Національний аерокосмічний університет «Харківський авіаційний інститут»
Міністерство освіти і науки України

Національний аерокосмічний університет «Харківський авіаційний інститут»
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах рукопису

АЛЬ-СЕНАЙХ РАЕД ЖИХАДОВИЧ

УДК 004.8:528.8:621.396.96

ДИСЕРТАЦІЯ

МЕТОДИ АДАПТИВНОЇ ФІЛЬТРАЦІЇ СПЕКЛ-ШУМУ ТА АПРІОРНОГО ОЦІНЮВАННЯ ЯКОСТІ РАДІОЛОКАЦІЙНИХ ЗОБРАЖЕНЬ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ

172 Телекомунікації та радіотехніка
(шифр і назва спеціальності)

17 Електроніка та телекомунікації
(галузь знань)

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Р. Ж. Аль-Сенайх
(підпис, ініціали та прізвище здобувача)

Науковий керівник
Рубель Олексій Сергійович, кандидат технічних наук, доцент

Харків – 2026

АНОТАЦІЯ

Аль-Сенайх Р. Ж. Методи адаптивної фільтрації спекл-шуму та апріорного оцінювання якості радіолокаційних зображень на основі нейронних мереж. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії з галузі знань 17 «Електроніка та телекомунікації» за спеціальністю 172 «Телекомунікації та радіотехніка». – Національний аерокосмічний університет «Харківський авіаційний інститут», Харків, 2026.

Дисертаційну роботу присвячено вирішенню **актуальної науково-прикладної задачі** підвищення якості та автоматизації процесу адаптивної фільтрації радіолокаційних зображень із синтезованою апертурою (РСА) шляхом розроблення методів апріорного прогнозування показників якості, прямої класифікації параметрів класичних фільтрів і гібридних нейромережевих архітектур з просторово-адаптивною обробкою. **Об'єктом дослідження є** процес придушення мультиплікативного спекл-шуму в зображеннях РСА; **предметом дослідження є** методи апріорного прогнозування мір якості, вибору параметрів фільтрації та адаптивної фільтрації спекл-шуму на РСА-зображеннях на основі нейронних мереж..

Мета дослідження полягає у підвищенні якості та автоматизації фільтрації РСА-зображень шляхом розробки методів адаптивної фільтрації спекл-шуму та апріорного прогнозування метрик і параметрів на основі трансферного навчання. Відповідно до мети поставлено завдання, структурно пов'язані з розділами дисертації: аналітичний огляд і постановка задачі (розділ 1), формування напівсинтетичних даних та методи прогнозування якості і параметрів фільтрації (розділ 2), розроблення гібридного методу VGG16-FrostNet (розділ 3), експериментальна верифікація, порівняльний аналіз і визначення напрямів застосування результатів (розділ 4).

У роботі використано методи математичного моделювання та цифрової обробки сигналів для опису спекл-шуму й диференційованої моделі фільтра

Фроста; методи машинного та глибокого навчання для класифікації параметрів фільтрації й апріорного прогнозування метрик якості; механізми трансферного навчання та просторово-канальної уваги для попіксельної адаптації; метрики PSNR, WSNR, SSIM, MS-SSIM, FSIM, HaarPSI, EPI, FOM та ENL для кількісної верифікації результатів.

У першому розділі систематизовано класичні локально-статистичні фільтри, нейромережеві методи фільтрації спекл-шуму та парадигми трансферного навчання. У другому розділі розроблено конвеєр генерації напівсинтетичних навчальних даних, адаптовано DenseNet-121 для прогнозування абсолютних значень п'яти метрик якості після застосування шести класичних фільтрів (середній $R^2 = 0{,}9732$, максимум $R^2 = 0{,}9972$ для Gamma MAP) та запропоновано MobileNetV2-класифікатор оптимального розміру вікна фільтра Лі (Accuracy 87,17 %, Balanced Accuracy 88,94 %, Top-2 Accuracy понад 99 %).

У третьому розділі створено VGG16-FrostNet: класичний фільтр Фроста подано як диференційований шар нейронної мережі, а модуль CBAM прогнозує просторово-змінну карту коефіцієнтів демпфування A_{map}. На тестовому наборі з 382 зображень 1024×1024 пікселів метод досяг PSNR = 34,18 дБ, SSIM = 0,9700 ± 0,0152, EPI = 0,8903 ± 0,0338, FOM = 0,8340; під час навчання використовувалися випадкові кропи 128×128, а час прямого проходу — близько 40,66 мс для патча 1024×1024 на NVIDIA RTX 4090. У четвертому розділі обґрунтовано переваги парадигми Hard Coupling та окреслено перспективи крос-сенсорної обробки.

Наукова новизна полягає в **запропонованому методі** прогнозування абсолютних мір якості РСА-зображень без еталонного зображення для оцінювання конкретного вхідного знімка на етапі інференсу і ручного розрахунку ознак, **удосконаленні методу** апріорного вибору параметрів фільтрації як прямої багатокласової класифікації, а також у **подальшому розвитку** адаптивної фільтрації спекл-шуму через поєднання VGG16, CBAM і математичної моделі фільтра Фроста. На відміну від підходів, що оцінюють відносне покращення після фільтрації, запропоновані моделі формують прогноз якості або керуючого

параметра до запуску фільтра. Положення новизни співвіднесено з результатами розділів 2 і 3 та узагальнено у висновках.

Практичне значення одержаних результатів полягає в тому, що запропоновані моделі забезпечують автоматизований вибір оптимальних параметрів фільтрації спекл-шуму РСА-зображень без повторних прогонів фільтрів, що знижує обчислювальні витрати, усуває ризики помилкових налаштувань та оптимізує конвеєри обробки даних дистанційного зондування. Результати впроваджено у виробничу практику компанії P-Product, Inc. у вигляді програмних прототипів модулів апріорного прогнозування якості, класифікації параметрів фільтра Лі та адаптивної фільтрації спекл-шуму VGG16-FrostNet. За матеріалами дисертації опубліковано 6 наукових праць: 3 статті у фахових виданнях України (2 категорії «Б», 1 категорії «А», індексується у Scopus) та 3 публікації у матеріалах міжнародних конференцій; відповідність публікацій розділам — у Додатку А.

Ключові слова: РСА-зображення, спекл-шум, придушення шуму, фільтр Фроста, фільтр Лі, VGG16, MobileNetV2, CBAM, трансферне навчання, глибоке навчання, Sentinel-1, прогнозування якості, деспеклізація.

SUMMARY

Al-Senaikh R. Z. Methods of Adaptive Speckle Filtering and A Priori Quality Assessment of Radar Images Based on Neural Networks. – A qualifying scientific thesis submitted as a manuscript.

The thesis for a degree of Doctor of Philosophy (PhD) in the field of knowledge 17 «Electronics and Telecommunications» in specialty 172 «Telecommunications and Radio Engineering». – National Aerospace University «Kharkiv Aviation Institute», Kharkiv, 2026.

The thesis addresses the **important scientific and practical task** of improving the quality and automation of adaptive filtering for Synthetic Aperture Radar (SAR) images through the development of methods for a priori quality prediction, direct classification of classical filter parameters, and hybrid neural network architectures with spatially-adaptive processing. **The object of the study is** the process of suppressing multiplicative speckle noise in SAR images; **the subject of the study is** methods for a priori prediction of quality measures, selection of filtering parameters, and adaptive despeckling of SAR images based on neural networks.

The aim of the thesis is to improve the quality and automation of SAR image filtering by developing methods of adaptive despeckling and a priori prediction of metrics and parameters based on transfer learning. In accordance with this aim, the tasks are structurally linked to the thesis chapters: analytical review and problem statement (Chapter 1), formation of semi-synthetic data and methods for quality and filtering-parameter prediction (Chapter 2), development of the hybrid VGG16-FrostNet method (Chapter 3), experimental verification, comparative analysis, and determination of application directions for the obtained results (Chapter 4).

The study uses methods of mathematical modelling and digital signal processing to describe speckle noise and the differentiable Frost filter model; machine and deep learning methods for classification of filtering parameters and a priori prediction of quality metrics; transfer learning and spatial-channel attention mechanisms for pixel-

wise adaptation; and PSNR, WSNR, SSIM, MS-SSIM, FSIM, HaarPSI, EPI, FOM, and ENL metrics for quantitative verification.

The first chapter systematises classical filters, neural-network despeckling methods, transfer-learning paradigms and objective quality metrics. The second chapter develops a semi-synthetic data-generation pipeline, adapts DenseNet-121 for prediction of five quality metrics after six classical filters (average $R^2 = 0.9732$; maximum $R^2 = 0.9972$ for Gamma MAP), and proposes a MobileNetV2 classifier of the optimal Lee filter window size (87.17 % Accuracy, 88.94 % Balanced Accuracy, over 99 % Top-2 accuracy).

The third chapter develops VGG16-FrostNet, where the Frost filter is reformulated as a differentiable neural layer and CBAM predicts the damping coefficient map A_{map} per pixel. On the 382-image test set of 1024×1024 pixels the method achieves PSNR = 34.18 dB, SSIM = 0.9700 ± 0.0152, EPI = 0.8903 ± 0.0338 and FOM = 0.8340; during training, random 128×128 crops were used, and inference time is about 40.66 ms per 1024×1024 patch on NVIDIA RTX 4090. The fourth chapter substantiates the Hard Coupling paradigm and outlines prospects for cross-sensor processing.

The scientific novelty consists in proposing transfer-learning prediction of absolute SAR quality measures without a reference image for evaluating a specific input image at the inference stage or manual feature engineering, **improving** a priori selection of Lee filter parameters as direct multiclass classification, and **further developing** adaptive SAR despeckling through the analytical combination of VGG16, CBAM and the Frost model. Unlike approaches that estimate only relative quality improvement after filtering, the proposed models predict the expected quality or control parameter before filter execution, which is important for automated large-scale satellite-data processing pipelines. The novelty statements are correlated with the results of Chapters 2 and 3 and summarised in the conclusions.

The practical significance lies in the fact that the proposed models enable automated selection of optimal despeckling parameters for SAR images without

repeated filter runs, reducing computational costs and the risks of incorrect parameter settings in remote sensing data processing pipelines. The results are introduced into the production practice of P-Product, Inc. as software prototypes of modules for a priori quality prediction, Lee filter parameter classification, and VGG16-FrostNet adaptive despeckling. According to the thesis material, 6 scientific publications have been issued: 3 articles in Ukrainian professional journals (2 in category B and 1 in category A, the latter indexed in Scopus) and 3 publications in proceedings of international scientific and practical conferences; the correspondence between the publications and thesis chapters is provided in Appendix A.

Keywords: SAR images, speckle noise, noise suppression, Frost filter, Lee filter, VGG16, MobileNetV2, CBAM, transfer learning, deep learning, Sentinel-1, quality prediction, despeckling

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Al-Senaikh R., Rubel O. Predicting Filtered Image Quality Using Transfer Learning on Sentinel-1 Speckle Noise with DenseNet-121 // Ukrainian Journal of Remote Sensing. — 2025. — Vol. 12, No. 4. — P. 4–15. — DOI: 10.36023/ujrs.2025.12.4.293.

2. Al-Senaikh R., Rubel O. Predicting Optimal Lee Filter Window Size for Sentinel-1 SAR Images Using Transfer Learning on MobileNetV2 // Вісник Хмельницького національного університету. Серія: Технічні науки. — 2026. — Т. 361, № 1. — С. 26–36. — DOI: 10.31891/2307-5732-2026-361-3.

3. Al-Senaikh R., Rubel O. Development of a hybrid method VGG16-FrostNet for adaptive despeckling of synthetic aperture radar (SAR) images using attention mechanism and differentiable Frost filter // Technology audit and production reserves. — 2026. — Т. 2, № 2(88). — С. 25–33. — DOI: 10.15587/2706-5448.2026.358316.

4. Аль-Сенайх Р. Ж., Рубель О. С. Мінімізація ризиків неправильного встановлення параметрів фільтрів для покращення якості РСА-зображень // Information, Functional and Cybersecurity: матеріали міжнар. студ. конф. з кібербезпеки (SCIFiC 2025) (м. Харків, 27–28 листопада 2025 р.). — Харків: НАУ «ХАІ», 2025. — С. 124–125. — DOI: 10.13140/RG.2.2.32148.16001.

5. Al-Senaikh R. Z., Rubel O. S. A hybrid approach to adaptive SAR despeckling using VGG16-FrostNet and a HaarPSI-based Frost baseline // Organization of scientific research in modern conditions '2026: SWorld-US Conference proceedings (Seattle, WA, USA, March 21, 2026). — Seattle: ProConference, 2026. — Vol. 1, Issue usc36-00. — P. 44–51. — DOI: 10.30888/2709-2267.2026-36-00-029.

6. Аль-Сенайх Р. Ж., Рубель О. С. Порівняльний аналіз парадигм інтеграції класичних адаптивних фільтрів у нейромережеві архітектури для деспеклізації радіолокаційних зображень // Перспективні мережні та комп'ютерні технології (ПерСиК 2026): матеріали 17-ї міжнар. студ. наук.-техн. конф. (23 квітня 2026 р.).

ЗМІСТ

АНОТАЦІЯ 2

SUMMARY 5

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ 8

ЗМІСТ 9

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ 17

ВСТУП 19

РОЗДІЛ 1. АНАЛІТИЧНИЙ ОГЛЯД МЕТОДІВ ПРИДУШЕННЯ СПЕКЛ-ШУМУ В РСА-ЗОБРАЖЕННЯХ, ОЦІНЮВАННЯ ЯКОСТІ ТА ПІДХОДІВ ТРАНСФЕРНОГО НАВЧАННЯ 28

1.1 Фізична природа та моделі формування спекл-шуму в РСА-зображеннях ... 28

1.1.1 Базові принципи технології РСА 28

1.1.2 Фізика виникнення спекл-шуму 30

1.1.3 Математична і статистична модель (Single-look) 31

1.1.4 Мультиплікативна природа шуму 32

1.1.5 Багатопоглядова обробка (Multi-looking) та зміна статистики 34

1.1.6 Вплив спекл-шуму на алгоритми комп'ютерного зору (Проблематика) 36

1.2 Класичні методи локальної статистики 37

1.2.1 Локальний коефіцієнт варіації та MMSE-фільтри (Лі, Куана) 38

1.2.2 Фільтр Фроста (експоненційна вагова функція) 41

1.2.3 MAP-фільтри (Gamma MAP) 43

1.2.4 Проблематика фіксованого вікна та ручного налаштування 44

1.3 Базові нейромережеві методи фільтрації спекл-шуму 46

1.3.1 Перехід до Data-Driven підходів (CNN) та Residual Learning 46

1.3.2 Архітектури DnCNN та U-Net у задачах фільтрації 48

1.3.3 Напівкеровані методи (Self-supervised, Noise2Noise) через відсутність еталонів 50

1.3.4 Обмеження чисто "чорноящикових" моделей 52

1.4 Метрики та підходи до оцінювання якості фільтрації 54

1.4.1 Базові піксельні метрики та їх обмеження (MSE, PSNR) 54

1.4.2 Багатомасштабні та структурні індекси (SSIM, MS-SSIM, FSIM) 56

1.4.3 Перцептивні метрики (HaarPSI, MDSI) та індекси контурів (EPI, FOM) 58

1.4.4 Специфічні радарні показники (ENL) та оцінювання без еталону (No-Reference) 60

