



KAPITEL 6 / CHAPTER 6⁶

PROCESSING OF MULTICHANNEL REMOTE SENSING DATA USING NEURAL NETWORKS

DOI: 10.30890/2709-2313.2023-18-03-002

Вступ

Багатоканальні зображення дистанційного зондування, що отримуються різноманітними типами датчиків та систем (мульти- та гіперспектральних, багатопоглядових, радіолокаційних багатополяризаційних), широко використовуються в даний час для багатьох застосувань у картографуванні земельних покривів, сільському та лісовому господарстві, екологічному моніторингу тощо [1-4]. Для того, щоб видобути корисну інформацію з таких даних, часто необхідно виконати кілька етапів попередньої обробки отриманих зображень. Під обробкою можна розуміти широкий спектр операцій, що включають суміщення зображень між собою та з топографічною мапою чи цифровою моделлю рельєфу, оцінку характеристик шумів або викривлень, придушення шуму, стиснення з втратами, класифікацію, розпізнавання та оцінку параметрів об'єкта тощо [4-10].

На кожному з етапів використовуються відповідні методи та цифрові алгоритми обробки інформаційних даних. Втім, є й певні спільні сучасні тенденції. Однією з них є широке використання навчених нейронних мереж (НМ) з різною структурою та машин опорних векторів (МОВ). Зокрема, вони застосовуються на етапах пошуку подібних блоків та суміщення зображень [11], оцінки статистичних характеристик завад та якості первинних (отриманих) зображень [12-15], фільтрації та прогнозування її ефективності [16, 17], стиснення з втратами [18, 19], розробці метрик для цих застосувань [20, 21], класифікації зображень дистанційного зондування [22, 23].

Таке широке застосування зумовлене кількома корисними властивостями нейронних мереж. По-перше, вони здатні навчатися, тобто забезпечувати бажаний вихід по вхідних даних, що аналогічні тим, які використовувались в процесі навчання. Така можливість в даний час підтримується великою кількістю доступних даних дистанційного зондування та інших типів зображень, які можна використовувати в навчанні. Так, тільки система Сентинель забезпечує дані для кожної ділянки поверхні Землі практично кожні 6 днів, а ця система є далеко не єдиною, яка формує зображення, що розповсюджуються розробниками систем.

⁶*Authors: Lukin Volodymyr Vasilyovych, Uss Mykhaylo Leontiyovych, Vasilyeva Iryna Karlivna, Proskura Galyna Anatoliyivna, Rubel Oleksii Sergiyovych, Rubel Andrii Sergiyovych, Benoit Vozel*



По-друге, нейромережі здатні виконувати суттєво нелінійні апроксимації, на що не здатні деякі інші алгоритми обробки. НМ можуть об'єднувати вхідні дані різного походження і суттєво різного масштабу. Нарешті, вони можуть самі виявляти корисні ознаки в даних. Все це сприяє як зацікавленості у використанні НМ, так і багатьом ситуаціям, коли НМ забезпечують кращі результати у порівнянні з іншими методами.

Метою українсько-французького проекту в рамках програми «Дніпро», що виконувався в 2021 році (саме його результати викладені у даному розділі), була розробка методів та алгоритмів ефективної обробки зображень на основі машинного навчання та інших сучасних засобів для багатоканальних даних дистанційного зондування. Дослідження знаходяться на перетині таких галузей, як «Інформаційні та комунікаційні технології» та «Раціональне природокористування». Розглянуті методи, які можна використати для різних етапів обробки даних ДЗ.

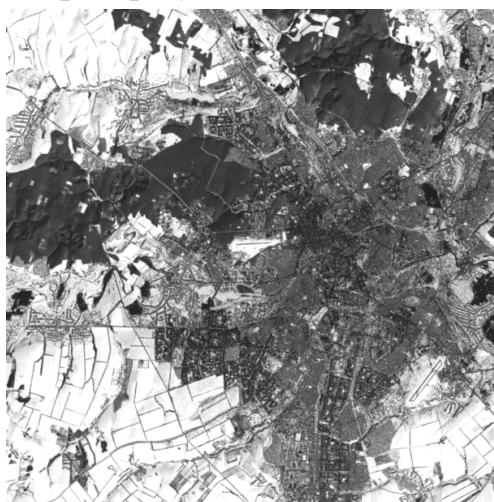
6.1. Міра подібності з додатковими властивостями для мультимодальних зображень ДЗ

Розглянемо проблему навчання ефективної міри подібності (МП) для зображень мультимодального дистанційного зондування (ДЗ). Бажано мати одну МП, ефективну для різних комбінацій режимів (мод) зондування [24-26]. Спочатку ми розглянемо вплив балансування навчального набору даних на ефективність МП і продемонструємо, що можна покращити загальну ефективність МП. Однак це поліпшення спостерігається лише для деяких комбінацій режимів. Щоб впоратися з цією проблемою, ми пропонуємо включити інформацію про порівнювані режими як додатковий вхід до запропонованої згорткової нейронної мережі (convolutional neural network - CNN). Завдяки цій додатковій інформації була покращена ефективність МП для всіх комбінацій режимів. Підвищення ефективності МП доведено для реальних даних з платформ Sentinel-2, Landsat-8, Hyperion, SIR-C і Sentinel-1, ASTER Global DEM 2 і ALOS World 30m (глобальні цифрові моделі рельєфу (ЦМР, digital elevation map– DEM)), а також для комбінацій режимів, включаючи суміщення двох оптичних (О-О) зображень, оптичного та радіолокаційного (О-Р) зображень, оптичного зображення та ЦМР (О-ЦМР) і радіолокаційного зображення і ЦМР (Р-ЦМР), ефективність запропонованої МП на основі CNN порівняно з існуючими МП.



Реєстрація зображення передбачає вимірювання (оцінювання) подібності між фрагментами двох зображень [27]. У дистанційному зондуванні зображення можна отримати різними способами, включаючи оптичні [28], радіолокаційні зображення [29] або карти цифрової моделі висот. Знайти подібність між різними модальностями (наприклад, О-Р або Р-ЦМР) є найскладнішою ситуацією через значні структурні відмінності одних і тих самих фрагментів поверхні Землі, представлених у різних модальностях [30].

Для практичного використання зручно мати єдину міру подібності, яка б була однаково ефективною для будь-якої комбінації модальностей. Однак існуючі МП зазвичай мають найкращу ефективність лише для певних комбінацій модальностей і можуть бути перевершені спеціально розробленою комбінацією МП [31]. Наприклад, в О-О випадку дескриптор MIND [32] є ефективнішим, ніж SIFT [29, 33]; однак SIFT перевершує MIND для О-Р випадку [31] (рис. 1).



а)



б)

а) оптичне зображення системи Landsat 8; б) радіолокаційне зображення системи SIR-C.

Рисунок 1 – Приклад суміщення оптичного та радіолокаційного зображень



МП, що базуються на навчанні, можна тренувати для певної комбінації модальностей [36] або незалежним від модальності способом. Наприклад, у [37] автори запропонували мультимодальну МП DLSSM (Deep Localization Similarity Measure) у формі CNN, яка перевершує існуючі МП, розроблені емпірично для О-О, О-Р, О-ЦМР та Р-ЦМР випадків суміщення. Однак така загальна модель може бути перевершена моделями, навченими для певної комбінації модальностей (наприклад, моделлю, навченою підтримувати лише О-О випадок).

Нижче нами досліджено два підходи до підвищення ефективності універсальної МП, що базується на навчанні. Перший підхід пов'язаний з кращим балансуванням набору навчальних даних. Цей підхід дозволяє покращити площу під кривою МП (area under the curve - AUC) [38] приблизно на 1,8% порівняно з нашим попереднім результатом для незбалансованого набору даних. Однак цей приріст не є рівномірним серед комбінацій режимів: у випадках О-О суміщення балансування набору даних не призводить до покращення AUC. Щоб впоратися з цією проблемою, ми пропонуємо інший підхід, який полягає у використанні інформації про режим як додатковий вхід CNN. Завдяки такому підходу нам вдалося отримати додаткові 1,2% у значенні AUC для розробленої універсальної МП.

Для вивчення впливу інформації про модальність на МП, що базується на навчанні, будемо використовувати міру подібності DLSSM, нещодавно розроблену авторами роботи [37], як базову. Ідея такої МП полягає в тому, що точність локалізації, тобто точність вимірювання просторового зсуву між фрагментами опорного зображення (reference image - RI) і зображення, що суміщується (template image - TI), може використовуватися як значення МП. Більш конкретно, DLSSM оцінює коваріаційну матрицю похибок вектора зсуву і використовує визначник цієї матриці як значення МП. Таким чином, спільно вирішуються завдання дискримінації та локалізації. У [37] було показано, що цей підхід дає певні переваги порівняно з класичними функціями втрат, які використовуються для навчання CNN вирішенню задачі узгодження (пошуку відповідності) блоків.

Структура DLSSM CNN показана на рисунку 2. DLSSM приймає вхідне зображення розміром $32 \times 32 \times 2$ пікселі і продукує вектор зсуву між RI і TI, а також три елементи коваріаційної матриці похибки оцінки зсуву. Частина DLSSM, яка здійснює видобування вектору ознак, складається з шести шарів двовимірної згортки з активацією ReLU і шарів максимального об'єднання. Потім вектор ознак обробляється повнозв'язним шаром і подається на три блоки, що



відповідають за оцінку вектору зсуву, діагональних значень матриці коваріації та коефіцієнта кореляції між елементами помилки вектору зсуву відповідно. DLSSM навчається за допомогою спеціальної функції втрат, яка поєднує триплетну функцію втрат і регресію з урахуванням невизначеності (див. [37] для детального опису).

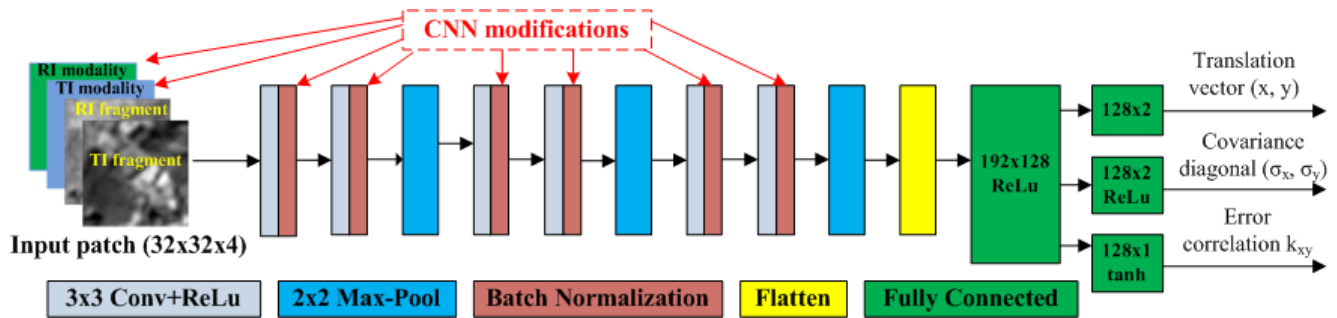


Рисунок 2 – Структура DLSSM CNN

Запропоновані нами модифікації позначені червоним кольором

Для навчання МП, придатної для різних пар модальностей, потенційно може використовуватися одне джерело інформації, а саме власна модальність фрагментів RI та TI. Якщо для кожної пари модальностей використовується окрема модель, використання знань про модальність є тривіальним: пара фрагментів RI/TI подається безпосередньо в МП, що відповідає модальності даних.

Нами пропонується додавати модальність як фрагментів RI, так і TI як додаткових входів CNN, розширюючи кількість входних каналів з двох до чотирьох. Оскільки модальність є якісним атрибутом, її не можна використовувати безпосередньо як вхід CNN. Тому ми кодуємо кожну модальність цілим числом: 0 для оптичного, 1 для радара, 2 для ЦМР. Як для RI, так і для TI створюється додатковий канал розміром 32×32 , рівномірно заповнений номером, що відноситься до модальності RI або TI. В результаті кількість входних каналів збільшується з двох до чотирьох (рис. 2).

Інший спосіб вплинути на процес навчання при створенні універсальної МП – збалансувати навчальні дані. Наш навчальний набір мультимодальних зображень ДЗ в роботі [37] був розбалансований щодо О-О випадків: при навчанні використовувалось 75% О-О фрагментів і лише 25% фрагментів відповідали іншим парам модальностей. З одного боку, це відбувається природно, оскільки оптичні датчики характеризуються більш високою просторовою роздільною здатністю та охопленням зображення. З іншого боку, щоб компенсувати вказаний недолік, нами було реалізовано алгоритм



випадкового вибору даних, який відбирає вибірки з імовірністю, обернено пропорційною кількості пар модальностей. В результаті навчальні міні-пакети містять усі пари модальностей з однаковою ймовірністю.

Версія DLSSM CNN в [37] не має шарів пакетної нормалізації (batch normalization - BN) [39]. Причина полягає в тому, що ця згорткова нейромережа не є надто глибокою, щоб шари BN мали значний вплив. Тренування з шарами BN і без них продемонстрували схожі результати. Ситуація змінюється, коли до входу CNN додається модальність. Чотириканальне вхідне зображення стає неоднорідним із двома каналами зображення, нормованими на нульове середнє значення та одиничну дисперсію, а також двома вхідними параметрами категорійної модальності. У цьому випадку нами було виявлено, що шари BN є необхідними для компенсації такої неоднорідності. Усі результати в експериментальній частині були отримані з версією DLSSM CNN, яка включає шари BN після всіх згорткових шарів.

Для навчання всіх CNN, що виконують пошук подібних блоків були використані ті ж 18 зареєстрованих пар мультимодальних зображень, що й у [37]. Ці пари охоплюють такі випадки суміщення: видимого до інфрачервоного, О-Р, О-ЦМР і Р-ЦМР. Ми використовуємо термін «оптичний» як для зображень видимого діапазону, так і для інфрачервоних модальностей. Далі ми згрупуємо всі ці випадки. Для оптичної модальності дані надходять з платформ Sentinel-2, Landsat-8 і Hyperion, радіолокаційної модальності – з платформ SIR-C і Sentinel-1, ЦМР – від систем ASTER Global DEM 2 і ALOS World 30m формування глобальних ЦМР.

Області зображення, що відповідають чотирьом парам модальностей, знаходяться в таких пропорціях: 75% для видимого та інфрачервоного, 9% для оптичного та радіолокаційного, 8% для оптичного та ЦМР та 8% для радіолокаційного діапазону та ЦМР. Як було запропоновано раніше, під час навчання зі збалансованими даними вибірки відбираються випадковим чином з кожної пари модальностей з імовірністю, обернено пропорційною до кількості даних. Загалом із вищезгаданих реєстраційних справ було зібрано 2 700 000 пар блоків розміром 32×32 пікселі. Ці пари були випадковим чином розподілені між наборами для навчання (75%) і для перевірки (25%).

Тестові дані збиралися з іншого набору з 16 зареєстрованих мультимодальних пар, які охоплюють ті самі випадки реєстрації (докладний опис дивись у [33]). Пари модальності рівномірно представлені в тестовому наборі (всього 100 000 пар блоків з 50% схожими і 50% несхожими).

Запропоновані CNN навчаються за допомогою оптимізатора Adam [40],



початкова швидкість навчання $2 \cdot 10^{-4}$ і ослаблення 10^{-5} . Навчання займає 750 епох, кожна епоха складається з 5000 кроків. Розмір пакету встановлений на 32. Ми використовуємо прості політики аугментації, включаючи перемикування RI та TI (як зображень, так і модальності), обертання RI та TI на 90, 180 та 270 градусів та горизонтальне переміщення RI та TI.

Запропоновану модифікацію МП порівнюємо з базовим варіантом DLSTM, а також із МП зі схожою структурою CNN, але натренованим з функціями втрат петлевого та триплетного відношень [41] (названих як DSM, Deep Similarity Measure), нейромережною МП L2-Net CNN [42] та емпірично розробленими МП, включаючи MIND [34], НОРС [30] та SIFT-ОСТ [31]. Результати представлені в таблиці 1.

Перша версія DLSTM CNN досягала значення AUC 84,07% у загальному випадку. Модель, запропонована в цьому дослідженні з шаром BN, показала дещо нижче значення AUC 83,67% (див. рис. 3). Можливо, значення швидкості навчання слід оптимізувати для версії CNN з BN-шарами. Більш детально для окремих пар модальностей: О-О випадок характеризується найвищою AUC 88,26%. Для решти випадків значення AUC значно нижчі - від приблизно 79,3 до 83%.

