



# KAPITEL 1 / CHAPTER 1<sup>1</sup>

## SOME ASPECTS OF PROCESSING OF MULTICHANNEL REMOTE SENSING DATA USING NEURAL NETWORKS

DOI: 10.30890/2709-2313.2023-19-02-001

### Вступ

Дистанційне зондування (ДЗ) у даний час переважним чином здійснюється в багатоканальному режимі, тобто за допомогою різноманітних мульти- та гіперспектральних систем, багатопоглядових та багатополяризаційних радіолокаційних комплексів авіаційного та космічного базування. Засоби ДЗ широко використовуються для екологічного моніторингу та картографування земельних покривів, для вивчення стану сільськогосподарських полів та лісів, у гляціології та метеорології тощо [1-4]. З одного боку, дані багатоканального ДЗ потенційно здатні забезпечити корисні дані для територій великих розмірів і, до того ж, їх часто отримують з певною періодичністю [3]. З іншого боку, видобування корисної інформації з даних багатоканального ДЗ пов'язане з необхідністю вирішувати широке коло задач, які насамперед зумовлені великим об'ємом даних, наявністю завод та спотворень, енергетичними та обчислювальними обмеженнями, особливо на борту супутників. Тож доводиться виконувати обробку даних, яка може включати широкий спектр операцій та бути розподіленою між бортовими та наземними засобами [4-10].

Для сучасних засобів ДЗ характерною рисою є переважне використання цифрових алгоритмів обробки інформаційних даних. Одним із сучасних засобів є нейронні мережі (НМ) різної складності та структури. Раніше нами було показано, яким чином НМ можна застосовувати на етапі пошуку подібних блоків та суміщення зображень [11], оцінки статистичних характеристик завод та якості первинних (отриманих) зображень [12-15], фільтрації та прогнозування її ефективності [16, 17], стиснення з втратами [18, 19], розробці метрик якості з еталоном для цих застосувань [20, 21], класифікації зображень дистанційного зондування [22, 23]. В цьому розділі головну увагу приділено розробці метрик якості без еталону, розвитку методів прогнозування характеристик стиснення (в тому числі для зображень, що спотворені інтенсивними заводами) та розпізнавання. Було також продовжено роботи, що стосуються класифікації

<sup>1</sup>Authors: Lukin V.V., Ieremeiev O.I., Vasilyeva I.K., Makarichev V.O., Rubel O.S., Li F., Chernova G.A., Kovalenko B.V., Benoit V.



багатоканальних зображень для різних ступенів їх стиснення. На всіх цих етапах нами було використано або можуть бути застосовані нейронні мережі.

Перспективність їх застосування, яку підтверджено результатами на тестових та реальних даних різних багатоканальних сенсорів, зумовлене певними перевагами нейронних мереж у порівнянні з традиційними підходами. Насамперед, треба відмітити, що НМ здатні навчатись. Для цього потрібні дані досить великого об'єму та певний час і обчислювальні ресурси. Але на даний час все це не є проблемою. По-перше, існують різні бази зображень [24-26] і їх кількість невпинно збільшується. Доступними є дані від багатьох багатоканальних систем ДЗ, як-то Сентінель, яка формує зображення однієї й тієї ж ділянки поверхні кожні 6 днів. Навчання НМ зазвичай виконується офлайн, тож тривалість навчання потім ніяким чином не впливає на швидкість роботи вже навченої НМ. Обчислювальні ресурси можуть бути обмежені на борту легких супутників, але на Землі вони зазвичай не є проблемою у зв'язку з існуванням кластерів та засобів хмарних обчислень.

Тож метою спільного українсько-французького проекту, що виконувався в 2022 році, була розробка методів та алгоритмів ефективної обробки зображень на основі машинного навчання для багатоканальних даних дистанційного зондування. Дослідження знаходиться на перетині таких галузей, як «Інформаційні та комунікаційні технології» та «Раціональне природокористування». Отримані результати відображені в спільних роботах [27-31], що опубліковано в 2022 році.

## **1.1. Аналіз якості первинних зображень ДЗ**

У цьому підрозділі аналізується можливість оцінки якості первинних (отриманих безпосередньо сенсором) зображень із застосуванням метрик якості з еталоном та без еталону. Актуальність такої задачі зумовлена двома аспектами. По-перше, отримані зображення не є ідеальними і вони можуть бути спотворені внаслідок кількох факторів. По-друге, еталону, тобто зображення без спотворень, може не бути, у такому випадку використати метрики з еталоном неможливо. По-третє, бажано розуміти (приймати рішення), чи є необхідним виконувати якусь обробку (фільтрацію, реставрацію) отриманих зображень в цілому або їх компонент (наприклад, зображень у деяких каналах, що спотворені



інтенсивними завадами).

Тож спочатку розглянемо деякі результати, що стосуються метрик якості з еталоном. Зазначимо, що спотворення первинних даних ДЗ можуть бути неминучими через принцип роботи системи формування зображень (прикладом є спекл на багатоканальних радіолокаційних зображеннях [32, 33]) або необхідності трансформувати дані ДЗ тим чи іншим способом перед їх візуалізацією та кінцевою обробкою (приклади - стиснення зображень до їх передачі від супутників до наземних центрів збору, обробки та розповсюдження даних ДЗ [10] або інтерполяція зображень [34]). Зображення ДЗ часто візуалізуються та аналізуються експертами [35, 36]. Тому візуальна якість зображення є важливою, тож слід використовувати адекватні метрики, що характеризують візуальну якість зображень ДЗ.

На даний час запропоновано багато метрик візуальної якості (visual quality metric – VQM) зображень ДЗ, які, зокрема, проаналізовані у джерелах [20, 37-40] і відносяться як до класу метрик з еталоном, так і без еталону. Між тим, відомо, що немає загальноприйнятої універсальної метрики. Це пояснюється багатьма причинами. По-перше, є численні типи зображень ДЗ, включаючи одноканальні (схожі на зображення у градаціях сірого), кольорові або триканальні зображення та мульти- та гіперспектральні зображення, для яких застосовуються спеціальні способи візуалізації даних (наприклад, у псевдокольорах). Цілком ймовірно, що різні VQM є оптимальними або квазіоптимальними залежно від застосувань. По-друге, залежно від типу зображень ДЗ домінуючими можуть бути різні типи спотворень. Наприклад, спекл, який є специфічним типом мультиплікативного шумоподібного феномену, є основним фактором, що погіршує якість радіолокаційних зображень із синтезованою апертурою (synthetic aperture radar - SAR). Просторово корельований шум може з'явитися внаслідок геометричних перетворень, які використовуються для реєстрації зображень в єдиній системі просторових координат. По-третє, немає встановленого жорсткого зв'язку між якістю зображень ДЗ та ефективністю вирішення подальших завдань обробки даних, таких як класифікація, виявлення об'єктів тощо. По-четверте, немає баз даних спотворених зображень, яку можна було б використовувати для аналізу та розробки метрик.

Якщо немає загальноприйнятої метрики візуальної якості, то варто проаналізувати застосовність існуючих метрик та розробити нові метрики [20]. Такий підхід було реалізовано раніше в нашій роботі [20]. Було використано базу



даних TID2013 [25] – з цієї бази даних беруться кольорові зображення з типами спотворень, характерними для даних дистанційного зондування. Хоча такий підхід має багато обмежень, він дозволив знайти найкращі метрики відповідно до їх рангової кореляції із середніми оцінками, отриманими людьми – учасниками експериментів (mean opinion score - MOS). Зверніть увагу, що жодна з елементарних метрик не дає значення рангової кореляції Спірмена (Spearman rank order correlation coefficient – SROCC) більше 0,9 для всіх типів спотворень, присутніх у базі TID2013. Тому один із способів полягає у використанні двох або більше елементарних метрик в аналізі якості зображень. Однак у разі спільного використання варто використовувати такі метрики, кореляція яких не дуже близька до одиниці, тобто метрики, що відрізняються одна від іншої своїми властивостями. Раніше в [20] було розроблено комбіновану метрику на основі тренуваної НМ. У такій НМ, щоб забезпечити простоту оцінки якості, доцільно застосувати обмежену кількість вхідних елементарних метрик. Вони повинні бути досить ефективними і, одночасно, взаємодоповнювати одна одну, тобто їх кореляція не повинна бути занадто високою. Таким чином, необхідно мати апріорну інформацію про властивості крос-кореляції елементарних метрик. Проведемо відповідний аналіз властивостей VQM для зображень із спотвореннями, характерними для даних ДЗ.

Нагадаємо основні результати аналізу, проведеного в [20]. База даних TID2013 містить 25 тестових кольорових зображень з 24 типами та 5 рівнями спотворень. Нас цікавили дві підмножини типів спотворень. Перша називається «Noise» і включає зображення з такими типами спотворень: адитивний гаусівський шум (№1), адитивний шум у кольорових компонентах (№2), просторово корельований шум (№3), замаскований шум (№4), високочастотний шум (№5), імпульсний шум (№6), шум квантування (№7), розмиття за Гаусом (№8), фільтрація шуму у зображенні (№9), мультиплікативний шум з гаусовим розподілом (№19), стиснення із втратами зображень з шумом (№21). Інший піднабір, який називається «Actual», включає зображення із зазначеними вище спотвореннями № 1, 3, 4, 5, 6, 8, 9, а також №10 (стиснення JPEG), №11 (стиснення JPEG2000), №19 і №21. Крім того, в [20] також було запропоновано розглядати спільно зображення для обох підмножин.

Значення SROCC для всіх типів спотворень і розглянутих підмножин представлені в таблиці 1. Дані наведено для 22 елементарних метрик, які дають  $SROCC > 0,9$  для підмножин “Noise” і “Actual” (загалом ми розглянули 50



елементарних метрик, посилання можна знайти в [41]).

Аналіз даних, наведених у таблиці 1, свідчить про наступне.

По-перше, кілька елементарних метрик забезпечують приблизно однакову ефективність для характеристики візуальної якості триканальних зображень з типами спотворень, які мають місце в зображеннях ДЗ.

По-друге, ці елементарні метрики належать до різних груп (сімейств), таких як метрики на основі PSNR (наприклад PSNR-NA і PSNR-HVS-M) і метрики на основі SSIM (наприклад ADD SIM і FSIM). Між тим, існують і ефективні елементарні метрики, які не належать до цих сімейств. Це означає, що, з одного боку, існує багато ефективних елементарних метрик (хоча жодна з них не є досконалою), а з іншого боку, вони базуються на різних принципах і, таким чином, можуть доповнювати одна одну. Фактично ця здатність була доведена результатами розробки комбінованих метрик в [20]. Було показано, що навчена нейронна мережа, яка використовує близько 20 елементарних метрик як вхідні дані, може забезпечити SROCC для підмножини «Noise&Actual» приблизно 0,965, тобто значно вище, ніж SROCC для найкращої елементарної метрики.

Інша проблема з елементарними метриками та характеристиками якості зображення полягає в тому, що діапазони варіації значень метрик є різними. Деякі метрики виражаються в дБ і змінюються в широких межах, інші змінюються в межах від 0 до 1. Більшість елементарних показників зростають, якщо рівень спотворень зменшується, хоча бувають й інші випадки. Лінеаризація MOS часто використовується для порівняння ефективності метрики [20]. Приклад такої лінеаризації наведено на рисунку 1. Діаграма розсіювання для кожної елементарної метрики показує пари значення метрики (після лінеаризації) та MOS для кожного розглянутого тестового зображення, типу спотворення та рівня.

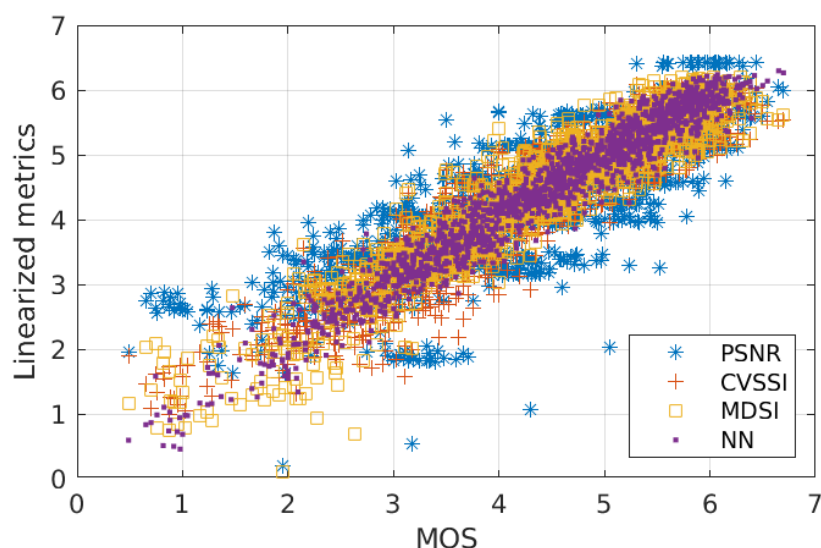
В ідеалі це має бути чисто лінійне відношення  $\text{Linearized metric} = \text{MOS}$  або, принаймні, набір точок, розташованих близько та компактно до цієї лінії. Однак, як видно, багато точок для метрики пікове відношення сигнал-шум (peak signal-to-noise ratio – PSNR) розташовані далеко від цієї лінії.

Деякі очевидні викиди спостерігаються навіть для найкращих елементарних показників візуальної якості, як-то MDSI та CVSSI. Комбінована метрика на основі нейронної мережі (neural network - NN) забезпечує більшу компактність точок діаграми розсіювання. Таким чином, попередній аналіз чітко вказав на необхідність кореляційного аналізу елементарних метрик.



**Таблиця 1 – Значення SROCC для 22 найкращих метрик (відповідно до результатів для підмножини (піднабору) “Noise&Actual”)**

Метрика	Всі типи спотворень	Noise	Actual	Noise & Actual
MDSI	0,8897	0,9275	0,9387	0,9374
CVSSI	0,8090	0,9248	0,9350	0,9341
MCSD	0,8045	0,9224	0,9326	0,9323
PSNRHA	0,8198	0,9230	0,9388	0,9322
GMSD	0,8004	0,9211	0,9314	0,9318
PSNRHMAm	0,8541	0,9221	0,9387	0,9315
PSIM	0,8926	0,9189	0,9309	0,9303
PSNRHAy	0,7794	0,9184	0,9272	0,9275
PSNRHVS	0,6536	0,9172	0,9257	0,9263
PSNRHMA	0,8137	0,9151	0,9343	0,9250
IGM	0,8023	0,9099	0,9220	0,9227
PSNRHMAy	0,7570	0,9107	0,9209	0,9226
VSI	0,8967	0,9101	0,9258	0,9218
SR-SIM	0,8076	0,9070	0,9211	0,9206
HaarPSI	0,8730	0,9063	0,9168	0,9190
ADM	0,7861	0,9113	0,9201	0,9189
PSNRHVSM	0,6246	0,9061	0,9175	0,9188
FSIMc	0,8510	0,9022	0,9150	0,9164
ADD GSIM	0,8310	0,9023	0,9151	0,9159
IQM2	0,7955	0,8995	0,9103	0,9122
ADD SSIM	0,8023	0,9008	0,9119	0,9120
FSIM	0,8011	0,8969	0,9108	0,9117

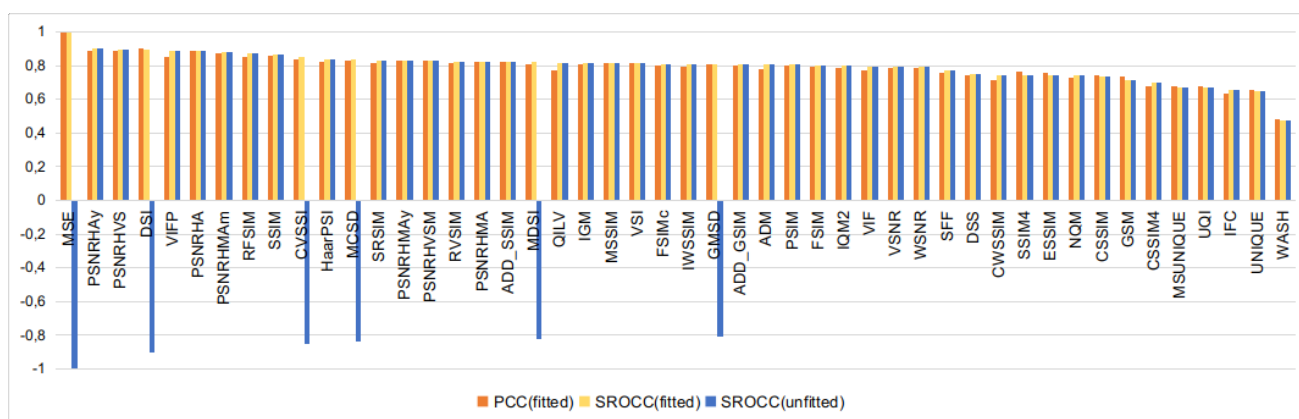


**Рисунок 1 – Діаграма розсіювання (скатерограма) лінеаризованих метрик від MOS для трьох елементарних метрик (PSNR, CVSSI та MDSI) та комбінованої метрики із застосуванням НМ**

Розпочнемо наш аналіз із розгляду кореляції між традиційною метрикою



PSNR та іншими елементарними метриками. Враховуючи нелінійність усіх показників, значення SROCC були визначені для мінімізації його впливу. Вони представлені на рисунку 2. Аналіз даних показує наступне. По-перше, є метрики візуальної якості, які сильно корелюють із PSNR (значення SROCC перевищують 0,8). В основному це метрики на основі PSNR (такі як PSNR-HVS, PSNR-HVS-M, WSNR та деякі інші). Більшість інших метрик (зокрема, більшість з тих, що базуються на SSIM) мають рангову кореляцію з PSNR близько 0,7. Це також стосується трьох основних метрик у таблиці 1, хоча SROCC для них становить близько -0,7 (це тому, що всі три показники знижуються, якщо якість покращується). У той же час існують інші елементарні метрики, такі як WASH (що має SROCC близько 0,3). Зверніть увагу, що ця метрика має дуже низьку ефективність для підмножини Noise&Actual.



**Рисунок 2 – Значення рангової кореляції SROCC між PSNR та іншими елементарними метриками**

Розглянемо тепер кореляційні властивості для елементарних метрик, що належать до одного сімейства. У таблиці 2 представлені дані для набору метрик на основі PSNR. Дані в таблиці 2 (а також далі в таблицях 3 і 4) розміщені в порядку спадання значень SROCC для елементарної метрики з MOS.

Аналіз результатів в таблиці 2 показує наступне. По-перше, для всіх метрик на основі PSNR їх кореляція з PSNR є високою – найменший SROCC дорівнює 0,79 для метрик WSNR і VSNR. По-друге, значення SROCC також значно вищі для всіх пар метрик візуальної якості – усі інші пари метрик показують кращі результати, ніж згадані. По-третє, існують пари метрик, для яких SROCC перевищує 0,9: WSNR і PSNR-HVS, WSNR і PSNR-HVS-M, PSNR-HVS і



**Таблиця 2 – Значення SROCC для елементарних метрик з MOS  
та між собою**

№		MOS	PSNRHA	PSNR HMAm	PSNRHVS	PSNR HMA	PSNR HVSM	WSNR	VSNR	PSNR	MSE
1	PSNRHA	0,93	1	1	0,97	0,99	0,96	0,93	0,91	0,89	-0,89
2	PSNRHMAm	0,93	1	1	0,97	0,99	0,96	0,93	0,91	0,88	-0,88
3	PSNRHVS	0,93	0,97	0,97	1	0,96	0,98	0,95	0,93	0,90	-0,90
4	PSNRHMA	0,93	0,99	0,99	0,96	1	0,97	0,94	0,91	0,83	-0,83
5	PSNRHVSM	0,92	0,96	0,96	0,98	0,97	1	0,97	0,93	0,83	-0,83
6	WSNR	0,90	0,93	0,93	0,95	0,94	0,97	1	0,90	0,79	-0,79
7	VSNR	0,88	0,91	0,91	0,93	0,91	0,93	0,90	1	0,79	-0,79
8	PSNR	0,83	0,89	0,88	0,90	0,83	0,83	0,79	0,79	1	-1
9	MSE	-0,83	-0,89	-0,88	-0,90	-0,83	-0,83	-0,79	-0,79	-1	1

PSNR-HVS-M, PSNR-NA і PSNR-HMA тощо. Це означає, що можна аналізувати такі метрики спільно, але немає особливих очікувань отримати користь від їх спільного використання. Якщо ми маємо справу з характеристикою якості для кольорових або триканальних зображень, здається розумним використовувати PSNR-NA, яка має високі значення SROCC для розглянутих підмножин спотворень і великі значення SROCC для інших метрик на основі PSNR.