1.5 Трансферне навчання та його застосування 62

1.5.1 Концепція Transfer Learning та Domain Adaptation 62

1.5.2 Попереднє навчання на оптичних даних (ImageNet) для задач PCA 65

1.5.3 Подолання доменного зсуву (Domain Shift) 66

1.5.4 Стратегії Fine-tuning 68

1.6 Гібридна парадигма адаптивної фільтрації спекл-шуму 70

1.6.1 Синергія математичних моделей та нейромереж (Model-based Deep Learning) 70

1.6.2 Глибоке розгортання алгоритмів (Deep Unrolling) 72

1.6.3 Прогнозування параметрів класичних фільтрів за допомогою мереж 74

1.6.4 Обґрунтування розробки VGG16-FrostNet 76

1.7 Висновки до розділу 1 77

РОЗДІЛ 2. МЕТОДИ АПРІОРНОГО ПРОГНОЗУВАННЯ ПОКАЗНИКІВ ЯКОСТІ ПІСЛЯ ФІЛЬТРАЦІЇ ТА ПАРАМЕТРІВ КЛАСИЧНИХ ФІЛЬТРІВ ДЛЯ ДАНИХ SENTINEL-1 НА БАЗІ ГЛИБОКИХ НЕЙРОМЕРЕЖ....79

2.1 Формалізація задачі апріорного прогнозування....79

2.1.1 Традиційний підхід до оцінювання якості: post-filtering та потреба в еталонах....79

2.1.2 Концепція апріорного оцінювання: переваги та обчислювальна ефективність....80

2.1.3 Математична постановка задачі прогнозування абсолютних метрик якості....80

2.1.4 Обґрунтування вибору цільових метрик (PSNR, WSNR, SSIM, MS-SSIM, FSIM)....81

2.1.5 Формалізація задачі вибору параметрів фільтрації (на прикладі вікна Лі)....82

2.1.6 Постановка задачі як багатокласової класифікації замість регресії....82

2.1.7 Роль глибоких ознак у виявленні текстурних властивостей....83

2.1.8 Обґрунтування застосування трансферного навчання для РСА-зображень....83

2.1.9 Загальна архітектура пропонованого рішення....84

2.2 Генерація напівсинтетичного навчального корпусу....84

2.2.1 Проблема відсутності парних РСА-даних для навчання....84

2.2.2 Вибір оптичних зображень Sentinel-2 (NIR діапазон) як базису....85

2.2.3 Алгоритм гістограмного узгодження оптичних даних із РСА-розподілами....85

2.2.4 Математична модель динамічного додавання мультиплікативного спекл-шуму....86

2.2.5 Процедура нарізки на фрагменти (патчі) та аугментація даних 89

2.2.6 Конвеєр обчислення еталонних цільових змінних (label generation) 89

2.2.7 Аналіз репрезентативності та статистичних властивостей створеного датасету 90

2.3 Прогнозування метрик якості на основі адаптованої DenseNet-121 91

2.3.1 Особливості та переваги архітектури DenseNet-121 91

2.3.2 Адаптація вхідних шарів для одноканальних PCA-зображень.....92

2.3.3 Модифікація класифікаційної голови для прогнозування 5 метрик 94

2.3.4 Стратегія та гіперпараметри повного донавчання 94

2.4 Прогнозування оптимального розміру вікна фільтра Лі на базі MobileNetV2 95

2.4.1 Вплив розміру вікна на компроміс між згладжуванням та збереженням деталей 95

2.4.2 Аналіз недоліків двоетапних регресійних моделей вибору вікна 96

2.4.3 Архітектура MobileNetV2: баланс між точністю та обчислювальною ефективністю 96

2.4.4 Формування міток класів (розмірів вікна) шляхом повного перебору ...97

2.4.5 Адаптація MobileNetV2 до задачі 5-класової класифікації 98

2.4.6 Стратегія навчання та боротьба з дисбалансом класів.....99

2.5 Експериментальні результати та їх обговорення 99

2.5.1 Постановка експерименту та статистичні критерії оцінювання 99

2.5.2 Аналіз точності прогнозування (R^2) для різних фільтрів.....101

2.5.3 Оцінка стабільності прогнозів для різних метрик якості.....105

2.5.4 Аналіз розподілу помилок прогнозування та відсутності перенавчання 107

2.5.5 Візуальне порівняння передбаченої та фактичної якості на тестових знімках 109
2.5.6 Оцінка обчислювальної ефективності апріорного вибору фільтрів 110
2.5.7 Аналіз загальної та збалансованої точності класифікації розміру вікна 110
2.5.8 Дослідження матриці помилок (Confusion Matrix) 113
2.5.9 Порівняння прямої класифікації з двоетапними підходами 115
2.5.10 Візуалізація результатів: карти класифікації на ландшафтах 116
2.5.11 Вплив обраного вікна на збереження просторових деталей 116
2.5.12 Оцінка швидкодії інференсу MobileNetV2 у реальних умовах 117
2.5.13 Синтез результатів: взаємозв'язок метрик та розміру вікна 118
2.5.14 Аналіз обмежень підходів (глобальна vs локальна адаптація) 119
2.5.15 Роль прогнозування в оптимізації конвеєрів супутникових даних 119
2.5.16 Порівняння енергоефективності предикторів та перебору 120
2.5.17 Перспективи інтеграції методів у хмарні платформи ДЗЗ 120
2.5.18 Підсумкова дискусія щодо доцільності апріорного прогнозування 121
2.6 Висновки до розділу 2 121
2.6.1 Підсумки щодо прогнозування метрик 121
2.6.2 Підсумки щодо класифікації розміру вікна 121
РОЗДІЛ 3. ГІБРИДНИЙ МЕТОД АДАПТИВНОЇ ФІЛЬТРАЦІЇ СПЕКЛ-ШУМУ З ІНТЕГРАЦІЄЮ ДИФЕРЕНЦІЙОВАНОГО ФІЛЬТРА ФРОСТА ТА МЕХАНІЗМУ УВАГИ (VGG16-FROSTNET) 123
3.1 Диференційована математична модель фільтра Фроста 123
3.1.1 Обґрунтування необхідності наскрізно-навчальних методів 123
3.1.2 Мультиплікативна модель спекл-шуму для РСА-зображень 124

3.1.3 Класичний фільтр Фроста: принцип роботи та обмеження 125

3.1.4 Формалізація задачі нелінійного відображення з параметрами 126

3.1.5 Проблема числових нестабільностей (underflow) у графі обчислень ... 126

3.1.6 Модифікація математичної моделі: заміна квадрата на першу ступінь 127

3.1.7 Стиснення динамічного діапазону та чисельна стабільність градієнтів 128

3.1.8 Роль просторово-змінного коефіцієнта як компенсатора модифікації 128

3.1.9 Аналітичне рівняння диференційованого фільтра для ковзного вікна 129

3.1.10 Значення підходу для оптимізації через зворотне поширення похибки 130

3.2 Архітектура VGG16-FrostNet 130

3.2.1 Використання переднавченої мережі VGG16 як базової архітектури .. 133

3.2.2 Модуль просторово-канальної уваги (CBAM) 135

3.2.3 Прогнозування карти коефіцієнтів та резидуальна гілка 137

3.3 Експериментальна оцінка гібридного методу 144

3.3.1 Підготовка навчального набору даних (Синтез) 147

3.3.2 Налаштування гіперпараметрів тренування 151

3.3.3 Абляційне дослідження та бенчмаркінг 154

3.3.4 Верифікація на реальних даних Sentinel-1 167

3.4 Практичні аспекти впровадження гібридної архітектури в системи обробки 170

3.4.1 Інтеграція в хмарні платформи та конвеєри попередньої обробки 170

3.4.2 Програмний прототип SAR Denoising Lab (P-Product, Inc.) 171

Висновки до розділу 3 173

РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ, ПОРІВНЯЛЬНИЙ АНАЛІЗ І ЗАСТОСУВАННЯ РЕЗУЛЬТАТІВ У КОНВЕЄРАХ ОБРОБКИ РСА-ДАНИХ 174

4.1 Організація експериментальних досліджень 175

4.1.1 Вибір мультиспектральних даних Sentinel-2 як оптичних еталонів 175

4.1.2 Математичне моделювання мультиплікативного спекл-шуму 175

4.1.3 Базові метрики оцінки: відношення сигнал/шум та структурна подібність 176

4.1.4 Оцінювання перцептивної якості: індекси FSIM та HaarPSI 176

4.1.5 Оцінка збереження контурів: Edge Preservation Index (EPI) 176

4.1.6 Апаратне забезпечення та обчислювальні фреймворки 177

4.1.7 Налаштування гіперпараметрів навчання нейромережевих моделей 177

4.1.8 Стратегія крос-валідації та запобігання витоку даних 178

4.1.9 Формування тестової вибірки для бенчмаркінгу 178

4.1.10 Огляд базових моделей для порівняльного аналізу 179

4.2 Порівняльний аналіз парадигм інтеграції 179

4.2.1 Парадигми слабкого зв'язку та розгортання ітеративних алгоритмів 180

4.2.2 Парадигми жорсткого зв'язку та керованої фільтрації 182

4.2.3 Вибір оптимальної парадигми 184

4.3 Експериментальна оцінка гібридного методу VGG16-FrostNet 186

4.3.1 Умови експерименту для архітектури VGG16-FrostNet 186

4.3.2 Результати оцінювання за метриками PSNR та SSIM 186

4.3.3 Порівняння з оптимізованим бейзлайном за перцептивними метриками 186

4.3.4 Оцінка за допомогою індексу збереження контурів (EPI) 187

4.3.5 Аналіз згенерованої карти коефіцієнтів демпфування 187

4.3.6 Стійкість гібридної моделі до варіацій рівня шуму (ENL) 188

4.3.7 Візуальний аналіз результатів на специфічних ландшафтах 188

4.3.8 Підсумок щодо подолання проблеми ручного налаштування 188

4.4 Мінімізація ризиків за допомогою класифікації оптимального вікна 190

4.4.1 Ризики неправильного вибору параметрів фільтрації для картографування 190

4.4.2 Класифікація оптимального вікна за допомогою MobileNetV2 190

4.5 Висновки до розділу 4 191

ВИСНОВКИ 193

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 195

ДОДАТКИ 218

ДОДАТОК А. СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ ТА ВІДОМОСТІ ПРО АПРОБАЦІЮ 218

ДОДАТОК Б. АКТИ ВПРОВАДЖЕННЯ 220

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ

У дисертаційній роботі застосовано такі скорочення і умовні позначення:

ДЗЗ	дистанційне зондування Землі
РСА	радіолокатор із синтезованою апертурою
Adam	Adaptive Moment Estimation (адаптивний алгоритм оптимізації)
ADMM	Alternating Direction Method of Multipliers
AWGN	Additive White Gaussian Noise (адитивний білий гаусівський шум)
CBAM	Convolutional Block Attention Module
CNN	Convolutional Neural Network (згорткова нейронна мережа)
DCT	Discrete Cosine Transform (дискретне косинусне перетворення)
DenseNet	Densely Connected Convolutional Network (густозв'язана згорткова нейронна мережа)
DenseNet-121	Densely Connected Convolutional Network, 121 шар
DnCNN	Denoising Convolutional Neural Network (згорткова нейромережа для придушення шуму)
ENL	Equivalent Number of Looks (еквівалентне число поглядів)
EPI	Edge Preservation Index (індекс збереження країв)
FCN	Fully Convolutional Network (повністю згорткова мережа)
FOM	Pratt's Figure of Merit (міра якості контурів Пратта)
FrostNet	нейромережева архітектура на основі диференційованого шара Фроста
FSIM	Feature Similarity Index Measure (індекс подібності за ознаками)
GAN	Generative Adversarial Network (генеративно-змагальна мережа)
GPU	Graphics Processing Unit (графічний процесор)
HaarPSI	Haar Wavelet-based Perceptual Similarity Index
HVS	Human Visual System (зорова система людини)
ImageNet	великомасштабна база даних зображень для навчання згорткових архітектур

IW GRD	Interferometric Wide Swath, Ground Range Detected (режим зйомки та формат продукту Sentinel-1)
MAE	Mean Absolute Error (середня абсолютна похибка)
MAP	Maximum A Posteriori (максимум апостеріорної ймовірності)
MDSI	Mean Deviation Similarity Index
MLP	Multi-Layer Perceptron (багатошаровий персептрон)
MMSE	Minimum Mean Square Error (мінімум середньоквадратичної похибки)
MobileNetV2	Mobile Networks v2 (легка згорткова архітектура)
MSE	Mean Squared Error (середньоквадратична похибка)
MS-SSIM	Multi-Scale SSIM (багатомасштабний індекс структурної подібності)
NIR	Near Infrared (ближній інфрачервоний діапазон Sentinel-2)
PC	Phase Congruency (фазова конгруентність)
PSNR	Peak Signal-to-Noise Ratio (пікове відношення сигнал/шум, дБ)
ResNet	Residual Network (залишкова нейронна мережа)
RMSE	Root Mean Squared Error (середньоквадратична помилка)
SAR	Synthetic Aperture Radar (англ. відповідник РСА)
SRAD	Speckle Reducing Anisotropic Diffusion
SSIM	Structural Similarity Index Measure (індекс структурної подібності)
VGG16	Visual Geometry Group network, 16 шарів
VRAM	Video Random Access Memory (відеопам'ять)
WSNR	Weighted Signal-to-Noise Ratio (зважене відношення сигнал/шум)

ВСТУП

Обґрунтування вибору теми дослідження. За останні роки спостерігається значне зростання інтересу до систем дистанційного зондування Землі (ДЗЗ). Це зумовлено потребою в геопросторовій інформації для моніторингу довкілля, відстеження змін клімату, контролю сільськогосподарських угідь та оцінювання наслідків стихійних лих [1]. У відповідь розвиваються масштабні програми спостереження Землі, зокрема Copernicus [2], що забезпечує відкритий доступ до даних. Особливе місце серед систем ДЗЗ займають радіолокаційні системи із синтезованою апертурою антени (РСА), зокрема супутники Sentinel-1. Порівняно з оптичними сенсорів, вони використовують мікрохвильовий діапазон, формуючи зображення високої роздільної здатності крізь хмарний покрив і незалежно від освітлення [3]. Дані Sentinel-1 у режимі IW GRD широко використовуються в операційних конвеєрах [4]. Зростання обсягів РСА-даних ставить потребу розробки ефективних автоматизованих методів їх обробки.

Попри переваги систем РСА, формування радіолокаційних зображень супроводжується виникненням викривлень, що обмежують ефективність їх подальшого аналізу. Найбільш критичним фактором є наявність інтенсивного шумоподібного явища — спекл-шуму. Його природа зумовлена когерентним характером радіолокаційних сигналів: хвилі, відбиті від безлічі елементарних розсіювачів, інтерферують між собою, створюючи зернисту структуру [5]. Порівняно з адитивним шумом оптичних зображень, спекл має мультиплікативний характер: дисперсія шуму зростає пропорційно рівню сигналу. Крім того, він має негаусову щільність розподілу та просторову кореляцію [6]. Присутність спекл-шуму не лише ускладнює дешифрування зображень операторами, але й значно деградує роботу алгоритмів комп'ютерного зору, знижуючи точність процедур сегментації та класифікації. Тому якісна фільтрація спекл-шуму є обов'язковим етапом попередньої обробки.