Таблиця 1 – Значення AUC×100% для трьох порівнюваних моделей і чотирьох пар модальностей: О-О, О-Р, О-ЦМР, Р-ЦМР і всі модальності разом (усі). Найкращі результати позначені жирним шрифтом

Модель	Всі	О-О	О-Р	О-ЦМР	Р-ЦМР
SIFT-ОСТ	65,86	65,78	73,51	58,97	68,21
НОРС	70,67	78,41	70,16	67,43	67,26
MIND	72,32	85,15	70,31	68,61	64,51
L2-Net	60,65	71,21	55,50	61,85	55,41
DSM	80,66	87,77	79,32	76,74	76,20
DSM, трипл. втрати	83,46	90,18	80,44	81,19	81,14
DLSTM, базова версія	83,67	88,26	82,32	79,30	83,05
DLSTM, О-О	79,21	90,26	74,07	74,35	74,14
DLSTM, збалансоване навчання	85,46	88,67	85,25	84,84	84,05
Модифік. DLSTMmod, збалансоване навчання	86,64	89,20	85,36	86,16	85,57

Серед порівнюваних методів МП DSM з триплетними втратами працює близько до базової моделі DLSTM, решта методів демонструють гіршу



ефективність. Порівняння DSM з триплетними втратами і DLSM ілюструє складність навчання МП, придатного для різних комбінацій модальностей: обидві моделі мають подібні значення AUC в загальному випадку; перший є кращим у випадках суміщення О-О зображень та оптичного зображення до ЦМР; другий є кращим для О-Р суміщення.

Давайте порівняємо між собою базову модель DLSM, спеціалізовану модель, навчену спеціально для О-О суміщення, модель DLSM, навчену з балансуванням навчальних даних, і модель DLSM з розширеним входом, позначену як DLSMmod. У нашому аналізі використовуються два добре відомі критерії: робоча характеристика приймача (receiver operating curve - ROC) і AUC [38].

Модель, навчена спеціально для О-О випадку, демонструє, як очікувалося, найвище значення AUC для О-О режиму суміщення, тобто 90,26% (на 2% вище, ніж для базової моделі). В інших випадках значення AUC знаходяться між значеннями для MIND та базової моделі.

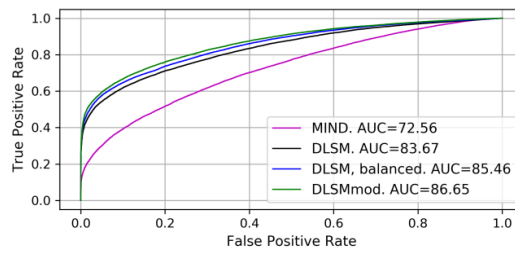
Після збалансування набору навчальних даних AUC для загального випадку збільшується приблизно на 1,8% до 85,55%. Значення AUC для О-Р, О-ЦМР і Р-ЦМР суміщення збільшилися на 1,0-5,5% в залежності від типу суміщення. При цьому значення AUC для О-О випадку практично не змінилося. Таким чином, балансування наборів даних покращує якість МП лише для деяких випадків, не покращуючи для інших.

Після додавання модальності як додаткового вхідного параметра AUC в загальному випадку збільшується на додаткові 1,2% і досягає 86,64%. Порівняно з первинною версією приріст AUC є значним і дорівнює 3%. Важливо, що нейромережа DLSMmod дозволила збільшити AUC для всіх випадків реєстрації в порівнянні з DLSM зі збалансованими навчальними даними. Основний приріст спостерігається для випадку О-ЦМР суміщення, для якого значення AUC зросло з 84,84 до 86,16%. Для випадку О-О суміщення значення AUC зросло до 89,2%. Це на 1% краще, ніж для первинної версії, але на 1% менше, ніж у моделі, спеціально підготовленої для О-О випадку.

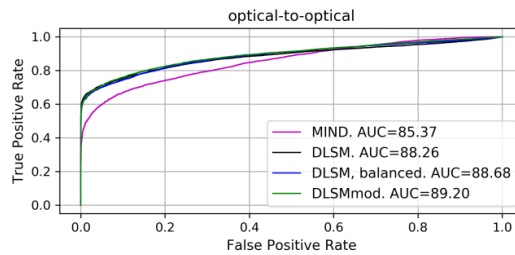
На рисунку 4 наведені суміщені оптичне та радіолокаційне зображення. Їх аналіз показує, що зображення суттєво відрізняються між собою. Втім, МП на основі нейромереж дозволяють знайти подібні блоки та виконати суміщення.

Таким чином, розглянуто проблему навчання ефективної міри подібності, яка підходить для мультимодальних зображень ДЗ. Виявлено, що три модифікації здатні суттєво покращити якість навченої МП: балансування вхідних даних, додавання індикаторів модальності як додаткових вхідних

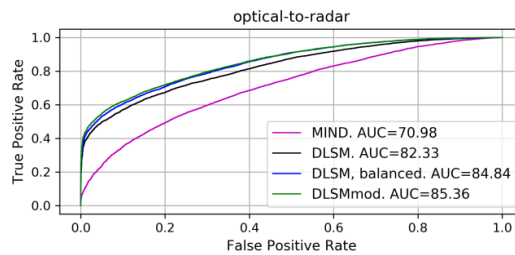
каналів та використання шарів пакетної нормалізації.



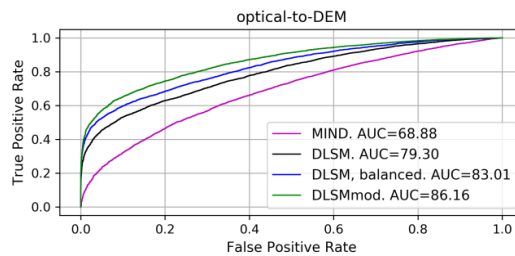
а)



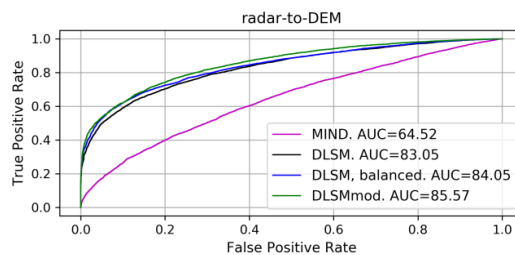
б)



в)



г)



д)

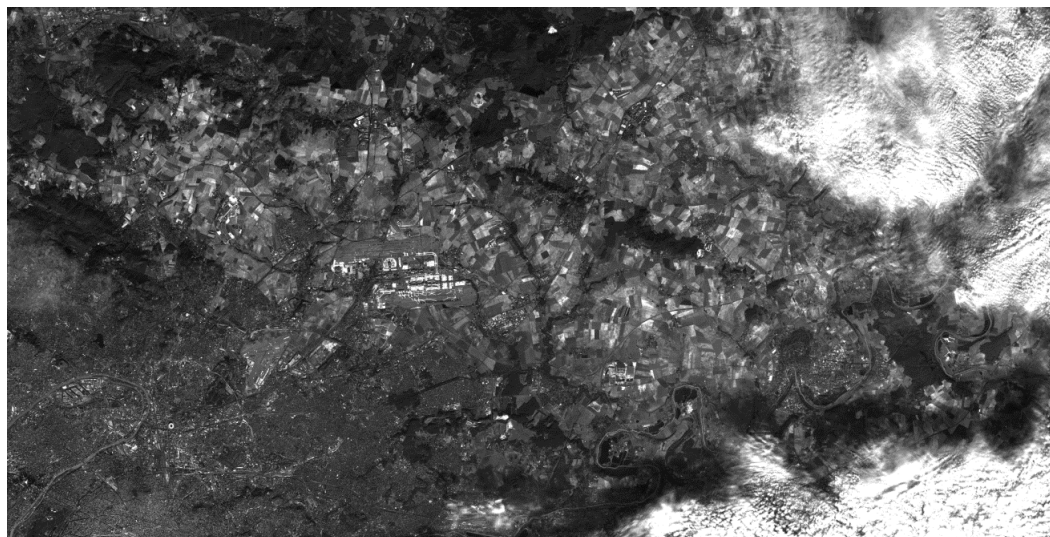
а) загальний випадок; б) O-O випадок; в) O-P випадок;
 г) O-ЦМР випадок; д) P-ЦМР випадок.

Рисунок 3 – ROC-криві для трьох МП

Просте збалансування даних здатне покращити якість МП лише для деяких



комбінацій модальностей. Безпосереднє надання інформації про модальність на вхід CNN додатково покращує значення AUC для всіх комбінацій модальностей. Однак додавання інформації про модальність робить вхідні дані неоднорідними та знижує ефективність навчання CNN. Введення шарів BN допомогло вирішити цю проблему навчання.



a)



б)

a) оптичне зображення; б) радіолокаційне зображення.

Рисунок 4 – Приклад суміщення зображень



6.2. Прогнозування якості даних ДЗ, що базується на застосуванні нейронних мереж

6.2.1. Прогнозування ефективності фільтрації зображень

Результат використання даних ДЗ для вирішення кінцевих задач залежить від багатьох чинників: розрізняювальної здатності системи, ефективності методів обробки даних на кінцевих етапах тощо. Але насамперед він залежить від якості первинних даних та ефективності їх попередньої обробки, якщо вона використовується (якщо якість первинних даних є такою, що попередня обробка є доцільною). Для багатьох сучасних систем ДЗ якість отриманих зображень є такою, що у їх фільтрації немає необхідності. Наприклад, це стосується більшості компонент гіперспектральних та мультиспектральних даних ДЗ [43-45]. Втім, є типи зображень, для яких якість є невисокою через різні причини. Такими, наприклад, є радіолокаційні зображення, на яких завжди присутній спекл [46, 47] (дивись приклади на рис. 5).

Проблема спеклу розглядалася більше 40 років тому в працях Дж.-С. Лі, В. Фроста та інших [48-50]. З тих пір було запропоновано велику кількість фільтрів для видалення спеклу, які базуються на різних принципах [48-54]. «Найстаріші» методи видалення спеклу [48, 49] використовують ковзні вікна та м'яку локальну адаптацію. Багато з них пізніше були модифіковані та вдосконалені [53, 55]. Особлива увага придушенню спеклу на РСА-зображеннях приділялася методам на основі ортогональних перетворень, це було приблизно 15-20 років тому [48, 52]. Були застосовані різні перетворення, такі як вейвлети, дискретне косинусне перетворення та інші. Пізніше основний інтерес дослідників привернули нелокальні методи придушення шумів [53, 56]. В останні роки апробовані білатеральна фільтрація [58], обробка на основі анізотропної дифузії [59] та інші підходи. Також стали популярними методика на основі глибокого навчання [60].

Незважаючи на всі ці досягнення та здобутки в сенсі покращення ефективності видалення спеклу, досі не існує єдиної думки, який фільтр є найкращим і які його параметри. З цього приводу наведемо наступні зауваження:

а) ефективність фільтра залежить від багатьох факторів, включаючи параметри, що використовуються; параметри можна змінювати та встановлювати залежно від типу фільтра; ці параметри включають розмір ковзного вікна [48], пороги [53, 54, 61], розмір блоку [61], параметри перетворень, що стабілізують дисперсію [53, 54], і кількість блоків, оброблених спільно в рамках нелокального підходу до видалення спеклу [53, 56];

б) ефективність придушення спеклу значною мірою залежить від

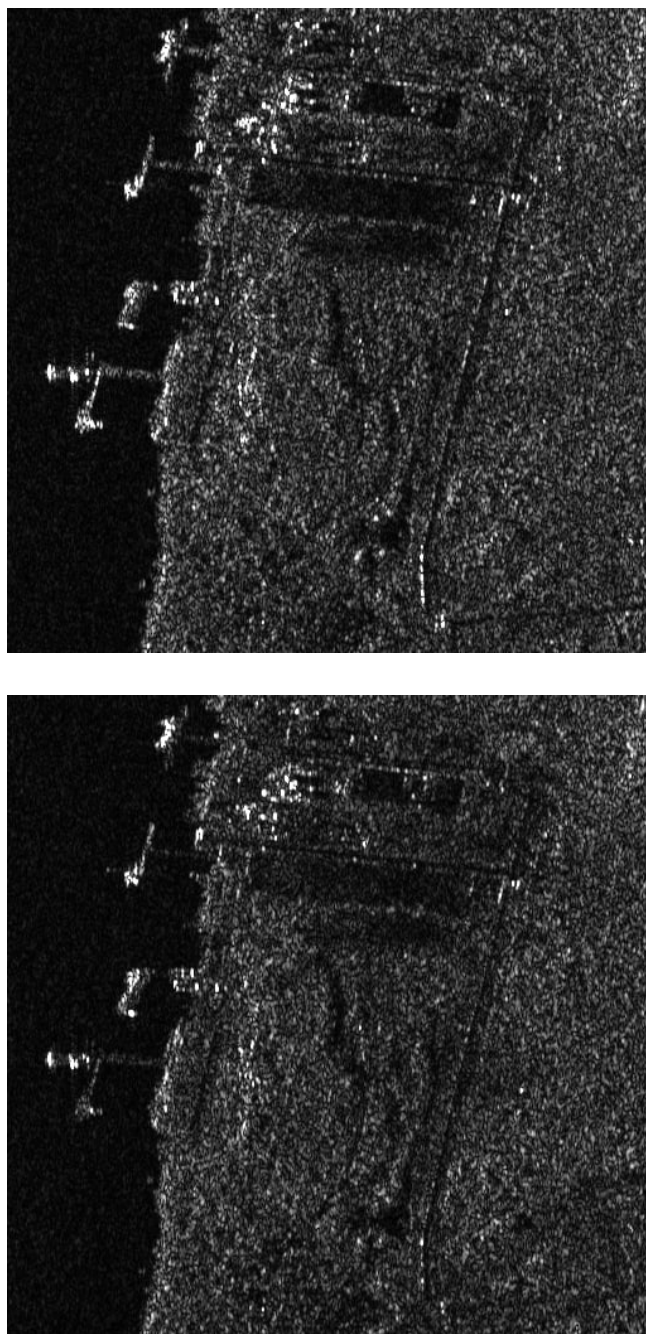


Рисунок 5 – Приклади радіолокаційних зображень з різними поляризаціями, що спотворені спеклом високої інтенсивності

властивостей зображення; для зображень простіших структур (які містять великі однорідні області) зазвичай досягається краща ефективність порівняно із зображеннями із складними структурами (які містять багато меж, невеликих об'єктів і текстур) [60-62];

в) властивості спеклу також впливають на ефективність фільтра; існують фільтри, застосовні до спеклу з щільністю ймовірності, близькою до Гауссової, але для деяких інших фільтрів такого обмеження немає; просторова кореляція спекла (і шуму загалом) відіграє ключову роль в ефективності його придушення



[65, 66]; це означає, що просторова кореляція спекла повинна бути відома заздалегідь або попередньо оцінена [67], а потім врахована при виборі фільтра та/або його параметрів;

г) ефективність фільтра може бути оцінена за допомогою різних кількісних критеріїв; для зменшення шуму в РСА-зображеннях зазвичай використовують пікове відношення сигнал/шум (PSNR) та ефективну кількість поглядів [53-56], хоча також застосовуються інші критерії; зокрема, стало популярним використання візуальних показників якості [68-70]; методи придушення спеклу також можна охарактеризувати з точки зору ефективності класифікації зображень після обробки [71-74]; метрика SSIM [75] стала популярною в програмах обробки даних дистанційного зондування, але ця метрика явно не найкраща метрика візуальної якості [76-78] серед тих, які розроблені на поточний момент.

Через ці труднощі доступні програмні засоби (пакети) для обробки РСА-зображень зазвичай мають обмежений набір застосовуваних фільтрів. Наприклад, такі програмні пакети, як ESA SNAP toolbox, ENVI тощо, мають набір фільтрів, таких як Фроста, Лі, модифікований Лі та деякі інші [79, 80]. Загальною особливістю цих фільтрів є те, що вони мають обмежену кількість параметрів, які необхідно встановити. Це розмір ковзного вікна, дисперсія мультиплікативного шуму та/або деякі інші параметри. Навіть у цьому випадку користувач повинен або мати знання (досвід) щодо налаштування параметрів, або спробувати доступні варіанти, перш ніж отримати остаточні результати.

Зверніть увагу, що «оптимальне» налаштування параметра навіть для заданого типу фільтра залежить від властивостей зображення та шуму [81-83]. Для зображень зі складнішою структурою та/або менш інтенсивним шумом більше уваги на етапі налаштування параметрів слід приділяти збереженню меж/деталей/текстур, і, навпаки, для зображень із простішою структурою та/або більш інтенсивним шумом, придушенню шуму. Можливі два підходи автоматичної або напівавтоматичної реалізації цієї «стратегії». Одну можна розглядати як глобальну адаптацію, де швидко аналізуються властивості того зображення, яке підлягає фільтрації, і відповідно встановлюються параметри (наприклад, пороги [83]) фільтра (відносно деяких правил або алгоритмів). Інший підхід [55, 82] стосується локальної адаптації, де параметри (наприклад, розмір ковзного вікна [55] або пороги [82]) змінюються відповідно до певного алгоритму. Зауважимо, що в обох випадках або глобально, або локально виконується деякий простий попередній аналіз. Рішення приймаються на основі деяких прогнозів ефективності фільтрації, глобальної або локальної [53, 54,



82-85].