У таблиці 3 наведено дані для сімейства метрик на основі SSIM. Аналіз показує, що, по-перше, єдиною метрикою, для якої SROCC із SSIM є меншим, ніж 0,7, є CWSSIM (0,67). По-друге, єдиною метрикою, для якої SROCC із SSIM перевищує 0,9, є MSSIM. Інші метрики на основі SSIM мають високу рангову кореляцію з SSIM – переважно в межах 0,8–0,9.

Ще один висновок полягає в тому, що всі метрики на основі SSIM сильно корельовані. Зокрема, для метрики FSIMc (кольоровий варіант FSIM) усі значення SROCC перевищують 0,83, більшість значень SROCC перевищує 0,9, а SROCC для FSIMc та FSIM практично дорівнює одиниці (через це дані для FSIM не наведено в таблиці 3). Подібні властивості мають місце для метрики PSIM. Пам'ятаючи про високі значення SROCC для цих показників і MOS (див. дані в таблиці 1), доцільно використовувати метрики FSIMc або PSIM з розглянутого сімейства.

Таким чином, ми залишили метрики PSNR-NA, FSIMc і PSIM для подальшого аналізу та додали хороші метрики з таблиці 1 (ті, які мають SROCC з MOS більше 0,9). Отримані значення SROCC наведені в таблиці 4. Під час аналізу ми ігноруємо знаки SROCC, оскільки вони вказують лише на збільшення або зменшення метрики, якщо якість зображення покращується. Оскільки всі метрики, представлені в таблиці 4, є «хорошими», вони мають високі значення





SROCC між собою. Наприклад, абсолютне значення SROCC для MDSI та інших метрик не менше, ніж 0,963. Значення SROCC для PSNR-НА та інших показників перевищують 0,94.

**Таблиця 3 – Значення SROCC для метрик на основі SSIM між собою**

№	Метрика	MOS	PSIM	SRSIM	FSIMc	ADD SSIM	IW SSIM	CS SIM	MS SIM	RF SIM	ES SIM	RV SIM	CW SSIM	SSIM
1	PSIM	0,93	1	0,98	0,98	0,96	0,95	0,97	0,94	0,93	0,94	0,91	0,84	0,80
2	SRSIM	0,92	0,98	1	0,99	0,97	0,97	0,97	0,96	0,95	0,93	0,95	0,84	0,85
3	FSIMc	0,92	0,98	0,99	1	0,97	0,98	0,98	0,97	0,93	0,94	0,95	0,83	0,85
4	ADDSSIM	0,91	0,96	0,97	0,97	1	0,98	0,96	0,97	0,91	0,91	0,91	0,81	0,87
5	IWSSIM	0,89	0,95	0,97	0,98	0,98	1	0,97	0,99	0,92	0,92	0,95	0,81	0,89
6	CSSIM	0,89	0,97	0,97	0,98	0,96	0,97	1	0,96	0,90	0,95	0,91	0,82	0,80
7	MSSIM	0,89	0,94	0,96	0,97	0,97	0,99	0,96	1	0,91	0,93	0,94	0,79	0,92
8	RFSIM	0,89	0,93	0,95	0,93	0,91	0,92	0,90	0,91	1	0,87	0,95	0,80	0,83
9	ESSIM	0,86	0,94	0,93	0,94	0,91	0,92	0,95	0,93	0,87	1	0,89	0,83	0,81
10	RVSIM	0,84	0,91	0,95	0,95	0,91	0,95	0,91	0,94	0,95	0,89	1	0,78	0,88
11	CWSSIM	0,81	0,84	0,84	0,83	0,81	0,81	0,82	0,79	0,80	0,83	0,78	1	0,67
12	SSIM	0,78	0,80	0,85	0,85	0,87	0,89	0,80	0,92	0,83	0,81	0,88	0,67	1

**Таблиця 4 – Значення SROCC для вибраних метрик між собою**

№	Метрика	MOS	MDSI	CVSSI	PSNRHA	GMSD	MCSD	PSIM	IGM	VSI	HaarPSI	FSIMc	ADD GSIM
1	MDSI	-0,937	1	0,976	-0,963	0,978	0,975	-0,985	-0,972	-0,988	-0,974	-0,985	-0,977
2	CVSSI	-0,934	0,976	1	-0,943	0,987	0,992	-0,974	-0,961	-0,96	-0,966	-0,972	-0,977
3	PSNRHA	0,932	-0,963	-0,943	1	-0,943	-0,945	0,959	0,951	0,966	0,965	0,952	0,941
4	GMSD	-0,932	0,978	0,987	-0,943	1	0,994	-0,983	-0,968	-0,963	-0,972	-0,975	-0,982
5	MCSD	-0,932	0,975	0,992	-0,945	0,994	1	-0,981	-0,966	-0,958	-0,973	-0,973	-0,985
6	PSIM	0,93	-0,985	-0,974	0,959	-0,983	-0,981	1	0,974	0,978	0,98	0,983	0,983
7	IGM	0,923	-0,972	-0,961	0,951	-0,968	-0,966	0,974	1	0,964	0,977	0,979	0,97
8	VSI	0,922	-0,988	-0,96	0,966	-0,963	-0,958	0,978	0,964	1	0,972	0,983	0,966
9	HaarPSI	0,919	-0,974	-0,966	0,965	-0,972	-0,973	0,98	0,977	0,972	1	0,986	0,974
10	FSIMc	0,916	-0,985	-0,972	0,952	-0,975	-0,973	0,983	0,979	0,983	0,986	1	0,983
11	ADDGSIM	0,916	-0,977	-0,977	0,941	-0,982	-0,985	0,983	0,97	0,966	0,974	0,983	1

Інше спостереження полягає в тому, що існують пари елементарних метрик, для яких SROCC наближається до одиниці. Це стосується метрик CVSSI і MCSD, GMSD і MCSD. Враховуючи це, пропонуємо залишити для подальшого аналізу та спільного використання наступні сім елементарних показників:

– MDSI, оскільки забезпечує найбільший MOS для розглянутих типів спотворень (див. дані в таблиці 1 або в таблиці 4);

– PSNR-НА, як є найкращою метрикою на основі PSNR, яка не має занадто високої кореляції з іншими метриками;

– GMSD, як представлення пари GMSD і MCSD, що мають дуже високу кореляцію між собою;



- PSIM, як є однією із найкращих метрик на основі SSIM, яка не має надто високої кореляції з іншими метриками згідно даних в таблиці 4);
- VSI, яка не має надто високої кореляції з іншими метриками (табл. 4);
- HaarPSI, яка маю високу кореляцію з MOS (табл. 1), але не надто сильно корелює з іншими метриками;
- FSIMс, яка є ще одним хорошим представником метрик на основі SSIM, але яка не надто сильно корелює з іншими метриками.

Як вже зазначалось, використання нейронних мереж є одним із можливих способів об'єднання (спільної обробки) набору елементарних метрик. Його ефективність залежить від кількох факторів: кількості входів, структури НМ, функції активації нейронів, стратегії навчання тощо. Ще одним фактором є попередня обробка елементарних метрик, зокрема, лінеаризовані вони чи ні? Результати, представлені в [20], показують можливий вплив більшості з цих факторів. Наше завдання тут полягає не в тому, щоб провести таке широке дослідження. Натомість ми хотіли б побачити результати об'єднання фіксованої кількості елементарних метрик (у нашому випадку сім) на основі проведеного аналізу їх кореляційних властивостей.

Для вивчення впливу структури НМ нами було проаналізовано шість варіантів, описаних у таблиці 5 у колонці «Шари» (наведено кількості нейронів у прихованих шарах). Конфігурація шарів нейронних мереж визначається, виходячи з обсягу доступних вхідних даних і складності завдання.

Підготовку даних було проведено згідно з результатами роботи [20]. Наприклад, підгонка (фітінг) для обраного набору елементарних метрик не використовується, оскільки завдання нормалізації даних ефективно виконується шляхом вибору сигмоїдної функції активації.

Була обрана мережа прямого зв'язку з двома або трьома прихованими шарами, як показано в таблиці 5. Щоб забезпечити максимальну точність навчання, спотворені зображення були динамічно розподілені між підмножинами навчання та верифікації у співвідношенні 70% до 30%. Для кожної конфігурації було навчено 50 мереж. Найкращі результати відповідно до отриманих для них значень SROCC наведено в таблиці 5. Кількість епох визначається динамічно кількістю максимальних помилок перевірки під час зменшення помилки зворотного поширення.

Аналіз даних у таблиці 5 показує, що є певне покращення ефективності



комбінованих показників порівняно з найкращими елементарними метриками. Немає великої різниці між НМ з різною структурою, хоча варіант №2 здається кількісно найкращим.

**Таблиця 5 – Дані для комбінованих метрик з 7 входами на основі НМ**

№	Шари та кількість нейронів в них	SROCC (навчання)	SROCC (верифікація)
1	[7, 4]	0,9496	0,9514
2	[7, 4, 1]	0,9499	0,9516
3	[2, 2]	0,9483	0,9496
4	[4, 4]	0,9507	0,9494
5	[7, 7]	0,9460	0,9559
6	[7, 7, 7]	0,9493	0,9487

Тим часом, досягнуті значення SROCC (приблизно 0,95) у будь-якому разі трохи гірші (менші), ніж досягнуті в [20] (де розглядались варіанти для більшої кількості елементарних метрик (близько 20), що використовуються як вхідні дані НМ). Таким чином, можна знайти компроміс між ефективністю і простотою НМ в залежності від пріоритету вимог на практиці.

Тож проведено аналіз метрик візуальної якості з еталоном із застосуванням до зображень із спотвореннями, характерними для даних дистанційного зондування. Показано, що існує багато метрик, які працюють досить добре. Кореляційні властивості для цих показників проаналізовані за допомогою SROCC. Показано, що SROCC є високим для багатьох можливих пар показників. Це дозволило рекомендувати найкращі метрики для сімейств на основі PSNR і SSIM. Також показано, що нейронні мережі, які використовують сім елементарних метрик як вхідні дані можуть забезпечити SROCC з MOS близько 0,95, тобто на дуже високому рівні.

Подивимось тепер, чи працездатен принцип розробки комбінованих метрик якості для метрик без еталону. Перший крок у цьому напрямку було зроблено в нашій роботі [42]. Показано, що найкраще (найбільше) значення SROCC для елементарних метрик якості без еталону для всіх типів спотворень в базі TID2013 сягає 0,435 для метрики CORNIA [43], а для підмножини типів спотворень «Noise&Actual» - 0,722 для метрики IL-NIQE [44]. Об'єднання 5 елементарних метрик за допомогою НМ дозволяє отримати SROCC з MOS близько 0,82, десяти елементарних метрик - SROCC з MOS близько 0,88, а двадцяти - SROCC з MOS близько 0,92. Зрозуміло, що комбіновані метрики вимагають послідовного чи



паралельного обчислення досить великої кількості елементарних метрик перед їх об'єднанням за допомогою НМ, але позитивний ефект від використання такого підходу є очевидним.

Втім, виникає питання – чи є підхід до розробки нових метрик візуальної якості на основі об'єднання кількох елементарних метрик нейронною мережею універсальним? Отримання відповіді на таке питання є також актуальним у зв'язку з розвитком нових технологій, зокрема, отримання зображень з дронів та безпілотних літальних апаратів (БПЛА).

Зазначимо, що в останні десятиліття сфера застосування дронів та інших безпілотних літальних апаратів швидко розширюється. Більшість із них містять камери; тому потреба в аналізі та обробці візуальних даних, зокрема зображень, значно зростає. З одного боку, БПЛА в основному використовують камери оптичного діапазону; тому застосовними є існуючі рішення цифрової обробки зображень [3, 45, 46]. З іншого боку, мобільність і автономність цих систем можуть накладати значні обмеження, і ці фактори також слід враховувати.

Ключовими проблемами застосування цифрової обробки зображень у цьому застосуванні є:

а) БПЛА потребують адаптивного інтегрованого підходу до придушення наявного шуму, розмиття внаслідок руху та інших типових спотворень, які лише частково можна компенсувати стабілізацією камери;

б) для передачі даних з літаючих дронів використовуються бездротові технології; дальність і надійність передачі даних визначають одну з ключових характеристик БПЛА – дальність польоту; у цьому сенсі ефективність обробки та стиснення даних високої роздільної здатності для передачі по каналу зв'язку з обмеженою пропускну здатністю є вирішальною;

в) дані, отримані на кінцевому пристрої, крім зберігання та більш складної постобробки, можуть бути використані в різних додатках; серед них все більшого поширення набувають завдання на основі машинного навчання, такі як виявлення, розпізнавання тощо [45, 47, 48].

Спільним для всіх трьох цих проблем є потреба в точних методах оцінки якості зображення та параметрів спотворення. Це необхідно для коректного застосування шумозпридушення і усунення інших спотворень, а також використання деяких інших методів для підвищення візуальної якості зображення. Крім того, ефективне стиснення з втратами необхідно в тому випадку, коли при стисненні видаляється лише непомітна частина інформації. У



цих областях вже досягнуто певних результатів обробки зображень, отриманих з БПЛА [49-53]. Тим не менш, необхідні методи, які можуть точно оцінити візуальну складову і визначити оптимальні параметри для подальших методів обробки зображень.

Завдання оцінки якості зображення (image quality assessment - IQA) зазвичай вирішується візуальними метриками якості. Для підвищення їх точності широко використовуються основні властивості людського сприйняття. Як і вище, будемо розглядати метрики без еталону та можливість їх застосування як інструменту для багатьох завдань, що виконуються за допомогою БПЛА [54-56]. Ідея полягає в тому, щоб обрати (обґрунтувати) підмножину типів спотворень, що є типовими для зображень, що отримуються з БПЛА, та підвищити точність оцінювання їх якості за допомогою існуючих метрик шляхом їх спільного використання. Певна успішність цього підходу підтверджена вище в даному розділі та в роботах [20, 42]. Пріоритетом розробки є забезпечення високої точності оцінки візуальної якості при збереженні задовільної обчислювальної складності.

Особливості використання дронів та БПЛА можуть призвести до значної кількості різноманітних спотворень зображення під час його отримання, обробки, стиснення та передачі по каналу зв'язку. У зв'язку з цим проектування комбінованої метрики вимагає наявності баз даних тестових зображень, які дозволяють моделювати такі ситуації. В результаті аналізу багатьох різних баз даних зображень нами, як і для попередніх досліджень, була обрана база TID2013. Відмінною особливістю цієї бази зображень є те, що вона містить 24 типи різноманітних спотворень, у тому числі такі унікальні, як бітові помилки при передачі стиснених зображень. Нагадаємо, що TID2013 містить 25 еталонних зображень, які були спотворені 24 типами спотворень на 5 рівнях інтенсивності, загалом 3000 тестових зображень.

З переліку типів спотворень, що змодельовані для кольорових зображень в базі, нами були вибрані наступні, які можуть бути присутніми для зображень, отриманих з БПЛА:

– адитивний гаусівський шум (№ 1 та 2), який є основною моделлю для представлення більшості фізичних процесів, які викликають шум; його присутність є особливо вираженою в умовах недостатнього освітлення;

– просторово корельований шум (№3) є характеристикою оптичних зображень через використання фільтра Байєра або його модифікацій на датчиках;



проявлення цього фактору значно збільшується при цифровому зумі;

– імпульсний шум (№6) може бути проявом мертвих пікселів або інших причин (помилки кодування/декодування);

– шум квантування (#7) може виникати під час отримання та трансформації зображення.

– розмиття (№8) є одним із найбільш актуальних спотворень через рух і вібрацію БПЛА;

– усунення шумів (№9) є проявом шумопридушення, вбудованого в більшість камер;

– стиснення (№ 10 та 11) є типовим кроком у ланцюжку обробки зображень для зменшення надмірності даних;

– помилки передачі (№ 12 і 13) є типовими для передачі по бездротових каналах зв'язку, особливо на великі відстані;

– зміни яскравості, контрастності та насиченості (№ 16-18) дозволяють імітувати зміни умов освітлення в різні періоди часу доби та погодні умови;

– мультиплікативний шум (№ 19) є актуальним, оскільки шум датчика здебільшого залежить від сигналу;

– комфортний шум (№ 20) дозволяє імітувати деякі артефакти обробки та стиснення зображення;

– стиснення із втратою шуму зображень (№ 21) є типовим прикладом реальної ситуації, коли стискається зображення з деяким шумом;

– хроматична аберація (№ 23) є результатом заломлення світла в оптиці камери.

Загалом наведені вище типи спотворень дозволяють врахувати переважну більшість типів шумів і спотворень, які можуть виникнути на зображеннях, отриманих БПЛА. Ці 18 типів спотворень дають разом 2250 тестових зображень із набору даних TID2013, тобто більшість зображень у цій базі. Крім того, перелік типів спотворень лише частково співпадає з підмножинами Noise та Actual, що розглядалися раніше. Тож давайте насамперед проаналізуємо ефективність існуючих метрик візуальної якості без еталону на цій підмножині зображень. Оскільки наше завдання — забезпечити високу точність оцінки, то включено максимально можливу кількість різних метрик. Значення SROCC для всієї бази даних TID2013 і вибраної підмножини наведено в таблиці 6.

Як видно з результатів у таблиці 6, найкращу кореляцію демонструє метрика



IL-NIQE, але значення SROCC для неї (0,492 для всіх та 0,529 для вибраних 18 типів спотворень) є досить низькими. Слід зазначити, що в таблиці 6 наведено абсолютні значення SROCC.

Наступний крок у розробці комбінованої метрики стосується правильного вибору елементарних метрик (наведених у таблиці 6) як вхідних даних нейронної мережі. Є кілька можливих рішень цієї проблеми, але не всі вони здійснимі або дають ефективне рішення:

1) повний перебір варіантів на практиці неможливий, оскільки навіть для 5 або 10 вхідних метрик потрібно буде розрахувати відповідно  $1,6 \times 10^8$  і  $2,7 \times 10^{16}$  комбінацій;

2) вибір найкращої метрики з високими показниками SROCC або виключення подібної метрики з високими значеннями крос-кореляції, досліджені в [27], показали недостатню ефективність у вирішенні цієї проблеми;

3) водночас у роботі [20] було апробовано підхід із використанням регуляризації LASSO та доведено його ефективність; це стандартний підхід до зменшення складності моделі в машинному навчанні; особливістю застосування регуляризації LASSO є те, що в результаті обмежень вона дозволяє визначити найменш важливі ознаки та встановити для них нульові вагові коефіцієнти; таким чином, фактично виключаються відповідні метрики.