Для придушення спекл-шуму традиційно використовуються методи локально-статистичної просторової фільтрації. До найвідоміших класичних

підходів належать адаптивні фільтри Лі [7], Фроста [8], Куана [9] та Gamma MAP [10]. Ці алгоритми базуються на ймовірнісних моделях мультиплікативного шуму та локальній статистиці (середньому та дисперсії) в межах ковзного вікна. Їхня перевага полягає у просторовій адаптивності: вони забезпечують інтенсивне згладжування в однорідних областях і зберігають перепади яскравості на контурах об'єктів. Проте широке застосування цих фільтрів має обмеження — їхня ефективність критично залежить від вибору керуючих параметрів, насамперед розміру вікна та константи демпфування. Некоректний вибір параметрів залишає високий рівень шуму або спричиняє розмиття дрібних деталей. В умовах конвеєрного оброблення масивів РСА-даних ручне підбирання параметрів для кожної сцени є неефективним і неможливим, що актуалізує проблему автоматизації налаштування фільтрів.

Разом з тим, застосування фільтрації не завжди гарантує позитивний ефект для кінцевих завдань користувача. Під час обробки урбанізованих територій або областей із високим рівнем текстурності будь-яке згладжування може призвести до втрати важливої інформації [11]. У таких випадках доцільніше відмовитися від придушення шуму, щоб зберегти цілісність даних і зекономити обчислювальні ресурси [12]. Саме тому в сучасних сервісно-орієнтованих архітектурах ДЗЗ стрімко розвивається перспективний напрямок — апріорне оцінювання потенційної ефективності фільтрації. Це передбачає розрахунок очікуваних показників якості обробки безпосередньо за початковим зашумленим зображенням. Таке прогнозування дозволяє системам динамічно визначати, чи потрібна фільтрація та які параметри забезпечать компроміс між придушенням шуму і збереженням деталей.

У працях, присвячених прогнозуванню ефективності придушення спекл-шуму для РСА-зображень Sentinel-1, обґрунтовано концепцію використання штучних нейронних мереж з локальними статистичними ознаками для оцінювання доцільності фільтрів без перебору варіантів [13], [14]. Подальші роботи цього колективу спрямовані на безпосередній вибір керуючих параметрів

класичних фільтрів методами машинного навчання: прогнозування оптимального розміру вікна для фільтра Лі [15] та вибір параметрів фільтра Фроста за показниками ефективності [16]. Ці дослідження підтверджують можливість автоматизації налаштування фільтрів. Водночас існуючі підходи орієнтовані переважно на експертно розроблені просторово-статистичні ознаки та вирішення задач прогнозування відносного приросту якості. Крім того, вибір параметрів реалізується через багатоетапні регресійні схеми, що призводить до накопичення похибок. Для мікросервісних платформ потрібні більш досконалі методи, здатні прогнозувати абсолютні метрики якості (PSNR, SSIM, FSIM) і забезпечувати багатокласову класифікацію параметрів. Також існуючі моделі адаптують параметри глобально для всієї сцени, ігноруючи потребу в попіксельній адаптації фільтрації.

Паралельно останніми роками спостерігається значний прогрес у використанні методів глибокого та трансферного навчання для наскрізної фільтрації спекл-шуму [17]. Завдяки згортковим нейронним мережам ці моделі демонструють високі показники відновлення зображень і здатність самостійно вилучати складні ознаки [18]. Однак масове впровадження суто нейромережевих фільтрів у критичні системи стримується низкою недоліків. Більшість з них функціонує як неінтерпретовані «чорні ящики», що знижує довіру до результатів. По-друге, чимало стандартних архітектур для обробки оптичних зображень не враховують мультиплікативну фізичну природу спекл-шуму. Унаслідок цього вони можуть генерувати непередбачувані артефакти. Тому найбільш перспективним є створення гібридних архітектур, що гармонійно поєднують математичні моделі просторових фільтрів зі здатністю нейромереж до аналізу просторового контексту та реалізації попіксельної адаптації параметрів.

Отже, у цифровій обробці РСА-даних сформувалася об’єктивна суперечність. Існує нагальна потреба в автоматизації, високій якості та прозорій інтерпретованості методів фільтрації для інтеграції у хмарні конвеєри моніторингу, проте наявний інструментарій має суттєві обмеження: традиційні

адаптивні фільтри вимагають ручного налаштування або базуються на недосконалих процедурах прогнозування, тоді як методи глибокого навчання потерпають від відсутності аналітичних моделей і залишаються «чорними ящиками». У зв’язку з цим, **актуальним науково-прикладним завданням** дисертаційної роботи є розробка прогнозуючих і гібридних методів попередньої обробки РСА-зображень. Завдання передбачає створення методів прогнозування абсолютних мір якості та методів прямої класифікації для вибору параметрів класичних фільтрів на базі трансферного навчання. Не менш важливою складовою є розробка гібридних нейромережевих архітектур, які об'єднують моделі адаптивної фільтрації з механізмами просторово-канальної уваги для автономної попіксельної адаптації обробки.

Об’єкт дослідження – процес придушення мультиплікативного спекл-шуму в зображеннях радіолокатора з синтезованою апертурою (РСА).

Предмет дослідження – методи апріорного прогнозування мір якості, вибору параметрів фільтрації та адаптивної фільтрації спекл-шуму РСА-зображень на основі нейронних мереж.

Мета і завдання дослідження. Мета дослідження полягає у підвищенні якості та автоматизації процесу фільтрації РСА-зображень шляхом розробки методів адаптивної фільтрації спекл-шуму та апріорного прогнозування метрик і параметрів на основі трансферного навчання.

Відповідно до поставленої мети в дисертаційній роботі сформульовано і вирішено такі основні **завдання:**

1. Проаналізувати існуючі методи придушення спекл-шуму та систематизувати парадигми інтеграції класичних фільтрів у нейромережеві архітектури.

2. Розробити метод формування наборів даних (псевдоеталонів) шляхом імітаційного моделювання мультиплікативного спекл-шуму на основі оптичних знімків інфрачервоного діапазону.

3. Розробити методи апріорного прогнозування метрик якості та автоматичного вибору керуючих параметрів фільтрації (на прикладі фільтра Лі) на основі трансферного навчання.

4. Розробити гібридний метод просторово-адаптивної фільтрації спекл-шуму (VGG16-FrostNet), що поєднує диференційовану модель фільтра Фроста з механізмами просторово-канальної уваги.

5. Провести експериментальні дослідження та порівняльний аналіз ефективності запропонованих методів на створених напівсинтетичних масивах даних.

Методи дослідження. Для вирішення поставлених завдань у дисертаційній роботі було використано методи математичного моделювання та цифрової обробки сигналів для побудови моделі формування спекл-шуму і розробки диференційованої моделі фільтра Фроста, методи машинного та глибокого навчання для класифікації параметрів фільтрації й апріорного прогнозування метрик якості зображень, теорію трансферного навчання та механізми просторово-канальної уваги для попіксельної просторової адаптації в гібридних нейромережевих архітектурах, методи об'єктивного оцінювання якості зображень для кількісної верифікації результатів придушення шуму зі збереженням контурів.

Наукова новизна результатів, отриманих автором, полягає у наступному:

– **вперше запропоновано** метод апріорного прогнозування абсолютних значень мір якості PCA-зображень після придушення спекл-шуму на основі трансферного навчання, який відрізняється від відомих тим, що не потребує еталонного зображення для оцінювання конкретного вхідного знімка на етапі інференсу і ручного розрахунку ознак, витягуючи текстурні характеристики безпосередньо з вхідних даних;

– **удосконалено** метод апріорного вибору параметрів фільтрації (на прикладі фільтра Лі) завдяки формулюванню задачі як прямої багатокласової класифікації, що дозволило усунути накопичення похибок двоетапного

регресійного прогнозування та безпосередньо прогнозувати оптимальний розмір вікна;

– **одержав подальший розвиток** метод адаптивної фільтрації спекл-шуму РСА-зображень шляхом аналітичного об’єднання трансферного навчання на основі мережі VGG16 (Visual Geometry Group) та просторово-канальної уваги (CBAM) із точною математичною моделлю фільтра Фроста, що дозволяє динамічно прогнозувати карту коефіцієнтів демпфування як внутрішню змінну тензора стану та оптимізувати модель наскрізно, забезпечуючи керований відгук із збереженням меж.

Особистий внесок здобувача. Усі основні наукові результати, викладені в дисертації, отримано автором самостійно. Усі статті та матеріали конференцій опубліковано у співавторстві з науковим керівником (Рубель О. С.), якому належить постановка задач, обговорення результатів та загальне консультування. У спільних працях здобувачеві належать такі результати.

У статті [1] здобувачем розроблено метод апріорного прогнозування абсолютних значень мір якості (PSNR, WSNR, SSIM, MS-SSIM, FSIM) РСА-зображень Sentinel-1 після придушення спекл-шуму, адаптовано архітектуру густозв'язаної нейромережі DenseNet-121 для одноканальних даних, створено напівсинтетичний корпус зображень із динамічним гамма-розподіленим шумом та проведено експериментальні дослідження з шістьма класичними фільтрами. У статті [2] здобувачем запропоновано та програмно реалізовано метод прямого прогнозування оптимального розміру вікна фільтра Лі як задачі багатокласової класифікації на базі переднавченої архітектури MobileNetV2, розроблено конвеєр генерації міток на основі повного перебору та проведено порівняльний аналіз із існуючими двоетапними підходами. У статті [3] здобувачем розроблено гібридний метод VGG16-FrostNet для просторово-адаптивної фільтрації спекл-шуму, інтегровано диференційовану математичну модель фільтра Фроста у граф нейромережі, застосовано механізм просторово-канальної уваги (CBAM) для

прогнозування карти коефіцієнтів демпфування та сформовано складену функцію втрат для наскрізного навчання моделі зі збереженням меж об’єктів.

У матеріалах конференції [4] здобувачем постановлено задачу мінімізації ризиків неправильного встановлення розміру вікна фільтра Лі для PCA-зображень Sentinel-1, обґрунтовано перехід від емпіричного підбору та вичерпного перебору до прямої багатокласової класифікації параметра з використанням трансферного навчання MobileNetV2 та стратегії повного дообучення мережі для одноканальних знімків; здобувачем підготовлено та представлено очну доповідь, у якій показано, що автоматизація вибору параметрів знижує ймовірність втрати деталей і сприяє стабільному збереженню меж об’єктів. У доповіді [5] здобувачем проведено порівняльний аналіз запропонованого гібридного підходу VGG16-FrostNet із базовим фільтром Фроста, оптимізованим за метрикою HaarPSI, наочно продемонстровано переваги нейромережевого просторово-адаптивного прогнозування коефіцієнтів демпфування над глобально налаштованим еталоном і обґрунтовано роль наскрізного навчання гібридної моделі для збереження структурних ознак при фільтрації спекл-шуму. У матеріалах конференції [6] здобувачем розроблено таксономію з п’яти парадигм інтеграції класичних адаптивних фільтрів у нейромережеві архітектури (слабкий зв’язок, розгортання ітеративних алгоритмів, plug-and-play, жорсткий диференційований зв’язок, guided filtering), побудовано порівняльну матрицю за критеріями інтерпретованості, навчальності, обчислювальної складності та просторової адаптивності, сформульовано рекомендації щодо вибору парадигми залежно від вимог до прозорості моделі та попіксельної адаптації та підготовлено очну доповідь на конференції «Перспективні мережні та комп’ютерні технології» (ПерСиК 2026).

Апробація матеріалів дисертації. Основні положення та результати дисертаційної роботи доповідалися й обговорювалися на наукових та науково-практичних конференціях: Міжнародній студентській конференції з кібербезпеки «Information, Functional and Cybersecurity» (SCIFiC 2025) (м. Харків, 27–28

листопада 2025 р.); Міжнародній науково-практичній конференції «Organization of scientific research in modern conditions '2026» (м. Сієтл, США, 21 березня 2026 р.); 17-й Міжнародній студентській науково-технічній конференції «Перспективні мережні та комп'ютерні технології» (ПерСиК 2026) (23 квітня 2026 р.), а також на наукових семінарах кафедри інформаційно-комунікаційних технологій ім. О. О. Зеленського Національного аерокосмічного університету «Харківський авіаційний інститут» у 2024–2026 рр.

Публікації за темою дисертації. Основні наукові результати дисертації повністю розкривають її мету та завдання дослідження і повною мірою висвітлені здобувачем у 6 наукових працях. Серед них опубліковано 3 наукові статті у провідних фахових виданнях України (2 категорії «Б» та 1 категорії «А», з яких остання індексується у міжнародній наукометричній базі даних Scopus). Окрім цього, результати досліджень представлено у 3 працях апробаційного характеру (матеріали та тези доповідей на міжнародних і всеукраїнських науково-практичних конференціях).

Структура та обсяг дисертації. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел і додатків. За поточною редакцією рукопис містить 20 рисунків, 14 таблиць (3 службові та 11 експериментальних у розділах 2–3) та список використаних джерел із 163 найменування. Остаточні відомості про кількість сторінок основного тексту, рисунків, таблиць, джерел і додатків підлягають уточненню після фінальної верстки дисертації у форматі DOCX/PDF.

Зв'язок роботи з науковими програмами, планами, темами, грантами. Дисертаційна робота виконувалася відповідно до планів науково-дослідних робіт кафедри інформаційно-комунікаційних технологій ім. О.О. Зеленського Національного аерокосмічного університету «Харківський авіаційний інститут» у рамках науково-дослідної теми (підрозділ «Фільтрація та стиснення мультимедійного контенту»).

Практичне значення одержаних результатів: запропоновані моделі забезпечують автоматизований вибір оптимальних параметрів фільтрації спекл-шуму РСА-зображень без повторних прогонів фільтрів, що знижує обчислювальні витрати, усуває ризики помилкових налаштувань та оптимізує конвеєри обробки даних дистанційного зондування. Апріорне оцінювання метрик якості дозволяє ухвалювати рішення про доцільність фільтрації ще до її застосування в інтелектуальних хмарних сервісах і системах геопросторового моніторингу.

Метод прогнозування розміру вікна фільтра Лі на базі MobileNetV2 придатний для використання у периферійних системах із обмеженою обчислювальною продуктивністю завдяки компактному розміру моделі. Предиктор метрик якості на базі DenseNet-121 дозволяє оперативно ранжувати параметри без запуску самого фільтра. Отримані програмні реалізації можуть використовуватися як модулі попередньої обробки знімків Sentinel-1 у виробничих конвеєрах підприємств на етапах, що передують сегментації та картографуванню.

Наукові та практичні результати роботи впроваджено у виробничу практику компанії P-Product, Inc.: програмні прототипи запропонованих методів використовуються в спільних науково-дослідних роботах з автоматизації попередньої обробки даних дистанційного зондування; акт впровадження наведено в Додатку Б.

РОЗДІЛ 1. АНАЛІТИЧНИЙ ОГЛЯД МЕТОДІВ ПРИДУШЕННЯ СПЕКЛ-ШУМУ В РСА-ЗОБРАЖЕННЯХ, ОЦІНЮВАННЯ ЯКОСТІ ТА ПІДХОДІВ ТРАНСФЕРНОГО НАВЧАННЯ

Цей розділ присвячений огляду фізичної природи спекл-шуму в радіолокаційних зображеннях та систематичному аналізу існуючих методів його придушення. Розглядаються класичні локально-статистичні фільтри, методи глибокого та трансферного навчання, а також метрики для об'єктивного оцінювання якості фільтрації. На основі виявлених прогалин та недоліків існуючих парадигм формулюються конкретні задачі для подальших досліджень у дисертаційній роботі.