Як варіант, розглянемо фільтр Лі та розробимо алгоритм встановлення його ковзного вікна глобально (для всього зображення або, принаймні, для його великих фрагментів). У нас є три основні гіпотези. По-перше, завдяки правильному налаштуванню розміру вікна фільтра часто можна досягти достатнього покращення ефективності видалення спеклу у порівнянні з випадком «середнього» налаштування (наприклад, 7×7 пікселів для будь-яких випадків). По-друге, прогнозування ефективності фільтра для різних розмірів вікон можна зробити досить легко і точно, а потім прийняти рішення про оптимальний розмір. По-третє, підхід, розроблений на основі двох попередніх гіпотез, було частково адаптовано до властивостей розглянутого класу РСА-зображень, у даному випадку зображень Sentinel-1 з кількістю поглядів приблизно рівною п'яти. Між тим, ми вважаємо, що запропонований підхід є загальним і, після модифікацій, може бути застосований до інших типів РСА-зображень та фільтрів для придушення спеклу.

Пояснимо, чому ми розглядаємо локальний статистичний фільтр Лі (або просто фільтр Лі) – це добре відомий фільтр, вихід для якого можна ефективно обчислити і який широко використовується в обробці РСА-зображень. У цьому дослідженні розглядаються зображення Sentinel-1, оскільки ці дані є відкрито доступними та вони отримуються з високою періодичністю, тобто їх доводиться постійно обробляти.

Ще однією особливістю нашого дослідження є те, що ми спираємося на результати, отримані в наших попередніх роботах [54, 83, 85]. У роботі [54] показано, що ефективність фільтра (за кількома критеріями) при обробці РСА-зображень Sentinel-1 можна точно прогнозувати за допомогою навченої нейронної мережі (neural network - NN). У [83] показано, що можлива глобальна оптимізація параметрів фільтра на основі прогнозування ефективності фільтрації. Нарешті, як впливає з результатів, наведених у [85], прогнозування ефективності фільтрації для фільтру Лі з розміром ковзного вікна 5×5 можливе з високою точністю.

Новизна дослідження (порівняно з нашою попередньою роботою [85] та іншими подібними попередніми дослідженнями [53, 54]) складається з двох аспектів. По-перше, показано, що ефективність фільтрації можна прогнозувати для різних розмірів вікон фільтра Лі із задовільною точністю. По-друге, продемонстровано, що на основі такого прогнозу можна з високою ймовірністю приймати правильні рішення щодо оптимального розміру вікна. Методологія нашого дослідження включає аналіз властивостей зображення та шуму для РСА-

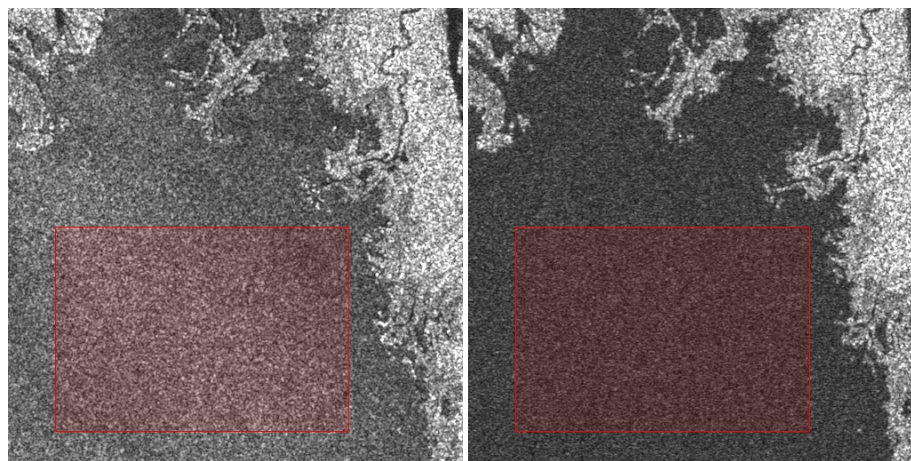


зображень, отриманих системою Sentinel-1, статистичний аналіз оптимальних розмірів вікна залежно від використовуваної метрики ефективності фільтрації, дизайн нейромережі та її тестування для змодельованих даних з аналізом точності прогнозування, перевірку запропонованого підходу для реальних РСА-зображень.

Модель зображення та шуму, яка нами використовується, спирається на загальну інформацію про властивості РСА-зображень та спеклу [49] та на доступну інформацію щодо характеристик спеклу у зображеннях Sentinel-1 [86, 87]. Поширеним є припущення, що спекл є чисто мультиплікативним. Експерименти підтвердили, що це припущення є правильним як для поляризації VV (вертикальна-вертикальна), так і для VH (вертикальна-горизонтальна) для даних радара Sentinel-1 [64, 86]. Відносна дисперсія спеклу σ_{μ}^2 приблизно дорівнює 0,05 як для поляризації VV, так і для VH [87]. Щільність розподілу ймовірності (ЩРЙ) спеклу не є «строго» гаусовою, але досить близькою до нього. Іншою важливою властивістю є те, що спекл на зображеннях Sentinel-1 є просторово корельованим. Це було доведено експериментами, проведеними в роботах [54, 86, 87] для обох поляризацій. Приклади фрагментів зображення розміром 512×512 пікселів наведені на рисунку 6.

Просторову кореляцію можна виявити та охарактеризувати різними способами: а) шляхом візуалізації та аналізу двовимірної (2D) просторової автокореляційної функції або її основних перерізів; б) шляхом аналізу спектру потужності Фур'є; в) шляхом аналізу просторових спектрів для інших ортогональних перетворень; всі мають бути визначені в однорідних областях зображення або оцінені за допомогою деяких стійких методів, здатних усунути або мінімізувати негативний вплив вмісту інформації зображення на отримані оцінки. Такий аналіз було проведено, результати якого представлені в [54, 86, 87]. Доведено, що спекл є просторово корельованим, і продемонстровано можливість моделювання спекла з тими ж характеристиками, що й у зображеннях Sentinel-1 [84].

Для кількісного аналізу ефективності фільтрації нами використано три типи метрик, визначених для змодельованих зображень. Спочатку ми визначаємо значення метрик для первинних зображень із шумом, які були отримані шляхом штучного введення спеклу з вищезгаданими властивостями до безшумних зображень. По-друге, ми обчислюємо значення показників після фільтрації. По-третє, ми оцінюємо «покращення» метрики завдяки видаленню спеклу, визначене як $IM = M_f - M_{inp}$, де M_f і M_{inp} є значеннями метрик після та до придушення спеклу.



а)

б)

а) VV поляризація; б) VH поляризація.

Рисунок 6 – Фрагменти зображень Sentinel-1 SAR розміром 512×512 пікселів із двома поляризаціями VV і VH з вибраними вручну квазіоднорідними областями, позначеними червоними прямокутниками

Для подальшого аналізу ми вирішили використати три показники: звичайне пікове відношення сигнал/шум (PSNR), PSNR-HVS-M [88] (пікове відношення сигнал-шум з урахуванням системи людського зору (HVS) і маскуванню (M)), а також версію індексу схожості ознак (FSIM) у відтінках сірого [89]. PSNR є стандартною метрикою, яка часто використовується в аналізі. Інші дві метрики є показниками візуальної якості, які є одними з найкращих для опису якості зображення у відтінках сірого (одноканального). Оскільки наразі не існує універсальних показників візуальної якості, ми вважаємо за краще використовувати й одночасно аналізувати два показники візуальної якості, які базуються на різних принципах. Крім того, властивості цих показників добре вивчені. Наприклад, PSNR і PSNR-HVS-M обидва виражені в дБ, і їх більші значення мають відповідати кращій якості. Пороги видимості спотворень для цих показників встановлені в [90]. Відомо також, що можна помітити різницю в якості оброблених зображень приблизно на 0,5 дБ і більше. Покращення PSNR понад 6 дБ та PSNR-HVS-M понад 4 дБ необхідні, щоб з високою ймовірністю констатувати, що візуальна якість PCA-зображення була покращена завдяки фільтрації. Метрика FSIM коливається в межах від 0 до 1 і має бути більше 0,99, щоб показати, що шум або спотворення невидимі [90]. Це дуже рідко трапляється для PCA-зображень, як первинних (зашумлених), так і фільтрованих. Через нелінійність FSIM важко сказати, яким має бути покращення за рахунок фільтрації, щоб гарантувати, що оброблене зображення має кращу візуальну якість, ніж відповідне первинне.



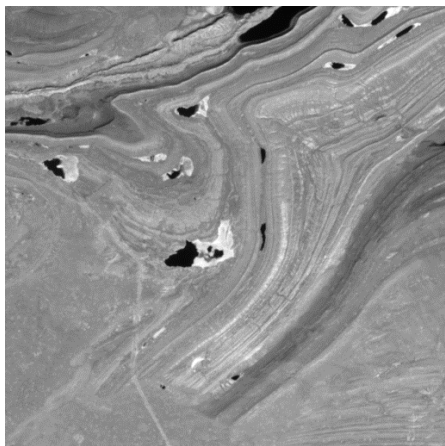
Нижче ми наведемо кілька прикладів вихідних даних фільтра Лі. Оскільки не існує загальноприйнятих безшумних і зашумлених тестових РСА-зображень, звичайною практикою є створення деяких штучних безшумних зображень або використання деяких практично безшумних зображень, отриманих іншими датчиками. У роботі [64] нами було використано компонентні зображення з каналів №5 і №11 мультиспектрального сенсора Sentinel-2. На рисунках 7-9 представлені деякі приклади. Зверніть увагу, що нами розглядається чотири розміри ковзного вікна: 5×5 , 7×7 , 9×9 і 11×11 пікселів.

На рисунку 7 наведено приклад ситуації, коли оброблене зображення має середню складність. Фільтри Лі з усіма чотирма розмірами ковзного вікна покращують усі три розглянуті показники (порівняно із первинним зображенням, значення показників розміщуються під відповідними зображеннями). Найкращий (найбільший) PSNR забезпечує ковзне вікно 9×9 пікселів, хоча PSNR для вікон 7×7 і 11×11 пікселів дуже близькі.

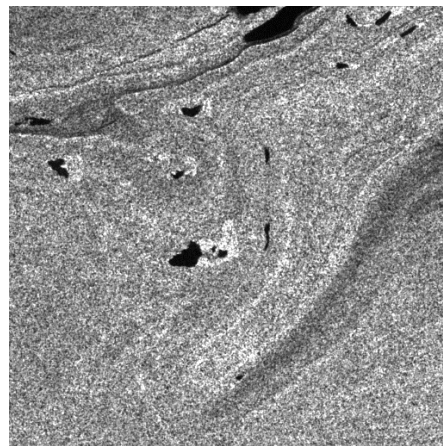
Те ж саме стосується метрики PSNR-HVS-M, хоча результати для вікон 7×7 і 5×5 дуже близькі. Для FSIM найкращий розмір вікна – 7×7 пікселів. Результати для вікна 5×5 найгірші за всіма критеріями, і, справді, придушення спеклу є недостатнім. Щодо інших трьох вихідних зображень, то суб'єктивні думки щодо їх якості можуть відрізнятися.

На рисунку 8 представлений «граничний» випадок, коли зображення майже однорідне. У цьому випадку значення показників суттєво зростають (покращуються), якщо розмір ковзного вікна збільшується. Найбільші покращення спостерігаються для вікна 11×11 пікселів, для PSNR і PSNR-HVS-M вони досягають відповідно майже 14 дБ і 12 дБ, чітко показуючи, що вікно 11×11 є найкращим вибором. Це добре узгоджується з інтуїтивними очікуваннями, оскільки саме ефективне придушення спеклу є основною вимогою для обробки в однорідних областях зображення, і така властивість досягається для фільтрів з великими ковзними вікнами.

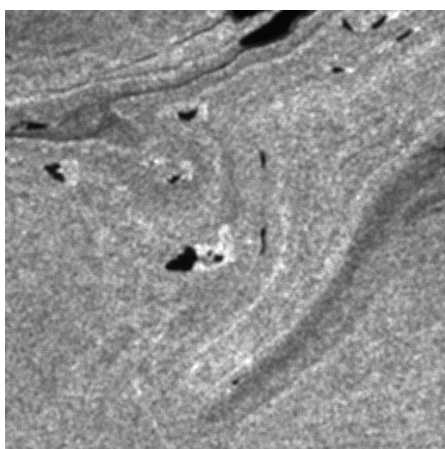
Інший «граничний» випадок показаний на рисунку 9. Тестове зображення має складну структуру (є текстурним). Завдяки цьому найкращі результати забезпечує ковзне вікно 5×5 пікселів за всіма проаналізованими метриками. Однак важко судити, чи покращилася якість зображення завдяки фільтрації чи ні, хоча покращення показників є позитивними. Якщо розмір вікна збільшується, якість вихідного зображення зменшується. Це пов'язано з тим, що збереження меж/деталей/текстур є основною властивістю фільтра в розглянутому випадку, а, як відомо, більше ковзне вікно зазвичай призводить до гіршого збереження зображення і, отже, гіршої візуальної якості.



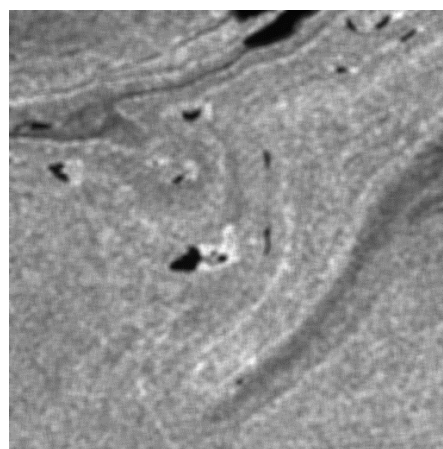
а)



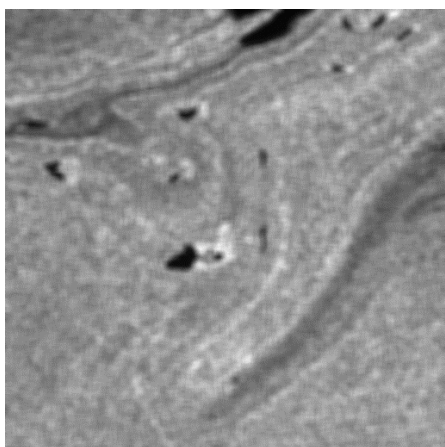
б) PSNR = 22,593 дБ,
PSNR-HVS-M = 19,7998 дБ,
FSIM = 0,70613



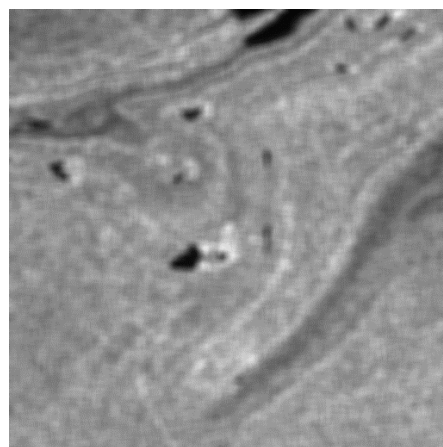
в) 5×5, PSNR = 29,3327 дБ,
PSNR-HVS-M = 24,4202 дБ,
FSIM = 0,83859



г) 7×7, PSNR = 30,2209 дБ,
PSNR-HVS-M = 25,4839 дБ,
FSIM = 0,86763



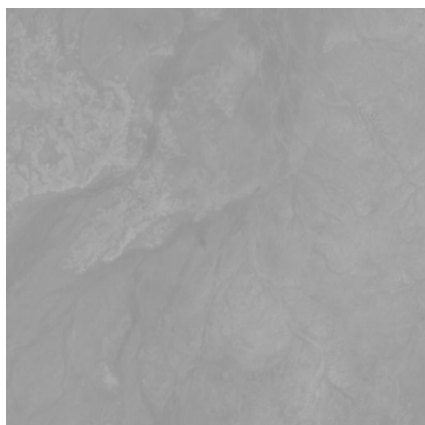
д) 9×9, PSNR = 30,3219 дБ,
PSNR-HVS-M = 25,5236 дБ,
FSIM = 0,8635



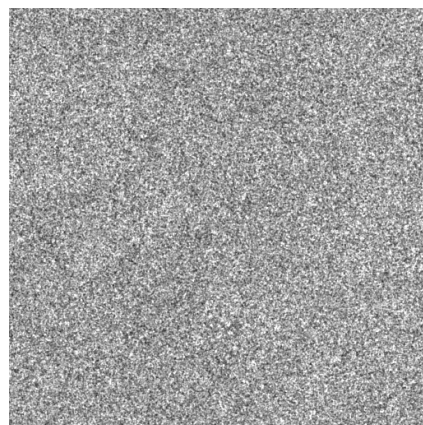
е) 11×11, PSNR = 30,1219 дБ,
PSNR-HVS-M = 25,2094 дБ,
FSIM = 0,84565

а) чисте; б) зашумлене; в-е) фільтровані зображення.