Щоб відобразити вплив кількості елементарних метрик, що використовуються, на показники навченої нейронної мережі, ми використовуємо кілька комбінацій елементарних метрик, визначених за допомогою LASSO, у діапазоні значень від мінімальних 3-5 до всіх 44 метрик. Параметри LASSO були обрані таким чином, щоб отримати ненульові ваги для заданої кількості метрик. Всього нами розглянуто 10 варіантів. Списки метрик для комбінацій з менш, ніж 20 метриками, представлені в таблиці 7.

Незважаючи на популярність нейронних мереж, їх використання в області оцінки якості зображення має деякі обмеження. По-перше, різноманітність і розмір наборів даних обмежені, оскільки можна застосовувати лише бази зображень, що містять значення MOS. У нашому випадку використовується 2250 зображень з TID2013.

Слід зазначити, що через обмежену кількість рівнів викривлення та різноманітність еталонних зображень можна припустити, що деякі тестові зображення мають унікальні властивості, і їх розподіл у навчальній або тестовій підмножині може вплинути на точність навченої нейронної мережі.



**Таблиця 6 – Значення SROCC для метрик візуальної якості без еталону для всієї бази TID2013 та підмножини типів спотворень, що аналізується (UAV)**

№	Метрика	SROCC		№		SROCC	
		Всі	UAV			Всі	UAV
1	ILNIQE [43]	0,492	0,529	23	DIQU [74]	0,240	0,251
2	CORNIA [44]	0,435	0,521	24	SDQI [75]	0,224	0,248
3	HOSA [57]	0,471	0,515	25	DIPIQ [76]	0,140	0,209
4	CDIIVINE[58]	0,373	0,448	26	MLV [77]	0,201	0,195
5	BLIINDS2 [59]	0,395	0,425	27	FISHBB [78]	0,145	0,152
6	BRISQUE[60]	0,367	0,416	28	JNBM [79]	0,141	0,152
7	BIQI [61]	0,405	0,409	29	DESIQUE[80]	0,069	0,150
8	SSEQ [62]	0,341	0,406	30	GMLOG [81]	0,109	0,139
9	NIQE [63]	0,313	0,403	31	NIQMC [82]	0,113	0,124
10	QAC [64]	0,372	0,379	32	ARISM [83]	0,145	0,109
11	SISBLIM SM [65]	0,318	0,360	33	CPBDM [84]	0,112	0,109
12	LPSI [66]	0,395	0,357	34	LSSn [71]	0,168	0,105
13	LPCSI [67]	0,323	0,354	35	PSS [71]	0,022	0,087
14	SISBLIM SFB [65]	0,336	0,348	36	LSSs [71]	0,114	0,084
15	DIIVINE [68]	0,344	0,343	37	ARISMC [83]	0,138	0,081
16	BIBLE [69]	0,281	0,333	38	PSI [85]	0,001	0,075
17	OGIQA [70]	0,276	0,327	39	SMETRIC[86]	0,097	0,074
18	BPRI [71]	0,229	0,313	40	FISH [78]	0,052	0,041
19	TCLT [72]	0,233	0,308	41	NRPWN [87]	0,016	0,039
20	SISBLIM WFB [65]	0,293	0,301	42	NMC [88]	0,054	0,033
21	MSGF [73]	0,244	0,274	43	BLUR [89]	0,008	0,020
22	SISBLIM WM [65]	0,239	0,265	44	NJQA [90]	0,100	0,007

**Таблиця 7**

Кількість метрик	Назви метрик
4	ARISM, CORNIA, DIPIQ, ILNIQE
5	ARISM, CORNIA, DIPIQ, ILNIQE, LPCSI
7	ARISM, CORNIA, DIPIQ, ILNIQE, LPCSI, MLV, NIQMC
10	ARISM, CORNIA, DIPIQ, ILNIQE, LPCSI, MLV, MSGF, NIQE, NIQMC, PSS
16	ARISM, CDIIVINE, CORNIA, DIPIQ, GMLOG, HOSA, ILNIQE, JNBM, LPCSI, MLV, MSGF, NIQE, NIQMC, PSI, PSS, TCLT

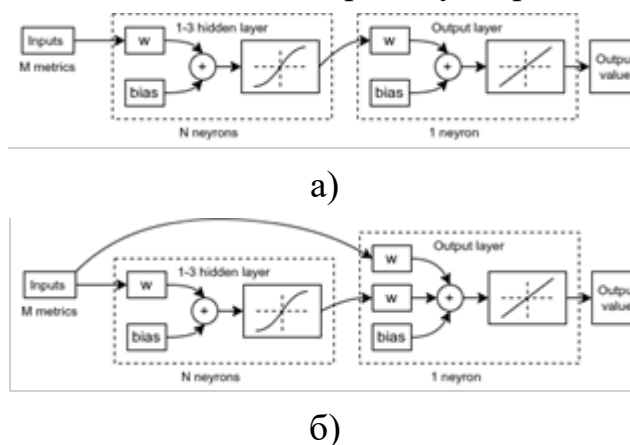
Як показують отримані результати, це важливий, але не єдиний фактор.





Тому неможливо точно вибрати, які зображення повинні бути в кожному з цих наборів. Щоб забезпечити наближення результату до оптимального, для кожної конфігурації нейронної мережі навчання слід повторювати кілька разів із випадковим розподілом зображень на навчальну та тестову підмножини. Під час проведених експериментів нами було виконано понад 100 повторень. По-друге, вибір типу нейронної мережі може істотно вплинути на кінцеву ефективність.

Нижче розглядаються два типи нейронних мереж: мережі прямого зв'язку та каскадні мережі, які мають нелінійний зв'язок між шарами, оскільки результуюче значення кожного шару, включаючи вхідний, впливає на всі наступні шари. Їх блок-схеми наведені на рисунку 3, де  $M$  – кількість елементарних метрик, а  $N$  – кількість нейронів у шарі.



а) мережа прямого розповсюдження; б) каскадна мережа.

**Рисунок 3 – Структури нейронних мереж**

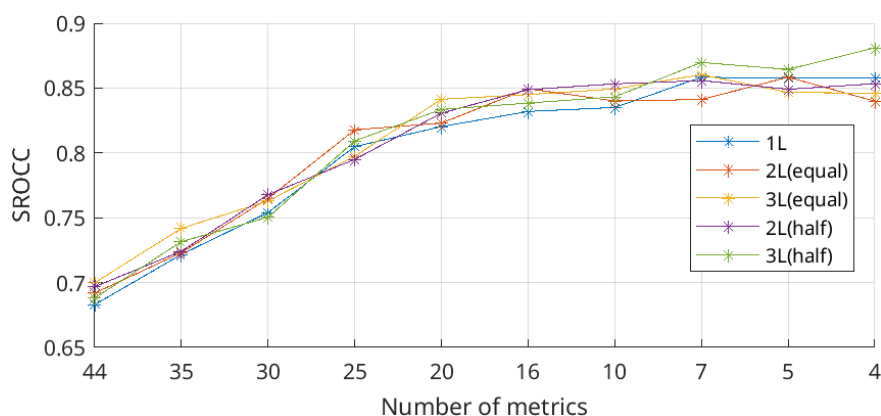
Крім того, ефективність нейронної мережі також визначається її структурою (кількістю прихованих шарів і кількістю нейронів у кожному з них). Оскільки вже вказано значну кількість факторів, що впливають на ефективність кінцевої нейронної мережі, на попередньому етапі аналізу використовується кілька основних конфігурацій. Більш точна конфігурація нейронної мережі буде визначена на завершальному етапі створення комбінованої метрики. На цьому етапі використовуються варіанти структури нейронної мережі з 1-3 прихованими шарами. Для кожного з них є два варіанти кількості нейронів у кожному шарі: 1) у всіх шарах воно дорівнює кількості вхідних метрик ( $N = M$ ), і 2) для кожного наступного шару, починаючи з другого, число нейронів ділиться на два ( $N_1 = M$ ,  $N_2 = M/2$ ,  $N_3 = M/4$ ). Всього варіантів всього 5, оскільки для одношарової мережі вони ідентичні.

В якості функції активації використовується сигмоїдна функція, яка



дозволяє, незалежно від діапазону значень використовуваної метрики, отримати після одного прихованого шару значення у фіксованому діапазоні  $[0, 1]$ . Ця процедура дозволяє фактично реалізувати вбудовану підгонку лінійної функції та нормалізацію значення. Загалом цей етап передбачає побудову 10 000 варіантів нейронних мереж (2 типи мережі  $\times$  5 конфігурацій  $\times$  100 повторень  $\times$  10 комбінацій метрик).

Проаналізуємо результати, отримані після навчання всіх цих нейронних мереж. Основна залежність полягає в тому, що точність комбінованої метрики зростає зі збільшенням кількості використовуваних елементарних метрик. Максимум досягається для всіх 44 показників. Графік залежності SROCC від кількості метрик наведено на рисунку 4 на прикладі для нейронних мереж із максимальними частотами SROCC серед повторів кожної конфігурації для нейромережі.



**Рисунок 4 – Графік залежності SROCC від кількості метрик**

На підставі цих результатів можна зробити декілька висновків. Таким чином, використання штучної НМ для поєднання метрик є ефективним рішенням для застосувань БПЛА, оскільки навіть варіант з мінімальною кількістю входів (4) значно перевищує за точністю максимальний результат серед елементарних метрик ( $SROCC = 0,53$ ). Розглянуті 5 конфігурацій структур НМ дають подібні показники. Графік на рисунку 4 дозволяє дати деякі рекомендації щодо вибору структури ШНМ в залежності від вимог і обмежень розв'язуваної задачі. Наприклад, якщо необхідно забезпечити максимальну простоту і досить високу ефективність, розумним вибором буде комбінована метрика з 5 елементарних, її результат досягає  $SROCC = 0,74$ , що значно вище, ніж для 4 метрик, а подальше підвищення точності з числом вхідних параметрів уповільнюється. Якщо точність є пріоритетом, то можуть бути корисними варіанти з 10 або 16 входами.

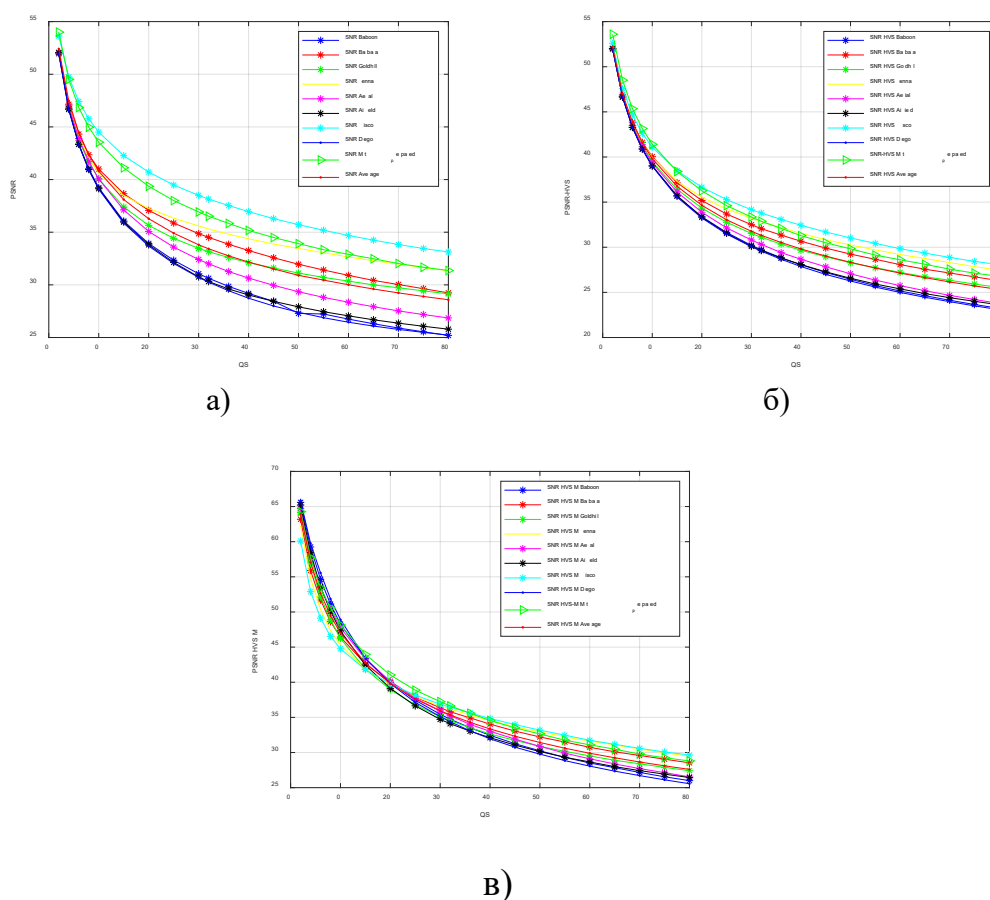


Їх точність досягає 0,82 - 0,84 (згідно SROCC). Крім того, точність на рівні 0,85 практично не залежить від кількості вхідних елементарних метрик.

## 1.2. Прогнозування якості даних ДЗ із застосуванням нейронних мереж

### 1.2.1. Прогнозування якості стиснених зображень

За останній час нами було проведено комплекс досліджень, що стосувались прогнозування параметрів стиснення з втратами. Їх актуальність зумовлена двома факторами. По-перше, для кожного зображення залежність метрики, що характеризує спотворення (наприклад, PSNR чи PSNR-HVS чи PSNR-HVS-M), від параметру, що контролює стиснення, як-то крок квантування (quantization step - QS) чи bpp (bits per pixel) є індивідуальною, про що свідчать криві на рисунку 5 для трьох метрик та кодеру AGU на основі дискретного косинусного перетворення (ДКП).



а) залежності для метрики PSNR; б) залежності для метрики PSNR-HVS;  
в) залежності для метрики PSNR-HVS-M.

**Рисунок 5 – Типові залежності метрик від кроку квантування**



По-друге – на практиці досить часто виникає задача забезпечити задану якість стиснутого зображення у відповідності до обраної метрики для конкретного зображення.

Як витікає з аналізу залежностей на рисунку 5, це неможливо зробити, використовуючи фіксований крок квантування (чи інший параметр, що керує стисненням (ПКС)). Наприклад, як витікає з аналізу залежностей на рисунку 5,а), PSNR=35 дБ може спостерігатись для кроку квантування від приблизно 18 до приблизно 58 в залежності від властивостей конкретного зображення (насамперед його складності). Тож на практиці необхідно вирішити задачу забезпечення бажаної якості стиснутого зображення (у відповідності до обраної метрики). При цьому задачу треба вирішити таким чином, щоб значення метрики забезпечувалось із задовільною похибкою, і все це виконувалось досить швидко.

Тут необхідно пояснити сенс термінів «задовільна похибка» та «досить швидко». Для кожної конкретної метрики задовільна похибка – це різна величина. Наприклад, для метрик PSNR, PSNR-HVS та PSNR-HVS-M бажане значення доцільно забезпечувати з похибкою не більше 1 дБ. Досить швидко означає з мінімальною кількістю ітерацій або зовсім без ітерацій.

В цьому сенсі існує декілька можливих підходів. По-перше, існують ітеративні процедури [91]. Вони забезпечують задану точність, але для них кількість ітерацій апріорно невідома і може бути досить великою. Є двоетапні процедури, що базуються на стисненні на першому етапі з ПКС, що визначається осередненою залежністю, декомпресії, визначенні скоректованого ПКС та стисненні з його використанням [92-94]. Ці методи дозволяють суттєво підвищити точність забезпечення бажаної якості після другого стиснення у порівнянні з точністю для першого етапу, але все одно точність у певних випадках може бути незадовільною. Ще одна група методів базується на прогнозуванні [95, 96]. Для них у певній кількості блоків зображення оцінюються характеристики зображення, що пов'язані із складністю зображення, а потім для даного ПКС характеристики спотворень прогнозуються за допомогою апроксимуючих формул. Відповідні процедури досить швидкі, але й для них бажану точність вдається досягти не завжди.

Одна з причин полягає в тому, що один параметр (наприклад, відсоток коефіцієнтів ДКП в блоках, які після квантування з заданим кроком обертаються в нулі) не є достатньо інформативним для того, щоб прогнозувати належну точність. В таких випадках зазвичай звертаються до спільного використання



кількох параметрів, які можна розрахувати швидко і які можна об'єднати за допомогою машин опорних векторів, багатопараметрових апроксиматорів чи НМ [97]. Аналогічні ідеї реалізовані при прогнозуванні ефективності фільтрації [98, 99]. Тож ми маємо позитивний досвід підвищення точності прогнозування за рахунок спільного використання кількох параметрів, що описують характеристики стиснення та конкретного зображення. Спробуємо використати цей досвід для прогнозування характеристик спотворень, що вносяться для стиснення з втратами.

Найпростіший та найпоширеніший параметр, що характеризує спотворення, - це середньоквадратична похибка (СКП, mean square error – MSE), що жорстко пов'язана з PSNR. Для кодерів на основі ДКП (AGU [100] та ADCTC [101]) відомо [95, 97], що при малих значеннях QS значення  $MSE \approx QS^2/12$ , а при великих значеннях ( $QS > 15$  для зображень у 8-бітному представленні) маємо  $MSE < QS^2/12$ , але конкретне значення залежить від складності зображення.

Таким чином, зрозуміло, що крок квантування має бути одним із вхідних параметрів НМ, що призначена прогнозувати СКП спотворень. Інший параметр чи інші параметри мають характеризувати зображення, що підлягає стисненню. Цей параметр чи параметри повинні мати високу швидкодію розрахунку, бути інформативними та «доповнювати» один одного. Одним із параметрів, що, безумовно, характеризує зображення, є ентропія, про що свідчать результати робіт [102, 103]. Навіть проста класифікація вхідних зображень на три групи за ентропією [103] дозволила суттєво покращити точність двоетапних процедур забезпечення бажаної якості стиснених зображень.

Втім, є й інші параметри, які можуть нести інформацію про складність зображення та його властивості. Такими, зокрема, є відношення меж (edge ratio - ER), контраст (contrast - C), енергія (energy - E), кореляція по строках та стовпцях (row-column correlation – RCC). Тож можна використовувати, принаймні, до шести вхідних параметрів (всі вони розраховуються дуже швидко, до того ж, це можна робити паралельно). Можливі й простіші комбінації вхідних параметрів, наприклад, крок квантування, ентропія і ще один з чотирьох параметрів, що перелічені вище. Питання полягає в тому, наскільки ефективним виявиться такий підхід та яким є пріоритет вимог до точності прогнозування та обмеження по складності і швидкодії.

На даний час нами отримані первинні результати, які свідчать про перспективність такого підходу. Використана найпростіша нейронна мережа



(перцептрон) з кількістю входів, що дорівнює кількості параметрів, що використовуються, та одним виходом, який прогнозує значення СКП.

Точність такого прогнозування можна характеризувати різними способами. Один із найбільш ілюстративних – діаграма розсіювання (скатерограма) прогнозованих та істинних значень із вписаними в них апроксимуючими кривими. Кореляція  $R$  між прогнозованими та істинними значеннями (в ідеалі дорівнює одиниці) теж може характеризувати точність прогнозування. Крім того, для навчених нейромереж зазвичай надають характеристики ефективності для даних, для яких проводилось навчання (training)), а також для тестових (test) та валідаційних (validation) даних. Бажано, щоб характеристики для тестових та валідаційних даних наближались до характеристик даних в результаті навчання.

На рисунку 6 наведені дані для трьох вхідних параметрів: кроку квантування, етропії та контрасту. Горизонтальна вісь – дійсне значення, а вертикальна – вихід НМ, тобто прогнозоване значення. На кожній діаграмі розсіювання біля вертикальної вісі наведено апроксимуючу (вписану) функцію, яка є лінійною. Над діаграмою розсіювання вказано, для яких даних (використаних при навчанні, при тестуванні або при валідації, All означає для всіх даних разом) вони отримані, а також наведено значення коефіцієнта кореляції. Скатерограми отримані для великої кількості зображень в градаціях сірого, стиснутих з різними значеннями кроку квантування.