1.1 Фізична природа та моделі формування спекл-шуму в РСА-зображеннях

1.1.1 Базові принципи технології РСА

Технологія радіолокації із синтезованою апертурою (РСА) є ключовим інструментом для сучасного дистанційного зондування Землі (ДЗЗ). У порівнянні з традиційними оптичними або інфрачервоними сенсорами, які залежать від сонячного освітлення або власного теплового випромінювання об'єктів і чутливі до погодних умов, РСА є активною системою мікрохвильового діапазону. Вона випромінює власні електромагнітні імпульси та реєструє відлуння, відбите від земної поверхні. Мікрохвильовий діапазон (зокрема, X-, C- та L-діапазони) дозволяє радіосигналам проникати крізь хмарний покрив, туман, дим та інтенсивні опади, забезпечуючи можливість цілодобового та всепогодного моніторингу [1].

Основною проблемою традиційної радіолокації є досягнення високої просторової роздільної здатності. Для її покращення вздовж шляху (азимута) потрібно використовувати антену неприпустимо великих розмірів. Технологія РСА вирішує цю проблему за рахунок руху платформи-носія (супутника або літака). Радар випромінює серію імпульсів і когерентно реєструє відбиті сигнали

(зберігаючи інформацію як про амплітуду, так і про фазу) під час переміщення. Спеціальна математична обробка отриманих даних (доплерівський аналіз історії фаз) дозволяє синтезувати віртуальну антену (апертуру), довжина якої еквівалентна відстані, яку пролетів носій за час опромінення цілі. Це дозволяє досягти просторової роздільної здатності порядку кількох метрів або навіть десятків сантиметрів, що є порівнянним з оптичними знімками високої роздільної здатності.

Сучасний етап розвитку ДЗЗ багато в чому визначається успіхами європейської програми Copernicus [2] та її ключової компоненти — супутникової місії Sentinel-1. Ці супутники оснащені C-діапазонним радаром і функціонують в інтересах відкритої науки та глобального моніторингу. Основний режим роботи Sentinel-1 над сушею — Interferometric Wide swath (IW) Ground Range Detected (GRD) — забезпечує оптимальний компроміс між широкою смугою огляду (250 км) та просторовою роздільною здатністю (приблизно 10×10 м). Завдяки відкритості та регулярності оновлення даних, знімки Sentinel-1 стали індустріальним стандартом для вирішення завдань сільськогосподарського моніторингу, відстеження змін лісового покриву, картування наслідків стихійних лих (повеней, зсувів ґрунту) та аналізу міської інфраструктури [4]. Враховуючи значні обсяги інформації, які щодня генерують системи РСА, на перший план виходить проблема створення повністю автоматизованих та надійних алгоритмів попередньої обробки зображень.

Практична значущість таких даних підтверджується прикладними роботами з класифікації сільськогосподарських угідь за Sentinel-1, комбінованого використання РСА та оптичних знімків для реагування на надзвичайні ситуації, а також розгортання операційних сервісів на базі TerraSAR-X; у всіх цих сценаріях якість попередньої обробки безпосередньо визначає інформативність кінцевого продукту [19–24].

1.1.2 Фізика виникнення спекл-шуму

Попри характерні переваги РСА, когерентна природа зондувального сигналу має серйозний побічний ефект — формування інтенсивного шумоподібного явища, відомого як спекл-шум (від англ. speckle — цятка, гранула) [3]. Це явище не є "шумом" у класичному розумінні, оскільки воно не спричинене тепловими флуктуаціями в приймальній апаратурі чи завадами в каналі зв'язку. Спекл має суто детерміновану електромагнітну природу і є результатом складної взаємодії випромінювання із шорсткою земною поверхнею.

Фізичний механізм виникнення спеклу можна пояснити на макроскопічному рівні. Роздільна здатність сучасних супутникових РСА-систем (наприклад, 10 метрів для Sentinel-1) залишається значно більшою за довжину хвилі зондувального сигналу (близько 5.5 см для С-діапазону). У межах кожного елемента роздільної здатності (пікселя) знаходиться велика, часто випадкова кількість елементарних розсіювачів. Цими розсіювачами можуть бути окремі гілки та листя дерев, грудки землі, каміння, елементи будівель, хвилі на поверхні води. Оскільки відстань від антени радара до кожного з цих дрібних розсіювачів є різною (з точністю до часток довжини хвилі), відбиті від них парціальні електромагнітні хвилі повертаються до приймача з випадковим розподілом фаз.

Антена радара когерентно підсумовує всі ці відбиті хвилі. Векторне додавання сигналів з випадковими фазами призводить до інтерференції. У тих пікселях, де відбиті сигнали переважно збігаються за фазою (конструктивна інтерференція), результуюча амплітуда значно підсилюється, і піксель на зображенні виглядає аномально яскравим (біла цятка). Якщо ж фази парціальних хвиль розподілені так, що вони взаємно гасять одна одну (деструктивна інтерференція), результуючий сигнал є слабким, і піксель виглядає темним (чорна пляма) [5]. Цей процес відбувається випадково і незалежно для кожного сусіднього елемента роздільної здатності. У результаті навіть фізично ідеально однорідна ділянка поверхні (наприклад, голе сільськогосподарське поле або спокійна водойма), яка має постійну ефективну площу розсіяння (ЕПР), на РСА-

зображенні набуває хаотичної "зернистої" (гранульованої) текстури. Ця зернистість маскує справжні деталі земного покриву і критично ускладнює дешифрування знімків.

1.1.3 Математична і статистична модель (Single-look)

Для того щоб розробляти ефективні алгоритми придушення спекл-шуму, необхідно мати точну математичну модель цього явища. Традиційно формування результуючого радіолокаційного сигналу в межах одного пікселя моделюється як задача випадкових блукань у комплексній площині (random walk model) [6].

Нехай комплексний сигнал z, прийнятий антеною від одного елемента роздільної здатності, є сумою N парціальних відблисків від елементарних розсіювачів:

$$z = x + jy = \sum_{k=1}^{N} A_k \exp(j\phi_k) \tag{1}$$

де A_k — амплітуда відбиття k-го розсіювача, ϕ_k — його фаза, x та y — дійсна (in-phase, I) та уявна (quadrature, Q) складові результуючого сигналу, j — уявна одиниця.

За умови, що поверхня є макроскопічно шорсткою (перепади висот перевищують довжину хвилі), фази ϕ_k можна вважати незалежними випадковими величинами, рівномірно розподіленими на інтервалі $[-\pi, \pi]$. Амплітуда A_k і фаза ϕ_k також є статистично незалежними. Окрім цього, припускається, що кількість елементарних розсіювачів N є достатньо великою. За таких умов, згідно з центральною граничною теоремою теорії ймовірностей, дійсна складова x та уявна складова y результуючого комплексного сигналу є незалежними гауссівськими випадковими величинами з нульовим середнім значенням і однаковою дисперсією $\sigma^2 / 2$, де σ^2 — середня інтенсивність відбитого сигналу (сподівана радіолокаційна відбивна здатність ділянки).

Амплітуда результуючого сигналу $A = \sqrt{(x^2 + y^2)}$ в однопоглядовому (single-look) зображенні підпорядковується розподілу Релея. Щільність ймовірності (Probability Density Function, PDF) для амплітуди A має вигляд:

$$P_A(A) = \frac{2A}{\sigma^2}\exp\left(-\frac{A^2}{\sigma^2}\right), \quad A \geq 0 \tag{2}$$

В інтенсивнісному (intensity) представленні, яке частіше використовується для візуалізації та аналізу, значення пікселя I дорівнює квадрату амплітуди ($I = A^2 = x^2 + y^2$). У цьому випадку I підпорядковується експоненційному розподілу з PDF:

$$P_I(I) = \frac{1}{\sigma^2}\exp\left(-\frac{I}{\sigma^2}\right), \quad I \geq 0 \tag{3}$$

Основною властивістю експоненційного розподілу є те, що його математичне сподівання $E[I]$ дорівнює σ^2, а дисперсія $Var(I)$ дорівнює квадрату математичного сподівання: $Var(I) = (E[I])^2 = \sigma^4$. Відповідно, стандартне відхилення σ_I дорівнює середньому значенню ($\sigma_I = E[I]$). Локальний коефіцієнт варіації $C_I = \sigma_I / E[I]$ для повністю розвиненого спекл-шуму в однопоглядовому зображенні інтенсивності дорівнює 1. Це свідчить про дуже високий рівень флуктуацій: розмах варіацій яскравості пікселів досягає 100% від середнього рівня корисного сигналу, що пояснює різку візуальну деградацію знімків [3; 6].

1.1.4 Мультиплікативна природа шуму

Із наведених вище статистичних викладок (зокрема, з пропорційності стандартного відхилення до середнього значення) випливає найважливіша особливість спекл-шуму: він є принципово неадитивним. В оптичних та інфрачервоних зображеннях домінуючим є термальний або квантовий шум, який

зазвичай моделюється як адитивний білий гауссівський шум (AWGN). Адитивний шум не залежить від корисного сигналу і просто додається до нього: $I = R + n$, де R — ідеальний сигнал, а n — шумовий компонент.

Водночас, спекл-шум має мультиплікативний характер і описується як добуток істинної радіолокаційної відбивної здатності сцени (Radar Cross Section, R) та шумової компоненти (speckle, u):

$$I(x, y) = R(x, y) \cdot u(x, y) \tag{4}$$

де x, y — просторові координати пікселя.

Мультиплікативна модель означає, що рівень "зашумленості" (абсолютні відхилення яскравості) прямо пропорційний силі відбитого сигналу. Для темних ділянок земної поверхні, таких як гладкі водні об'єкти (які дзеркально відбивають імпульс від радара і тому здаються чорними) або асфальтовані дороги, флуктуації яскравості є мінімальними. Проте для яскравих цілей — міської забудови, лісових масивів або схилів гір, орієнтованих на радар, — спекл-шум стає дуже інтенсивним, досягаючи максимальних амплітуд [6].

Така поведінка шуму робить класичні методи цифрової обробки зображень малопридатними. Якщо застосувати до РСА-знімка звичайний лінійний фільтр низьких частот (наприклад, розмиття за Гауссом), який ефективно придушує AWGN, він призведе до несприятливих наслідків. В однорідних яскравих областях він згладить спекл, але одночасно він повністю знищить високочастотну просторову інформацію — розмиє різкі контури, змішає дрібні об'єкти та "з'їсть" точкові цілі. Тому розробка алгоритмів фільтрації спекл-шуму вимагає спеціалізованого апарату, здатного враховувати не лише нестаціонарний мультиплікативний характер завади, але й адаптуватися до локального контексту: інтенсивно згладжувати там, де шум домінує (однорідні ділянки), і зберігати оригінальні пікселі там, де присутні цінні структурні деталі (межі, точкові відбивачі) [5; 6].

1.1.5 Багатопоглядова обробка (Multi-looking) та зміна статистики

Оскільки рівень шуму в однопоглядовому зображенні є занадто високим для комфортного аналізу, на практиці, при наземній обробці сирого сигналу (формування продуктів L1, зокрема GRD) застосовується процедура багатопоглядової обробки (multi-looking). Суть цієї процедури полягає у поділі загального спектра доплерівських частот на кілька піддіапазонів. Кожен такий піддіапазон використовується для формування окремого, незалежного "погляду" (look) на ту саму ділянку поверхні [3].

Через випадковий характер спекл-шуму він дещо відрізняється в кожному з цих незалежних поглядів. Результуюче зображення формується шляхом некогерентного усереднення (додавання інтенсивностей) цих окремих поглядів. Якщо ми усереднюємо L незалежних поглядів, ми отримуємо багатопоглядове (multi-look) зображення:

$$I_{ML} = \frac{1}{L}\sum_{i=1}^{L} I_i \tag{5}$$

де I_i — інтенсивність i-го однопоглядового зображення.

Закон великих чисел стверджує, що таке усереднення зменшує дисперсію випадкової величини. І дійсно, у результаті багатопоглядової обробки статистичний розподіл шуму змінюється. Інтенсивність у багатопоглядовому зображенні вже підпорядковується не експоненційному, а гамма-розподілу, щільність ймовірності якого має вигляд:

$$P_I(I) = \frac{L^L I^{L-1}}{R^L \Gamma(L)} \exp\left(-\frac{L \cdot I}{R}\right), \quad I \geq 0 \tag{6}$$

де R — істинна відбивна здатність (математичне сподівання), $\Gamma(L)$ — гамма-функція, а L — кількість незалежних поглядів.

Властивості гамма-розподілу суттєво відрізняються від експоненційного. Його математичне сподівання залишається $E[I] = R$, але дисперсія зменшується в L разів: $Var(I) = R^2 / L$. Відповідно, локальний коефіцієнт варіації для однорідної ділянки багатопоглядового зображення тепер дорівнює:

$$C_v = \frac{\sqrt{\mathrm{Var}(I)}}{E[I]} = \frac{1}{\sqrt{L}} \tag{7}$$

Тобто, багатопоглядова обробка дозволяє зменшити рівень спекл-шуму в $\sqrt{L}$ разів. На практиці сусідні пікселі PCA-зображення не є повністю статистично незалежними (через надмірну дискретизацію та обробку сигналів), тому замість теоретичного L використовують емпіричний параметр — еквівалентне число поглядів (Equivalent Number of Looks, ENL). Для продуктів Sentinel-1 режиму IW GRD значення ENL зазвичай становить близько 4.3–4.4. Це означає, що шум частково придушений вже на етапі формування продукту. Тим не менш, платою за багатопоглядову обробку є зниження просторової роздільної здатності зображення (у L разів по азимуту). Крім того, навіть після multi-looking залишковий рівень шуму ($C_v \approx 0.48$ для $ENL \approx 4.4$) залишається достатньо високим, щоб створювати серйозні перешкоди для алгоритмів обробки на пізніших етапах [25; 26].

Окремі дослідження, присвячені оцінюванню спекл-характеристик Sentinel-1 та поляриметричних даних, показують, що практична оцінка ENL і параметрів спеклу суттєво залежить від типу продукту, поляризації та просторової кореляції шуму, а тому не може розглядатися як суто формальна константа алгоритму [25–29].

1.1.6 Вплив спекл-шуму на алгоритми комп'ютерного зору (Проблематика)

Попри застосування методів багатопоглядової обробки, наявність інтенсивного спекл-шуму залишається основним бар'єром для масового впровадження сучасних алгоритмів автоматичного дешифрування (Machine Learning / Deep Learning) у процеси аналізу РСА-даних [30].

По-перше, мультиплікативний шум значно знижує ефективність алгоритмів сегментації об'єктів. Більшість класичних методів (наприклад, Watershed, регіональне нарощування, K-means) базуються на припущенні про просторову зв'язність і відносну однорідність об'єктів. Гранульована текстура спеклу руйнує цю однорідність, спричиняючи масову "надсегментацію" (over-segmentation), коли суцільне поле або водойма розпадаються на тисячі дрібних помилкових сегментів. По-друге, значна дисперсія значень яскравості всередині одного семантичного класу ускладнює роботу статистичних класифікаторів і спричиняє перекриття гістограм різних класів [31–34].

Цей ефект підтверджується прикладними роботами з виявлення цілей, детектування контурів, сегментації міських сцен та класифікації сільськогосподарських угідь, де якість фільтрації спекл-шуму прямо впливає на точність подальшого аналізу [31–36].