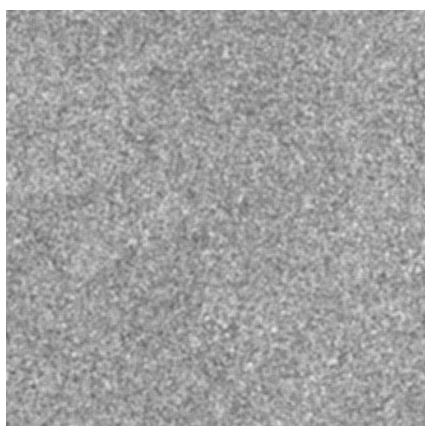
Рисунок 7 – Приклад обробки зображення середньої складності



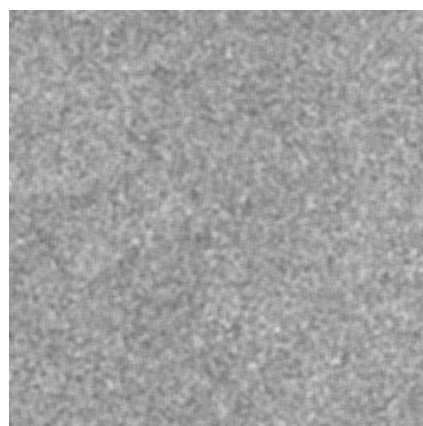
а)



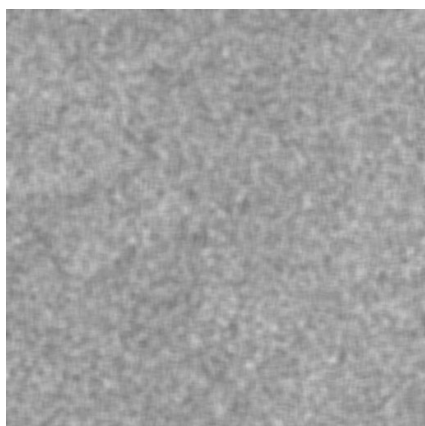
б) PSNR = 19,5309 дБ,
PSNR-HVS-M = 16,6694 дБ,
FSIM = 0,34833



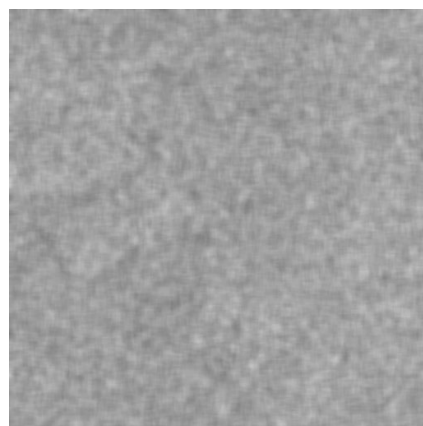
в) 5×5, PSNR = 27,3809 дБ,
PSNR-HVS-M = 21,9575 дБ,
FSIM = 0,56475



г) 7×7, PSNR = 29,883 дБ,
PSNR-HVS-M = 24,5688 дБ,
FSIM = 0,6979



д) 9×9, PSNR = 31,8134 дБ,
PSNR-HVS-M = 26,6571 дБ,
FSIM = 0,78908



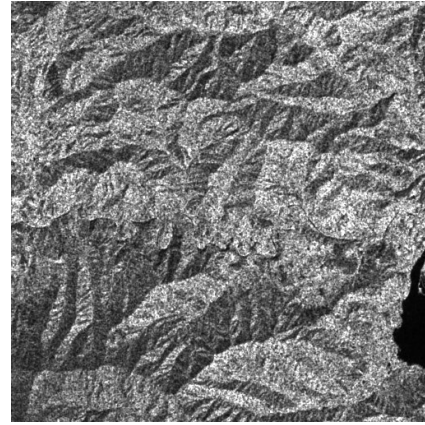
е) 11×11, PSNR = 33,3519 дБ,
PSNR-HVS-M = 28,3137 дБ,
FSIM = 0,84361

а) чисте; б) зашумлене; в-е) відфільтровані зображення.

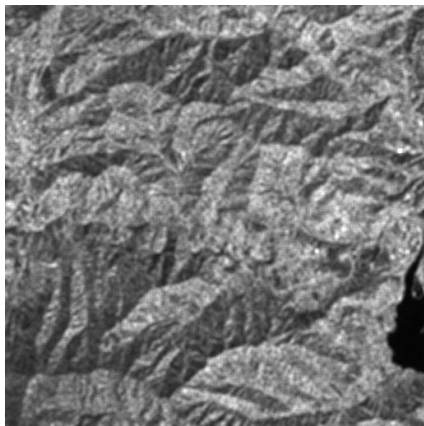
Рисунок 8 – Приклад обробки зображення низької складності



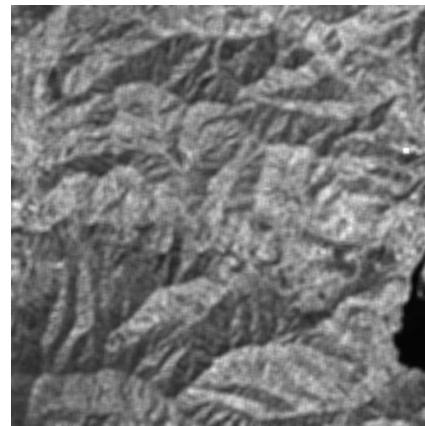
а)



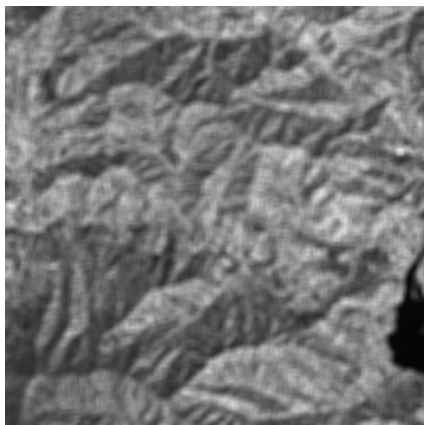
б) PSNR = 32,8715 дБ,
PSNR-HVS-M = 30,4083 дБ,
FSIM = 0,92843



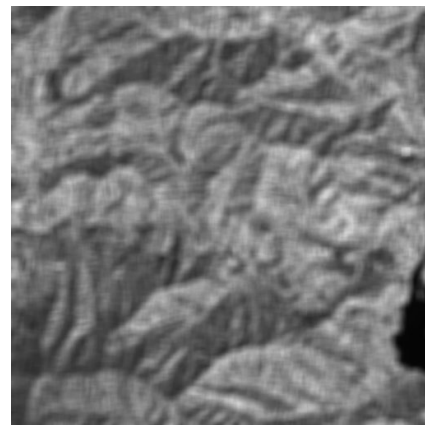
в) 5×5, PSNR = 36,4243 дБ,
PSNR-HVS-M = 32,6792 дБ,
FSIM = 0,95045



г) 7×7, PSNR = 35,4384 дБ,
PSNR-HVS-M = 31,2733 дБ,
FSIM = 0,92181



д) 9×9, PSNR = 34,6121 дБ,
PSNR-HVS-M = 30,001 дБ,
FSIM = 0,88302



е) 11×11, PSNR = 33,9963 дБ,
PSNR-HVS-M = 29,1406 дБ,
FSIM = 0,84598

а) чисте; б) зашумлене; в-е) відфільтровані зображення.

Рисунок 9 – Приклад обробки зображення високої складності



Наведені приклади підтверджують, що оптимальний розмір ковзного вікна сильно залежить від вмісту зображення та використовуваного показника якості. Загальні тенденції наступні. По-перше, для зображень складної структури слід застосувати менше ковзне вікно. По-друге, для візуальних показників якості оптимальний розмір ковзного вікна або такий самий, як оптимальний розмір відповідно до PSNR, або трохи менший.

Це пояснюється двома фактами: а) для показників візуальної якості збереження меж/деталей/текстур «важливіше», ніж придушення шуму в однорідних областях; б) краще збереження меж/деталей/текстури зазвичай забезпечується фільтрами з меншими розмірами вікон (для того самого типу фільтра). По-третє, відмінності у значеннях показників для фільтрів з різними ковзними вікнами можуть бути суттєвими. Наприклад, FSIM у прикладі 1 змінюється від 0,87 до 0,84, PSNR у прикладі 2 варіюється від 27,4 дБ до 33,4 дБ, PSNR-HVS-M у прикладі 3 варіюється від 32,7 дБ до 29,1 дБ. Це показує, що доцільно застосувати оптимальний розмір вікна.

Оскільки вище ми розглядали лише «граничні» випадки, для яких очевидна необхідність вибору (визначення) оптимального розміру вікна, ми також провели додаткове дослідження. Спочатку було визначено ймовірності для кожного розміру вікна, що він є оптимальним. Для досягненн цієї мети було використано 8100 тестових зображень розміром 512×512 пікселів. Після додавання спеклу було застосовано фільтрацію з чотирма розмірами ковзного вікна та розраховано значення показників. Для кожного показника ми визначили, скільки разів кожен розмір вікна давав найкращі результати. Дані, що отримані для всіх трьох розглянутих метрик, наведені на рисунках 10-12. Графік на рисунку 10 показує, що, відповідно до PSNR, вікна 5×5 і 11×11 є оптимальними частіше, ніж вікна 7×7 і 9×9 пікселів.

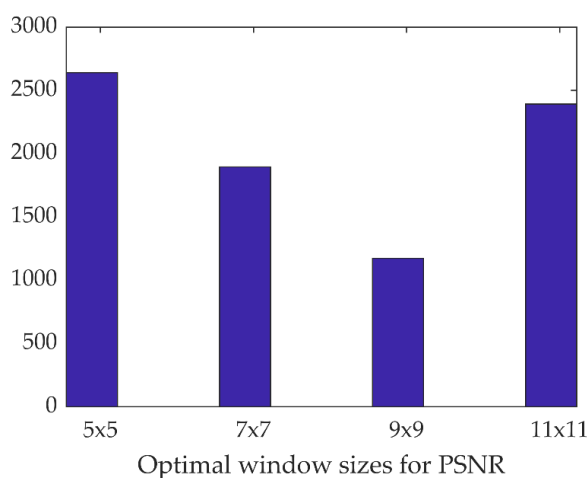


Рисунок 10 – Кількості тестових зображень з різними оптимальними розмірами вікна відповідно до метрики PSNR

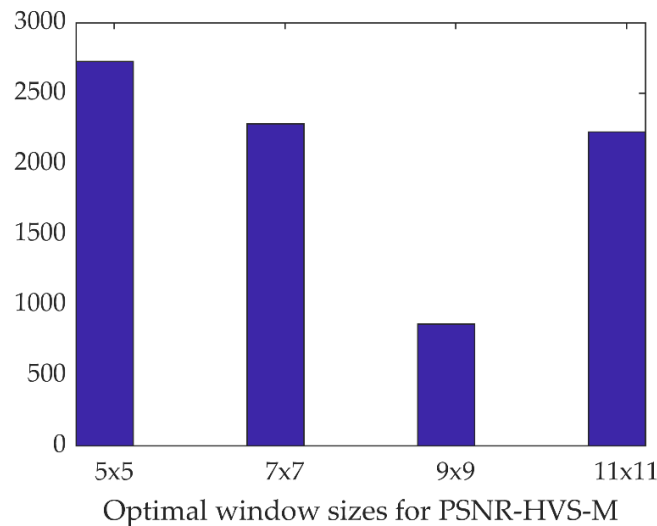


Рисунок 11 – Кількості тестових зображень з різними оптимальними розмірами вікна відповідно до метрики PSNR-HVS-M

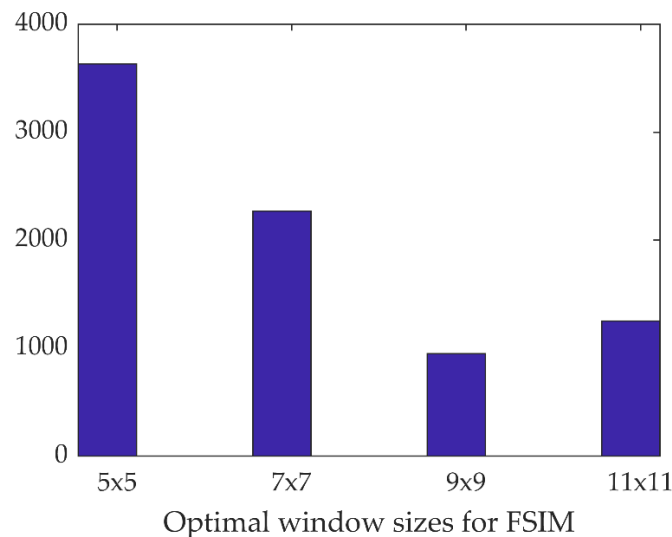


Рисунок 12 – Кількості тестових зображень з різними оптимальними розмірами вікна відповідно до метрики FSIM

Між тим, аналіз графіків на рисунках 9 і 12 демонструє, що, відповідно до PSNR-HVS-M і FSIM, вікна 5×5 і 7×7 кращі частіше, ніж вікна 9×9 і 11×11. Причини, чому це відбувається, були пояснені раніше. Зверніть увагу, що цілком ймовірно, що в деяких випадках оптимальними можуть бути вікна 3×3 або 13×13. Проте наша мета тут – довести, що різні розміри вікон можуть бути оптимальними залежно від властивостей зображення/шуму та критеріїв ефективності фільтрації.

Глобальну адаптацію розміру вікна варто проводити, якщо вигода, що отримується, є високою. Цю вигоду (виграш) можна визначити по-різному. Ми розрахували два наступні параметри

$$\Delta PSNR = Max_{PSNR} - (PSNR_5 + PSNR_7 + PSNR_9 + PSNR_{11}) / 4 \quad (1)$$



$$\Delta PSNR - HVS - M = \text{Max}_{PSNR-HVS-M} - (PSNR - HVS - M_5 + PSNR - HVS - M_7 + PSNR - HVS - M_9 + PSNR - HVS - M_{11}) / 4 \quad (2)$$

де Max_{PSNR} і $\text{Max}_{PSNR-HVS-M}$ є максимальними значеннями (з чотирьох доступних) вихідних PSNR і PSNR-HVS-M відповідно, а нижній індекс відноситься до розміру ковзного вікна (PSNR₅ означає PSNR для вікна 5×5).

Гістограма значень $\Delta PSNR$ представлена на рисунку 13. Вона має моду для $\Delta PSNR$ приблизно 0,4 дБ. Мінімальне значення становить близько 0,1 дБ, а максимальне значення становить близько 3,7 дБ. Це означає, що вигода від правильного підбору оптимального розміру вікна може бути досить великою.

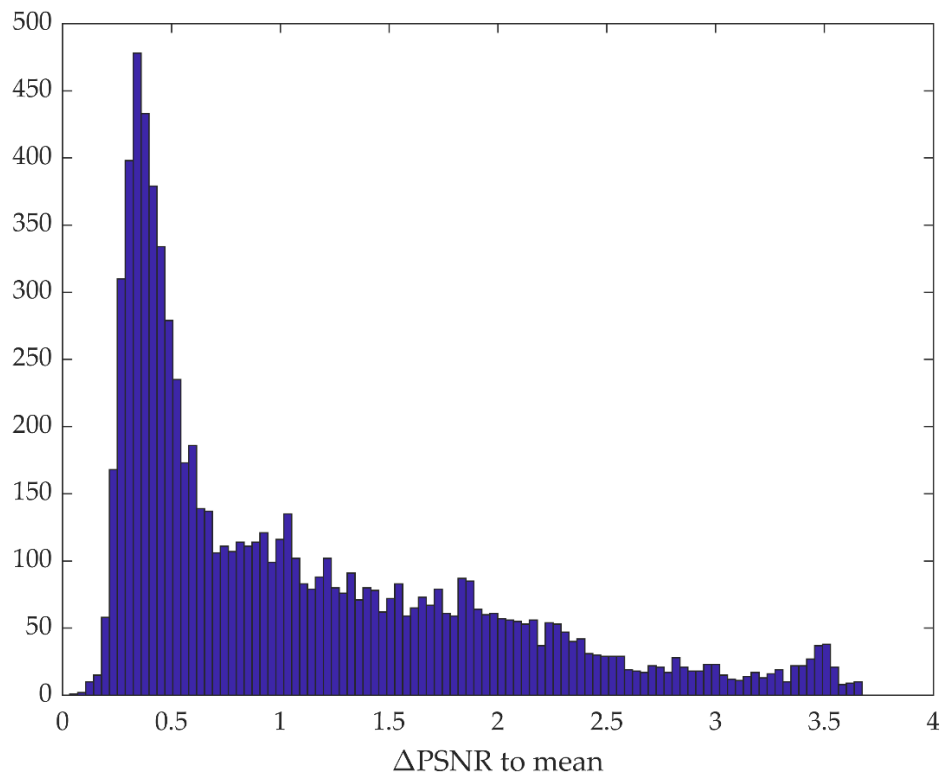


Рисунок 13 – Гістограма $\Delta PSNR$ для тестових зображень

Аналогічно на рисунку 14 (2.10) представлено гістограму $\Delta PSNR-HVS-M$. Мода розподілу становить близько 0,5 дБ, мінімальна вигода (виграш) близька до 0,2 дБ, а максимальна - майже 5 дБ. Таким чином, перша гіпотеза, висунута вище, доведена і варто вибирати оптимальний розмір вікна. Щоб показати, що це можливо, вивчимо, як прогнозувати ефективність фільтрації.