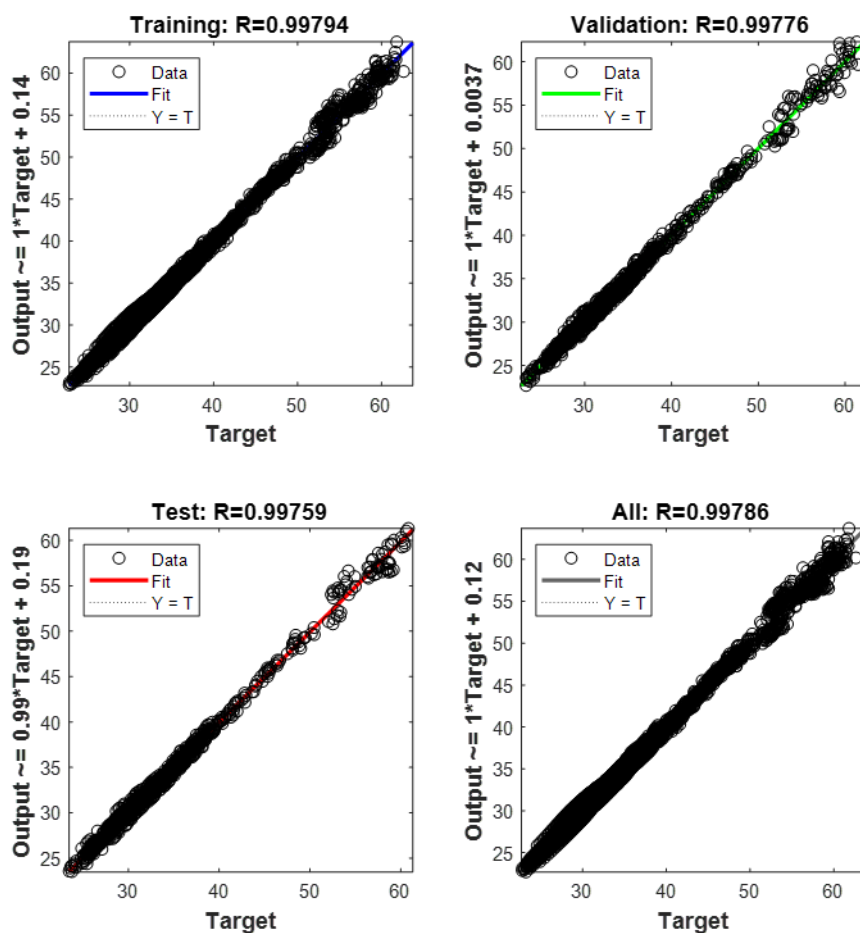
Аналіз даних дозволяє зробити кілька висновків:

а) лінійна залежність між прогнозованим та істинним значеннями є безсумнівною, прогнозовані значення у середньому трохи більше за істинні, але несуттєво, цю зміщеність можна ігнорувати, бо вона менше одного відсотка;

б) точки діаграми розсіювання розташовані дуже щільно та без викидів; значення коефіцієнта кореляції дуже близькі до одиниці; все це свідчить про можливість прогнозування з високою точністю;

в) значення  $R$  для результатів, отриманих під час тестування та валідації, лише трохи гірші за результат, отриманий для даних під час навчання; тож, з одного боку, навчання виконано якісно, і, з другого боку, можна бути впевненим, що на практиці не виникнуть непередбачені ситуації.

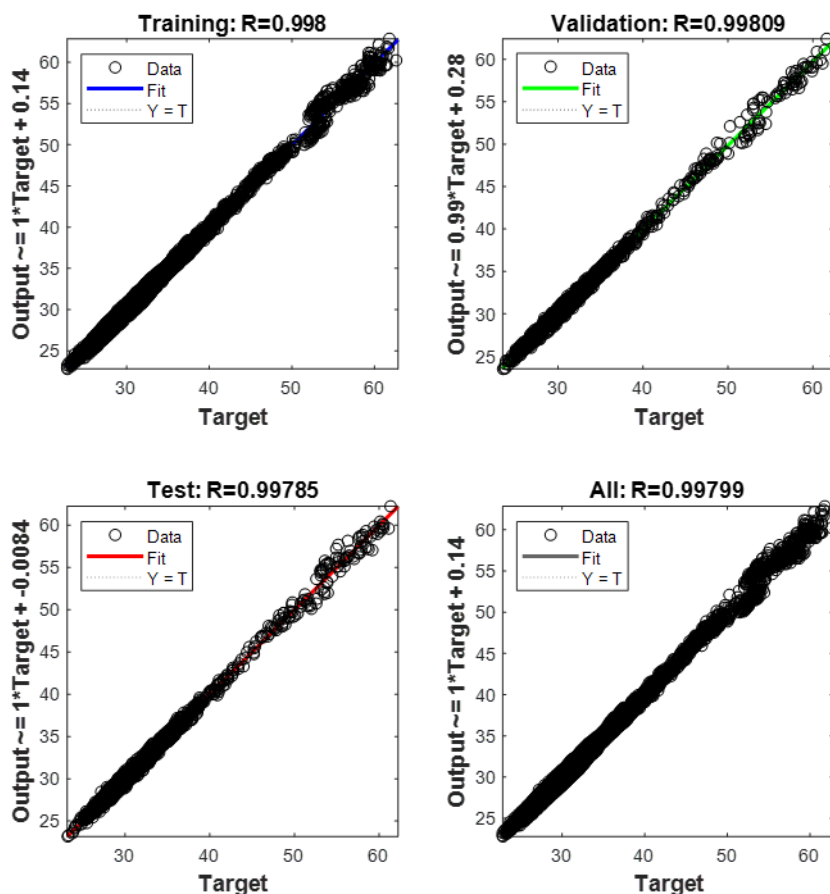
Наведені результати показують перспективність використання кількох параметрів та їх об'єднання нейронною мережею для покращення прогнозування СКП. Дисперсія похибок прогнозування знаходиться на рівні 0,1-0,2, що краще, ніж для двохетапних методів. Втім, наведені вище дані не відповідають на



**Рисунок 6 – Результати навчання та тестування НМ з трьома вхідними параметрами: кроком квантування, ентропією та контрастом**

запитання, чи є проаналізована підмножина вхідних параметрів оптимальною (такою, що забезпечує найвище значення коефіцієнту кореляції). Тому нами були розглянуті усі можливі комбінації з трьох вхідних параметрів, фіксуючи крок квантування як обов'язковий вхідний параметр, від якого СКП внесених спотворень залежить в першу чергу. Тож наведемо деякі результати. На рисунку 7 показані діаграми розсіювання для комбінації крок квантування, ER, E. Як можна побачити, основні тенденції такі ж, як і для даних на рисунку 6 – результати для тестування трохи гірші, ніж для навчання, спостерігається очевидна лінійність залежностей та компактне (скупчене) розташування точок уздовж вписаних ліній. Втім, значення R трохи кращі, ніж у попередньому випадку.

Ще трохи кращі результати спостерігаються для комбінації вхідних параметрів крок квантування, ентропія, кореляція по строках та стовпцях. Відповідні результати представлені на рисунку 8. Як можна бачити, для всіх

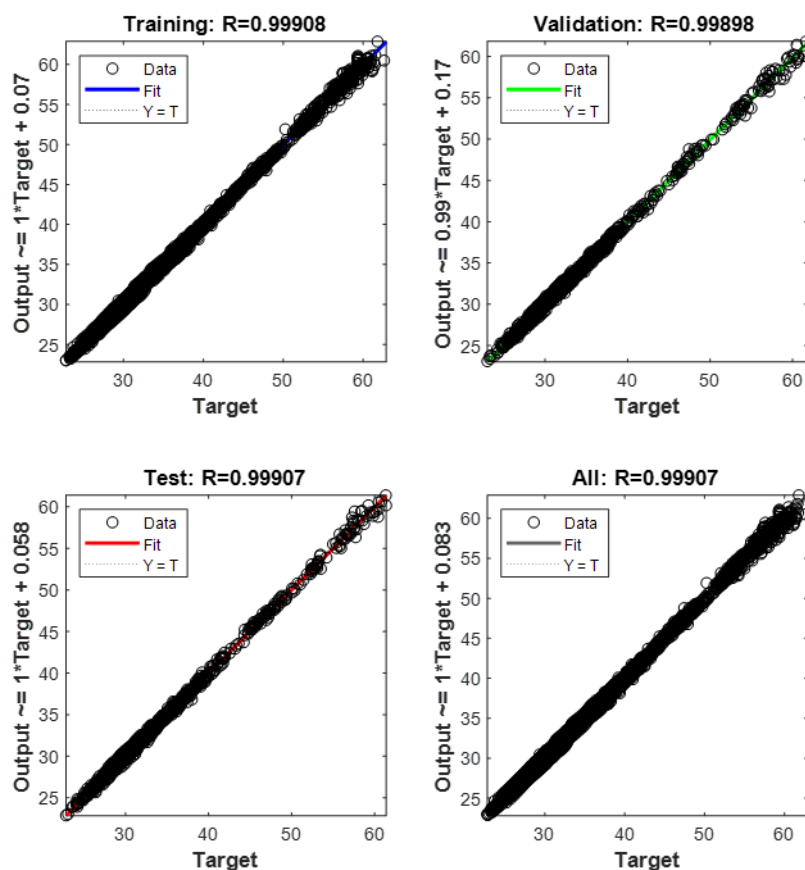


**Рисунок 7 – Результати навчання та тестування НМ з трьома вхідними параметрами: кроком квантування, ER та E**

трьох типів даних (навчання, тестування та верифікації) значення R або дуже близькі до 0,999, або перевищують цей поріг. Це свідчить про дуже високу точність прогнозування такою нейронною мережею. Втім, можна використовувати ще і більшу кількість вхідних параметрів.

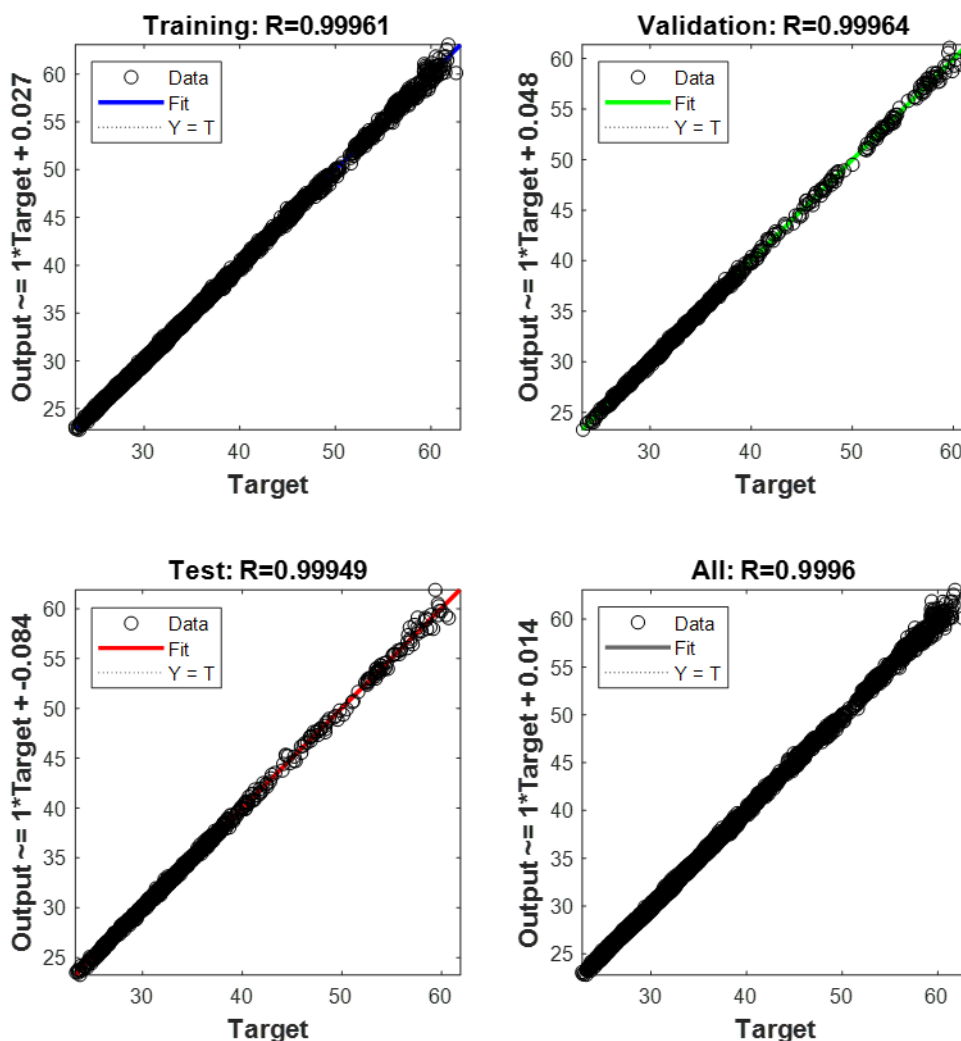
Приклад даних, отриманих для чотирьох вхідних параметрів (крок квантування, ентропія, контраст та відношення меж (ER)), наведено на рисунку 9. Як можна бачити із порівняння цих даних з тими, що отримані для трьох вхідних параметрів, то чотири вхідні параметри забезпечують ще кращі результати, дозволяючи досягти R близько 0,9996. Нами були перевірені деякі випадки для п'яти вхідних параметрів та випадок використання усіх шести перелічених вище вхідних параметрів. Для них досягнуто R близько 0,9996. Тож немає особливого сенсу використовувати 5 чи 6 вхідних параметрів, достатньо використати чотири чи найкращу комбінацію з трьох вхідних параметрів, покращуючи тим самим швидкодію та зменшуючи складність прогнозування.





**Рисунок 8 – Результати навчання та тестування НМ з трьома вхідними параметрами: кроком квантування, ентропією, кореляцією по строках та стовпцях**

Таким чином, показано, що спільне використання кроку квантування та двох чи трьох нескладних в обчисленні параметрів, що характеризують складність зображення (чи його певні властивості) з їх об'єднанням дуже простою нейронною мережею дозволяє дуже точно прогнозувати СКП внесених спотворень (або, автоматично) PSNR. Це означає, що отримуючи результати прогнозування для кількох значень кроку квантування (інші вхідні параметри при цьому не змінюються), можна легко та швидко знайти таке значення кроку квантування, що забезпечить задане СКП. Ми також сподіваємось, що таке прогнозування може бути корисним і для прогнозування модифікованих СКП, які відповідають метрикам PSNR-HVS та PSNR-HVS-M.



**Рисунок 9 – Результати навчання та тестування НМ з чотирма вхідними параметрами: кроком квантування, ентропією, контрастом та відношенням меж (ER)**

### ***1.2.2. Стиснення зображень, що спотворені завадами, із прогнозуванням існування оптимальної робочої точки та параметрів у ній***

Загалом для методів стиснення зображень з втратами компроміс між коефіцієнтом стиснення (КС) та внесеними спотвореннями залежить від багатьох факторів: а) яке застосування розглянутого зображення та як воно далі обробляється (аналізується, інтерпретується) [104-106]; б) який кодер (метод стиснення) використовується і наскільки легко варіювати та контролювати параметри стиснення; в) які властивості (вміст, складність) зображення, що підлягає стисненню [106]; г) які основні вимоги до стиснення та їх пріоритет.



Щодо фактора №1 – стиснення з втратами не повинно призводити до надто значного погіршення якості зображення щодо його візуального сприйняття людьми або класифікації чи виявлення країв/об'єктів. Відповідні дослідження були проведені для прогнозування лише помітних спотворень [107] або інших показників прийнятної деградації стиснених зображень [104, 106]. Щодо фактора №2 – різні методи стиснення використовують різні механізми зміни та контролю якості та коефіцієнту стиску (compression ratio – CR) стисненого зображення. Наприклад, JPEG2000 [108] і SPIHT [109] дозволяють забезпечити бажаний CR (або біт на піксель (BPP)) із широкими межами варіації якості стисненого зображення залежно від його складності [96]. У свою чергу, сучасний кодер BPG (better portable graphics) [110, 111] використовує параметр Q, який дозволяє контролювати пікове співвідношення сигнал/шум (PSNR) між вихідним і стисненим зображеннями в широких межах [112]. Це хороша властивість, яка привернула нашу увагу саме до цього кодера, хоча PSNR і не є адекватним показником якості зображення [113]. Щодо фактора №3 – для будь-якого кодера складніше стиснути більш складні за структурою зображення із забезпеченням відповідної якості для даного CR. Крім того, наявність шуму в зображеннях, що підлягають стисненню, додатково ускладнює завдання стиснення зображення [114-117]. Нижче ми розглянемо завдання стиснення з втратами зображень, пошкоджених шумом. Таке стиснення має певні особливості, зокрема ефект фільтрації шуму, вперше виявлений у роботах [114, 116] та можливе існування оптимальної робочої точки (ОРТ), яка спостерігалася для багатьох різних методів стиснення на основі малохвильових або дискретного косинусного перетворення (DCT). Основна мета дослідження, результати якого викладені нижче, — показати, що існування ОРТ і параметрів стиснення в ній можна спрогнозувати для кодера BPG з досить високою точністю. Нарешті, що стосується фактора №4, зазвичай потрібно забезпечити якомога більш високу якість стиснених зображень для даного CR, якомога швидше виконати стиснення тощо. Деякі з цих вимог є суперечливими і їх важко задовольнити водночас. Нижче ми зосередимося на забезпеченні відповідної якості стиснених зображень із якомога більшим CR.

Шум може з'явитися на зображеннях через багато факторів [118-120], і він може бути видимим або невидимим. Ми зосереджуємося на випадку видимого шуму, який виникає, якщо PSNR первинного зображення менше певного порогу (близько 35 дБ). Крім того, ми зосереджуємося на моделі найпростішого



адитивного білого гаусового шуму (additive white Gaussian noise - AWGN) [119], залишаючи більш складні моделі шуму [120] для майбутніх досліджень.

За цією моделлю маємо

$$I_{ij}^{noisy} = I_{ij}^{true} + n_{ij}, \quad (1)$$

де  $I_{ij}^{noisy}$  - значення зображення, спотвореного шумом, в  $ij$ -му пікселі;

$I_{ij}^{true}$  - істинне значення в  $ij$ -му пікселі;

$n_{ij}$  - значення AWGN з нульовим математичним очікуванням та дисперсією  $\sigma^2$ .

Нами припускається, що дисперсія шуму апріорно відома або оцінена автоматичними засобами [120, 121] із високою точністю.

Традиційними метриками для опису якості первинного зображення з шумом є середньоквадратична похибка (СКП, MSE) та PSNR, що розраховуються як

$$MSE^n = \sum_{i=1}^I \sum_{j=1}^J (I_{ij}^n - I_{ij}^{true})^2 / (IJ),$$

$$PSNR^n = 10 \log_{10} \left( \frac{255^2}{MSE^n} \right) = 10 \log_{10} \left( \frac{255^2}{\sigma^2} \right), \quad (2)$$

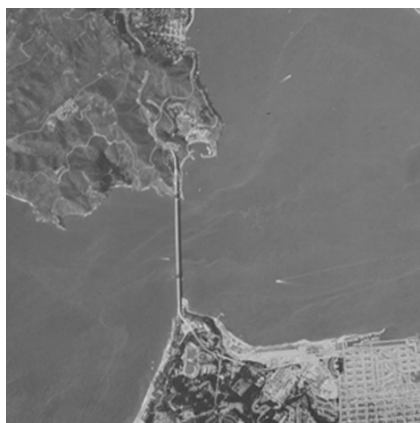
де  $I, J$  визначають розміри зображення та припускається, що зображення представлено як 8-бітні дані. В такому випадку шум, щоб його було видно, повинен мати дисперсію близько 20 для зображень з простою структурою (див. приклад на рис. 10,а)) та приблизно 30...40 для зображень із складною структурою (див. приклади на рис. 10,б та 10,в)) [122, 123]. Це пояснює, чому під час моделювання нами розглядалися значення дисперсії більші за 40.