Вирішення цієї проблеми можливе лише через якісну цифрову фільтрацію. Однак, перед дослідниками постає глобальна методологічна криза. Існує гостра потреба зберегти всі дрібні текстурні деталі і границі об'єктів (наприклад, інфраструктуру в містах), які можуть бути безповоротно знищені жорсткими фільтрами. Водночас оцінка ефективності самих фільтрів об'єктивно ускладнена: через фізичну природу процесу зондування не існує ідеальних, безшумних РСА-зображень (еталонів), з якими можна було б порівняти результат фільтрації. Цей факт — відсутність достовірних Ground Truth (еталонних) даних для реальних радіолокаційних сцен — унеможливлює повноцінне використання стандартних підходів контрольованого навчання (supervised learning) при навчанні глибоких нейронних мереж і стимулює пошук альтернативних стратегій: починаючи від

класичної локально-статистичної фільтрації до застосування трансферного навчання та синтетичних наборів даних, згенерованих, зокрема, на основі узгоджених оптичних чи інфрачервоних знімків високої роздільної здатності [18; 30].

Саме цьому шляху еволюції методів придушення спекл-шуму та подолання цих принципових обмежень присвячені наступні підрозділи даного розділу.

1.2 Класичні методи локальної статистики

Історично першим і найбільш дослідженим підходом до вирішення проблеми придушення мультиплікативного шуму в радіолокаційних зображеннях стало використання просторових адаптивних фільтрів [6]. Після детального математичного моделювання природи спекл-шуму стало очевидним, що традиційні лінійні фільтри низьких частот (такі як середньоарифметичний або гауссівський), які ефективно згладжують адитивний гауссівський шум (AWGN), є абсолютно непридатними для обробки РСА-даних. Оскільки спекл-шум пропорційний до інтенсивності сигналу, лінійне згладжування однаково впливає як на однорідні ділянки (де це корисно), так і на контури об'єктів (де це критично шкідливо). Це призводить до незворотної втрати високочастотної просторової інформації, розмиття меж, знищення дрібних деталей та «злиття» точкових цілей [5].

Для подолання цієї ключової проблеми в 1980-х роках було розроблено цілий клас адаптивних просторових фільтрів [7; 8]. Їхня головна ідея полягає в тому, що ступінь згладжування (вага центрального пікселя відносно його сусідів) має змінюватися динамічно, попіксельно, залежно від статистичних характеристик зображення всередині ковзного просторового вікна. Ці методи отримали назву методів локальної статистики. Вони не вимагають попереднього навчання на великих масивах даних і спираються виключно на математичні припущення щодо розподілу шуму та корисного сигналу.

Поза базовими Lee/Frost/Kuan-підходами у суміжній літературі активно розвивалися також edge-preserving, wavelet- та орієнтовані на дискретне косинусне перетворення (Discrete Cosine Transform, DCT) алгоритми радарної й загальної фільтрації спекл-шуму, які додатково підкреслили важливість просторової адаптації, локального контексту та компромісу між згладжуванням і збереженням структур [37–47].

1.2.1 Локальний коефіцієнт варіації та MMSE-фільтри (Лі, Куана)

Ключовим параметром, який керує поведінкою більшості класичних адаптивних фільтрів, є локальний коефіцієнт варіації C_I (іноді його називають індексом гетерогенності або коефіцієнтом розсіювання). Він обчислюється як відношення локального стандартного відхилення σ_I до локального середнього значення інтенсивності $\bar{I}$ в межах заданого ковзного вікна (наприклад, 3×3 або 5×5 пікселів):

$$C_I = \frac{\sigma_I}{\bar{I}} \tag{8}$$

Значення C_I слугує індикатором текстурності сцени. Для ідеально однорідної поверхні (гомогенної ділянки), де варіативність яскравості пікселів зумовлена виключно наявністю спекл-шуму, теоретичне значення цього коефіцієнта C_v залежить виключно від еквівалентного числа поглядів (Equivalent Number of Looks, ENL) і обчислюється як $C_v = 1 / \sqrt{ENL}$ (для зображень інтенсивності). Якщо розрахований локальний коефіцієнт варіації $C_I \approx C_v$, ділянка класифікується як однорідна, і до неї застосовується максимальне просторове усереднення. Якщо ж $C_I \gg C_v$, це свідчить про наявність у вікні контурів, текстур або яскравих точкових цілей (гетерогенна ділянка). У такому випадку фільтр повинен мінімізувати або повністю зупинити згладжування, щоб зберегти цінні структурні деталі [3; 7].

Одним із найвідоміших базових алгоритмів у цій категорії є **фільтр Лі (Lee filter)**, запропонований Чонг-Сен Лі у 1980 році [7]. Цей фільтр базується на критерії мінімуму середньоквадратичної похибки (Minimum Mean Square Error, MMSE). Оскільки мультиплікативна модель шуму є нелінійною і складною для прямої аналітичної оптимізації, Лі використав розкладання в ряд Тейлора першого порядку для лінеаризації моделі навколо локального середнього значення.

У результаті математичних перетворень, рівняння фільтра Лі зводиться до простої лінійної комбінації локального середнього значення $\bar{I}$ та оригінального значення центрального пікселя I:

$$\hat{R} = \bar{I} + W \cdot (I - \bar{I}) \tag{9}$$

де $\hat{R}$ — оцінене "чисте" (деспеклізоване) значення радіолокаційної відбивної здатності, а W — адаптивна вагова функція, що визначається як:

$$W = \frac{\mathrm{Var}(I) - \bar{I}^2 \sigma_v^2}{\mathrm{Var}(I)} \tag{10}$$

де $Var(I)$ — локальна дисперсія інтенсивності у просторовому вікні, а σ^2_v — дисперсія шуму. Для гомогенної ділянки в межах класичної багатопоглядової моделі приймають $\sigma^2_v = 1/L$; у практичних PCA-продуктах це співвідношення інтерпретують через еквівалентне число поглядів ENL, тобто $\sigma^2_v \approx 1/ENL$.

Механіка фільтра Лі є простою: в однорідних областях $Var(I)$ наближається до дисперсії шуму, вагова функція W прямує до 0, і оцінка $\hat{R}$ стає рівною локальному середньому (максимальне згладжування). На контрастних краях об'єктів локальна дисперсія стрімко зростає, W наближається до 1, і фільтр

повертає оригінальне значення пікселя I, фактично вимикаючись і запобігаючи розмиттю [7].

Подальшим удосконаленням цього підходу став **фільтр Куана (Kuan filter)** [9]. Порівняно з Лі, який використав наближення Тейлора, Куан розробив більш строгу нестаціонарну модель середнього та дисперсії (Non-stationary Mean, Non-stationary Variance — NMNV). Фільтр Куана перетворює мультиплікативну модель шуму на еквівалентну модель адитивного шуму, залежного від сигналу. Його оцінка зберігає структуру MMSE-фільтрації, але ваговий коефіцієнт виражається через відношення локального та шумового коефіцієнтів варіації:

$$\hat{R} = \bar{I} + W_{\text{Kuan}} \cdot (I - \bar{I}), \quad W_{\text{Kuan}} = \frac{1 - C_u^2/C_I^2}{1 + C_u^2}, \quad C_u^2 = 1/L. \tag{11}$$

де C_I — локальний коефіцієнт варіації у вікні, а C_u — коефіцієнт варіації мультиплікативного шуму. Попри більш глибоке теоретичне обґрунтування, на практиці (особливо для багатопоглядових РСА-зображень Sentinel-1) результати роботи фільтрів Лі та Куана є візуально та метрично дуже схожими.

Водночас інтерес до цього класу методів зберігається і в сучасних роботах: запропоновано модифікації фільтра Куана для обробки зображень, отриманих космічними радіолокаційними системами [48]. Це підтверджує, що підходи локальної статистики залишаються релевантною базою для подальших удосконалень і порівняльних експериментів.

Основною проблемою фільтрів Лі та Куана є використання жорсткого прямокутного вікна для обчислення локальної статистики. Рівномірне усереднення всередині квадратного масиву пікселів призводить до появи специфічних "блочних" артефактів на відфільтрованому зображенні, особливо поблизу лінійних структур та контрастних меж, оскільки пікселі по різні боки межі рівноцінно впливають на розрахунок дисперсії [7; 9].

1.2.2 Фільтр Фроста (експоненційна вагова функція)

Для вирішення проблеми блочних артефактів і забезпечення більш природного, згладженого переходу між однорідними областями та контурами, у 1982 році був розроблений **фільтр Фроста (Frost filter)** [8]. Цей алгоритм має особливе значення в контексті даної дисертаційної роботи, оскільки саме його математична модель була інтегрована в архітектуру запропонованої глибокої нейронної мережі VGG16-FrostNet для створення диференційованого шару адаптивної фільтрації.

Фільтр Фроста також базується на критерії MMSE, але замість простого усереднення в межах прямокутного вікна він використовує просторову згортку (convolution) з експоненційно спадною кругово-симетричною ваговою функцією. Математичною основою фільтра є припущення, що істинна радіолокаційна відбивна здатність сцени є не просто стаціонарним процесом, а моделюється як двовимірний авторегресійний процес (autoregressive process). Це означає, що сусідні пікселі мають певну просторову кореляцію, яка експоненційно зменшується зі збільшенням відстані [8].

Оцінка "чистого" пікселя у фільтрі Фроста обчислюється як зважена сума всіх пікселів всередині ковзного просторового вікна:

$$\hat{R} = \sum_{(x,y)\in\Omega} m(x,y) \cdot I(x,y) \tag{12}$$

де вагові коефіцієнти для кожного сусіднього пікселя розраховуються за формулою:

$$m(x,y) = K_1 \cdot \exp(-K_2 \cdot C_I^2 \cdot d(x,y)) \tag{13}$$

У класичній формулі Фроста закладено глибокий адаптивний механізм:
– $d(x,y)$ — просторова відстань (найчастіше евклідова $\sqrt{(x^2 + y^2)}$) від центрального пікселя вікна до поточного сусіда з координатами (x,y);
– C_I — вже знайомий локальний коефіцієнт варіації, обчислений у вікні (він виступає індикатором наявності контурів), а квадрат C^2_I підсилює реакцію фільтра на різкі локальні перепади;
– K_2 — константа демпфування (damping factor), що керує загальною швидкістю спадання експоненти і, відповідно, "агресивністю" згладжування;
– K_1 — нормувальний множник, який розраховується так, щоб сума всіх вагових коефіцієнтів $m(x,y)$ у вікні дорівнювала одиниці, що запобігає глобальній зміні яскравості сцени.

Ця архітектура забезпечує значно плавнішу просторову адаптацію. В гомогенних областях (наприклад, на водній поверхні) значення C_I є дуже малим. Відповідно, показник експоненти близький до нуля, вагові коефіцієнти $m(x,y)$ майже не залежать від відстані, і фільтр діє як широкий ізотропний фільтр низьких частот (рівномірне розмиття). Навпаки, на краях об'єктів або поблизу точкових цілей локальна дисперсія, а отже і C_I, різко зростає. Експонента починає спадати дуже швидко, і значущу вагу (близьку до 1) отримує лише сам центральний піксель (де відстань $d = 0$), тоді як вплив усіх сусідів множиться на 0. Отже, фільтр зберігає різкість контуру [8].

Відсутність жорстких "обрізаних" меж рівномірного вікна дозволяє фільтру Фроста уникати блочних артефактів і генерувати візуально більш природні РСА-зображення. Однак обчислення експоненти для кожного пікселя в кожному вікні робить його обчислювально набагато складнішим за фільтр Лі. Крім того, якість роботи фільтра Фроста критично залежить від ручного вибору параметра демпфування K_2 та розміру вікна, що є його головним вразливим місцем [16].

Саме ця класична формула надалі була використана як теоретична відправна точка для побудови гібридного методу VGG16-FrostNet. Водночас безпосереднє включення оригінального запису в нейромережевий граф виявилося

чисельно нестабільним, тому в розділі 3 буде введено окрему диференційовану модифікацію, у якій квадрат локального коефіцієнта варіації замінено на першу ступінь із відповідною компенсацією через просторово-змінний параметр демпфування.

1.2.3 MAP-фільтри (Gamma MAP)

Третім базовим підходом до локально-статистичної фільтрації є метод максимізації апостеріорної ймовірності (Maximum A Posteriori — MAP). Його концепція полягає в пошуку такого значення відбивної здатності, яке максимізує ймовірність, тобто ймовірність того, що істинне значення є заданим, за умови, що радар виміряв зашумлене значення інтенсивності. Згідно з теоремою Байєса, ця задача зводиться до максимізації добутку функції правдоподібності (likelihood), що описує модель шуму, та апріорного розподілу самої відбивної здатності сцени.

Найбільш досконалим представником цього класу є фільтр **Gamma MAP** (GMAP) [10]. Його головна інновація полягає у використанні специфічних апріорних статистичних розподілів: передбачається, що як спекл-шум (через багатопоглядову обробку), так і сама істинна радіолокаційна відбивна здатність земної поверхні (наприклад, текстура рослинності або міської забудови) підпорядковуються гамма-розподілу. Це припущення є значно точнішим для реальних багатопоглядових даних (таких як продукти GRD місії Sentinel-1), ніж спрощені гауссівські наближення.

Фільтр Gamma MAP реалізує адаптивну тримодальну стратегію обробки, жорстко розмежовуючи пікселі на три класи за допомогою двох порогів на базі локального коефіцієнта варіації C_I: нижнього порогу $C_v = 1/\sqrt{ENL}$ (шум) та верхнього порогу $C_{max} = \sqrt{(1 + 2/ENL)}$ (геометричні структури):

1. Однорідні ділянки ($C_I \leq C_v$): варіативність яскравості у вікні повністю пояснюється спекл-шумом. Фільтр виконує просте середньоарифметичне усереднення, максимально придушуючи заваду.
2. Текстуровані ділянки ($C_v < C_I \leq C_{max}$): варіативність є результатом суміші шуму

та природної макротекстури сцени (наприклад, крони дерев). Застосовується безпосередньо рівняння MAP-оцінки, яке є розв'язком квадратного рівняння:

$$\hat{R} = \frac{(\alpha - L - 1)\bar{I} + \sqrt{(\alpha - L - 1)^2 \bar{I}^2 + 4\alpha L I \bar{I}}}{2\alpha} \quad (14)$$

де L — кількість поглядів (ENL), а параметр порядку для гамма-розподілу сцени $\alpha = (1 + C^2_v) / (C^2_I - C^2_v)$.
3. Точкові цілі та контури ($C_I > C_{max}$): варіативність надто велика, щоб пояснити її текстурою. Вважається, що в центрі вікна знаходиться ізольований яскравий об'єкт (будівля, корабель, щогла) або різка границя. Щоб не зруйнувати його, фільтр повністю вимикається: $\hat{R} = I$.

Таке чітке розділення дозволяє фільтру Gamma MAP дуже ефективно зберігати точкові цілі та кутові відбивачі, які MMSE-фільтри (Лі та Фроста) часто "розмазують", оскільки їхня вагова функція наближається до одиниці асимптотично, а не жорстко відсікається [10].

1.2.4 Проблематика фіксованого вікна та ручного налаштування

Попри глибоке теоретичне обґрунтування та повсюдне впровадження у стандартні геоінформаційні системи (наприклад, SNAP від Європейського космічного агентства [4]), класичні фільтри локальної статистики зіткнулися з об'єктивною «стелею» ефективності. У сучасних умовах автоматизованої масової обробки супутникових знімків ці фільтри демонструють низку системних недоліків, які роблять їх використання вузьким «пляшковим горлом» для наступних алгоритмів автоматизованого дешифрування [11].