Вище зазначалося, що наш підхід базується на прогнозуванні ефективності фільтрації. Ми припускаємо, що існує метод та/або пристрій, який дозволяє спрогнозувати ефективність фільтру за деякими критеріями (наприклад, за допомогою метрики або кількох показників). Потім можна оцінити (спрогнозувати) ефективність фільтру для набору значень параметрів фільтру

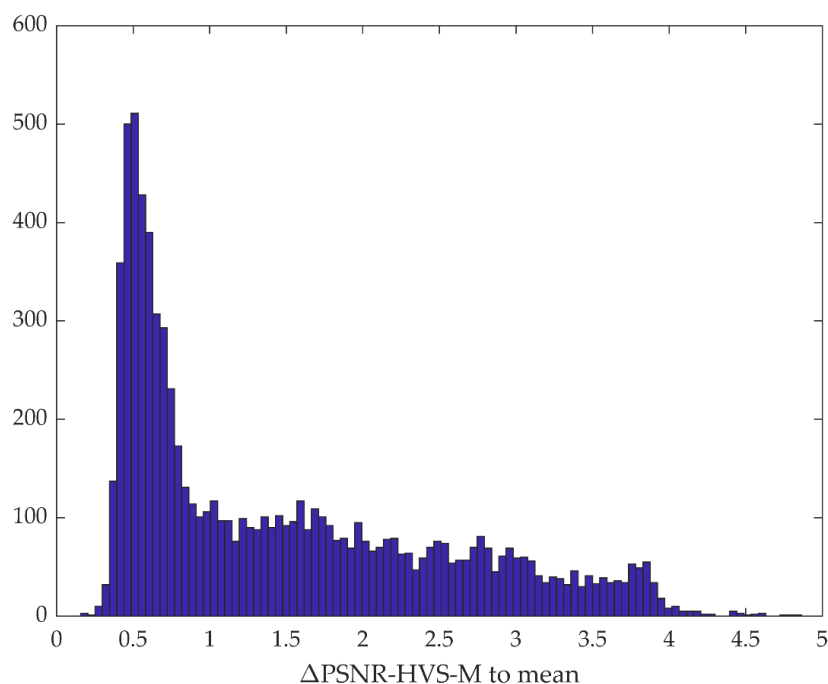


Рисунок 14 – Гістограма Δ PSNR-HVS-M для тестових зображень

(наприклад, розміру вікна для фільтру Лі). На основі цього прогнозу можна прийняти рішення, яке значення розглянутого параметра встановити, щоб отримати «оптимальний результат» для даного зображення. Ядром цього підходу є нейронна мережа-предиктор, натренований в автономному режимі для тестових зображень, які мають приблизно такі ж властивості зображень та шуму, що й зображення з реального життя, які підлягають подальшій обробці.

Коротко пояснимо, що це означає, та вимоги до такого прогнозування. Наше перше припущення полягає в тому, що існує один або кілька параметрів, здатних адекватно характеризувати ефективність фільтрації. Цей аспект вже обговорювався, і далі ми будемо припускати, що покращення PSNR, PSNR-HVS-M і FSIM (позначені відповідно як IPSNR, IPHVSM та IFSIM) можна вважати адекватними показниками. Тоді ми припускаємо, що є один або кілька параметрів, які здатні характеризувати властивості зображення та шуму. Загалом, це можуть бути різні параметри [53, 54, 64, 65] з наступними основними вимогами до них: а) параметри мають бути інформативними, б) вони мають розраховуватись легко та швидко. Нарешті, має бути зв'язок між вибраними вихідними параметрами (прогнозованими показниками або метриками) та вхідними параметрами. Цей зв'язок повинен дозволяти оцінити вихідний параметр(и) за допомогою вхідного(их). Зв'язок може бути реалізований різними способами – як аналітичний вираз, як регресор або як більш складний інструмент, такий як машина опорних векторів (SVM) або нейронна мережа (NN).



Наш попередній досвід розробки та аналізу прогнозування ефективності фільтрів [53, 54, 64, 65] продемонстрував наступне:

– навіть один вхідний параметр (якщо він інформативний і враховує статистику шуму або спектр) здатний забезпечити досить точне прогнозування ефективності фільтрації для багатьох різних методів придушення шуму та критеріїв (метрик) [53, 64];

– спільне використання кількох вхідних параметрів, реалізованих як багатопараметрична регресія [64] або навчена нейромережа [64, 85], зазвичай призводить до суттєвого підвищення точності прогнозування за рахунок необхідних додаткових обчислень.

Було показано [85], що покращення багатьох показників для фільтру Лі з вікном 5×5 можна точно прогнозувати за допомогою навченої нейромережі. Згадаймо тут, що ефективність будь-якої нейронної мережі залежить від багатьох факторів, незалежно від того, які функції виконує NN (апроксимація, класифікація, розпізнавання тощо). Це такі фактори: а) структура та параметри нейромережі (наприклад, кількість прихованих шарів); б) використовувані вхідні параметри та їх кількість, в) тип і параметри функції активації; г) методика навчання НМ.

Важко проаналізувати вплив усіх цих факторів. Через це ми врахували деякі знання та досвід, отримані в наших попередніх роботах [69, 84, 85]. Зокрема, в [84] розглянуто різні набори вхідних параметрів. Спочатку було запропоновано чотири групи вхідних параметрів із загальною кількістю 28 параметрів, що з різних точок зору характеризують статистику ДКП-коефіцієнтів у блоках розміром 8×8 пікселів та статистику зображення. Потім з них було залишено 13 найбільш інформативних параметрів без втрати точності прогнозування, але з меншою складністю нейромережі та простотою обчислень.

Тож для прогнозування ефективності фільтру Лі застосовувались ті ж самі 13 вхідних параметрів. Частина з них може бути розрахована паралельно або конвеєрним способом. Істотне прискорення обчислень також пов'язане з тим [63], що зазвичай можна обробляти дані лише в 1000 блоках, розміщених випадковим чином, щоб отримати необхідну статистику з відповідною точністю.

Можуть використовуватися численні структури нейромереж. Багатошарова структура перцептрона (multi layer perceptron - MLP), представлена на рисунку 15 (2.11), добре зарекомендувала себе при виконанні подібних задач [64, 85].

Приклади зображень, які використовувались при навчанні нейромережі, наведені на рисунках 7-9. Підготувавши 8100 тестових зображень розміром 512×512 пікселів, ми поставили за мету «покрити» різні типи місцевості різної



складності, щоб представити широкий спектр можливих практичних ситуацій.

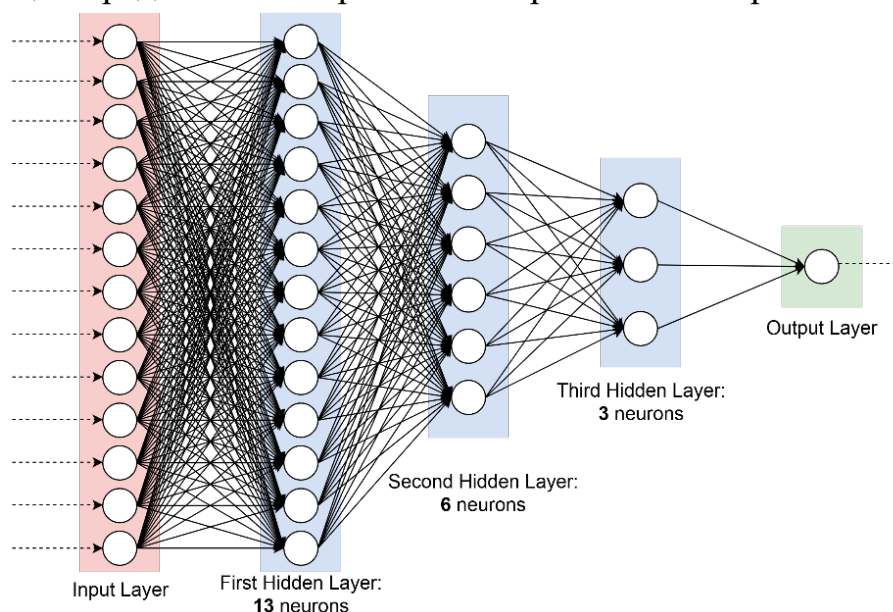


Рисунок 15 – Структура нейронної мережі

Пристрої прогнозування на основі MLP були навчені окремо для кожної з трьох метрик (IPSNR, IPVHSM, IFSIM). Як видно на рисунку 15, NN має три прихованих шари. Для всіх з них використовується функція активації гіперболічного тангенса (\tanh). Для вихідного шару використовується функція лінійної активації. MLP має 13 згаданих вище входів. Ці входні параметри повинні бути розраховані для кожного РСА-зображення, для якого необхідно передбачити ефективність фільтрації. NN була навчена за допомогою зворотного поширення похибки для байєсової регуляризації.

Процес навчання та верифікації НМ складається з чотирьох етапів. Цілі перших двох етапів, які стосуються перевірки по власних наборах даних, - з'ясувати остаточну архітектуру та кількість епох навчання. У свою чергу, етапи 3 і 4 стосуються перехресної оцінки для наборів даних і перевірки точності отриманих рішень з використанням даних, які не використовуються в процесі навчання. Сто високоякісних зображень (без хмар) із загальним розміром близько 5500×5500 пікселів було отримано з компонентів із високим відношенням сигнал-шум мультиспектральних даних ДЗ, отриманих Sentinel-2. Вони були взяті з каналу №5 з довжиною хвилі близько 700 нм і каналу №11 з довжиною хвилі близько 1600 нм. Використовуючи такі великі фрагменти, були отримані зображення розміром 512×512 пікселів (8100 зображень на кожен канал). Ці зображення розміром 512×512 пікселів були використані як безшумні (істинні) зображення, для яких було змодельовано спекл.

На попередньому етапі за допомогою зашумлених і відповідних істинних зображень були визначені значення метрик для входних (первинних) зображень.



Потім було отримано відфільтровані зображення та розраховано значення показників якості для них. Після отримання всіх цих даних було здійснено наступні дії для отримання та перевірки працездатності нейромережі. На етапі перевірки для власного набору даних було встановлено оптимальну кількість епох навчання на рівні 30. Набір даних поділено на дві нерівні частини, 80% зображень були використані для навчання, а 20% зображень, що залишилися, для перевірки. Отримані результати навчання є випадковими. Щоб усунути цей недолік, перевірка була повторена 1000 разів, використовуючи повну перестановку набору даних. Це дозволило отримати середньоквадратичну похибку (RMSE) і параметр R^2 [91] після усереднення для визначення параметрів нейромережі. Менша похибка RMSE і більше значення R^2 відповідають кращим рішенням.

Результати перевірки для власних наборів даних представлені в таблиці 2 (2.1), де були використані тестові зображення для каналу №11. Як бачимо, результати прогнозів дуже хороші. IPSNR і IPHVSM прогножуються з RMSE близької до 0,3 дБ; значення R^2 мають практично однакові й високі значення і знаходяться в межах від 0,976 до 0,989, що показує, що прогнозування, що проведено навченою NN, є досить точним.

Таблиця 2 – Результати перевірки для власного набору даних на тих самих зображеннях Sentinel-2, канал №11

Метрика, що прогнозується	Розмір ковзного вікна	RMSE	R^2
IPSNR	5×5	0,234	0,976
IPSNR	7×7	0,289	0,986
IPSNR	9×9	0,319	0,989
IPSNR	11×11	0,355	0,990
IPHVSM	5×5	0,208	0,966
IPHVSM	7×7	0,303	0,983
IPHVSM	9×9	0,351	0,988
IPHVSM	11×11	0,396	0,989
IFSIM	5×5	0,007	0,984
IFSIM	7×7	0,011	0,987
IFSIM	9×9	0,016	0,986
IFSIM	11×11	0,019	0,985

Якщо нейронна мережа навчається для одного набору даних, а потім застосовується до іншого набору, її характеристики можуть радикально погіршитися. Щоб перевірити цей момент, ми виконали перехресну перевірку.



Було проведено навчання для 6480 тестових зображень, а параметри, що характеризують точність, оцінені для інших 1620 зображень. Отримані результати наведено в таблиці 3.

Перехресна оцінка була виконана для даних Sentinel-2 з того ж каналу №11. Аналіз показує, що значення RMSE дещо збільшилися, а R^2 дещо зменшився порівняно з відповідними даними в таблиці 2. При цьому точність прогнозу залишається дуже високою.

Таблиця 3 – Результати перехресної перевірки для інших зображень Sentinel-2, канал №11

Метрика, що прогнозується	Розмір ковзного вікна	RMSE	R^2
IPSNR	5×5	0,263	0,966
IPSNR	7×7	0,328	0,980
IPSNR	9×9	0,361	0,985
IPSNR	11×11	0,399	0,986
IPHVSM	5×5	0,229	0,951
IPHVSM	7×7	0,338	0,975
IPHVSM	9×9	0,396	0,983
IPHVSM	11×11	0,446	0,985
IFSIM	5×5	0,008	0,981
IFSIM	7×7	0,013	0,983
IFSIM	9×9	0,017	0,982
IFSIM	11×11	0,021	0,980

Зазвичай нейронній мережі ще важче виконувати обробку даних, якщо набори, що використовуються в навчанні та перевірці, достатньо відрізняються. У нашому випадку це може статися, якщо NN навчається для тестових зображень, що складаються з даних в одному каналі зображень Sentinel-2, а потім застосовується до тестових зображень, що складаються з даних в іншому каналі. Щоб перевірити цей випадок, ми провели навчання NN на наборі даних Sentinel-2 у каналі №5, а потім провели перехресну оцінку набору даних із каналу Sentinel-2 №11. Отримані результати представлені в таблиці 4.

Аналіз отриманих результатів показує, що значення RMSE майже збігаються з відповідними значеннями в таблиці 3. Значення R^2 дещо знизилися порівняно з відповідними даними в таблиці 3. Проте прогноз досить точний. Це значно краще, ніж для предикторів, заснованих на одному вхідному параметрі [53] і двох вхідних параметрах [84]. Ми пов'язуємо цю перевагу з двома факторами. По-перше, НМ використовує більше вхідних параметрів, які



використовують інформацію про статистику зображення. По-друге, НМ використовує інформацію про спектральні властивості спеклу. Основні параметри, що характеризують точність прогнозування для версій фільтра Лі з різними розмірами вікна сканування, приблизно такі ж, як і для фільтра на основі ДКП, проаналізованого в роботі [54]. Таким чином, можна зробити висновок про певну загальність розглянутого підходу до прогнозування ефективності фільтрів.

Існує кілька способів адаптувати вікно фільтра для даного зображення до його вмісту на основі прогнозування:

- виконувати прогнозування лише для одного параметра, наприклад IPSNR, для всіх можливих розмірів ковзного вікна та для вибору розміру вікна, для якого прогнозований показник (наприклад, IPSNR) є найбільшим;

- спільно проаналізувати дві або три метрики (наприклад, IPSNR і IPHVSM або IPSNR, IPH-VSM і IFSIM) і прийняти рішення (ймовірно, для цього існує багато алгоритмів);

- отримати три рішення на основі окремого аналізу IPSNR, IPHVSM та IFSIM, як у першому пункті, а потім застосувати алгоритм більшості голосів іншого правила прийняття рішень.

Нижче ми зосередимося на способі, описаному в пункті 1, як найпростішому рішенні, залишаючи інші варіанти на майбутнє.