Якість первинного зображення (з шумом) також можна характеризувати іншим чином, зокрема за допомогою метрик візуальної якості. Це можуть бути зокрема, метрика [124] або MS-SSIM [125]. Так

$$PSNR - HVS - M^n = 10 \log_{10} \left( \frac{255^2}{MSE - HVS - M^n} \right),$$

де  $PSNR - HVS - M^n$  розраховується в наборі блоків розміром  $8 \times 8$  пікселів з урахуванням різної чутливості зору людини (human vision system - HVS) до викривлень на різних просторових частотах, а також маскуючий ефект [124].

Нас більше цікавлять інші метрики та їх залежність від параметрів стиснення. Зокрема, розглянемо залежності  $PSNR^c(Q)$  і  $PSNR-HVS-M^c(Q)$ , де ці



а)



б)



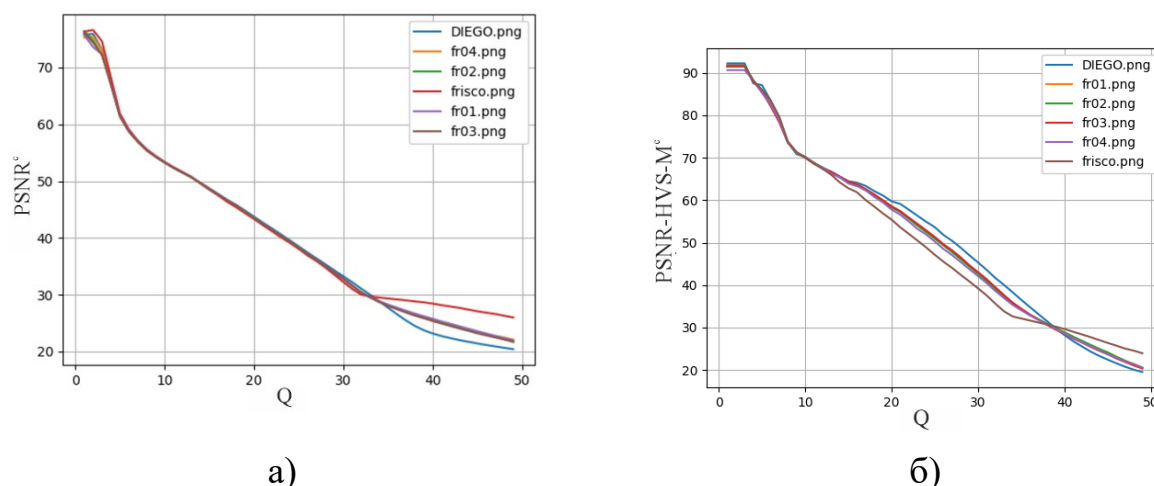
в)

а) Frisco; б) Diego; в) Fr01.

### Рисунок 10 – Приклади зображень різної складності

метрики розраховані для зображень, стиснених VPG з різною  $Q$ , і відповідних оригінальних (зашумлених) зображень. Приклади таких залежностей для шести тестових зображень (зображення Fr02, Fr03 і Fr04 мають приблизно таку ж саму складність, що й тестове зображення Fr01 на рисунку 10,в)), взятих із статті [112], наведені на рисунку 11.

Як бачимо, на цих кривих можна визначити три основні області. Перша – для  $Q < 9$ , де залежності є нелінійними, а якість стиснутих зображень характеризується дуже високими значеннями  $PSNR^c$  і  $PSNR - HVS - M^c$ , що означає, що спотворень майже немає. Потім, для  $Q$  приблизно від 9 до  $Q=33$ , ми маємо практично лінійне зниження  $PSNR^c$  (приблизно як  $63-Q$ ) і майже лінійне зниження  $PSNR - HVS - M^c$  (приблизно як  $83-1,33Q$ ). Ця частина кривих відповідає невидимим спотворенням (для  $Q$  близько 28 і менше) і видимим спотворенням для  $Q > 28$ . Видимі спотворення стосуються як фільтрації шуму, так



а)  $PSNR^c$  від  $Q$ ; б)  $PSNR - HVS - M^c$  від  $Q$ .

**Рисунок 11 – Залежності метрик від  $Q$  для кодеру VPG для шести тестових зображень, спотворених AWGN з дисперсією 64 ( $PSNR^n$  близько 30 дБ)**

і погіршення інформаційного контенту зображень. Третя область ( $Q > 33$ ) характеризується відхиленням поведінки залежностей від попередньої майже лінійної поведінки залежно від складності зображення.

Подібний аналіз було проведено для кількох інших значень дисперсії шуму. Межі першої ділянки залишаються практично незмінними. Права межа другої області залежить від дисперсії шуму. Значення  $Q$  для правої межі приблизно дорівнює  $Q_{rm} = 63 - PSNR_n$ . На практиці залежності  $PSNR^c$  і  $PSNR - HVS - M^c$  від  $Q$  можна отримати для будь-якого конкретного зображення, але наведені вище наближення можна використовувати для визначення початкової точки.

Теоретично, тобто в моделюванні, коли є зображення без шумів і до нього додається AWGN, а потім застосовується стиснення з втратами, також можна обчислити показники між стисненим зображенням і відповідним зображенням без шумів. Позначимо такі метрики як  $PSNR^{ct}$  і  $PSNR - HVS - M^{ct}$ . Приклади таких залежностей від  $Q$  для шести тестових зображень представлені на рисунках 12 і 13.

Як видно на рисунку 12, ОПТ спостерігаються для п'яти із шести тестових зображень. Для найпростішого тестового зображення структури Frisco ОПТ є більш очевидною, ніж для інших зображень. Для зображення найскладнішої структури Diego ОПТ відсутнє в обох випадках і залежності є монотонними. ОПТ «більш очевидні» для більшого  $\sigma^2$ . Крім того, для заданої дисперсії завад  $Q$  для ОПТ майже однакове для всіх зображень, для яких вона існує, і воно більше для



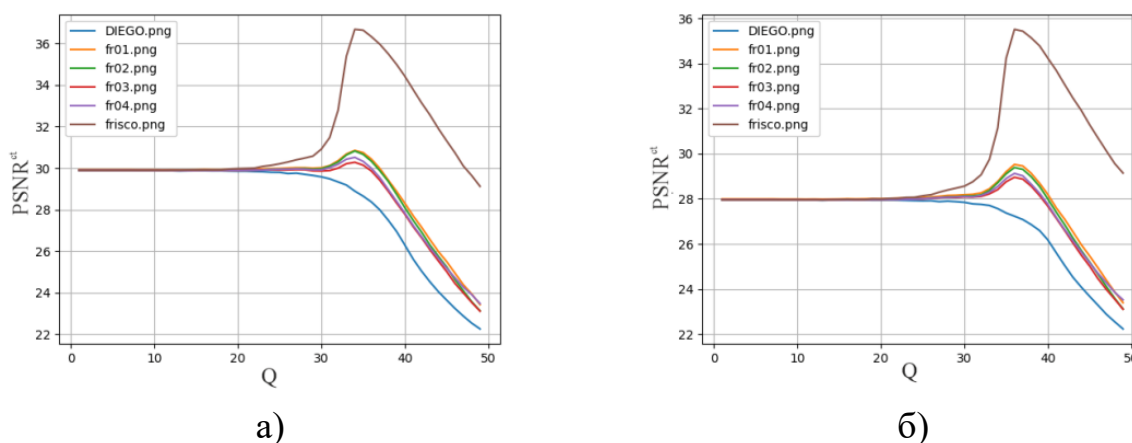
більшого  $\sigma^2$ . Ретельний аналіз показав, що формулу для  $Q_{OOP}$  можна записати як  $63 - Q_{OOP} = PSNR^n$  або, що еквівалентно,

$$Q_{OOP} = 14,9 + 20 \log_{10}(\sigma) \quad (3)$$

Тоді маємо  $Q_{OOP}$  близько 33 для  $\sigma=8$  (див. дані на рис. 12,а) і близько 35 для  $\sigma=10$  (див. дані на рис. 12,б).

Аналіз залежностей на рисунку 13 показує, що ОПТ також можна спостерігати для метрик візуальної якості (наприклад,  $PSNR-HVS-M$ ), хоча це трапляється рідше. ОПТ не є такою «очевидною», як для  $PSNR^{ct}$ . Вираз (3) приблизно дійсний для визначення  $Q_{OOP}$  і для цієї метрики.

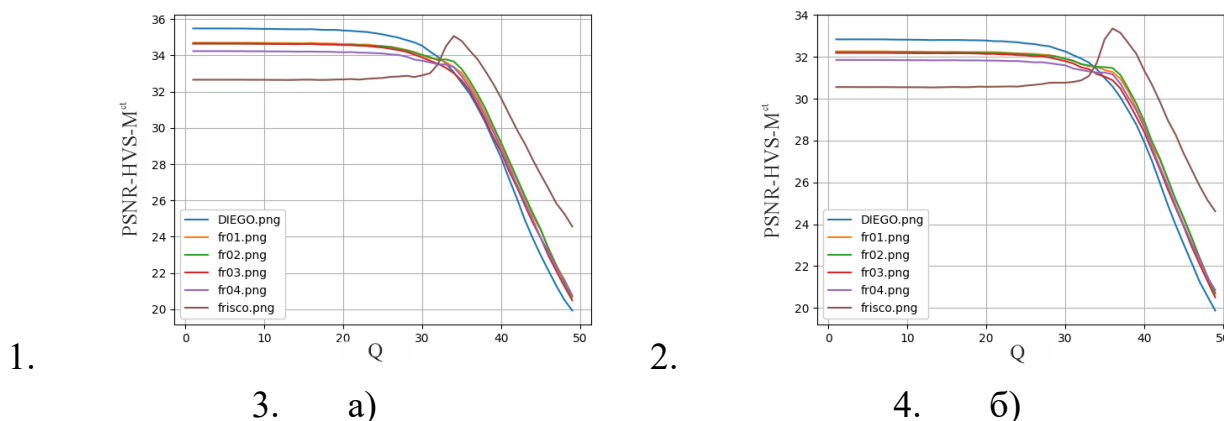
Узагальнюючи представлені результати та дані, наведені в роботах [112, 117], можна стверджувати, що ОПТ можна спостерігати, і, якщо це має місце, то доцільно стискати такі зображення в ОПТ.



а) дисперсія 64 ( $PSNR^n \approx 30$  дБ); б) дисперсія 100 ( $PSNR^n \approx 28$  дБ).

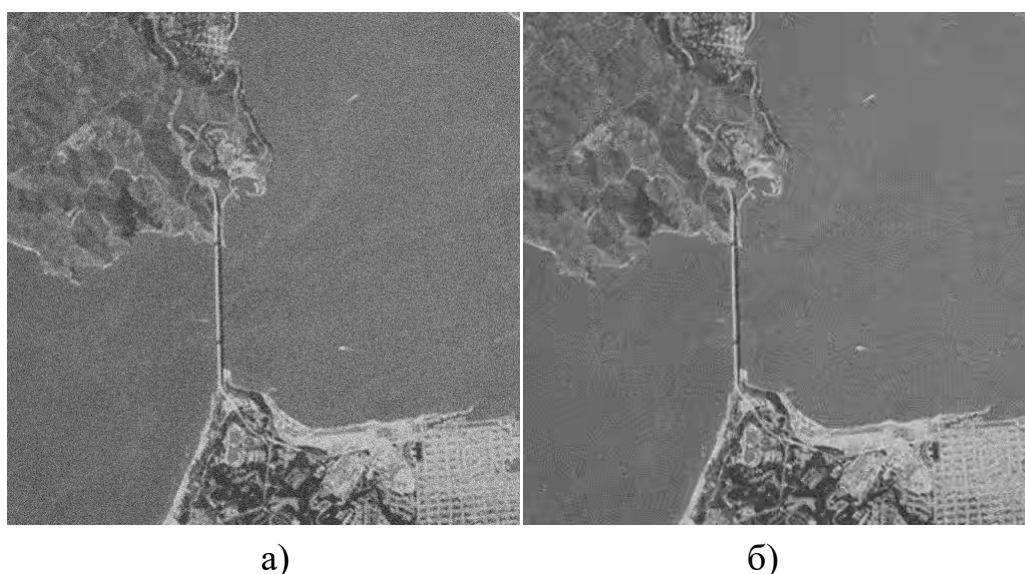
**Рисунок 12 – Залежності  $PSNR^{ct}$  від  $Q$  для кодеру VPG для шести тестових зображень з шумом**

Це дозволяє забезпечити якість стисненого зображення ближче до зображення без шумів порівняно з вихідним зображенням і забезпечити досить великий CR (рисунок 14 наводить приклад такої ситуації). Тим часом, якщо ОПТ для даного зображення не існує, то варто стиснути це зображення з  $Q$ , меншим, ніж  $Q_{OOP}$  (3), щоб уникнути внесення занадто великих спотворень.



а) дисперсія 64 ( $PSNR^n \approx 30$  дБ); б) дисперсія 100 ( $PSNR^n \approx 28$  дБ).

**Рисунок 13 – Залежності  $PSNR_{HVS} - M^{ct}$  від  $Q$  для кодеру VPG**



а) Зображення Frisco з шумом; б) зображення, стиснене в ОРТ.

**Рисунок 14 – Ілюстрація стиснення зображення в ОРТ, дисперсія шуму дорівнює 196**

На практиці немає зображення без шумів, і, отже, ані  $Metr^{ct}(Q)$ , ні  $Metr^n$  не можуть бути обчислені. Через це ми пропонуємо наступний підхід. Припустимо, що можна спрогнозувати (оцінити) різницю  $\Delta M = Metr^{ct}(Q_{OOP}) - Metr^n$ , маючи на увазі, що  $\sigma$  є апіорно відомим або попередньо оціненим з високою точністю, і, таким чином,  $Q_{OOP}$  також відомий відповідно до (3). Тоді, якщо прогнозований  $\Delta M$  є позитивним, висновок (рішення) полягає в тому, що ОРТ (згідно з розглянутою метрикою) існує, і навпаки.

Насправді ми використовуємо підхід, подібний до запропонованого нами





раніше в статті [116] для кодерів на основі ДКП, і перевіряємо, чи застосовний він для кодера VPG. Спочатку нагадаємо вимоги до прогнозування. Воно повинно бути досить швидким і досить точним. Говорячи «швидко», ми маємо на увазі, що воно має бути суттєво більш швидким за саме стиснення. Говорячи «достатньо точно», мається на увазі, що похибки прогнозування  $\Delta M$  повинні мати досить мале середньоквадратичне відхилення (root mean square error - *RMSE*), наприклад, менше 1 дБ для показників *PSNR* і *PSNR-HVS-M*, обидва виражені в дБ.

Загалом, підхід [116] припускає, що існує залежність між  $\Delta M$  і деяким вхідним параметром, який «характеризує» властивості зображення та шуму та може бути легко обчислений. Ця залежність відома (визначена) заздалегідь (офлайн) і доступна до моменту, коли потрібно спрогнозувати існування ОПТ для даного зображення, яке потрібно стиснути.

Як вхідні параметри, в [116] було запропоновано та протестовано декілька статистичних параметрів. Вони базуються на порівнянні амплітуд коефіцієнтів ДКП, розрахованих в блоках  $8 \times 8$  пікселів, з певними порогами, пов'язаними зі середньоквадратичним відхиленням шуму. Простота отримання таких вхідних параметрів пов'язана з двома аспектами. По-перше, ДКП у блоках  $8 \times 8$  пікселів є стандартною операцією в обробці зображень та відео [126], і її можна реалізувати дуже швидко. По-друге, ДКП не слід виконувати для всіх можливих положень блоку даного зображення. Досить взяти 300...1000 випадково розміщених блоків, щоб з належною точністю оцінити вхідний параметр (це стане зрозуміло пізніше з аналізу діаграм розсіювання, також це доведено в нашій роботі [29]).

На даний момент незрозуміло, який статистичний параметр розглянутого типу є найкращим для аналізованого застосування. Ми намагаємося лише показати, що запропонований підхід може працювати досить добре, залишаючи завдання оптимізації для майбутніх досліджень.

Подібно до роботи [116], ми розглядаємо два вхідних параметри –  $P_{2\sigma}$  і  $P_{2,7\sigma}$ . Перший параметром є ймовірність того, що амплітуди ДКП-коефіцієнтів менші за  $2\sigma$ . Другим параметром є ймовірність того, що амплітуди ДКП-коефіцієнтів перевищують  $2,7\sigma$ . Такі параметри виникли з теорії придушення шумів на основі ДКП [127]. Їх властивості стануть зрозумілі з подальшого аналізу.

Щоб отримати залежності для прогнозування, були отримані діаграми

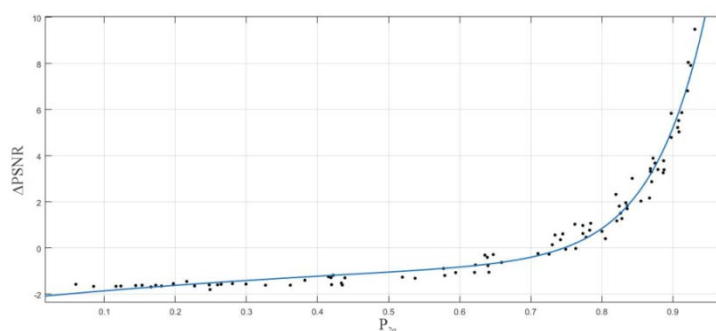


розсіювання  $\Delta M$  від вхідних параметрів. Один приклад представлено на рисунку 15 для  $\Delta PSNR$  і  $Q_{OOP}$ . Кожна конкретна точка діаграми розсіювання відповідає одному тестовому зображенню, спотвореному AWGN із заданою дисперсією шуму та стиснутому з  $Q_{OOP}$ , встановленим відповідно до (3). Використано 11 тестових зображень різної складності та 8 значень дисперсії шуму в межах від 0,25 до 400.

Аналіз даних на рисунку 15 показує наступне. По-перше,  $P_{2\sigma}$  змінюється в широких межах від майже нуля до майже одиниці. Малі значення  $P_{2\sigma}$  відносяться до зображень складної структури та/або досить малих дисперсій шуму (див. приклад на рисунку 16), на стиснутому зображенні спостерігаються згладжування країв/деталей). У свою чергу, великі значення  $P_{2\sigma}$  мають місце для зображень з простою структурою, спотворених досить інтенсивним шумом (дивись приклад на рисунку 14).