Першою і найголовнішою проблемою є залежність класичних методів від жорстко заданого розміру ковзного вікна обробки. Як було показано у попередніх працях [15], вибір розміру вікна (наприклад, 3×3, 5×5, 7×7 пікселів) є

критичним компромісом (trade-off), який не має універсально правильного рішення для цілої сцени. Мале вікно (наприклад, 3×3) ідеально зберігає дрібні просторові деталі, інфраструктурні об'єкти та контури, але збирає занадто мало статистичних даних (лише 9 пікселів) для ефективного усереднення шуму. Як наслідок, на однорідних ділянках (водні масиви, поля) залишається значний рівень залишкового гранульованого спеклу, що погіршує роботу алгоритмів сегментації.

Водночас велике вікно (7×7 або 9×9) дозволяє глибоко "вигладити" гомогенні області, забезпечуючи низьку дисперсію. Проте при наближенні до меж об'єктів у таке велике вікно одночасно потрапляють пікселі, що належать до абсолютно різних семантичних класів (наприклад, ліс та річка, або забудова і поле). Це порушує базове математичне припущення про локальну стаціонарність і призводить до значного розмиття контурів, знищення структурної цілісності міської забудови та "злиття" близько розташованих об'єктів. Жоден класичний фільтр не здатний самостійно змінити свій фізичний розмір вікна (рецептивне поле) в процесі сканування зображення [15].

Другою критичною проблемою є необхідність апріорного ручного налаштування керуючих параметрів. Оптимальний розмір вікна для фільтра Лі, коефіцієнт демпфування у фільтрі Фроста або пороги у Gamma MAP не є універсальними константами [16]. Вони кардинально залежать від двох факторів: технічних характеристик знімка (насамперед, значення ENL, яке варіюється залежно від режиму зйомки та попередньої обробки) та типу земного покриву. Наприклад, для моніторингу сільськогосподарських угідь доцільно застосовувати агресивне згладжування з вікном 7×7, тоді як для ідентифікації міської забудови використання вікна, більшого за 3×3, призведе до втрати критичної інфраструктури.

В умовах сучасного оперативного супутникового моніторингу (наприклад, щоденного аналізу тисяч сцен місії Sentinel-1 [49]) ручний візуальний підбір параметрів фільтрації для кожного знімка оператором є фізично неможливим і

економічно недоцільним. Застосування ж усереднених, «безпечних» параметрів за замовчуванням (найчастіше це вікно 3×3 у SNAP) призводить до того, що потенціал алгоритмів розкривається лише частково: якість фільтрації спекл-шуму залишається посередньою [12].

Звідси випливає природний перехід до задачі апріорного прогнозування ефективності фільтрації: у низці робіт для DCT- та багатоканальних фільтрів уже пропонувалися нейромережеві, регресійні та орієнтовані на зорову систему людини (Human Visual System, HVS) схеми оцінювання очікуваного виграшу, а також безеталонні критерії прийняття рішення щодо доцільності обробки [50–59].

Саме ця принципова нездатність класичних математичних моделей динамічно адаптувати своє рецептивне поле до семантичного контексту сцени та необхідність ручного калібрування параметрів стимулювали перехід до методів, керованих даними (data-driven), [13]. Розв'язання задачі автоматичного прогнозування оптимальних параметрів фільтрації (що вирішується у подальших розділах дисертації за допомогою класифікаційних згорткових мереж), а також інтеграція математичних моделей (таких як фільтр Фроста) всередину нейромережевих архітектур для наскрізної оптимізації, є єдиним шляхом подолання цих принципових обмежень.

1.3 Базові нейромережеві методи фільтрації спекл-шуму

1.3.1 Перехід до Data-Driven підходів (CNN) та Residual Learning

Історично склалося так, що переважна більшість класичних алгоритмів фільтрації спекл-шуму РСА-зображень, розглянутих у попередньому підрозділі, належала до категорії "model-driven" методів. Вони вимагали чіткого аналітичного формулювання математичної та статистичної моделі розсіяння електромагнітних хвиль, після чого завдання зводилося до пошуку оптимального розв'язку (наприклад, за критерієм MMSE або MAP). Попри формальну строгість, ці моделі зіткнулися з об'єктивною "стелею" якості: будь-яке математичне припущення (наприклад, про гамма-розподіл відбивної здатності земної поверхні)

неминуче є спрощенням реальної фізичної картини світу, а жорстко задане ковзне вікно не здатне динамічно реагувати на складні просторові структури (наприклад, міську забудову чи мережу доріг) [10].

З розвитком обчислювальних потужностей та накопиченням великих обсягів супутникових даних стався суттєвий зсув парадигми в бік "data-driven" підходів, основою яких стали глибокі згорткові нейронні мережі (Convolutional Neural Networks, CNN). Головна концептуальна перевага нейронних мереж полягає в їхній здатності автоматично вилучати складні, нелінійні ієрархічні ознаки безпосередньо з "сирих" даних, не вимагаючи від дослідника попереднього задання жорстких статистичних розподілів. Перші спроби застосування нейронних мереж для обробки зображень були зосереджені на задачах класифікації (зокрема на базі ImageNet [60; 61]), однак невдовзі цей інструментарій був адаптований для задач низькорівневого зору, включаючи шумопридушення.

Критичним етапом еволюції нейромережевих методів фільтрації стала поява концепції залишкового навчання (Residual Learning), вперше запропонованої в архітектурі залишкових мереж ResNet (Residual Networks) для вирішення проблеми затухання градієнтів у дуже глибоких мережах [62]. У контексті фільтрації спекл-шуму PCA-зображень залишкове навчання набуло специфічного і ефективного застосування. Замість того, щоб намагатися навчити нейронну мережу прогнозувати складне, багате на деталі та текстури «чисте» зображення (відбивну здатність сцени R), мережу навчають прогнозувати виключно шумову компоненту (speckle). Відповідно, очищене зображення отримується шляхом віднімання спрогнозованого шуму від оригінального зашумленого зображення [63].

Такий підхід пояснюється тим, що шумовий патерн, хоча і є випадковим, часто має більш передбачувану і однорідну статистичну структуру (наприклад, наближається до гауссівського розподілу після логарифмування), ніж сама сцена, що містить безліч непередбачуваних антропогенних та природних об'єктів.

Відокремлення задачі оцінки шуму дозволяє мережі набагато краще зберігати високочастотні просторові деталі: контури будівель, мости, берегові лінії, які інакше могли б бути розмиті в процесі агресивного прямого прогнозування сцени [63–65].

Однак мультиплікативна природа спекл-шуму в радіолокаційних зображеннях створює певну перешкоду: архітектури, розроблені для придушення адитивного білого гауссівського шуму (AWGN), не працюють «з коробки» на РСА-даних. Тому більшість ранніх CNN-базованих методів фільтрації спекл-шуму обов'язково включали етап гомоморфного перетворення. На вхід мережі подається не лінійна інтенсивність $I = R \cdot u$, а її логарифм $log(I) = log(R) + log(u)$, що перетворює шум на адитивний. Після того як мережа знімає адитивний шум $log(u)$ і прогнозує $log(R)$, виконується зворотне експоненційне перетворення. Попри успіх цього підходу (наприклад, у моделях SAR-CNN [66; 67]), логарифмічне перетворення має побічний ефект: воно суттєво змінює математичне сподівання шуму (яке перестає дорівнювати нулю) і може призводити до зсуву середньої яскравості відновленого зображення, що потребує додаткової компенсації.

1.3.2 Архітектури DnCNN та U-Net у задачах фільтрації

Значним кроком у застосуванні згорткових нейронних мереж для задач шумопридушення стала архітектура DnCNN (Denoising Convolutional Neural Network) [63]. Розроблена спочатку для оптичних зображень і придушення AWGN, DnCNN поєднала глибину архітектури VGG, механізм залишкового навчання (Residual Learning) та пакетну нормалізацію (Batch Normalization). Завдяки відмові від шарів субдискретизації (pooling) на користь збереження початкового просторового розміру тензора на кожному етапі, DnCNN довела здатність нейронних мереж перевершити за метриками якості класичні state-of-the-art алгоритми, такі як BM3D [68].

У сфері радіолокаційної зйомки архітектура DnCNN стала основою для створення спеціалізованої моделі SAR-CNN [66]. Розробники цієї мережі імплементували стратегію гомоморфного логарифмування, описану раніше, та навчили багатошарову згорткову мережу на великих базах напівсинтетичних даних, де оптичні зображення штучно контамінувалися шумом з розподілом, наближеним до спеклу багатопоглядових РСА-знімків. Аналогічний підхід був використаний у роботі [67], де CNN навчалася відновлювати просторові деталі на основі глибоких залишкових блоків. Ці алгоритми сформували новий базовий рівень для оцінки якості фільтрації спекл-шуму, продемонструвавши високі значення індексу SSIM та PSNR на тестових наборах.

Подальший розвиток цих ідей привів до нелокальних CNN-модифікацій, у яких самоподібні патчі використовуються не лише як евристика, а як структурний принцип побудови мережі для фільтрації спекл-шуму РСА [69].

Тим не менш, підхід на основі послідовних згорткових шарів (як у DnCNN) стикається зі значною проблемою "розширення рецептивного поля". Щоб зрозуміти глобальний контекст сцени та відрізнити масштабну макротекстуру від структурного шуму, нейронній мережі необхідно "бачити" велику ділянку зображення. У базових CNN рецептивне поле зростає лінійно зі збільшенням кількості шарів, що вимагає створення дуже глибоких (і обчислювально важких) архітектур, схильних до зникання градієнтів та перенавчання [63; 70].

Вирішення цієї проблеми прийшло з адаптацією архітектури типу U-Net, яка спочатку створювалася для медичної сегментації, але виявилася ідеальною для image-to-image трансляції. Архітектура U-Net представляє собою глибокий симетричний кодер-декодер (encoder-decoder). Етап кодування використовує серію згорток та шарів субдискретизації (max-pooling) для агресивного зменшення просторової роздільної здатності тензорів при одночасному збільшенні кількості каналів ознак (глибини). Це дозволяє моделі формувати глибоке, глобальне семантичне розуміння об'єктів сцени (озера, лісові масиви,

урбанізовані території) і при цьому експоненційно розширює рецептивне поле без надмірного зростання кількості параметрів [70; 71].

Однак проста архітектура кодер-декодер неминуче втрачає високочастотну інформацію під час шарів субдискретизації (pooling), що є критичним недоліком для задачі фільтрації спекл-шуму, де збереження різкості контурів та точкових відбивачів є пріоритетом. Ключовою ідеєю U-Net є введення пропускних з'єднань (skip connections), які безпосередньо конкатенують карти ознак з етапу кодування (високої роздільної здатності) до відповідних шарів етапу декодування. Отже, мережа отримує і "глобальний семантичний контекст" з глибоких шарів, і "точні просторові координати країв" з мілких шарів. Цей механізм дозволив U-Net-подібним моделям забезпечити високу стабільність у збереженні просторової структури [64; 72]. Альтернативним шляхом вирішення проблеми рецептивного поля стало використання дилатованих згорток (dilated convolutions), як це реалізовано в мережі SAR-DRN (Dilated Residual Network) [63]. Дилатовані згортки "проріджують" ядро, дозволяючи охоплювати більшу площу без збільшення кількості вагових коефіцієнтів та без використання деструктивних шарів субдискретизації.

1.3.3 Напівкеровані методи (Self-supervised, Noise2Noise) через відсутність еталонів

Попри ефективність та архітектурну зрілість зазначених CNN (таких як DnCNN, SAR-CNN, U-Net), їх практичне впровадження в оперативний моніторинг тривалий час стримувалося основною перешкодою: проблемою навчальної вибірки. У традиційній парадигмі контрольованого навчання (supervised learning) для оптимізації ваг мережі необхідні ідеальні пари: "зашумлене зображення" (вхід) та відповідне йому ідеально чисте "еталонне зображення" (Ground Truth, цільова змінна). Проте у радіолокаційному зондуванні Землі поняття "ідеально чистого" зображення фізично не існує: будь-яке когерентне вимірювання за своєю природою супроводжується спекл-шумом [18].

Щоб обійти це обмеження, дослідники спочатку використовували два обхідних шляхи. Перший шлях — створення напівсинтетичних датасетів на базі оптичних знімків. Для цього брали високоякісні аерофотознімки або оптичні супутникові дані (наприклад, Sentinel-2), і математично накладали на них штучний шум з розподілом Релея чи гамма-розподілом. Хоча моделі ідеально навчалися на таких напівсинтетичних парах, при спробі застосувати їх до реальних даних місії Sentinel-1 результати різко погіршувалися. Причина полягає в проблемі доменного зсуву (Domain Gap): статистичні, текстурні та геометричні властивості оптичного і радарного знімків кардинально відрізняються (наприклад, оптичні знімки не мають ефектів "накладання" або кутових відбивачів). Другий шлях полягав у використанні багаточасових серій (multitemporal averaging): усереднення десятків радарних знімків однієї території за рік створює псевдо-чисте зображення. Проте такий підхід вимагає абсолютної нерухомості об'єктів (відсутності сезонних змін, забудови, змін рівня води), що на практиці майже недосяжно для динамічних ландшафтів [18].

Важливим кроком у вирішенні цієї проблеми стала поява парадигми напівкерованого та самокерованого навчання (Self-supervised learning). Значний внесок здійснила архітектура Noise2Noise (N2N), яка математично довела, що нейронну мережу можна успішно навчити видаляти шум, маючи в якості цільової змінної не чисте зображення, а інше зашумлене зображення тієї ж самої сцени (з незалежною реалізацією шуму). Оскільки математичне сподівання шуму є нульовим (в адитивному еквіваленті), мережа вчиться ігнорувати некорельовані флуктуації і концентрується на вилученні корисного сигналу, збігаючись до того самого оптимуму, що й при навчанні на чистих даних [73].

Наступним кроком стала поява моделей типу Noise2Void та Speckle2Void [65]. Ці підходи дозволяють навчати мережу взагалі на одному єдиному зашумленому зображенні, не потребуючи навіть пари незалежних вимірювань. Ідея базується на концепції «сліпих плям» (blind-spot networks): в архітектурі мережі (через спеціальні маски) блокується доступ до центрального пікселя

рецептивного поля. Отже, мережа змушена прогнозувати значення центрального пікселя виключно на основі інформації з навколишнього просторового контексту (сусідніх пікселів). Оскільки спекл-шум вважається статистично незалежним (білим) у просторовому відношенні, мережа не може передбачити шум центрального пікселя з шуму сусідів, і змушена моделювати лише структурований корисний сигнал R. Для боротьби з просторовою кореляцією шуму, що виникає внаслідок попередньої обробки супутникових даних, використовуються вдосконалені схеми з вибіркою Бернуллі (Bernoulli-sampling-based learning), які дозволяють декоррелювати вхідні дані і зробити самокероване навчання стабільнішим [74].

Окрім самокерованого навчання, передовий рубіж досліджень також зосереджений на нелокальних алгоритмах та інтеграції механізмів уваги (Attention). Нелокальні згорткові мережі (Non-Local CNN) намагаються об'єднати ідею алгоритму Non-Local Means з глибоким навчанням, шукаючи самоподібні структури (патчі) по всьому зображенню для оптимізації ваг, а не лише в локальному вікні [64]. З появою Vision Transformers (ViT), моделі на базі Swin Transformers з механізмом багатовіконної само-уваги доводять свою здатність встановлювати нові рекорди (state-of-the-art) ефективності, захоплюючи глобальні просторові залежності з високою глибиною [75]. У комплексних оглядах останніх років підкреслюється, що комбінація самокерованого навчання, трансформерів та дифузійних моделей (DDPM) є визначальним трендом розвитку радіолокаційної обробки на наступне десятиліття [76; 77].