Рішення можна охарактеризувати різними способами. Нас цікавлять в основному два аспекти – яка ймовірність правильного рішення для нашого підходу і що станеться, якщо прийняте рішення є неправильним, тобто якщо буде вирішено використовувати неправильний розмір ковзного вікна. Ймовірності правильних рішень були визначені для перевірки власних наборів даних та оцінки між наборами даних. Для етапу перевірки власних наборів даних ймовірність правильного рішення приблизно дорівнює 0,9 для IPSNR, 0,918 для IPHVSM і 0,907 для IFSIM. Таким чином, була досягнута висока ймовірність правильних рішень.

Для перехресної оцінки набору даних у каналі (№ 11) ймовірність дорівнює 0,898 для IPSNR, 0,916 для IPHVSM та 0,901 для IFSIM. Для перехресної оцінки набору даних з каналом (№ 5) ймовірність правильного рішення становить 0,857 для IPSNR, 0,899 для IPH-VSM та 0,877 для IFSIM. Для перехресної оцінки набору даних з іншим каналом (№ 5) ймовірність правильного рішення для IPSNR дорівнює 0,857; для IPHVSM – до 0,899, а для IFSIM – до 0,877. Всі ймовірності досить високі.

Давайте тепер подивимося, що станеться, якщо буде прийнято неправильне рішення. Очевидно, що це має призвести до зниження ефективності фільтрації.



Отже, ми оцінили відмінності значень метрики для оптимального (максимально досяжного) значення метрики та значення, отриманого у випадку неправильного рішення. Розподіл таких відмінностей для IPSNR представлено на рисунку 16. Більшість відмінностей дуже малі (менше 0,2 дБ), тому помилкове рішення не є проблемою. Водночас є кілька (шість) випадків, коли різниця перевищує 0,5 дБ.

Аналогічно, на рисунку 17 представлено розподіл відмінностей між оптимальним IPHVSM та відповідними значеннями, отриманими у випадках неправильних рішень. Знову ж таки, більшість відмінностей дуже малі і не перевищують 0,2 дБ. Є лише два тестові зображення, для яких різниця перевищує 0,5 дБ. На рисунку 18 показані відмінності для IFSIM. В основному відмінності дуже невеликі (менше 0,005). Є лише три випадки, коли різниця перевищує 0,01.

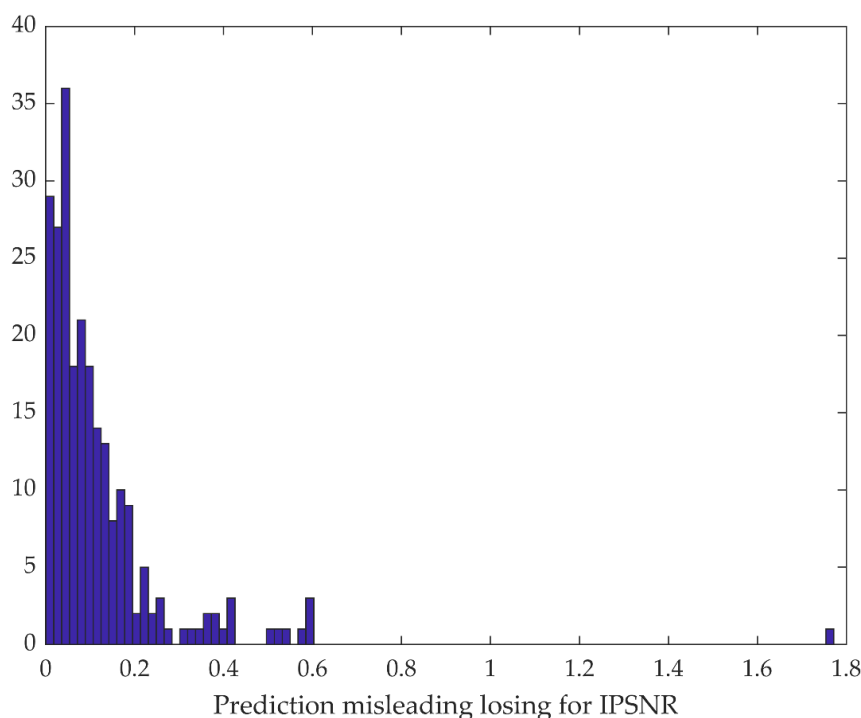


Рисунок 16 – Гістограма відмінностей між оптимальним і отриманим IPSNR у випадках неправильних рішень

Наведемо кілька прикладів прийняття правильних рішень (за будь-яким розглянутим показником). На рисунку 19 показано істинне зображення а), зображення зі спеклом б), оптимальний результат фільтрації для вікна 11×11 в) і вихід фільтра для вікна 5×5 , який, безумовно, не є найкращим вибором. Крім того, ми надаємо всі істинні та всі спрогнозовані значення метрики. У цьому прикладі всі істинні та відповідні спрогнозовані значення близькі один до одного. При цьому всі прогнозовані значення трохи перевищують відповідні істинні значення.

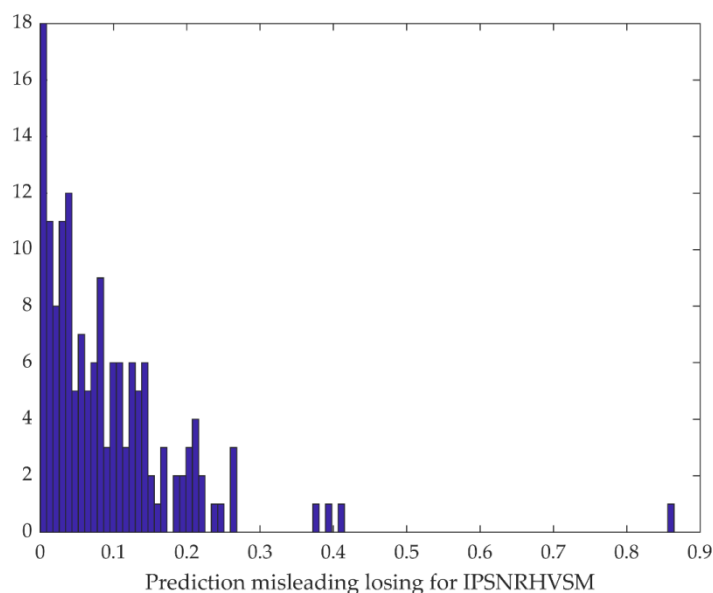


Рисунок 17 – Гістограма відмінностей між оптимальним і отриманим ІРНVM у випадках неправильних рішень

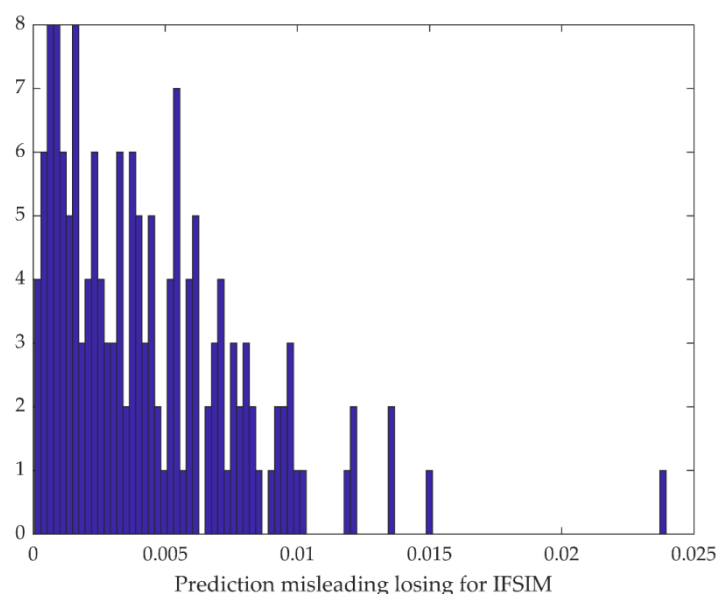
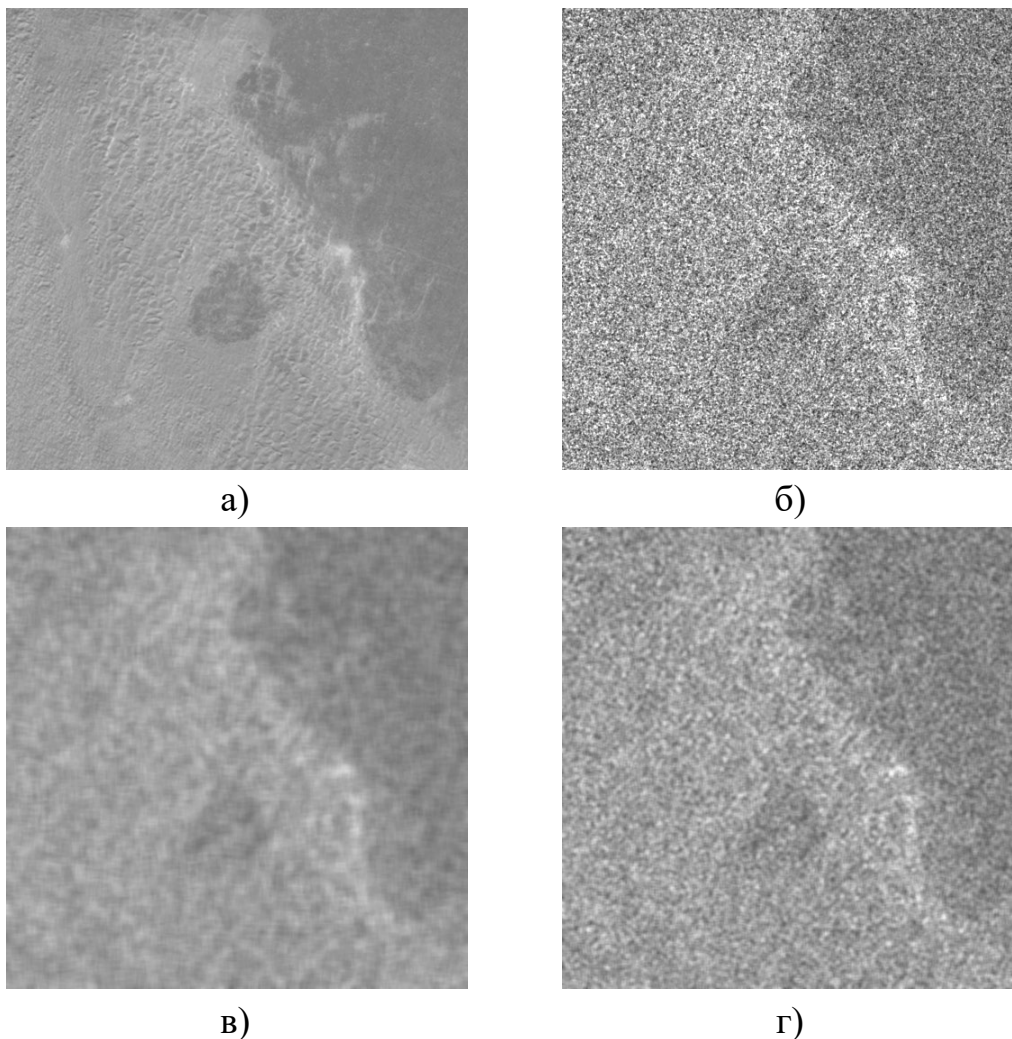


Рисунок 18 – Гістограма відмінностей між оптимальним і отриманим IFSIM у випадках неправильних рішень

На рисунках 20 і 21 показано ще два приклади. Для зображення на рисунку 20 вікно 7×7 є найкращим вибором відповідно до істинних і прогнозованих значень усіх трьох метрик, хоча вікно 9×9 також дає хороші результати. Для зображення на рисунку 21 вікно 5×5 є найкращим вибором за всіма трьома показниками, як істинними, так і прогнозованими. Використання вікна 9×9 призводить до надмірно згладженого результату. Зверніть увагу, що IFSIM може бути негативним, що вказує на погіршення якості зображення через фільтрацію.



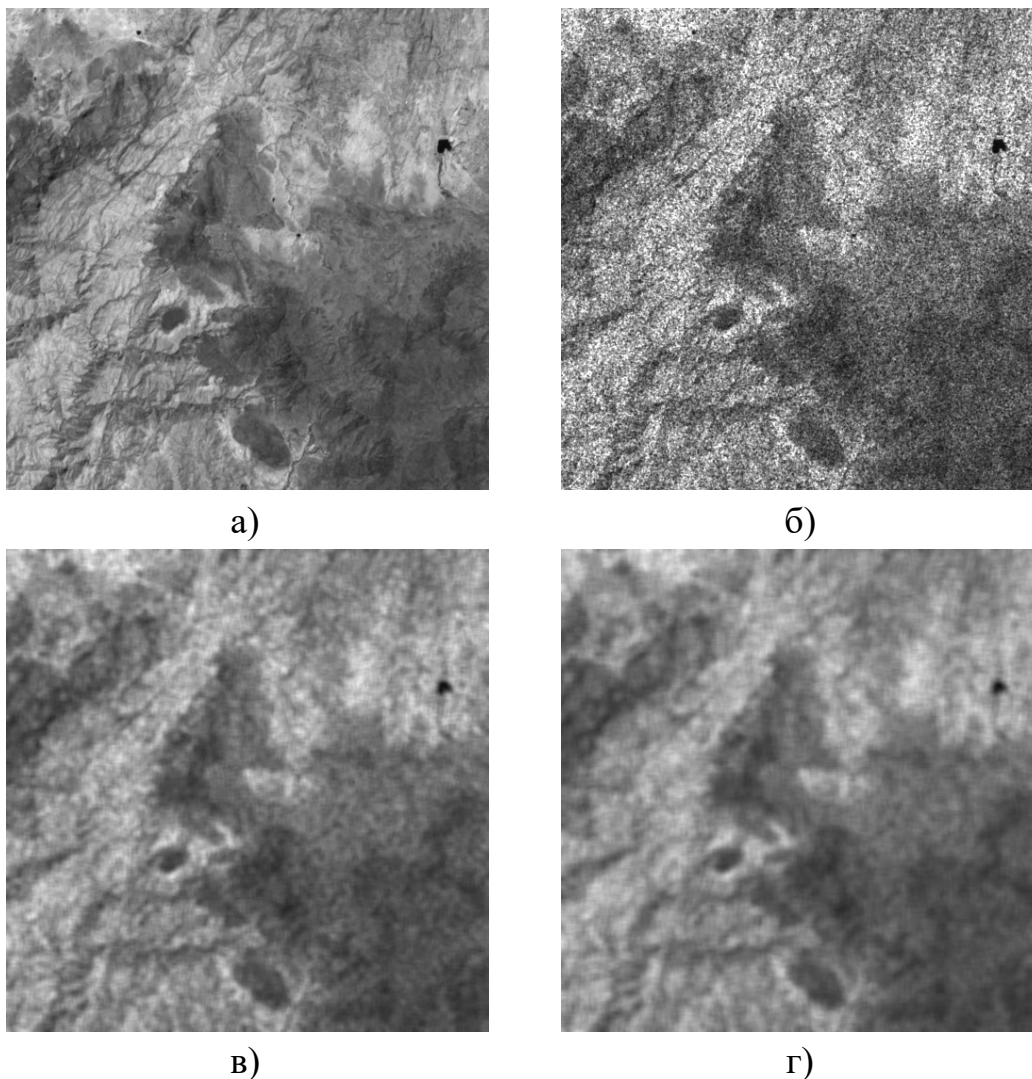
Істинні значення: IPSNR = 7,3178; 9,1137; 10,2396; 10,9724;
 IPHVSM = 4,9985; 6,9851; 8,2321; 9,0462;
 IFSIM = 0,1838; 0,2580; 0,2864; 0,2869.

Спрогнозовані значення: IPSNR = 7,3507; 9,3201; 10,5613; 11,3226;
 IPHVSM = 5,068; 7,2653; 8,6717; 9,5455;
 IFSIM = 0,1853; 0,2654; 0,3004; 0,3077;

*а) істинне зображення; б) зі спеклом; в) на виході фільтру 11×11;
 г) на виході фільтру 5×5.*

Рисунок 19 – Ілюстрація адаптації розміру ковзного вікна

На рисунку 22 показаний приклад неправильного рішення. Відповідно до прогнозованого IPSNR, необхідно використовувати вікно 9×9. Тим часом, відповідно до справжнього IPSNR, вікно 7×7 є найкращим вибором (прогнозований IPHVSM та IFSIM також на користь вікна 7×7). Однак у цьому випадку використання вікна 9×9 не призводить до значного негативного ефекту.



