi, et al. The Graph Neural Network Model. IEEE Transactions on Neural Networks. 2009. Vol. 20. pp. 61-80. DOI: <https://doi.org/10.1109/tnn.2008.2005605>.

106. Kipf T.N., Welling M. Semi-supervised classification with graph convolutional networks. arXiv:1609.02907. 2016. DOI: <https://doi.org/10.48550/arXiv.1609.02907>.

107. Velickovic P., Cucurull G., Casanova A., et al. Graph attention networks. stat. arXiv:1710.10903. 2017. Vol. 1050. DOI: <https://doi.org/10.48550/arXiv.1710.10903>.

108. Hajiramezanali E., Hasanzadeh A., Narayanan K., et al. Variational graph recurrent neural networks. Advances in neural information processing systems. 2019. Vol. 32. pp. 1-11.

109. Zhou J., Cui G., Hu S., et al. Graph neural networks: A review of methods and applications. AI Open. 2020. Vol. 1. pp. 57-81. DOI: <https://doi.org/10.1016/j.aiopen.2021.01.001>.

110. Wu Y., Lian D., Xu Y., et al. Graph Convolutional Networks with Markov Random Field Reasoning for Social Spammer Detection. Proceedings of the AAAI Conference on Artificial Intelligence. 2020. Vol. 34. pp. 1054-1061. DOI: <https://doi.org/10.1609/aaai.v34i01.5455>.

111. Wieder O., Kohlbacher S., Mйlaine Kuenemann Garon A., et al. A compact review of molecular property prediction with graph neural networks. Drug Discovery Today: Technologies. 2020. Vol. 37. pp. 1-12. DOI: <https://doi.org/10.1016/j.ddtec.2020.11.009>.

112. Wu Q., Zhang H., Gao X., et al. Dual Graph Attention Networks for Deep Latent Representation of Multifaceted Social Effects in Recommender Systems. The World Wide Web Conference. ACM,

New York, NY, USA. 2019. pp. 2091-2102. DOI: <https://doi.org/10.1145/3308558.3313442>.

113. Wang Z., Chen T., Ren J., et al. Deep Reasoning with Knowledge Graph for Social Relationship Understanding. arXiv:1807.00504. 2018. DOI: <https://doi.org/10.48550/arXiv.1807.00504>.

114. Yu B., Yin H., Zhu Z. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. arXiv:1709.04875. 2018. DOI: <https://doi.org/10.48550/arXiv.1709.04875>.

115. Khan N., Chaudhuri U., Banerjee B., et al. Graph convolutional network for multi-label VHR remote sensing scene recognition. Neurocomputing. 2019. Vol. 357. pp. 36-46. DOI: <https://doi.org/10.1016/J.NEUCOM.2019.05.024>.

116. Търкменлі І., Аптула Е., Каябол К. Superpixel based graph convolutional neural network for SAR image segmentation. Image and Signal Processing for Remote Sensing. 2021. Vol. 11862. pp. 148-153. DOI: <https://doi.org/10.1117/12.2599864>.

117. Gao L., Xiao S., Hu C., et al. Hyperspectral Image Classification Based on Fusion of Convolutional Neural Network and Graph Network. Recent Advances in Robotics and Intelligent Robots Applications. 2023. Vol. 13, no. 12. DOI: <https://doi.org/10.3390/app13127143>.

118. Tian S., Cao L., Kang L., et al. A Novel Hybrid Attention-Driven Multistream Hierarchical Graph Embedding Network for Remote Sensing Object Detection. Remote Sensing. 2022. Vol. 14, no. 19. DOI: <https://doi.org/10.3390/rs14194951>.

119. Wang R., Wang L., Wei X., et al. Dynamic Graph-Level Neural Network for SAR Image Change Detection. IEEE Geoscience and Remote Sensing Letters. 2022. Vol. 19. pp. 1–5. DOI: <https://doi.org/10.1109/lgrs.2021.3132167>.

120. Li D., Shen X., Yu Y., et al. GGM-Net: Graph Geometric Moments Convolution Neural Network for Point Cloud Shape Classification. IEEE Access. 2020. Vol. 8, pp. 124989–124998. DOI: <https://doi.org/10.1109/access.2020.3007630>.

121. Debes C., Merentitis A., Heremans R., et al. Hyperspectral and LiDAR Data Fusion: Outcome of the 2013 GRSS Data Fusion Contest. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2014. Vol. 7. pp. 2405-2418. DOI: [10.1109/JSTARS.2014.2305441](https://doi.org/10.1109/JSTARS.2014.2305441).

122. Yavorskyi, O. and Kussul, N. (2023) “Studying the multipartite graph representations with topological data analysis”, International

Scientific Technical Journal "Problems of Control and Informatics", 68(5), pp. 107–117. doi: 10.34229/1028-0979-2023-5-9.

123. Jia N., Tian X., Gao W., et al. Deep Graph-Convolutional Generative Adversarial Network for Semi-Supervised Learning on Graphs. *Remote Sensing*. 2023. Vol. 15, no. 12. DOI: <https://doi.org/10.3390/rs15123172>.

124. Farias A.D., Sun G. Data Mining and Machine Learning in Earth Observation—An Application for Tracking Historical Algal Blooms. In *CS & IT Conference Proceedings*. 2020 .Vol. 10. pp. 39-58. DOI: 10.5121/csit.2020.100204.

125. Lavreniuk M., Shumilo L., Yailymov B., Kussul N. Reviewing Deep Learning Methods in the Applied Problems of Economic Monitoring Based on Geospatial Data. *Cybern Syst Anal*. 2022. Vol. 58, pp. 1008–1020. DOI: 10.1007/s10559-023-00535-9.

126. Kussul, N., Kuzin, V., Shelestov, A. (2024). A Survey of Deep Learning for Remote Sensing, Earth Intelligence and Decision Making. In: Luntovskyy, A., Klymash, M., Melnyk, I., Beshley, M., Schill, A. (eds) *Digital Ecosystems: Interconnecting Advanced Networks with AI Applications*. TCSET 2024. Lecture Notes in Electrical Engineering, vol 1198. pp 81–109. Springer, Cham. [https://doi.org/10.1007/978-3-031-61221-3\\_5](https://doi.org/10.1007/978-3-031-61221-3_5)

127. Kussul, N., Shelestov, A., Basarab, R., Skakun, S., Kussul, O., & Lavrenyuk, M. (2015). Geospatial Intelligence and Data Fusion Techniques for Sustainable Development Problems. *ICTERI*, 1356, 196-203.