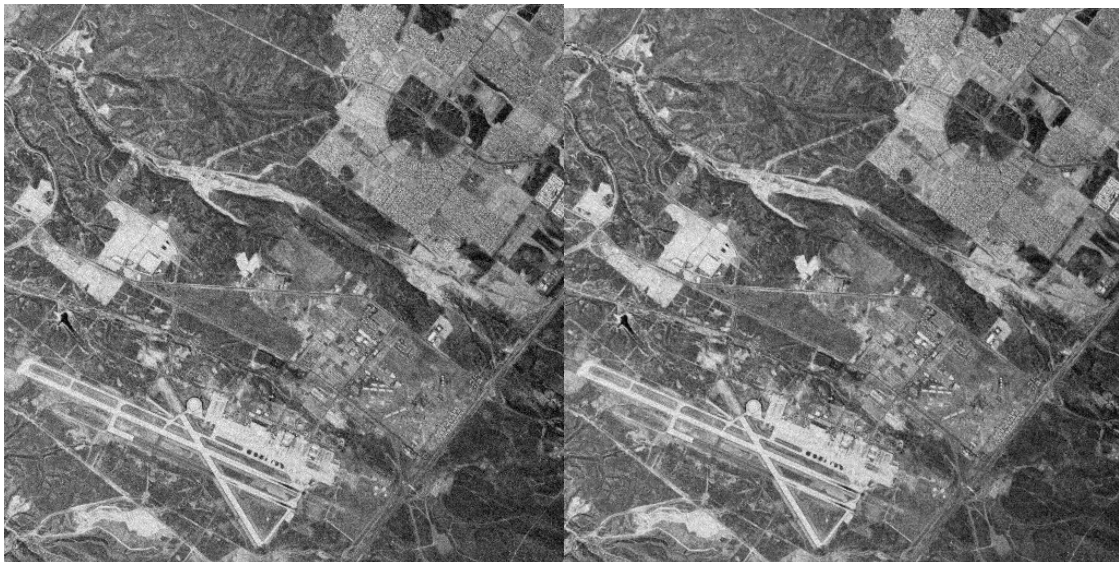
По-друге,  $\Delta PSNR$  змінюється в досить широких межах від приблизно -2 дБ (дивись дані на рисунку 12,а) для тестового зображення Diego) до приблизно +8 дБ (дивись дані на рисунку 12,б) для тестового зображення Frisco ).

По-третє, існує очевидна тенденція до збільшення  $\Delta PSNR$ , якщо  $P_{2\sigma}$  стає більшим. Розкид точок діаграми розсіювання для даного  $P_{2\sigma}$  є невеликим (здебільшого менше 1 дБ). Це означає, що можлива потенційно точна підгонка (фітінг) і, використовуючи підігнані криві, можна досить точно прогнозувати  $\Delta PSNR$  (нижче ці питання будуть обговорюватися більш детально).



**Рисунок 15 – Діаграма розсіювання  $\Delta PSNR$  від  $P_{2\sigma}$  та вписана крива**

По-четверте, якщо основним питанням є наявність ОПТ (а не прогнозування значення  $\Delta PSNR$  в ОПТ), то можливе просте правило. Можна вважати, що ОПТ для метрики  $PSNR$  існує з досить високою ймовірністю, якщо  $P_{2\sigma}$  перевищує 0,75.



а)

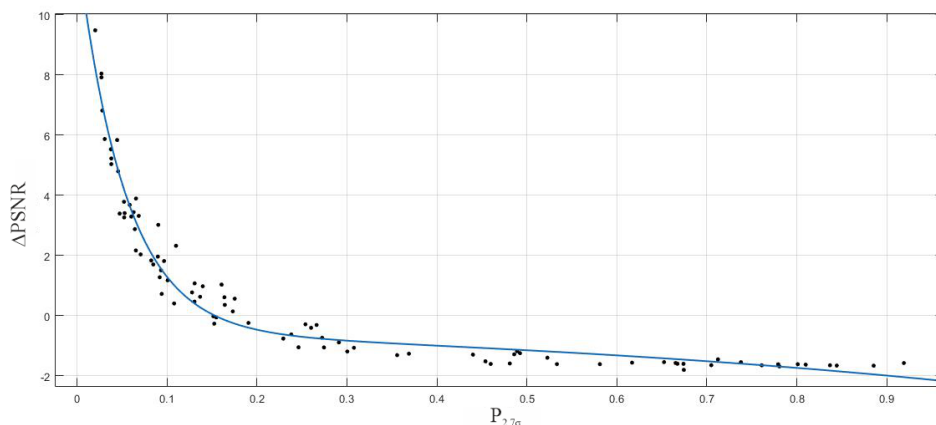
б)

а) зображення Diego, спотворене шумом;

б) це ж зображення після стиснення з  $Q_{009}$ .

**Рисунок 16 – Ілюстрація впливу стиснення для зображення із складною структурою, дисперсія завад дорівнює 64**

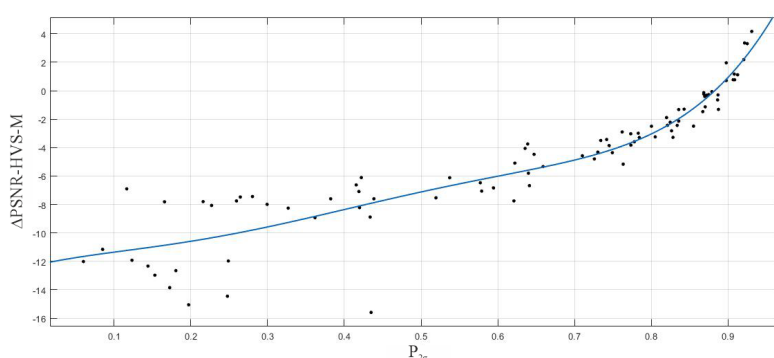
Розглянемо тепер інший вхідний параметр -  $P_{2,7\sigma}$ . Діаграма розсіювання  $\Delta PSNR$  від  $P_{2,7\sigma}$  представлена на рисунку 17. На відміну від діаграми розсіювання на рисунку 15, спостерігається тенденція до зменшення  $\Delta PSNR$ , якщо  $P_{2,7\sigma}$  збільшується. Це пояснюється тим, що  $P_{2\sigma}$  і  $P_{2,7\sigma}$  визначаються порізному (менші та більші за порогове значення відповідно). У цьому випадку правило для прийняття рішення про існування ОПТ можна сформулювати так: ОПТ існує з високою ймовірністю, якщо  $P_{2,7\sigma}$  менше, ніж 0,15. Таким чином, рішення можна прийняти для обох вхідних параметрів.



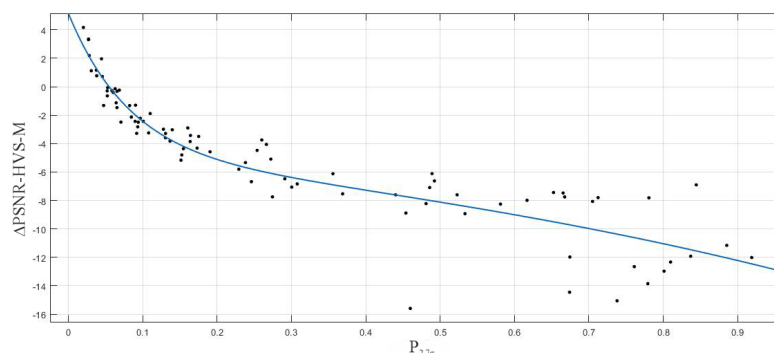
**Рисунок 17 – Діаграма розсіяння  $\Delta PSNR$  від  $P_{2,7\sigma}$  та вписана крива**



Тепер проаналізуємо, чи можна прогнозувати  $\Delta PSNR-HVS-M$ . Графіки розсіювання для  $\Delta PSNR-HVS-M$  від  $P_{2\sigma}$  і  $P_{2,7\sigma}$  зображені на рисунку 18. Діаграми розсіювання не такі компактні, як для  $\Delta PSNR$ . При цьому основні тенденції подібні до тих, що спостерігалися раніше.  $\Delta PSNR-HVS-M$  збільшується, якщо  $P_{2\sigma}$  збільшується, а  $P_{2,7\sigma}$  зменшується. ОПТ за  $\Delta PSNR-HVS-M$  спостерігається для меншої кількості зображень. Ще одна особливість полягає в тому, що  $\Delta PSNR-HVS-M$  змінюється в інших межах, ніж  $\Delta PSNR$ : приблизно від -15 дБ до приблизно 4 дБ. Відповідно до  $\Delta PSNR-HVS-M$  ОПТ спостерігається, якщо  $P_{2\sigma}$  перевищує 0,85 або  $P_{2,7\sigma}$  менше 0,08.



а)



б)

а) діаграма розсіювання для  $P_{2\sigma}$ ; б) діаграма розсіювання для  $P_{2,7\sigma}$ .

**Рисунок 18 – Діаграма розсіювання  $\Delta PSNR-HVS-M$  від різних параметрів та вписані криві  $P_{2\sigma}$  (а) та  $P_{2,7\sigma}$  (б) та вписані криві**

Може здатися, що можна спостерігати дуже велике зниження  $PSNR - HVS - M^{ct}$  через стиснення з втратами (як впливає з діаграм розсіювання на рисунку 18,  $\Delta PSNR-HVS-M$  може становити близько -10 дБ).

Однак це не фатальний випадок - це відбувається для  $P_{2\sigma}$  близько 0,2 ( $P_{2,7\sigma}$



близько 0,7), тобто для дуже складних зображень із шумом низької інтенсивності. Це означає, що  $PSNR - HVS - M^n$  може бути порядку 60 дБ, а  $PSNR - HVS - M^{ct}$  для  $Q_{OOP}$  становить 50 дБ.

$PSNR - HVS - M^n$  близько 60 дБ означає, що шум невидимий в первинному зображенні, а  $PSNR - HVS - M^{ct}$  близько 50 дБ означає, що спотворення також не видно. Таким чином,  $\Delta PSNR - HVS - M$  близько -10 дБ у таких випадках не є проблемою. Проблема полягає в тому, що для деяких зображень  $PSNR - HVS - M^n$  становить близько 30-40 дБ (тобто шум видно), а  $\Delta PSNR - HVS - M$  становить близько -2 дБ. Саме такі випадки мають місце для тестового зображення Дієго на рисунку 13, ілюстрація наведена на рисунку 16.

Нами ще не вивчено питання найкращого вписування кривої в діаграми розсіювання. З одного боку, приклади, показані на рисунках 15, 17 і 18, інтуїтивно показують, що підгонка може бути виконана ефективно. З іншого боку, вимоги до фітінгу та кількісні критерії ефективності підгонки не обговорювалися.

Підгонка кривої до діаграми розсіювання може бути виконана різними методами, використовуючи різні інструменти та критерії [128, 129]. Зокрема, Matlab і Excel пропонують хороші можливості для підгонки кривих. У наших дослідженнях ми використовували Matlab Curve Fitting Tool.

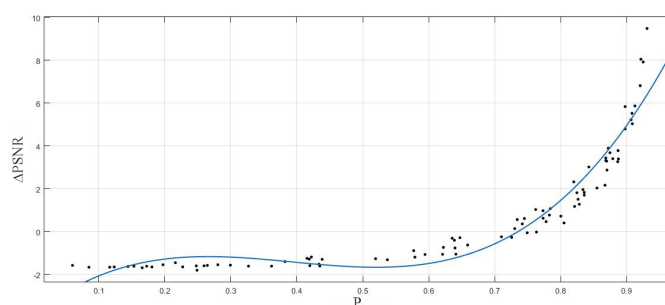
Якість підгонки зазвичай характеризується декількома кількісними параметрами (критеріями), такими як goodness-of-the-fit  $R^2$ ,  $RMSE$ , скоригований (adjusted)  $R^2$  тощо [128].  $R^2$  і скоригований  $R^2$  мають бути якомога більшими (наближаючись до одиниці), а  $RMSE$  має бути якомога меншим для гарного фітінгу. Регресія вважається дуже хорошою, якщо  $R^2$  і скоригований  $R^2$  перевищують 0,9.

У [116] було використано поліноміальну апроксимацію. Було показано, що поліноми четвертого і п'ятого порядку зазвичай забезпечують хорошу підгонку. Два приклади підгонки поліномів наведені на рисунку 19. Для даних на рисунку 19,а) маємо  $R^2 = 0,93$ ,  $RMSE = 0,71$ , скоригований  $R^2$  дорівнює 0,93, тобто формально результати досить хороші. Для полінома п'ятого порядку (рис. 19,б)) результати підгонки ще кращі ( $R^2 = 0,97$ ,  $RMSE = 0,47$ , скоригований  $R^2$  дорівнює 0,97). При цьому підігнані криві не є монотонними і можна спостерігати локальні максимуми. Ми не стверджуємо, що залежності повинні

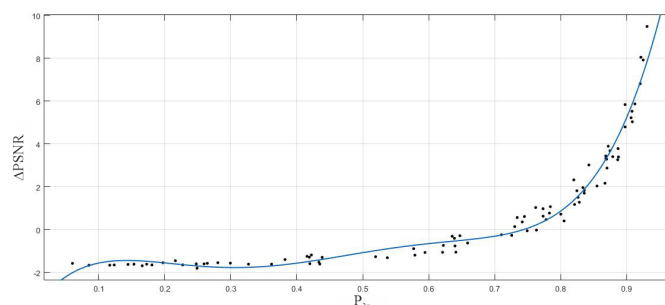


бути монотонними, але це виглядає так з візуального аналізу діаграм розсіювання.

Для діаграми розсіювання  $\Delta PSNR-HVS-M$  від  $P_{2\sigma}$  (див. рис. 20 ) отримано досить хороші результати підгонки за допомогою поліномів. Отримані криві для всіх поліномів є монотонними, і найкращу відповідність отримано для полінома п'ятого порядку ( $R^2 = 0,87$ ,  $RMSE=1,66$ , скоригований  $R^2$  дорівнює 0,86). Зауважте, що велике середньоквадратичне значення пов'язане в основному з розкидом даних діаграми розсіювання для  $P_{2\sigma} < 0,5$ . Це практично не впливає на рішення щодо існування ОПТ, яке, нагадаємо, приймаються для більших значень параметру, що аналізується.



а)



б)

а) поліном 3-го порядку; б) поліном 5-го порядку.

**Рисунок 19 – Приклади вписування поліномів різного порядку в діаграму розсіювання  $\Delta PSNR$  від  $P_{2\sigma}$**

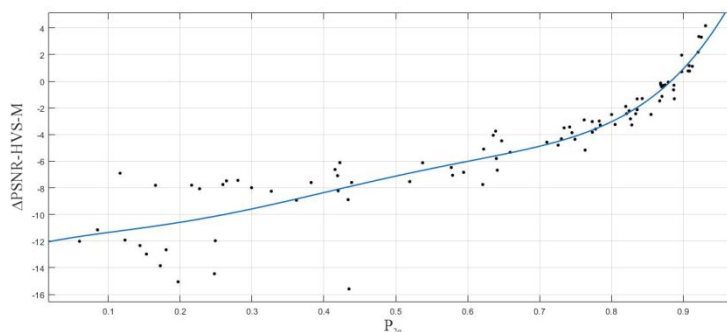
Зрозуміло, що підгонка за допомогою поліномів не є єдиним можливим варіантом. Інша опція – виконувати фітінг за допомогою експоненційних функцій як

$$f(x) = a * \exp(bx) + c \exp(dx), \quad (4)$$

для якої є 4 параметри, які треба визначити:  $a$ ,  $b$ ,  $c$  і  $d$ . Саме така підгонка використана для отримання кривих на рисунках 15, 17 і 18. Як можна бачити,



фітінг виконано якісно, а отримані криві є монотонними.



**Рисунок 20 – Приклад поліноміального фітінгу для діаграми розсіювання  $\Delta PSNR-HVS-M$  від  $P_{2\sigma}$  (поліном 5-го ступеня)**

Проаналізуємо також кількісні характеристики. Вони наведені в таблиці 8. Як бачимо, висновки, які можна зробити, такі ж, як і для підгонки полінома. Використання  $P_{2\sigma}$  як вхідного параметра забезпечує трохи кращу підгонку, ніж використання  $P_{2,7\sigma}$ . Діаграма для  $\Delta PSNR-HVS-M$  характеризується більшим розкидом даних і, таким чином, прогнозується гірше, ніж  $\Delta PSNR$ .

Враховуючи бажання мати монотонні апроксимуючі криві (4) та характеристики точності, що забезпечуються різними варіантами фітінгу, нами рекомендується використовувати наступні варіанти:

$$\Delta PSNR = 0,0008269 * \exp(9,755 * P_{2\sigma}) - 13,24 * \exp(-1,192 * P_{2\sigma}), \quad (5)$$

$$\Delta PSNR - HVS - M = -4,897 * \exp(-4,897 * P_{2\sigma}) + 10,07 * \exp(-12,12 * P_{2\sigma}). \quad (6)$$

Вище ми звернули увагу на точність апроксимацій, які суттєвим чином визначають точність прогнозу. Це, загалом, вірно, хоча точність прогнозування також визначається ще одним фактором, а саме, наскільки точно оцінюються вхідні параметри.

**Таблиця 8 – Характеристики фітінгу для суми двох експонент**

Параметри, що прогнозується	Вхідний параметр	$R^2$	Скоригований $R^2$	RMSE
$\Delta PSNR$	$P_{2\sigma}$	0,974	0,973	0,437
$\Delta PSNR$	$P_{2,7\sigma}$	0,963	0,962	0,523
$\Delta PSNR-HVS-M$	$P_{2\sigma}$	0,870	0,865	1,64
$\Delta PSNR-HVS-M$	$P_{2,7\sigma}$	0,863	0,859	1,68

Спеціальне дослідження [29] показало, що  $P_{2\sigma}$  для більшості зображень



можна визначити з похибками менше 0,01, якщо кількість аналізованих блоків порядку 300-1000. Як впливає з аналізу поведінки підігнаних кривих, такі похибки зазвичай не призводять до істотних змін (систематичних похибок) прогнозованих значень. Принаймні, такі похибки менші, ніж  $RMSE$  фітінгу.

Крім того, для підгонки діаграми розсіювання можна використовувати інші функції [29]. Зокрема, раціональні функції забезпечують приблизно такі ж результати фітінгу, як і сума двох експонент.

Проблемою залишається точність прогнозування  $\Delta PSNR-HVS-M$ , яку бажано покращити. Можна сподіватись, що точність прогнозування можна покращити за рахунок використання не одного, а двох чи більшої кількості вхідних параметрів та їх об'єднання навченою нейромережею.

Таким чином, показано, що стиснення із втратами зображень, що спотворені шумом, на основі VPG має певні особливості. Зокрема, OPT можна спостерігати за умови, що зображення досить просте, а шум досить інтенсивний. Якщо OPT існує, доцільно стиснути зображення в OPT. Якщо OPT не спостерігається,  $Q$  має бути меншим, ніж  $Q_{OOP}$ , щоб не вносити занадто великі спотворення. Показано, що можна спрогнозувати, чи існує OPT і, крім того, яка якість зображення для  $Q_{OOP}$ . Можливість існування OPT показано для двох метрик, і це можна прогнозувати досить легко та швидко. Прогноз використовує простий аналіз статистики ДКП-коефіцієнтів, визначених в обмеженій кількості блоків  $8 \times 8$  пікселів. Маючи такий прогноз, стає можливим прийняти правильне рішення, який  $Q$  встановити. В нашій роботі [29] проведено первинний аналіз випадків сигнально-залежного шуму та багатоканальних зображень і доведено, що OPT може існувати і для цих випадків. Є впевненість, що точність прогнозу для метрики  $PSNR-HVS-M$  можна покращити. Є також сподівання, що можна прогнозувати і покращення (чи погіршення) й інших метрик візуальної якості.

### **1.2.3. Прогнозування ймовірності правильної класифікації багатоканальних зображень ДЗ**

Слід враховувати, що процес розробки засобів обробки/аналізу даних має дослідницький характер. При вирішенні конкретних завдань обробки даних ДЗ часто виникає необхідність застосування багатоваріантного підходу, при якому використовуються різні методи отримання результатів з подальшим їх аналізом і прийняттям рішення про доцільність того чи іншого способу обробки, як





попереднього, так і спеціального. Однією з найскладніших задач обробки зображень є, як відомо, завдання аналізу об'єктів на зображеннях з подальшим їх розпізнаванням та ідентифікацією [130].