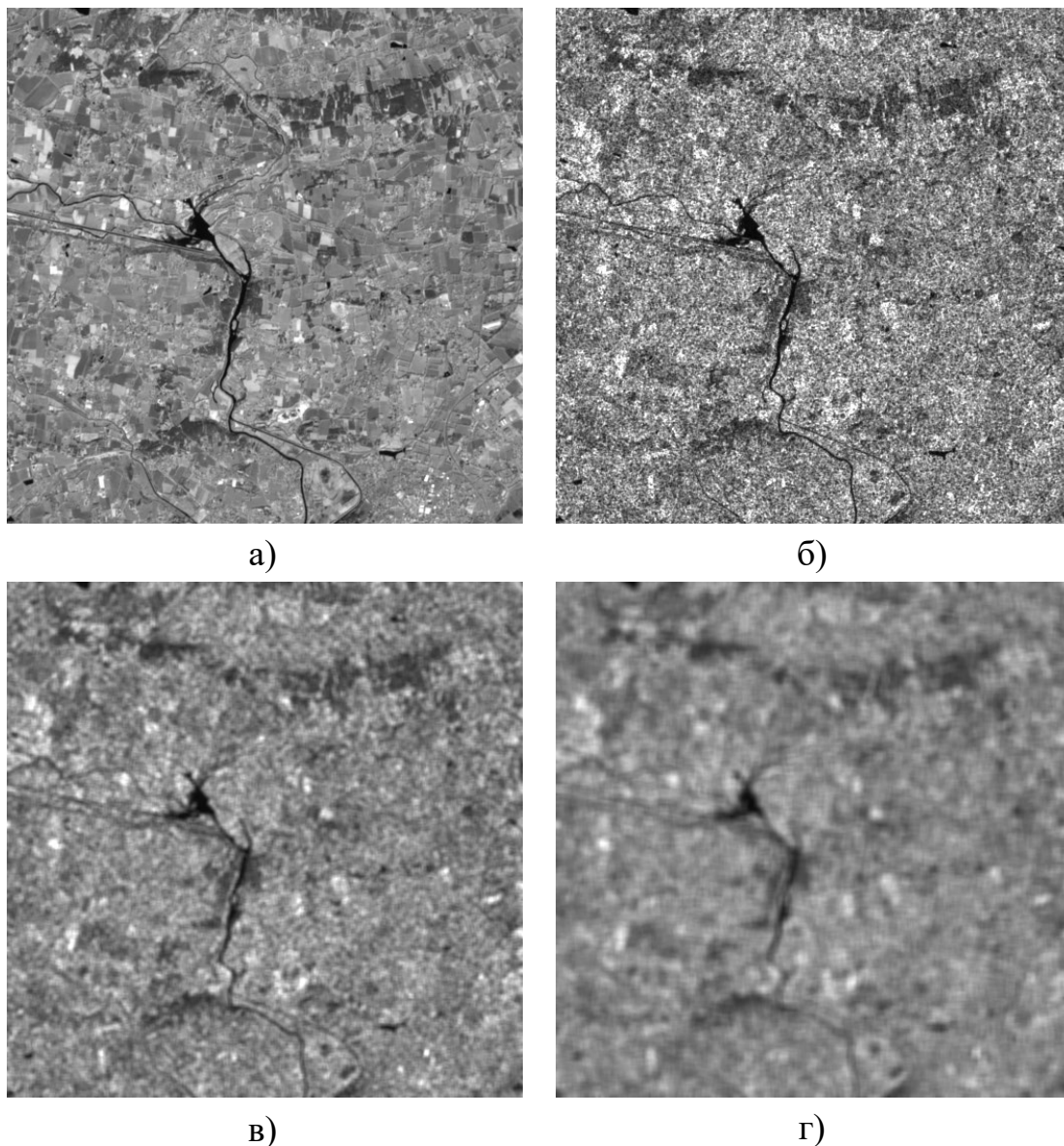
Істинні значення метрик: IPSNR = 6,3279; 6,8486; 6,7304; 6,431;
 IPHVSM = 4,283; 4,8979; 4,6221; 4,1777;
 IFSIM = 0,1028; 0,1074; 0,0801; 0,0452.

Спрогнозовані значення метрик:

IPSNR = 6,4152; 6,9188; 6,8015; 6,5136;
 IPHVSM = 4,4549; 5,1001; 4,9049; 4,5135;
 IFSIM = 0,1103; 0,123; 0,1042; 0,0763;

*а) істинне зображення; б) зі спеклом; в) на виході фільтру 7×7;
 г) на виході фільтру 9×9.*

Рисунок 20 – Ілюстрація адаптації розміру ковзного вікна



Істинні значення: IPSNR = 4,5826; 4,2411; 3,7214; 3,2524;

IPHVSM = 3,2512; 2,7054; 1,8067; 1,0943;

IFSIM = 0,0478; 0,0169; -0,0325; -0,0804;

Спрогнозовані значення: IPSNR = 4,6489; 4,3550; 3,8271; 3,2800;

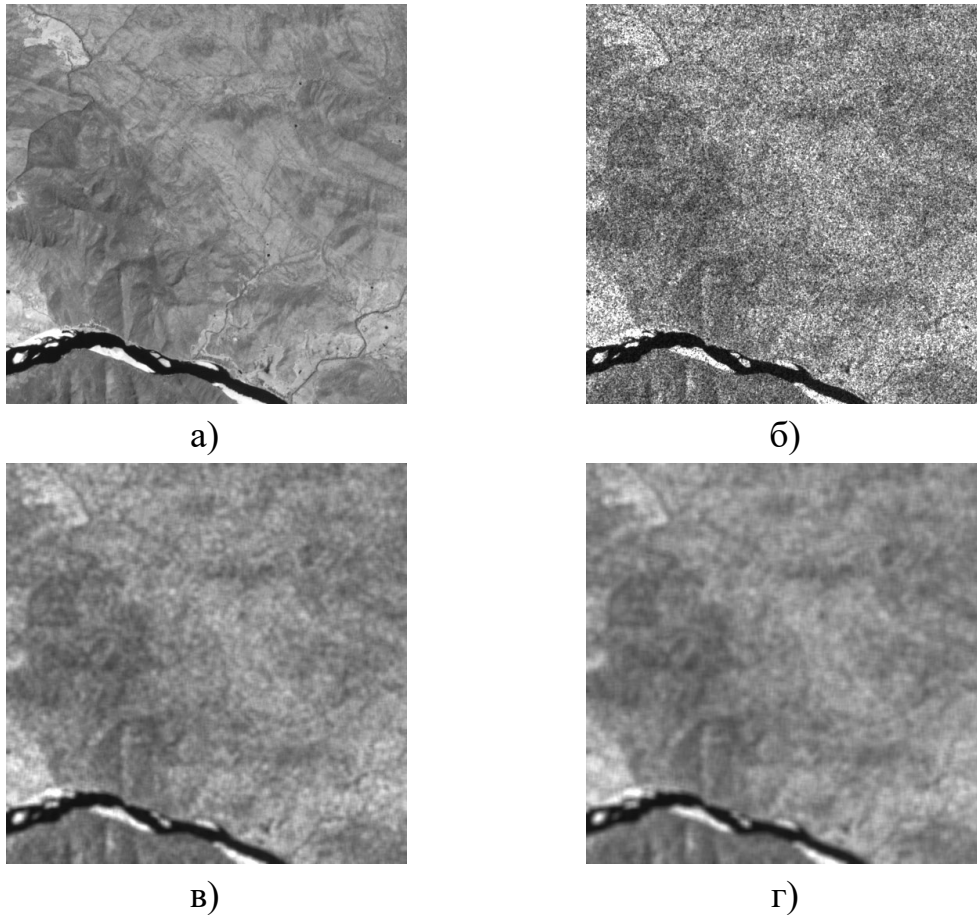
IPHVSM = 3,1585; 2,7088; 1,8091; 1,0398;

IFSIM = 0,0522; 0,023; -0,0259; -0,0725;

a) істинне зображення; б) зі спеклом; в) на виході фільтру 5×5;

г) на виході фільтру 9×9.

Рисунок 21 – Ілюстрація адаптації розміру ковзного вікна



Істинні значення метрик: IPSNR = 6,58; 7,3584; 7,3262; 6.9943;
 IPHVSM = 4,5256; 5,4396; 5,2939; 4,8355;
 FSIM = 0,1209; 0,1457; 0,1363; 0,1129;

Спрогнозовані значення метрик: IPSNR = 6,3703; 6,9498; 6,9727; 6,7566;
 IPHVSM = 4,4884; 5,2494; 5,0829; 4,7303;
 FSIM = 0,1193; 0,1393; 0,1264; 0,1024;

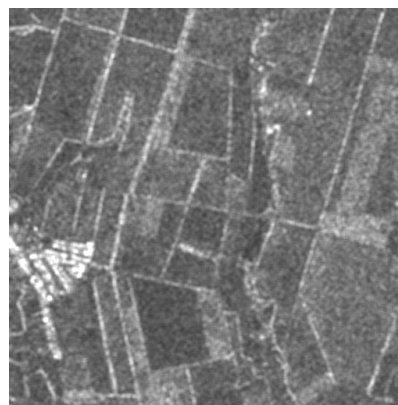
*a) істинне зображення; б) зі спеклом; в) на виході фільтру 7×7;
 г) на виході фільтру 9×9.*

Рисунок 22 – Ілюстрація неправильної адаптації розміру ковзного вікна

Наведемо приклади обробки реального PCA-зображення. На рисунку 23 наведено приклад реального зображення Sentinel-1 розміром 512×512 пікселів. Оскільки в даному випадку ми не маємо істинного зображення, ми можемо продемонструвати результати лише візуально та проаналізувати їх. Первинне зображення було оброблено фільтром Лі з розмірами вікна сканування 5×5 (б), 7×7 (в) і 9×9 (г) пікселів. Прогнозовані значення метрик наведено під відповідними виходами. На «їхню думку», розмір вікна 7×7 є найкращим і, на нашу думку, це добре корелює з візуальним аналізом.



а)


 б) прогнозовані IPSNR = 6,7235 дБ,
 IPHVSM = 4,5873 дБ,
 IFSIM = 0,1247

 в) прогнозовані IPSNR = 7,338 дБ,
 IPHVSM = 5,4404 дБ,
 IFSIM = 0,1478

 г) прогнозовані IPSNR = 7,252 дБ,
 IPHVSM = 4,8769 дБ,
 IFSIM = 0,1389

а) первинне зображення; б) на виході фільтру 5×5 ; в) на виході фільтру 7×7 ;
 г) на виході фільтру 9×9 .

Рисунок 23 – Ілюстрація адаптації розміру ковзного вікна для реального PCA-зображення

Таким чином, показано, що розмір ковзного вікна суттєво впливає на якість вихідних зображень. Тому його адаптивне налаштування видається доцільним. Ми показуємо, як цей розмір можна визначити для даного зображення за допомогою прогнозування ефективності фільтра, реалізованого нейронною мережею. Розглянуто головні аспекти розробки та навчання нейронної мережі. Високу точність прогнозу продемонстровано для трьох показників якості. Показано, що правильні рішення можуть бути прийняті з високою ймовірністю (перевищує 0,85). Вивчаються також випадки прийняття неправильних рішень. Показано, що в більшості ситуацій негативні результати таких рішень є незначними.



6.2.2. Стиснення та розпізнавання зображень

Обсяг отриманих даних ДЗ швидко збільшується через декілька факторів [92-96], зокрема: а) часті спостереження; б) тенденцію до проектування та застосування багатоканальних систем; в) бажання мати кращу просторову роздільну здатність тощо. Ці фактори викликають проблеми в передачі, зберіганні та поширенні даних, що призводить до необхідності стиснення зображення.

Відомо, що існують методи стиснення зображень без втрат і з втратами [97-100]. Часто стиснення без втрат неможливо використовувати через обмежений коефіцієнт стиснення (КС, compression ratio - CR). Тому необхідно застосувати стиснення з втратами. У цьому випадку існують певні вимоги до використовуваного методу. По-перше, слід знайти розумний компроміс між КС та досягнутою якістю стиснутих даних [99]. КС може бути занадто великим, щоб гарантувати, що корисну інформацію можна видобути з даних ДЗ з відповідною точністю та/або надійністю [101]. Таким чином, якість стиснутих зображень необхідно контролювати [101]. По-друге, також можуть бути вимоги з точки зору зелених технологій [102], стандартів [103], обчислювальної ефективності [99], простоти реалізації на борту [104] тощо.

Існує кілька важливих аспектів стиснення з втратами та контролю якості. По-перше, необхідно зрозуміти, що означає якість для розглянутої задачі і які критерії (метрики) можна і потрібно використовувати [105, 106]. Стиснуті дані можуть аналізуватись людьми, а тому можуть бути використані візуальні показники якості [105, 107]. В інших випадках основним завданням може бути класифікація стиснутих зображень, і тоді стиснення з втратами має бути охарактеризоване з точки зору точності класифікації [108-110]. По-друге, стиснення із забезпеченням бажаної якості має бути досить швидким і вимагати обмежених обчислень (особливо якщо стиснення здійснюється на борту). По-третє, рекомендації щодо налаштування параметрів мають бути простими та дійсними для зображень різної складності. По-четверте, КС має бути якомога більшим.

Зважаючи на це, ми розглянули окремий випадок трьохканальних зображень ДЗ, використовуючи як приклад дані Sentinel оптичного діапазону. Застосовано кодер AGU [111], який має як одно-компонентну, так і тривимірну (3D) [112] версії. Використовується класифікатор максимальної правдоподібності (ML) [105, 108], оскільки він є одним із найкращих класифікаторів. Також використано класифікатор на основі нейромереж.

Ефективність стиснення з втратами при застосуванні до даних ДЗ залежить



від багатьох факторів [105, 113, 114]: властивостей зображення та шуму, використовуваного кодера, завдання, яке потрібно розв'язати тощо. Зокрема, в кількох роботах [105, 109, 110, 115] показано, що стиснення з втратами може мати позитивний вплив на точність класифікації зображень ДЗ. Зазвичай це відбувається за кількох умов: а) складність зображення невисока, тобто є обмежена кількість класів, вони переважно представлені квазіоднорідними об'єктами великого розміру; б) КС не надто великий; в) є деяке придушення шуму через стиснення з втратами; г) класифікатор добре навчений. Однак, навіть якщо ці умови задоволені не повністю, точність класифікації зображень може залишатися майже однаковою для досить широкого діапазону значень КС. Це має місце, якщо внесені спотворення не є надто великими за певним критерієм.

Як згадувалося вище, ефективність також залежить від використовуваного кодера. На даний момент запропоновано безліч методів стиснення з втратами. Деякі з них можна застосовувати лише компонентно, інші спеціально розроблені для тривимірного (3D) випадку, тобто для кольорових та багатоканальних зображень. Більшість передових методик базуються на ортогональних перетвореннях, а саме на дискретному косинусному перетворенні, вейвлетах та деяких інших [116, 117]. Крім того, передові методи зазвичай використовують міжканальну кореляцію даних, властиву багатоканальним зображенням [118], щоб покращити ступінь стиснення без втрати якості зображення [112, 119, 120].

У роботі [105] запропоновано підхід до стиснення трьохканальних зображень із втратами, заснований на забезпеченні заданої візуальної якості. Цей підхід передбачає покомпонентне стиснення даних кодером AGU [111], який використовує ДКП у блоках 32×32 пікселів, сучасне кодування квантованих коефіцієнтів ДКП та автоматичний деблокінг після декомпресії. Перевагами цього кодера є наступні: він перевершує JPEG і JPEG2000, досить швидкий, а його характеристики контролюється кроком квантування (quantization step - QS), для якого можна передбачити як КС [121], так і такі показники якості, як PSNR і PSNR-HVS-M. Нагадаємо, що і PSNR, і PSNR-HVS-M виражаються в дБ, і більші значення цих показників відповідають кращій якості. Крім того, для обох метрик встановлені пороги невидимості спотворення (приблизно 36 дБ і 41 дБ відповідно) і саме ці пороги відповідають практичним випадкам, коли класифікація стиснених зображень практично так само ефективна, як і класифікація первинних (нестиснених, стиснених без втрат) зображень. Беручи до уваги ці переваги, варто проаналізувати, чи 3D-версія AGU може надати їм одночасно ще одну перевагу – забезпечення більшого КС порівняно з компонентною обробкою.



Особливості класифікатора максимальної правдоподібності (МП, maximum likelihood – ML), що застосовано нами, полягають у наступному. Його навчання проведено для нестиснених зображень, припускаючи, що такі дані є. Особливістю є те, що ознаки явно мають негауссовий розподіл [105]. Для опису класу спектральних ознак ми застосували апроксимацію гістограми за допомогою розподілу Джонсона (Johnson SB [122]), яка має чотири параметри, що налаштовуються, що дозволяє виконувати апроксимацію досить точно. Передбачалося, що в даних ДЗ будуть присутні чотири класи, а саме: Вода, Рослинність, Грунт без рослинності та Місто (Urban). Навчальні зразки мають розмір кількох тисяч пікселів. Перевірочні зразки мають у кілька разів більші розміри. Приклади зображень для стиснення та класифікації наведені на рисунку 24. Це триканальні мультиспектральні дані Sentinel-2 у кольоровому представленні, доступні на https://apps.sentinel-hub.com/eo-browser/?lat=46.45&lng=34.12&zoom=6&time=2019-11-03&preset=3_NDVI&datasource=Sentinel-2%20L1C). Це зображення м. Харкова (Україна) та сільської місцевості поблизу нього, зібрані наприкінці серпня 2019 року. Обидва фрагменти мають розмір 512×512 пікселів. Приклади відібрані для представлення даних ДЗ з простою структурою а) і складною структурою б).