1.3.4 Обмеження чисто "чорноящикових" моделей

Попри значні успіхи підходів, керованих даними (data-driven), та їх домінування в порівняльних оцінках за метриками PSNR та SSIM, застосування суто "чорноящикових" наскрізно-навчальних моделей в інженерній та науковій практиці розшифрування РСА-знімків стикається з низкою принципових обмежень та ризиків [17; 78].

По-перше, глибокі нейронні мережі за своєю природою не мають суворої фізичної інтерпретованості. У той час як поведінку класичного фільтра Фроста або Gamma MAP можна математично передбачити та пояснити для будь-якого вхідного значення дисперсії [10], рішення 100-шарової залишкової мережі є результатом взаємодії мільйонів нелінійних вагових коефіцієнтів. Мережа діє як функціональний апроксиматор, що просто мінімізує функцію втрат (loss function) на етапі навчання. Відсутність інтегрованих знань про фізику процесу когерентного розсіяння та статистичні властивості спеклу означає, що результати роботи мережі можуть стати непередбачуваними, щойно вона зіткнеться з "out-of-distribution" даними — наприклад, зображеннями з іншого супутника, іншої поляризації, або з іншим значенням еквівалентного числа поглядів (ENL).

По-друге, нейромережеві моделі, навчені на мінімізацію середньоквадратичної похибки (L_2 loss) або похибки абсолютних відхилень (L_1 loss), схильні до надмірного згладжування (over-smoothing) та генерації галюцинацій (створення неіснуючих структур). Глибока згорткова мережа здатна візуально зробити PCA-зображення надмірно згладженим (що добре для метрики PSNR), але при цьому вона може непомітно знищити цінні точкові відбивачі (point scatterers) або спотворити фазові характеристики, якщо вони не були достатньо репрезентовані у навчальній вибірці [17]. Для вирішення цієї проблеми дослідники часто змушені ускладнювати функції втрат, додаючи до них компоненти повної варіації (Total Variation) або дивергенції Кульбака-Лейблера (KL-divergence), щоб змусити мережу поважати первісну статистику зображення [78; 79].

Саме ці обмеження моделей типу "чорна скринька" стимулювали формування нової парадигми — гібридної фільтрації спекл-шуму. Замість того, щоб протиставляти строгі фізико-математичні моделі (такі як фільтр Фроста або Лі) сучасним нейронним мережам (CNN, ResNet), гібридний підхід пропонує їх глибоку інтеграцію. Математична модель виступає надійним каркасом, який гарантує дотримання статистичних закономірностей розсіяння та зберігає

контроль над процесом обробки. А нейронна мережа бере на себе роль адаптивного керувального модуля, завдання якого — здійснювати семантичний аналіз просторового контексту сцени і адаптивно розраховувати оптимальні керуючі параметри для математичної моделі в кожній точці зображення. Обґрунтування, розробка та практична реалізація такого гібридного підходу на базі класифікаційної архітектури VGG16 та математичної моделі фільтра Фроста і є центральним предметом дослідження у цій дисертаційній роботі.

1.4 Метрики та підходи до оцінювання якості фільтрації

1.4.1 Базові піксельні метрики та їх обмеження (MSE, PSNR)

Оцінювання якості придушення спекл-шуму є винятково складним та багатогранним завданням, оскільки радіолокаційне зображення поєднує в собі базові фізичні властивості розсіяння електромагнітних хвиль та випадкові інтерференційні флуктуації. Історично, найпоширенішими, найстарішими та найпростішими для обчислення метриками якості відновлених зображень у комп'ютерному зорі є середня квадратична похибка (Mean Squared Error, MSE) та пікове відношення сигнал-шум (Peak Signal-to-Noise Ratio, PSNR). Вони належать до класу метрик з повним еталоном (Full-Reference), тобто категорично вимагають наявності ідеального чистого зображення (Ground Truth) для порівняння з результатом обробки.[80; 81].

Метрика MSE обчислює усереднену суму квадратів різниць між відповідними пікселями відфільтрованого зображення та ідеального еталону. Для цифрового зображення розміром $H \times W$ пікселів вона визначається класичним математичним рівнянням:

$$\mathrm{MSE} = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(I_{\mathrm{filtered}}(i,j) - I_{\mathrm{clean}}(i,j)\right)^2 \tag{15}$$

Метрика PSNR, своєю чергою, є логарифмічним представленням метрики MSE. Вона обернено пропорційна логарифму похибки і характеризує максимальний рівень корисного сигналу відносно потужності внесених спотворень (шуму). Вона традиційно вимірюється у децибелах (дБ):

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}}\right) = 20\log_{10}\left(\frac{\mathrm{MAX}_I}{\sqrt{\mathrm{MSE}}}\right) \tag{16}$$

де MAX_I — максимально можливе значення яскравості пікселя у заданому динамічному діапазоні (наприклад, 255 для 8-бітного зображення або 65535 для 16-бітного формату, що часто використовується для збереження радіолокаційних амплітуд). Вищі значення PSNR вказують на математично менше загальне відхилення від еталону.

Попри їхню абсолютну математичну простоту та повсюдне використання як базових критеріїв оптимізації (як функції втрат типу L2 Loss) у переважній більшості архітектур глибокого навчання, MSE та PSNR мають суттєві недоліки. Головна проблема полягає в тому, що вони базуються на жорсткому, ізольованому попіксельному порівнянні абсолютних значень і повністю ігнорують просторову кореляцію пікселів та структурну цілісність сцени. Як наслідок, ці метрики вкрай погано корелюють із суб'єктивним візуальним сприйняттям якості людиною-оператором (дешифрувальником). Наприклад, зображення, яке було злегка і рівномірно розмите (з повною втратою критичних контурів об'єктів), може парадоксальним чином мати вищий (кращий) PSNR, ніж зображення, яке зберегло ідеальну різкість меж, але містить невеликий, ледь помітний рівень залишкового високочастотного шуму.[17; 81].

Подібний розрив між формальною похибкою та реальною інформативністю зображення давно відзначається і в ширшій літературі з кольорової, медичної та загальної цифрової обробки зображень, де саме ця проблема стимулювала перехід від суто піксельних критеріїв до перцептивно орієнтованих показників [82–84].

Для специфіки радіолокаційних зображень (РСА), де інформативність зосереджена не у візуальній красі, а в наявності чітких фізичних структур, цей недолік є фатальним. Згладжування яскравих точкових відбивачів (point scatterers) — які можуть займати лише один-два пікселі, але є життєво важливими ознаками для розпізнавання бронетехніки чи кораблів — майже не впливає на глобальне макро-значення MSE, оскільки їхній мізерний кількісний внесок розчиняється на тлі мільйонів інших пікселів фону (лісу, води, ґрунту). Тому оцінка якості фільтрації виключно за метрикою PSNR є абсолютно недостатньою і, як доведено багатьма дослідниками, часто призводить до помилкового вибору алгоритмів, що схильні до надмірного розмиття (over-smoothing) просторових деталей [17; 81].

1.4.2 Багатомасштабні та структурні індекси (SSIM, MS-SSIM, FSIM)

Для кардинального подолання обмежень піксельних метрик, наукова спільнота розробила структурний індекс подібності (Structural Similarity Index Measure, SSIM). Замість обчислення тривіальної абсолютної різниці помилок, SSIM намагається математично моделювати сприйняття структурної інформації, припускаючи, що зорова система людини (Human Visual System, HVS) еволюційно високо адаптована до вилучення структури сцени. SSIM порівнює локальні патерни нормалізованої яскравості (luminance), контрасту (contrast) та структури (structure) у локальному ковзному вікні:[81].

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{17}$$

де μ_x, μ_y — локальні середні значення двох порівнюваних вікон, σ_x, σ_y — локальні стандартні відхилення (оцінки контрасту), σ_{xy} — локальна коваріація (що відображає збереження структурних зв'язків), а C_1, C_2 — константи стабілізації, які запобігають діленню на нуль у темних ділянках зображення. Теоретичне значення SSIM лежить у діапазоні *[-1, 1]*, де 1 означає абсолютну ідентичність

зображень. Завдяки врахуванню коваріації, SSIM чутливо реагує на деградацію просторової структури (наприклад, зсув фази або розмиття контурів будівель) під час фільтрації.

Однак, оскільки сприйняття структурних деталей критично залежить від масштабу спостереження (відстані до зображення та його просторової роздільної здатності), базовий індекс SSIM було природним чином розширено до багатомасштабної версії — MS-SSIM (Multi-Scale SSIM). Цей підхід будує піраміду роздільних здатностей шляхом ітеративного низькочастотного фільтрування та зменшення масштабу, обчислюючи SSIM на кожному рівні. Формально MS-SSIM визначається як зважений добуток SSIM на п'яти масштабах:[85].

$$\text{MS-SSIM}(x,y) = [l_M(x,y)]^{w_M} \cdot \prod_{j=1}^{M} \left[c_j(x,y)\right]^{w_j} \left[s_j(x,y)\right]^{w_j}, \qquad (18)$$

де $l_M(x,y)$ — порівняння яскравості на найгрубшому масштабі, $c_j(x,y)$ та $s_j(x,y)$ — відповідно компоненти контрасту і структури на j-му рівні піраміди, а w_j — стандартні вагові коефіцієнти ($w_1 = 0.0448$, $w_2 = 0.2856$, $w_3 = 0.3001$, $w_4 = 0.2363$, $w_5 = 0.1333$). Кожен наступний рівень отримують зменшенням масштабу вдвічі. Значення MS-SSIM лежить у діапазоні $[0, 1]$, де 1 означає повну ідентичність. Завдяки багатомасштабному аналізу MS-SSIM збалансовано оцінює збереження як дрібних високочастотних деталей, так і макротекстури сцени.

Подальшим розвитком структурних підходів є метрика FSIM (Feature Similarity Index). Порівняно з SSIM, яка спирається на локальну статистику яскравості та дисперсії, FSIM оцінює подібність за двома низькорівневими ознаками, до яких чутлива зорова система: фазовою конгруентністю (Phase Congruency, PC) та градієнтом амплітуди (Gradient Magnitude, GM). FSIM агрегує локальні карти подібності за цими ознаками:[86].

$$\mathrm{FSIM} = \frac{\sum_{p\in\Omega} S_{\mathrm{PC}}(p) S_G(p) \mathrm{PC}_m(p)}{\sum_{p\in\Omega} \mathrm{PC}_m(p)}, \tag{19}$$

де $S_{PC}(p)$ — локальна подібність за фазовою конгруентністю, $S_G(p)$ — локальна подібність за градієнтом амплітуди, $PC_m(p) = \max(PC_1(p), PC_2(p))$ — вагова функція фазової конгруентності, Ω — множина пікселів зображення. Значення FSIM лежить у діапазоні $[0, 1]$, де 1 означає повну ідентичність. На локальному рівні $S_{PC}(p)$ та $S_G(p)$ є нормованими картами подібності, які порівнюють відповідно фазову конгруентність $PC_1(p)$, $PC_2(p)$ та градієнтні амплітуди $G_1(p)$, $G_2(p)$ двох зображень. Тому FSIM чутлива до збереження геометричної точності контурів антропогенних об'єктів на РСА-знімках.

1.4.3 Перцептивні метрики (HaarPSI, MDSI) та індекси контурів (EPI, FOM)

Для підвищення відповідності математичних показників діагностичній цінності радіолокаційних знімків застосовують спеціалізовані перцептивні індекси. Однією з таких метрик є HaarPSI (Haar wavelet-based Perceptual Similarity Index) [87], яка базується на двовимірному дискретному вейвлет-перетворенні Хаара. HaarPSI обчислює локальні карти подібності за високочастотними коефіцієнтами з вагуванням за низькочастотними компонентами:

$$\mathrm{HaarPSI}_{f_1,f_2} = l_\alpha^{-1}\left(\frac{\sum_x \sum_{k=1}^{2} HS_{f_1,f_2}^{(k)}[x] \cdot W_{f_1,f_2}^{(k)}[x]}{\sum_x \sum_{k=1}^{2} W_{f_1,f_2}^{(k)}[x]}\right)^2, \tag{20}$$

де $HS^{(k)}_{f1,f2}[x]$ — локальна карта подібності між відгуками вейвлетів Хаара для двох зображень, $W^{(k)}_{f1,f2}[x]$ — відповідна вагова карта, що підсилює внесок візуально значущих ділянок, $k \in \{1,2\}$ — горизонтальні та вертикальні компоненти, l_α — логістичне перетворення. Таким чином, HaarPSI порівнює не

самі пікселі, а подібність їх частотних представлень у базисі Хаара [87]. Значення HaarPSI ∈ *[0, 1]*, де 1 означає повну ідентичність.

MDSI (Mean Deviation Similarity Index) [88] агрегує локальну карту спотворень за принципом середнього відхилення:

$$\mathrm{MDSI} = \left(\frac{1}{|\Omega|} \sum_{p \in \Omega} |D(p) - \mu_D|^{\rho} \right)^{1/\rho}, \tag{21}$$

де $D(p)$ — комбінована локальна карта спотворень (градієнт, контраст, структура), μ_D — її середнє значення на множині пікселів Ω, ρ — показник агрегації. На відміну від більшості індексів подібності, менше значення MDSI відповідає вищій якості відновлення.

Сутність MDSI полягає в тому, що спочатку формується карта локальних спотворень $D(p)$, яка агрегує відмінності за градієнтами, контрастом і структурою, а потім оцінюється не середнє значення цієї карти, а її середнє відхилення від власного середнього μ_D. Завдяки цьому метрика чутлива не лише до сили спотворення, а й до його просторової нерівномірності [88].

Окрім глобальних індексів подібності, для РСА-зображень критично важливе локальне оцінювання збереження структурних меж. З цією метою застосовують контурні метрики EPI (Edge Preservation Index) та FOM (Pratt's Figure of Merit). EPI вимірює кореляцію між картами країв відфільтрованого та еталонного зображень:

$$E = \sqrt{g_x^2 + g_y^2 + \varepsilon}, \quad g_x = G_x * I, \quad g_y = G_y * I, \tag{22}$$

$$\mathrm{EPI} = \frac{\sum(E_f - \mu_{E_f})(E_c - \mu_{E_c})}{\sqrt{\sum(E_f - \mu_{E_f})^2 \sum(E_c - \mu_{E_c})^2}} \tag{23}$$

де E_f та E_c — карти країв, сформовані, наприклад, оператором Собеля через горизонтальний і вертикальний градієнти g_x, g_y, а μ_{Ef}, μ_{Ec} — їхні середні значення. EPI ∈ *[0, 1]*, де 1 означає повне збереження контурів. FOM додатково оцінює просторову точність виявлених країв відносно еталону:

$$\mathrm{FOM} = \frac{1}{\max(N_{\mathrm{detected}}, N_{\mathrm{reference}})} \sum_{i=1}^{N_{\mathrm{detected}}} \frac{1}{1 + \alpha d_i^2}, \tag{24}$$

де $N_{detected}$ та $N_{reference}$ — кількість виявлених та еталонних пікселів країв, d_i — відстань від i-го виявленого краю до найближчого еталонного, $\alpha = 1/9$ — ваговий параметр. FOM ∈ *[0, 1]*, де 1 відповідає ідеальному збігу контурів. FOM суворо штрафує просторовий зсув або потовщення меж — типову проблему надмірного згладжування у класичних фільтрах.