Складність застосування алгоритмів класифікації пов'язана з наявністю великої кількості алгоритмів і підходів до вирішення задачі розпізнавання образів [131]. Як правило, вибір найбільш прийняттого алгоритму (за певним критерієм якості) здійснюється евристичним або емпіричним шляхом. В останньому випадку цей вибір ґрунтується на порівнянні результатів для оцінюваних алгоритмів. Водночас реалізація зазначених алгоритмів, як і виконання серії тестових експериментів з розпізнавання образів, є трудомістким процесом, який потребує багато часу. Додатковим аспектом проблеми, що розглядається, є принципова можливість (або необхідність) вибору не тільки системи правил прийняття рішення про клас, а й вибору простору класифікаційних ознак. Розмірність простору ознак значною мірою визначає вартість вимірювання необхідних характеристик об'єктів, обчислювальну складність процедур навчання та прийняття рішень, а також надійність класифікації. Таким чином, ефективність (інформативність) виділених ознак може бути безпосередньо визначена ймовірністю правильного розпізнавання.

На практиці для оцінки ймовірностей правильного (чи помилкового) розпізнавання використовується статистичний підхід [132]. Це включає значну кількість процедур розпізнавання для різних еталонних зображень з використанням різних класифікаторів і різних комбінацій ознак. На відміну від статистичних оцінок, які базуються на аналізі всього обсягу наявних даних, прогнози можна робити на значно меншій кількості даних, що допускає неповноту та неточність наявної інформації [133, 134]. Тому особливу увагу слід приділяти розробці та дослідженню методів прогнозування ефективності класифікації для вибірки обмеженого розміру. При цьому передбачається, що спосіб формування цієї вибірки має бути таким, щоб мінімізувати витрати часу на прогнозування значення заданого критерію при гарантованій (заданій) достовірності цих прогнозів [135]. У нашій роботі [136] запропоновано метод прогнозування повної ймовірності правильної класифікації зображень на основі НМ або квазібайєсівського підходу за умов неповної апріорної інформації про класи. Однак цей метод не був достатньо перевірений. Тож розглянемо результати застосування запропонованого методу прогнозування достовірності класифікації триканальних зображень ДЗ.



У статистичній теорії розпізнавання класичним є байєсівський підхід. Він забезпечує оптимальний класифікатор, коли відомі щільності розподілу для всього набору класів, упорядковані ймовірності появи об'єктів з кожного класу та задана функція втрат, яка чисельно виражає суму узагальнених втрат, пов'язаних з помилковими рішеннями. Однак у реальних задачах розпізнавання часто немає достовірної інформації про повний набір класів, невідомі апріорні ймовірності класів, а також тип і параметри розподілів випадкових ознак класів.

Розглянемо випадок, коли кількість і тип класів відомі заздалегідь (або можуть бути встановлені на етапі попереднього аналізу даних). Це припущення дозволяє використовувати контрольовані алгоритми класифікації, зокрема, статистичний класифікатор максимальної правдоподібності (maximum likelihood - ML) або класифікатори на основі нейромереж [137]. На відміну від критерію Байєса, критерій ML не вимагає знання апріорних ймовірностей класів і функції втрат, дозволяє оцінити надійність рішень і легко узагальнюється на випадок багатьох класів [131, 138].

Як класифікаційні ознаки будемо приймати спектральні ознаки (наприклад, R, G, B у даних кольору), які формують тривимірну характеристику кожного пікселя аналізованого зображення, пов'язану із записаною спектральною відбивною здатністю об'єктів сцени в оптичному діапазоні спектра. Етап навчання класифікатора полягає у формуванні опорних описів класів на основі розмічених (вибраних) навчальних зразків. Враховуючи стохастичний характер класифікаційних ознак, описи класів є статистичними оцінками умовних щільностей розподілу ймовірностей  $\hat{f}_p(\vec{x} | a_k)$ ,  $\vec{x} = \{x_1, x_2, \dots, x_p\}$ , де  $p$  – кількість ознак (розмірність простору ознак),  $k = 1 \dots K$ ,  $K$  – кількість класів. Оскільки розподіли реальних багатоканальних даних часто є негаусовими і водночас характеризуються наявністю сильних кореляцій, для їх апроксимації доцільно використовувати багатовимірні  $S_b$ -розподіли Джонсона [138]. Методи оцінювання параметрів розподілу Джонсона є ітераційними і призводять до вирішення оптимізаційної задачі, метою якої є мінімізація повної середньоквадратичної похибки при апроксимації емпіричного розподілу прийнятою моделлю. З негаусовістю розподілу ознак зазвичай досить легко можуть впоратись і класифікатори на основі нейронних мереж.

Після визначення опису еталонного класу пікселі зображення призначаються до певного класу за допомогою прийнятого правила прийняття



рішень. При використанні процедури попиксельної класифікації кожен піксель з індексами  $(i, j)$  аналізується окремо. Значення вектору для поточного пікселя  $\vec{x}^*(i, j) = \{R_{ij}, G_{ij}, B_{ij}\}$  підставляються в аналітичну модель опорних класів або використовуються як входи класифікатора на основі нейромережі. Результати  $\hat{f}(\vec{x}^*(i, j) | a_k)$ ,  $k = 1 \dots K$  (або значення на  $K$  виходах нейромережі) порівнюють між собою і обирається максимальна оцінка; її індекс співпадає з індексом класу, до якого й відносять поточний піксель  $s(i, j)$ .

Відсоток правильно розпізнаних пікселів в опорному зображенні зазвичай вибирають у якості критерія якості (точності) класифікації. Статистичні оцінки матриці плутанини можна знайти як

$$\hat{P}_{uv} = \frac{1}{M} \sum_{i,j:s(i,j) \in a_k} c_{uv}(\vec{x}(i, j)), \quad (7)$$

де  $c_{uv}(\vec{x}(i, j)) = 1$ , якщо піксель  $s(i, j)$  належить до класу  $a_u$ , і класифікатор відносить його до класу  $a_v$ ,  $u, v = 1 \dots K$ , в протилежному випадку  $c_{uv}(\vec{x}(i, j)) = 0$ ,  $M$  – розмір контрольної вибірки (кількість відмічених пікселів у опорному зображенні).

Статистична оцінка сумарної ймовірності правильного розпізнавання за результатами класифікації еталонного зображення визначається з використанням обох оцінок ймовірностей правильного розпізнавання для кожного класу з повного набору  $A = \{a_k\}$ , та інформації про апріорні ймовірності  $P(a_k)$ ,  $k = 1, \dots, K$ :

$$\hat{P}_{total} = \sum_{k=1}^K P(a_k) \hat{P}_{kk}. \quad (8)$$

Якщо апріорні ймовірності  $P(a_k)$  невідомі, то вважають, що вони є однаковими:  $P(a_k) = 1/K$ . В інших випадках їх значення отримують з експертних оцінок чи попереднього аналізу опорних зображень. Традиційний підхід до оцінювання  $P(a_k)$  базується на припущенні, що відносна пропорція пікселів  $k$ -го класу в опорному зображенні відповідає апріорній ймовірності  $P(a_k)$ . Втім часто стається так, що контрольні вибірки є репрезентативними у тому сенсі, що вони добре визначають стохастичні властивості індивідуальних класів, але, взяті разом, вони не репрезентують апріорні ймовірності класів. На жаль, неправильні або суттєво неточні оцінки апріорних ймовірностей можуть призвести до підвищення ймовірності прийняття невірних рішень. В роботі [136] метод Монте-Карло [139] було запропоновано використовувати для прогнозування як



апріорних ймовірностей класів  $P(a_k)$ , так і ймовірностей їх правильного розпізнавання  $P_{kk}$ . Вибірка  $X$  вміщує пікселі  $\{s(i, j)\}_N$  тестового зображення з випадковими просторовими координатами  $(i, j)$ , де індекс у рядку  $i$  та у стовпці  $j$  незалежні та мають рівномірний розподіл:  $i \in [1, r]$ ,  $j \in [1, c]$ ,  $r \times c$  – розмір зображення,  $N$  – розмір вибірки. У такому випадку для розпізнавання використовується не все зображення, а тільки вибірка  $X$ , згенерована методом Монте Карло. Потім оцінюється точність класифікації для пікселів вибірки; в результаті отримують матрицю плутанини, діагональні елементи якої  $P_{kk}^*$  є прогнозами (оцінками) ймовірностей правильного розпізнавання класів. Точність таких прогнозів суттєво залежить від розміру вибірки  $N$  [140]. Чим більше кількість класів  $K$  і чим більше відрізняються їх апріорні ймовірності  $P(a_k)$ , тим більше повинно бути  $N$ . Якщо ймовірності  $P(a_k)$  – величини приблизно одного порядку, тоді мінімальний розмір вибірки можна розрахувати за емпіричною формулою  $N_{min} = 100K$ .

Прогнозована апріорна ймовірність  $k$ -го класу  $P^*(a_k)$  є відносною частотою, з якою класифікатор приймає рішення на користь цього класу під час аналізу вибірки  $X$ . Кожен піксель унікальним чином належить до одного з взаємно виключних класів. В той же час в [136] було зазначено, що прогнози, отримані від ряду зображень фіксованого розміру, застосовувані до урахування помилок, характеризуються більшою конвергенцією до відповідних статистичних оцінок для повного набору даних. Врахування помилок класифікації відноситься до гнучкої (fuzzy) класифікації, для якої кожен піксель (враховуючи неоднозначність рішення) може належати до кількох різних класів [141]. У відповідності до цієї неоднозначності можна перерахувати кількість пікселів для кожного класу у вибірці  $X$ .

Тоді відповідно до матриці плутанини належність кожного пікселя до  $k$ -го класу не є єдиним рішенням, замість цього є набір рішень  $\{\gamma_i\}$ ,  $1 \leq i \leq K$  з умовними ймовірностями  $P(\gamma_i | a_k)$ , тож, використовуючи загальну формулу ймовірностей, можна знайти ймовірність для класу  $a_k$  для випадку, що виходом класифікатора є рішення  $\gamma_i$ :

$$P(a_k | \gamma_i) = P(\gamma_i | a_k) / \sum_{v=1}^K P(\gamma_i | a_v) \quad (9)$$

Ймовірності  $P(a_k | \gamma_i)$  є мірами невизначеності рішень класифікатора.

Таким чином, якщо для “жорсткої” класифікації апріорна ймовірність класу



$a_k$  визначається відносною частотою рішень  $\gamma_k$  на користь тільки цього класу (тобто  $P(a_k) = P(\gamma_k)$ ), то при гнучкій (fuzzy) класифікації приймається до уваги той факт, що кожен піксель може бути віднесений до кожного класу з певним ступенем невизначеності. Позначимо прогноз (оцінку) апіорної ймовірності класу  $a_k$ , отриману в результаті “жорсткої” класифікації, як  $P_C^*(a_k)$ . Тоді прогноз на основі гнучкої класифікації можна визначити як:

$$P_F^*(a_k) = \sum_{i=1}^K P(a_k | \gamma_i) P_C^*(a_i). \quad (10)$$

Тож прогноз точності класифікації базується на прогнозуванні ймовірностей правильного розпізнавання класів та апіорних ймовірностей класів. Прогнози базуються на результаті класифікації вибірки обмеженого розміру, яка складається із пікселів з випадковими просторовими координатами. Прогнозоване значення повної ймовірності правильної класифікації розраховується за формулою

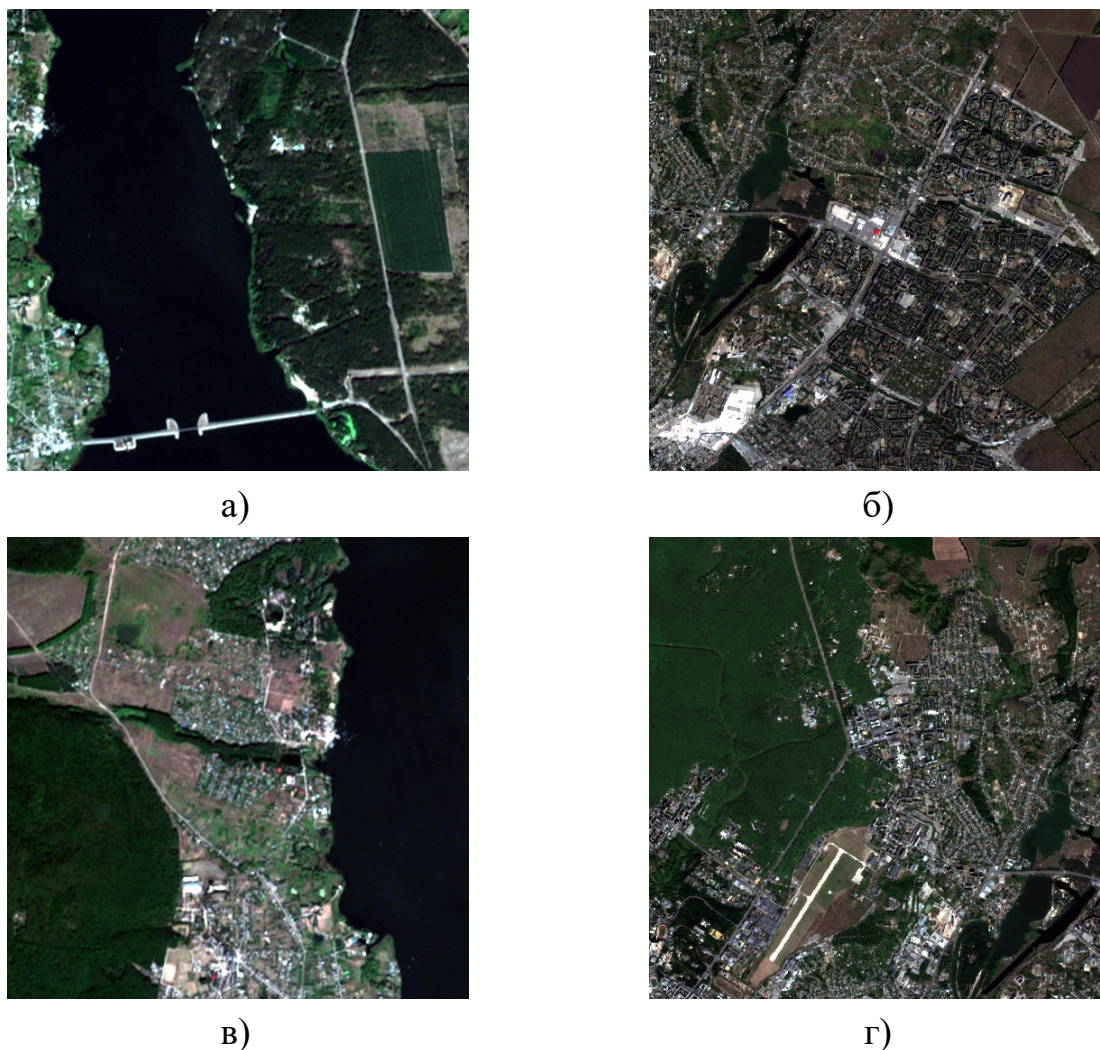
$$P_{total}^* = \sum_{k=1}^K P^*(a_k) P_{kk}^*. \quad (11)$$

Для перевірки методу прогнозування точності класифікації було використано чотири триканальних зображення (рис. 21) розміром  $512 \times 512$  пікселів. Ці зображення були отримані з мультиспектральних даних сенсора супутника Сентінель (Sentinel-2); використовувалися канали, близькі до довжин хвиль R, G, B оптичного діапазону. На знімках представлені фрагменти території Харкова (зображення SS2 і SS4) та його околиць (зображення SS1 і SS3), Україна. Шляхом візуального аналізу та використання карт відповідних регіонів було встановлено, що на кожному зображенні є чотири класи: 1 – Місто, 2 – Вода (водна поверхня), 3 – Рослинність, 4 – Грунт (без рослинності). На основі фактичних даних щодо спостережуваних ділянок на зображеннях були виділені відносно однорідні ділянки, що представляють кожен клас. Отримані області (помічені кольоровими масками) позначалися умовними кольорами, що відповідають певному класу: 1 – жовтий, 2 – синій, 3 – зелений, 4 – чорний. Позначені області були розділені на дві підмножини, які були використані для навчання та оцінки якості класифікатора. Кількість пікселів у навчальних вибірках становила близько  $(4 \dots 20) \times 10^3$ , розмір контрольних вибірок для кожного класу становив  $(7 \dots 50) \times 10^3$  пікселів.

Для аналізу застосовності методу прогнозу також необхідно було визначити еталонні значення для прогнозованих параметрів. Потім порівняння еталонних



значень і прогнозів дозволить оцінити точність прогнозу і зробити висновок про те, наскільки ця методика відповідає поставленим цілям.



а) SS1; б) SS2; в) SS3; г) SS4.

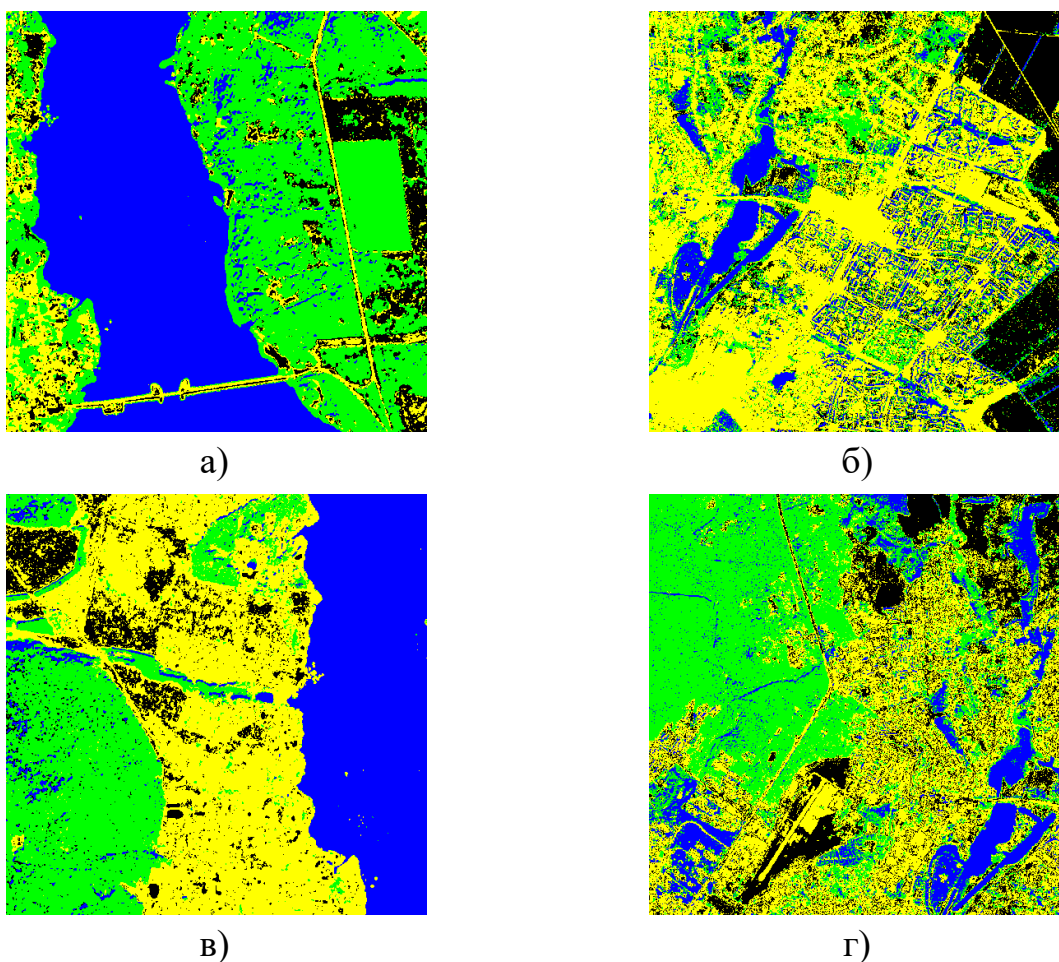
### Рисунок 21 – Реальні триканальні зображення для перевірки

Статистичні оцінки ймовірностей правильного розпізнавання класів  $\hat{P}_{kk}$  та статистичні оцінки апіорних ймовірностей класів  $\hat{P}(a_k)$  були використані як опорні значення для параметрів, що прогнозуються. Щоб розрахувати ці оцінки нами були використані результати попиксельної класифікації тестових зображень на рисунку 21. Отримані мапи класів, що позначені різними кольорами для чотирьох класів наведені на рисунку 22.