а) сільська місцевість; б) місто.

Рисунок 24 – Фрагменти розглянутих триканальних зображень

Для обох версій кодеру параметром, який контролює стиснення, є QS. Якщо зображення компонентів представлені у вигляді 8-бітових масивів даних, QS зазвичай становить близько десятків. Отже, проаналізуємо звичайні характеристики для компонентного та тривимірного стиснення, тобто залежність



КС та PSNR-HVS-M від QS. Оскільки ефективність залежить від складності зображення, дані для обох фрагментів зображення на рисунку 24 наведені відповідно в таблицях 4 і 5.

Аналіз даних з таблиць 4 і 5 дозволяє зробити наступні висновки:

а) КС дійсно залежить не тільки від QS, але і від складності зображення; для фіксованого QS КС приблизно в два рази більше для зображення простішої структури (табл. 4), ніж для зображення з текстурою (табл. 5);

б) КС і PSNR-HVS-M досить близькі для даного QS для всіх трьох компонентів для випадку покомпонентного стиснення триканальних зображень;

в) КС для 3D стиснення для даного зображення та заданого QS майже вдвічі більша, ніж для покомпонентного стиснення (де КС приблизно дорівнює КС, усередненому для трьох компонентів);

г) PSNR-HVS-M для даного QS для 3D та компонентного стиснення також майже однакові; це означає, що методологія налаштування QS для забезпечення бажаної якості, яка раніше розроблена для покомпонентного (у відтінках сірого) стиснення зображення [105], діє і для 3D-випадку; зокрема, для забезпечення стиснення з втратами на порозі невидимості спотворень (близько 41 дБ), необхідно встановити QS близько 17.

Таблиця 4 – Характеристики 3D і покомпонентного стиснення для фрагмента зображення на рисунку 24,а)

QS	3D стиснення		Червона компонента		Зелена компонента		Синя компонента	
	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M
5	6,49	51,0	5,3	51,8	4,1	52,8	3,4	53,5
10	12,34	44,5	8,48	45,9	6,93	45,6	5,82	45,5
15	18,51	41,1	11,37	42,7	9,53	41,8	8,37	41,2
20	25,33	38,9	14,15	40,4	12,01	39,4	10,84	38,4
25	32,36	37,6	16,92	38,5	14,41	37,4	13,39	36,4
30	39,95	35,9	19,54	36,9	16,88	35,9	16,03	34,8
35	47,94	34,8	22,14	35,7	19,35	34,5	18,59	33,5

Таким чином, для фіксованого QS візуально немає істотної різниці між зображеннями, стиснутими в 3D і компонентним способом. Для підтвердження цього на рис. 25 показано стиснуті зображення для QS=20 (для зображення на рис. 24,б)). Можна помітити деякі незначні відмінності порівняно із зображенням на рис. 24,б), але загалом візуальна якість зображень на рис. 25,а) та 25,б)



Таблиця 5 – Характеристики 3D і покомпонентного стиснення для фрагмента зображення на рисунку 24,б)

QS	3D стиснення		Червона компонента		Зелена компонента		Синя компонента	
	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M
5	3,41	51,9	2,30	53,9	2,14	55,5	2,21	53,9
10	5,65	44,9	3,14	47,3	2,89	47,2	3,02	46,7
15	8,43	41,1	3,93	43,5	3,62	42,7	3,78	32,8
20	11,62	38,6	4,71	40,5	4,39	39,5	4,52	40,0
25	14,41	36,8	5,55	38,2	5,22	37,2	5,30	37,8
30	17,33	35,4	6,42	36,4	6,10	35,3	6,13	36,0
35	20,39	34,3	7,39	34,7	7,07	33,8	7,02	34,5

однакового порядку.

Між тим, може бути цікаво дізнатися, чи впливають ці відмінності на точність триканальної класифікації МП-класифікатором.

Було б цікаво розглянути результати класифікації як для окремих класів, так і загалом. У таблицях 6 і 7 наведено дані для зображення на рисунку 24,а) для компонентного та 3D стиснення відповідно. Аббревіатура 1a_45 означає, що зображення було стиснуто з наданням PSNR-HVS-M=45 дБ. Залежності для класів різні. Клас Місто (Urban) розпізнається трохи краще для стиснутих зображень. Виявлення води стає дещо гіршим, якщо КС збільшується, а якість зображення знижується. На стиснутих зображеннях значно краще розпізнається клас Рослинність. Клас Грунт розпізнається з приблизно такою ж імовірністю правильної класифікації для компонентного стиснення, але він краще розпізнається на стиснутих зображеннях у випадку тривимірного стиснення.

Наведемо дані для зображення на рисунку 24,б). Вони надані у таблицях 8 та 9. Тут аббревіатура 1b_42, означає, що зображення на рисунку 24,б) було стиснуто з наданням PSNR-HVS-M=42 дБ. Залежності для класів і методів стиснення різні. Клас Місто (Urban) розпізнається трохи краще для зображень, стиснутих по компонентах (табл. 8), але гірше для 3D стиснутих зображень (табл. 9).

Класифікація води стає гіршою, якщо КС збільшується, а якість зображення знижується для обох методів стиснення. Клас Рослинність краще розпізнається на зображеннях, стиснутих по компонентах. Якщо застосовано тривимірне стиснення, то точність класифікації залишається приблизно на тому ж рівні.



а)

б)

а) покомпонентне стиснення; б) тривимірне стиснення.

Рисунок 25 – Результати стиснення з однаковим $QS=20$

Таблиця 6 – Ймовірності правильної класифікації для класів та загалом для зображення на рисунку 24,а), що стиснуте поканалньо

Класи	Первинне	1a_45	1a_42	1a_39	1a_36	1a_33	1a_30
Місто	0,883	0,89	0,89	0,891	0,892	0,894	0,893
Вода	0,997	0,996	0,996	0,995	0,994	0,992	0,990
Рослинність	0,911	0,955	0,961	0,965	0,972	0,978	0,984
Грунт	0,696	0,715	0,716	0,717	0,705	0,698	0,675
Загалом (P_{total})	0,872	0,889	0,891	0,892	0,891	0,891	0,885

Таблиця 7 – Ймовірності правильної класифікації для класів та загалом для зображення на рис. 24,а), що стиснуте тривимірним кодером

Класи	Первинне	1a_45	1a_42	1a_39	1a_36	1a_33	1a_30
Місто	0,883	0,894	0,896	0,90	0,908	0,919	0,938
Вода	0,997	0,996	0,996	0,997	0,997	0,997	0,994
Рослинність	0,911	0,965	0,970	0,978	0,981	0,989	0,994
Грунт	0,696	0,769	0,793	0,821	0,839	0,87	0,866
Загалом (P_{total})	0,872	0,906	0,914	0,924	0,931	0,944	0,948

Клас Грунт краще розпізнається на зображеннях, стиснутих у тривимірному режимі. Загалом ймовірність правильної класифікації для цього класу майже така ж, як і для первинних даних.



Таблиця 8 – Ймовірності правильної класифікації для класів та загалом для зображення на рис. 24,б), що стиснуте поканалньо

Класи	Первинне	1a 45	1a 42	1a 39	1a 36	1a 33	1a 30
Місто	0,423	0,437	0,448	0,463	0,486	0,499	0,509
Вода	0,863	0,81	0,799	0,793	0,783	0,774	0,764
Рослинність	0,97	0,99	0,992	0,992	0,993	0,994	0,994
Грунт	0,775	0,754	0,75	0,749	0,751	0,753	0,751
Загалом (P_{total})	0,758	0,748	0,747	0,749	0,753	0,755	0,754

Таким чином, ми розглянули та порівняли два підходи, компонентний та 3D, до стиснення із втратами триканальних зображень дистанційного зондування. Було ретельно досліджено два зображення, одне з яких має досить просту структуру, а інше – досить складну структуру.

Таблиця 9 – Ймовірності правильної класифікації для класів та загалом для зображення на рис. 24,б), що стиснуте тривимірним кодером

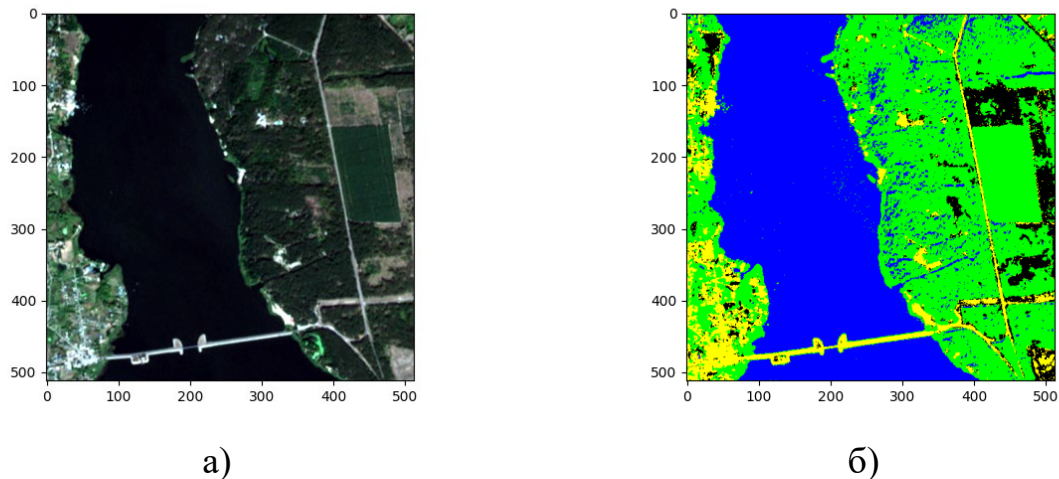
Класи	Первинне	1a 45	1a 42	1a 39	1a 36	1a 33	1a 30
Місто	0,423	0,431	0,418	0,406	0,396	0,379	0,367
Вода	0,863	0,816	0,798	0,793	0,771	0,744	0,749
Рослинність	0,97	0,931	0,943	0,96	0,97	0,974	0,97
Грунт	0,775	0,764	0,765	0,773	0,766	0,766	0,773
Загалом (P_{total})	0,758	0,736	0,731	0,733	0,726	0,716	0,715

Загальними висновками є наступні. По-перше, тривимірне стиснення дозволяє отримати приблизно вдвічі більший КС для зображень тієї ж якості. По-друге, для зображень, стиснутих без візуально помітних спотворень, точність класифікації приблизно така ж, як і для первинних даних. Ймовірність правильної класифікації може бути трохи кращою для стиснутих даних, ніж для первинних даних, якщо зображення досить просте і навпаки. По-третє, схоже, що тривимірне стиснення забезпечує кращі передумови для класифікації зображень простіших структур і гірші передумови для більш складних зображень структури. Однак цей висновок (припущення) слід перевірити на основі отримання та обробки додаткових зображень.

Трохи інше дослідження було проведено для стиснення зображень на основі атомарних функцій [123, 124] та з використанням двох інших триканальних зображень мультиспектральних даних ДЗ Сентінель, для класифікації яких використовувалась нейромережа. Зображення та результати їх класифікації для

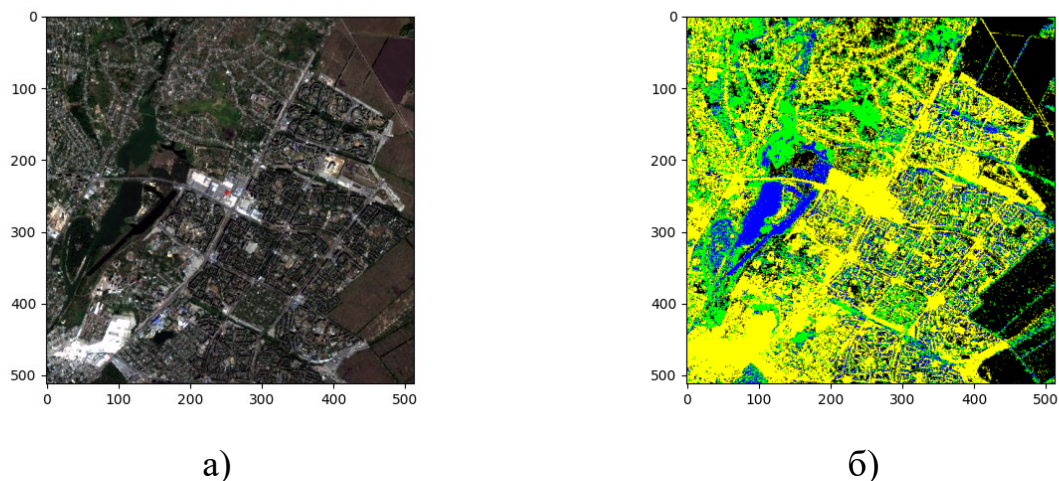


стиснених даних показані на рисунках 26 та 27. Синім кольором показаний клас Вода, жовтим – Місто, зеленим – Рослинність, чорним – Грунт. Кількісні результати зібрані в таблицях 10 та 11. Тут необхідно пояснити, що для метолів стиснення на основі атомарних перетворень рівень внесених втрат контролюється максимальним модульним відхиленням (maximal absolute deviation - MAD). Тому в таблицях наведені дані для широкого діапазону зміни цього параметру.



а) триканальне зображення; б) результат класифікації для стисненого зображення.

Рисунок 26 – Результат обробки зображення з простою структурою



а) триканальне зображення;
б) результат класифікації для стисненого зображення.

Рисунок 27 – Результат обробки зображення зі складною структурою

Як витікає з аналізу даних у таблицях 10 та 11, стиснення з втратами або практично не призводить до погіршення ймовірності правильної класифікації (як



це має місце для більш простого зображення, табл. 10), або призводить до невеликого погіршення (як це має місце для зображення більш складної структури, табл. 11). Умовою цього є стиснення з візуально непомітними втратами (саме таким у випадку, що розглядається, є стиснення з MAD, що не перевищує 40). Таким чином є можливість контрольованого стиснення з втратами, але без помітного погіршення ймовірності правильної класифікації.

**Таблиця 10 – Ймовірності правильної класифікації
для зображення на рисунку 26,а)**

Клас	Первинне	MAD=4	MAD=9	MAD=13	MAD=17	MAD=26	MAD=32
Місто	0,79	0,81	0,82	0,83	0,83	0,74	0,82
Вода	0,99	0,99	0,98	0,98	0,99	0,98	0,98
Рослинність	0,93	0,93	0,92	0,92	0,92	0,92	0,91
Ґрунт	0,76	0,79	0,78	0,79	0,78	0,75	0,75
Загалом	0,95	0,9489	0,9453	0,9479	0,9480	0,9388	0,9425

**Таблиця 11 – Ймовірності правильної класифікації
для зображення на рисунку 27,а)**

Клас	Первинне	MAD=5	MAD=9	MAD=13	MAD=17	MAD=30	MAD=39
Місто	0,93	0,93	0,92	0,92	0,91	0,91	0,91
Вода	0,71	0,73	0,69	0,71	0,74	0,69	0,66
Рослинність	0,64	0,65	0,63	0,63	0,64	0,62	0,60
Ґрунт	0,90	0,90	0,90	0,89	0,88	0,90	0,89
Загалом	0,8551	0,8618	0,8497	0,8488	0,8459	0,8463	0,8365



Висновки

Під час виконання НДР головним чином аналізувались можливості розробки методів обробки даних багатоканального дистанційного зондування на основі нейронних мереж. На основі проведеного аналізу та розробок можна зробити наступні висновки:

а) НМ можна використовувати під час пошуку подібних блоків у зображеннях, які необхідно представити в єдиній системі координат; застосування НМ у МП дозволяє підвищити її ефективність (AUC зростає на кілька сотих), один з таких варіантів запропоновано в наших роботах [31, 124];

б) нейромережі можна використовувати для прогнозування ефективності фільтрації, в тому числі РСА-зображень, для яких низька якість зумовлена спекл-шумом із досить складними статистичними та спектральними характеристиками;

в) нейромережі дають задовільні результати при класифікації триканальних зображень, що стиснені різними методами; при цьому достатньо, щоб внесені втрати не були занадто великими (візуально помітними); це можна досягти досить просто; при цьому доцільно використовувати тривимірні кодери [19-21].