Прогнозування характеристик точності класифікації базується на припущенні, що використання вибірки обмеженого розміру, якщо вона правильно сформована, дозволяє оцінити значення аналізуємих параметрів з



допустимими похибками. Прогнози можна вважати адекватними, якщо вибірка достатньою мірою відображає стохастичну природу класів. Як зазначалося вище, вибірка  $X$  формується з пікселів із випадковими індексами в рядках  $i$  стовпцях тестового зображення. Рівномірний розподіл просторових координат дозволяє очікувати, що при достатньому розмірі вибірки  $N$  вона повинна включати об'єкти всіх класів, а відносна частота появи у вибірці об'єкта  $k$ -го класу буде ефективною статистикою для  $P(a_k)$ . Для визначення необхідного розміру  $N$  випадкової вибірки  $X$  було проведено серію експериментів, у яких розмір вибірки змінювався в діапазоні  $N = 400, 600, \dots, 1200$ . Для кожного значення  $N$  кількість прогонів генераторів випадкових просторових координат  $i$  та  $j$  становило 20. Отримана при кожному прогоні вибірка, що складалася з  $N$  пікселів із випадковими координатами, надходила на вхід класифікатора.



а) SS1; б)SS2; в)SS3; г)SS4.

### Рисунок 22 – Отримані мапи класів

На основі результатів класифікації були знайдені прогнози ймовірностей правильної класифікації класів  $P_{kk}^*$ , що були осереднені за ансамблем реалізацій.



Потім нами були розраховані середні  $\delta_{mean}$  (по класам) та максимальні  $\delta_{max}$  відносні похибки (використовуючи статистичні оцінки  $\hat{P}_{kk}^*$ , отримані по результатах розпізнавання опорного зображення). Залежності середньої та максимальної відносної похибок прогнозованих ймовірностей правильного розпізнавання класів залежно від розміру вибірки  $N$  для чотирьох тестових зображень наведені відповідно на рисунках 23 та 24.

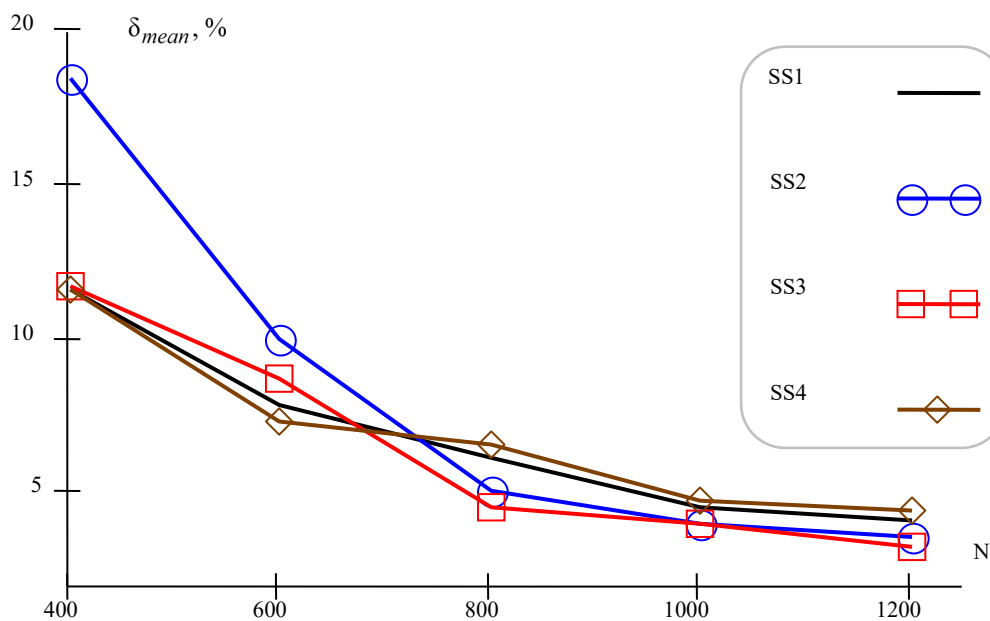


Рисунок 23 – Залежності середньої відносної похибки прогнозу  $P_{kk}^*$  від розміру вибірки

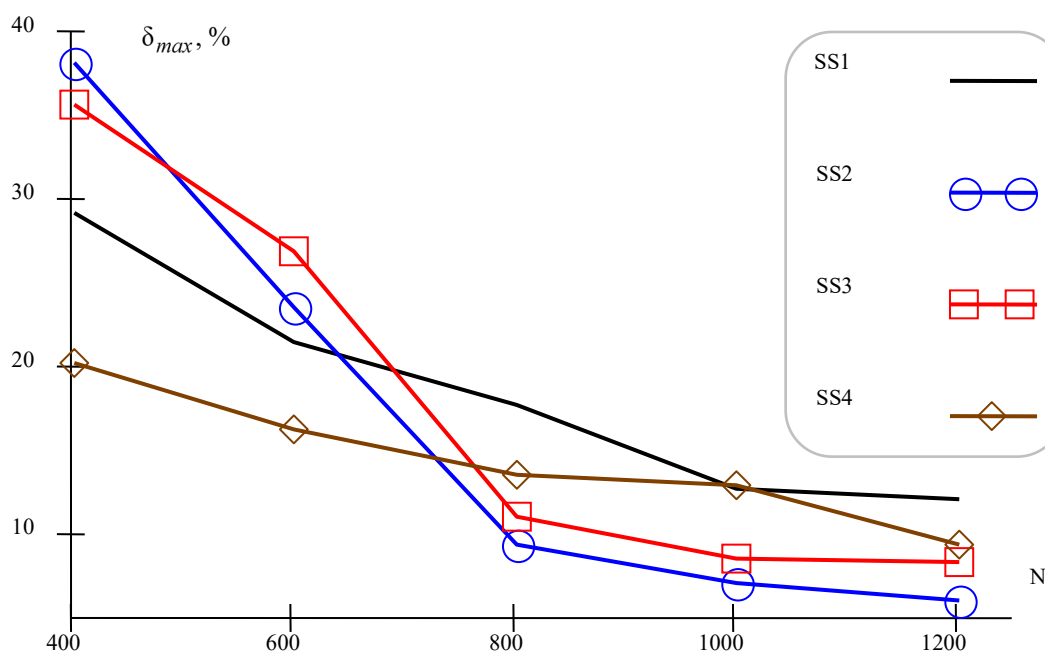


Рисунок 24 – Залежності максимальних відносних похибок прогнозування  $P_{kk}^*$  від розміру вибірки

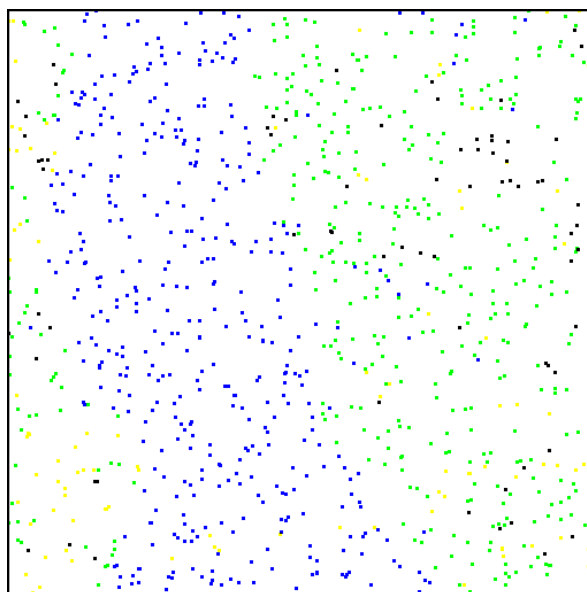




Ці залежності показують, що середня похибка прогнозу зменшується майже експоненціально зі збільшенням  $N$  і не перевищує 5% для  $N = 1000$ ; однак із збільшенням  $N$  до 1200 середня і максимальна похибки змінюються несуттєво. Результати формування вибірки випадкових пікселів на зображенні SS1 наведені на рисунку 25 (пікселі, що входять до вибірки, позначені білим кольором). На рисунку 26 показано результати класифікації для цих вибірових пікселів. Порівняння статистичних оцінок і прогнозів для ймовірностей правильного розпізнавання класів, а також сумарної ймовірності правильного розпізнавання (за умови однакових апіорних ймовірностей класів) наведено в таблицях 2.2 – 2.5.



**Рисунок 25 – Случайним чином вибрані пікселі в зображенні SS1 (позначені білим кольором)**



**Рисунок 26 – Результати класифікації вибраних вікселів в зображенні SS1**

Як впливає з наведених вище результатів, незважаючи на те, що для деяких



класів похибка прогнозування може сягати до 13%, прогноз загальної ймовірності правильного розпізнавання за класифікацією випадкової вибірки розміром  $N = 1000$  відрізняється від її оцінки згідно класифікації еталонного зображення не більше, ніж на 5%. Така точність цілком прийнятна для прогнозування ступеня розрізнення об'єктів із заданого набору класів на зображенні, що розглядається.

**Таблиця 9 – Ймовірності правильної класифікації класів та загалом для SS1**

Ймовірності правильної класифікації класів					$P_{total}$
Клас	Місто	Вода	Рослинність	Грунт	
Оцінка	0,706	0,997	0,88	0,846	0,857
Прогноз	0,682	0,992	0,865	0,739	0,819
Відносна похибка,%	3,40	0,50	1,70	12,65	4,43

**Таблиця 10 – Ймовірності правильної класифікації класів та загалом для SS2**

Ймовірності правильної класифікації класів					$P_{total}$
Клас	Місто	Вода	Рослинність	Грунт	
Оцінка	0,917	0,923	0,691	0,856	0,847
Прогноз	0,911	0,966	0,719	0,915	0,878
Відносна похибка,%	0,65	4,66	4,05	6,89	3,66

**Таблиця 11 – Ймовірності правильної класифікації класів та загалом для SS3**

Ймовірності правильної класифікації класів					$P_{total}$
Клас	Місто	Вода	Рослинність	Грунт	
Оцінка	0,912	0,995	0,895	0,706	0,877
Прогноз	0,958	0,992	0,871	0,647	0,867
Відносна похибка,%	5,04	0,30	2,68	8,36	1,14

**Таблиця 12 – Ймовірності правильної класифікації класів та загалом для SS4**

Ймовірності правильної класифікації класів					$P_{total}$
Клас	Місто	Вода	Рослинність	Грунт	
Оцінка	0,645	0,86	0,943	0,795	0,811
Прогноз	0,643	0,75	0,968	0,769	0,782
Відносна похибка,%	0,31	12,79	2,65	3,27	3,58



Розглянемо тепер результати прогнозування апіорних ймовірностей класів. Такі прогнози можуть виконуватись як на основі жорсткої, так і гнучкої класифікації пікселів вибірки  $X$ . В другому випадку треба спочатку знайти ймовірності  $P(a_k | \gamma_i)$ , які можна трактувати як міри належності заданого пікселя різним класам. Для таких обчислень використовують матриці плутанини, які можна отримати з результатів розпізнавання для вибірки  $X$ . Як приклад в таблиці 13 наведені первинні дані – матриця плутанини для вибірки випадкових пікселів зображення SS1, а результати розрахунків матриці умовних ймовірностей  $P(a_k | \gamma_i)$  наведені в таблиці 14.

Результати оцінювання та прогнозування для апіорних ймовірностей класів, а також для сумарної ймовірності правильної класифікації наведені в таблицях 15 – 18. Використовуються наступні позначення:  $\hat{P}_C(a_k)$  - оцінки для жорсткої класифікації опорного зображення;  $\hat{P}_F(a_k)$  - оцінки для гнучкої класифікації опорного зображення. Прогнози для  $N = 1000$  позначені як  $P_C^*(a_k)$  та  $P_F^*(a_k)$  відповідно для жорсткої та гнучкої класифікації. Символ  $\delta$  позначає відносну похибку прогнозування ймовірності  $P_{total}$ .

**Таблиця 13 – Матриця плутанини для зображення SS1**

Клас	Ймовірності рішень, $P(\gamma_i   a_k)$			
	Місто	Вода	Рослинність	Грунт
Місто	0,682	0	0,045	0,273
Вода	0	0,992	0,008	0
Рослинність	0,021	0,071	0,865	0,043
Грунт	0,217	0	0,043	0,739

**Таблиця 14 – Матриця ймовірностей відповідності рішень класам для зображення SS1**

Клас	Ймовірності рішень, $P(a_k   \gamma_i)$			
	Місто	Вода	Рослинність	Грунт
Місто	0,682	0	0,045	0,273
Вода	0	0,992	0,008	0
Рослинність	0,021	0,071	0,865	0,043
Грунт	0,217	0	0,043	0,739



**Таблиця 15 – Априорна ймовірність класів та загальна ймовірність  
правильної класифікації для зображення SS1**

Клас	Місто	Вода	Рослинність	Ґрунт	$P_{total}$	$\delta, \%$
$\hat{P}_c(a_k)$	0,096	0,411	0,396	0,096	0,908	1,87
$P_c^*(a_k)$	0,098	0,419	0,406	0,077	0,891	
$\hat{P}_F(a_k)$	0,122	0,391	0,388	0,099	0,901	2,0
$P_F^*(a_k)$	0,112	0,394	0,399	0,095	0,883	

**Таблиця 16 – Априорна ймовірність класів та загальна ймовірність  
правильної класифікації для зображення SS2**

Клас	Місто	Вода	Рослинність	Ґрунт	$P_{total}$	$\delta, \%$
$\hat{P}_c(a_k)$	0,512	0,123	0,154	0,211	0,87	2,41
$P_c^*(a_k)$	0,527	0,121	0,141	0,211	0,891	
$\hat{P}_F(a_k)$	0,429	0,122	0,211	0,238	0,855	2,81
$P_F^*(a_k)$	0,479	0,114	0,202	0,204	0,879	

**Таблиця 17 – Априорна ймовірність класів та загальна ймовірність  
правильної класифікації для зображення SS3**

Клас	Місто	Вода	Рослинність	Ґрунт	$P_{total}$	$\delta, \%$
$\hat{P}_c(a_k)$	0,383	0,286	0,242	0,089	0,911	1,21
$P_c^*(a_k)$	0,388	0,291	0,245	0,075	0,922	
$\hat{P}_F(a_k)$	0,294	0,272	0,266	0,168	0,894	1,45
$P_F^*(a_k)$	0,304	0,348	0,207	0,141	0,907	

**Таблиця 18 – Априорна ймовірність класів та загальна ймовірність  
правильної класифікації для зображення SS4**

Клас	Місто	Вода	Рослинність	Ґрунт	$P_{total}$	$\delta, \%$
$\hat{P}_c(a_k)$	0,333	0,121	0,364	0,182	0,807	0,99
$P_c^*(a_k)$	0,313	0,118	0,36	0,209	0,799	
$\hat{P}_F(a_k)$	0,326	0,155	0,304	0,215	0,801	1,75
$P_F^*(a_k)$	0,288	0,188	0,292	0,231	0,787	

Як видно з представлених результатів, прогнози сумарної ймовірності



правильного розпізнавання відрізняються від статистичних оцінок не більше, ніж на 3%. Можна відзначити, що оцінки та прогнози для нечіткої класифікації дають дещо нижчі значення  $P_{total}$ , ніж для «жорсткої» класифікації, але ці відмінності є незначними і в гіршому випадку не перевищують 2%.

Порівняння оцінок і прогнозів для ймовірності  $P_{total}$  відповідними значеннями, наведеними в таблицях 9 – 12 (ці значення отримано для рівних між собою апіорних ймовірностей класів), показує, що використання оцінок (прогнозів) для  $P(a_k)$  може призвести як до збільшення, так і до зменшення прогнозованих значень повної ймовірності правильного розпізнавання залежно від особливостей розподілу апіорних ймовірностей класів.

Таким чином, проблема аналізу якості вирішальних правил тісно пов'язана з роллю апіорної інформації. Оптимальна процедура прийняття рішень повинна забезпечувати мінімум середнього ризику (мінімізувати вартість узагальнених втрат, безпосередньо пов'язаних з ймовірностями помилок розпізнавання). Однак для оцінки середнього ризику необхідна повна апіорна інформація про класи. Реальні проблеми розпізнавання характеризуються наявністю значної невизначеності; зокрема, апіорні ймовірності появи об'єктів з кожного класу невідомі. Оцінки бажаних критеріїв можна отримати шляхом відповідної обробки навчальних і еталонних зображень, які являють собою великі масиви даних. На відміну від статистичних оцінок, які базуються на аналізі всього обсягу наявних даних, прогнози можна отримати на основі значно меншої кількості даних, що враховує неповноту та неточність наявної інформації.

У запропонованому методі прогнози критеріїв точності визначаються з вибірки пікселів, яка формується методом Монте-Карло. Цей метод перевірявся на чотирьох супутникових зображеннях розміром  $512 \times 512$  пікселів. Обсяг випадкової вибірки для прогнозів був зафіксований на рівні  $N = 1000$ . Розміри контрольних вибірок (загальна кількість позначених пікселів на еталонних зображеннях) для статистичних оцінок (еталонних значень критеріїв) варіювалися в діапазоні  $(60 \dots 160) \times 10^3$  пікселів. Результати порівняння прогнозів і статистичних оцінок критеріїв точності показують, що запропонований метод дозволяє отримати достатньо точні прогнози при значному скороченні аналізованих даних. Таким чином, отримані прогнози сумарної ймовірності правильного розпізнавання відрізняються від її статистичних оцінок не більш, ніж на 3%. Слід також зазначити, що достовірність прогнозів суттєво залежить



від репрезентативності та розміру використовуваних вибірок, тому особливу увагу слід приділяти вдосконаленню методів їх отримання.

## **Висновки**

Головну увагу в цьому розділі нами було приділено розробці та тестуванню методів обробки даних багатоканального дистанційного зондування на основі нейронних мереж. На основі проведеного аналізу можна зробити наступні висновки:

а) НМ дозволяють розробляти та оптимізувати так звані комбіновані метрики візуальної якості як для оцінювання якості з еталоном, так і без еталону, характеристики комбінованих метрик виявляються суттєво кращими, ніж характеристики найкращої з елементарних метрик навіть якщо кількість елементарних метрик є невеликою;

б) нейромережі можна використовувати для прогнозування характеристик стиснення з втратами, при цьому точність суттєво зростає навіть якщо кількість вхідних параметрів дорівнює три чи чотири, де один з параметрів керує стисненням, а інші описують складність зображення; також можна прогнозувати існування оптимальної робочої точки у випадку стиснення з втратами зображень, що спотворені шумом;

в) для випадку застосування нейромереж або інших класифікаторів до триканальних зображень можливо прогнозувати характеристики класифікації (наприклад, загальну ймовірність правильної класифікації чи ймовірності правильної класифікації для різних класів, використовуючи для цього вибірку даних невеликого розміру, наприклад 1000 пікселів.

Останні результати наших досліджень опубліковані у [142-144].