1.4.4 Специфічні радарні показники (ENL) та оцінювання без еталону (No-Reference)

Завершуючи огляд методів оцінювання якості, неможливо оминути принципову проблему практичної радіолокації: у реальних сценаріях експлуатації орбітальних систем ДЗЗ (наприклад, Sentinel-1) ідеального (безшумного) еталонного зображення не існує в принципі. Це робить неможливим пряме застосування метрик типу PSNR чи SSIM для оперативної оцінки якості на реальних даних [6]. Тому критичною складовою оцінки якості РСА-обробки є метрики без еталону (No-Reference/Blind Quality Metrics), які повністю спираються на відомі фізичні та статистичні властивості спеклу.

Абсолютно головним та найвідомішим індикатором ефективності придушення шуму є еквівалентне число поглядів (Equivalent Number of Looks, ENL) [3]. Цей показник обчислюється виключно на спеціально обраних, максимально однорідних (гомогенних) ділянках сцени (наприклад, спокійні водні поверхні водосховищ або рівні сільськогосподарські поля з однаковою шорсткістю) як відношення квадрата локального середнього значення інтенсивності до її дисперсії:

$$\mathrm{ENL} = \frac{\mu_{\mathrm{hom}}^2}{\sigma_{\mathrm{hom}}^2} \tag{25}$$

Що вище емпіричне значення ENL на гомогенній ділянці після обробки, то ефективнішим є локальне усереднення і краще придушено мультиплікативний спекл-шум. Однак, варто пам'ятати, що ENL у жодному разі не можна використовувати як єдиний, ізольований критерій якості. Причина проста: дуже високе значення ENL можна легко отримати шляхом тотального математичного розмиття зображення (наприклад, застосувавши вікно усереднення розміром 51×51 піксель). Проте при цьому всі корисні об'єкти, контури та інфраструктура будуть безповоротно знищені. Саме тому ENL завжди розглядається у жорсткій зв'язці з індексами збереження контурів (такими як EPI або Edge-Save Ratio) для досягнення оптимального компромісу (trade-off) між згладжуванням гомогенних ділянок та різкістю гетерогенних зон [3; 6].

Для комплексної та глибокої оцінки якості без еталону (No-Reference) дослідниками також розроблено специфічні статистичні індекси, такі як M-Index, які оцінюють залишкові флуктуації відносно локальних середніх значень. Ці індекси перевіряють, чи відповідає статистика видаленого (залишкового) сигналу теоретичному гамма-розподілу справжнього радіолокаційного шуму. Якщо видалений шум містить просторові структури (контури), це свідчить про недолік фільтра, який разом із шумом видалив частину корисного сигналу. Такий

комплексний багатовимірний підхід — поєднання сучасних перцептивних метрик (HaarPSI, FSIM, MDSI) для оцінки збереження структур та специфічних радарних показників (ENL, EPI) — є одним із найбільш обґрунтованих способів порівняльного оцінювання методів адаптивної фільтрації [89–91].

У суміжних роботах з оцінювання якості для задач шумопридушення та ДЗЗ розглядалися як перцептивні HVS-орієнтовані критерії, так і спеціалізовані нейромережеві або статистичні оцінювачі, що дозволяють краще пов'язати числові метрики з реальною інформативністю зображення [89–94].

Особливо близькою до задачі цієї дисертації є лінія досліджень, де прогнозується очікувана якість або оптимальна робоча точка під час стиску noisy images: від ранніх DCT- та ADCTC-підходів до сучасних BPG-орієнтованих моделей, що оцінюють спотворення, перцептивну якість та доцільність подальшої обробки без повного перебору всіх режимів [95–112].

Для задач стиску та подальшої аналітики зображень важливими виявилися також роботи з прогресивного та високоефективного кодування, оцінювання помітності спотворень, компресії мульти- і гіперспектральних даних та аналізу того, як стиск впливає на подальшу класифікацію і картографування [113–123].

Не менш важливою передумовою коректного прогнозування якості є надійне оцінювання статистики шуму; у загальній літературі для цього розроблено методи оцінювання параметрів сигнально-залежного, змішаного та просторово-корельованого шуму, зокрема для raw-, стоматологічних, мультиспектральних та гіперспектральних даних [124–133].

1.5 Трансферне навчання та його застосування

1.5.1 Концепція Transfer Learning та Domain Adaptation

Традиційне глибоке машинне навчання вимагає для побудови ефективних моделей наявності значних обсягів розмічених даних [134]. Проте у сфері обробки радіолокаційних зображень (РСА), зокрема для завдань фільтрації спекл-шуму або класифікації, формування достатньої кількості якісних еталонних даних часто

є технічно неможливим завданням. У таких умовах особливо актуальною стає парадигма трансферного навчання (Transfer Learning), яка дозволяє переносити знання, отримані моделлю під час розв'язання однієї базової задачі (Source Task) на великому наборі даних, на нову цільову задачу (Target Task) з обмеженою кількістю розмічених зразків [135].

Основна ідея трансферного навчання базується на гіпотезі про те, що ранні шари згорткових нейронних мереж (CNN) вивчають універсальні, низькорівневі ознаки зображень — такі як просторові градієнти, контури, базові геометричні примітиви та текстурні патерни [136]. Ці базові ознаки є спільними для широкого класу візуальних даних, незалежно від їхньої фізичної природи. Водночас глибокі шари формують специфічні, високорівневі семантичні абстракції, орієнтовані на конкретне завдання. Отже, замість навчання мережі "з нуля" (from scratch) з випадковою ініціалізацією ваг, трансферне навчання пропонує використовувати попередньо навчену (pre-trained) архітектуру як ефективний екстрактор ознак, донавчаючи лише її фінальні шари на нових цільових даних.

Тісно пов'язаним із трансферним навчанням є поняття доменної адаптації (Domain Adaptation) [134]. Доменна адаптація — це спеціальний підклас трансферного навчання, який зосереджується на проблемі, коли базова задача та цільова задача є ідентичними (або дуже схожими), але розподіли ймовірностей базового (Source Domain) та цільового доменів (Target Domain) суттєво відрізняються. У контексті обробки РСА-зображень ця ситуація є типовою: алгоритм, навчений придушувати адитивний гауссівський шум на звичайних оптичних фотографіях, зазнає значного падіння ефективності (performance degradation) при застосуванні до радіолокаційних знімків, уражених мультиплікативним спекл-шумом [137]. Процес адаптації покликаний мінімізувати цю розбіжність розподілів і зробити модель інваріантною до специфічних доменних спотворень.

Передумовою такої адаптації є також коректне зіставлення мультисенсорних ознак і статистик: у літературі з ДЗЗ це простежується як у

роботах із порівнянням мір подібності для мультисенсорних знімків, так і в задачах класифікації багатоканальних даних, де узгодження простору ознак є критичним ще до етапу глибокого донавчання [138; 139].

З формальної точки зору трансферне навчання можна класифікувати за трьома осями [134]: (i) індуктивне трансферне навчання (Inductive Transfer), коли базова й цільова задачі відрізняються, а розмічені зразки доступні в цільовому домені (типовий випадок прогнозування мір якості фільтрації спекл-шуму); (ii) трансдуктивне навчання (Transductive Transfer), що відповідає класичному сценарію доменної адаптації — задача та сама, але домени різні (наприклад, перенесення з Sentinel-2 на Sentinel-1); (iii) непідконтрольне трансферне навчання (Unsupervised Transfer), коли в цільовому домені немає міток узагалі. Для задач фільтрації спекл-шуму РСА найчастіше застосовують перші дві категорії, оскільки прогнозування якості та параметрів фільтра потребує хоча б невеликої кількості розмічених прикладів.

Серед сучасних підходів до доменної адаптації виділяють методи, що мінімізують статистичну дистанцію між розподілами ознак у джерельному й цільовому доменах. До них належать максимальне середнє розходження (Maximum Mean Discrepancy, MMD) та кореляційне вирівнювання (CORAL — CORrelation ALignment), які додаються до функції втрат як регуляризатор і змушують глибокі ознаки джерельних і цільових зразків мати близькі статистичні моменти [78]. Альтернативним напрямом є змагальна адаптація (Adversarial Domain Adaptation), реалізована, зокрема, в архітектурах DANN (Domain-Adversarial Neural Network) [137], де додатковий дискримінатор намагається відрізнити домени, а магістральна мережа — обманути його, формуючи доменно-інваріантні представлення. У контексті РСА такі підходи дозволяють використовувати велику кількість немічених реальних знімків Sentinel-1 разом із обмеженим розміченим напівсинтетичним корпусом.

1.5.2 Попереднє навчання на оптичних даних (ImageNet) для задач РСА

Найпоширенішою стратегією у сфері комп'ютерного зору є використання моделей, попередньо навчених на масштабному наборі оптичних зображень ImageNet [60], [61]. Цей набір даних містить понад мільйон зображень, розподілених за тисячею категорій, що дозволяє згортковим мережам (наприклад, VGG16 [140], ResNet [62], DenseNet [141]) сформувати розгалужений простір візуальних ознак. Попри суттєві фізичні відмінності між некогерентним оптичним зображенням та когерентним радіолокаційним відликом, дослідження доводять ефективність перенесення «оптичних» ознак у домен РСА [135].

Зокрема, базові просторові фільтри, сформовані на перших згорткових шарах попередньо навчених моделей, виявляються ефективними детекторами структурних меж на радіолокаційних знімках [136]. Використання таких екстракторів значно прискорює збіжність (convergence) процесу навчання у цільовому домені та слугує ефективним регуляризатором, що запобігає перенавчанню (overfitting) моделі на невеликих датасетах РСА. Особливо дієвим цей підхід виявляється при розв'язанні специфічних завдань параметричної оптимізації, таких як прогнозування метрик якості фільтрації [30] або класифікація оптимального розміру вікна для класичних адаптивних фільтрів (наприклад, фільтра Лі) [15].

З точки зору архітектури, застосування базових мереж (backbones), таких як MobileNetV2, дозволяє створювати компактні та обчислювально ефективні рішення. MobileNetV2 базується на інвертованих залишкових блоках (Inverted Residuals) та лінійних вузьких місцях (Linear Bottlenecks), що забезпечує значне зниження кількості параметрів та мультиплікативно-акумулятивних операцій (MACs). Це робить такі переднавчені моделі придатними для інтеграції в операційні конвеєри попередньої обробки великих масивів РСА-даних (наприклад, Sentinel-1 [49]), де вимоги до швидкодії та пам'яті є критичними.

Окрему групу досліджень становлять роботи з РСА-специфічного попереднього навчання (SAR-pretraining), коли базова мережа навчається не на

ImageNet, а на великих неанотованих корпусах самих радіолокаційних знімків з використанням самокерованих сигналів — контрастного навчання (SimCLR, MoCo), маскованого моделювання патчів (Masked Image Modeling) або відновлення зашумлених фрагментів [73; 75]. Такі базові ваги формують ознаки, що від початку враховують когерентну природу РСА-сигналу, і потім легко донавчаються на цільові задачі — фільтрацію спекл-шуму, сегментацію землекористування або класифікацію військових об'єктів. Для практичних завдань кафедри XAI, де доступні значні архіви Sentinel-1, такий підхід відкриває перспективу побудови базової РСА-моделі, що замінить ImageNet як стартову точку всіх подальших гібридних архітектур.

Важливим інженерним аспектом застосування переднавчених backbone у задачах фільтрації спекл-шуму є коректне вирівнювання динамічних діапазонів. Класичні ImageNet-моделі очікують на вході нормалізовані RGB-значення в діапазоні [0, 1] зі стандартизацією по ImageNet-середніх і дисперсіях. Радіолокаційні амплітуди Sentinel-1, навпаки, мають широкий 16-бітний динамічний діапазон і логнормальний розподіл інтенсивностей. Без узгодження статистик (логарифмічного стиснення, квантильної нормалізації або per-image standardization) попередньо навчені ваги швидко «зміщуються» у небажану зону активацій, нівелюючи переваги перенесення знань [15]. Тому в усіх практичних реалізаціях, описаних у роботах автора, попередній препроцесинг РСА-даних виконується саме з урахуванням специфіки обраної базової архітектури.

1.5.3 Подолання доменного зсуву (Domain Shift)

Ключовою перешкодою при прямому перенесенні моделей з оптичного домену в домен РСА є проблема доменного зсуву (Domain Shift) [134]. Цей феномен зумовлений глибокими фізичними відмінностями в процесах формування зображень. Оптичні зображення є зазвичай багатоканальними (RGB), характеризуються адитивною природою шуму та залежать від умов освітленості. Водночас РСА-зображення найчастіше є одноканальними (інтенсивність або

амплітуда відбитого сигналу), мають високий динамічний діапазон і завжди спотворені інтенсивним мультиплікативним спекл-шумом [6].

Пряме застосування моделі, навченої на оптичних даних (Source Domain), до РСА-знімків (Target Domain) призводить до непередбачуваних помилок [137]. Нейромережа, яка звикла аналізувати оптичні градієнти, інтерпретує висококонтрастні гранули спекл-шуму як корисні дрібні деталі або текстури, намагаючись їх зберегти або навіть підсилити. Для подолання доменного зсуву застосовується комплекс спеціалізованих математичних та алгоритмічних методів.

Першим етапом подолання зсуву є структурна адаптація архітектури: модифікація вхідних шарів мережі для обробки одноканальних тензорів замість триканальних. Наступним кроком є узгодження статистичних розподілів даних. Це реалізується за допомогою методів співставлення гістограм (Histogram Matching), які дозволяють трансформувати яскравісні характеристики еталонних оптичних зображень (наприклад, Sentinel-2 у ближньому інфрачервоному діапазоні) до специфічного негаусового розподілу амплітуд радіолокаційних знімків (наприклад, Sentinel-1) [15]. Завершальним, найважливішим етапом є синтез доменно-специфічного шуму. На узгоджені «чисті» зображення накладається синтетичний мультиплікативний спекл-шум із контрольованою просторовою кореляцією та заданим еквівалентним числом поглядів (ENL) [3]. Такий підхід дозволяє створити ідеальні парні датасети (напівсинтетичний РСА / ідеальний еталон) для ефективного донавчання моделей [30].

Сучасна методологія генерації напівсинтетичних РСА-датасетів передбачає кілька принципово важливих кроків [18; 78]. По-перше, моделювання нерівномірності спекл-шуму через випадкове семплювання параметра ENL з рівномірного або логарифмічно рівномірного розподілу в діапазоні, що відповідає різним режимам зйомки Sentinel-1 (IW, EW, SM). По-друге, додавання просторово корельованого шуму, який моделюється згорткою незалежного гамма-шуму з обмеженою імпульсною характеристикою фільтра попередньої обробки SLC-

знімків. По-третє, моделювання радіометричних спотворень (теплового шуму антени, нерівномірності розсіяння по куту падіння). Сумарно ці методи дозволяють зменшити розрив між напівсинтетичними та реальними РСА-даними, що підтверджується експериментально як вищою сталістю метрик при крос-доменному тестуванні.

Окрему підгрупу методів адаптації представляють підходи на основі переносу стилю (Style Transfer) та генеративно-змагальних мереж (Generative Adversarial Networks, GAN). Архітектури типу CycleGAN та pix2pixHD дозволяють «перевести» оптичне зображення у візуальний РСА-стиль без явної потреби у парних еталонах, формуючи близькі за статистикою синтетичні знімки [76]. У задачах фільтрації спекл-шуму такі підходи застосовуються переважно як допоміжний компонент розширення вибірки (data augmentation), оскільки повна заміна фізично коректних моделей шуму на «навчені статистичні апроксимації» залишається предметом дискусії в науковій спільноті через ризик внесення прихованих систематичних спотворень.

1.5.4 Стратегії Fine-tuning

Процес перенесення знань (Fine-tuning) вимагає обережного вибору стратегії оптимізації ваг попередньо навченої моделі [136]. Залежно від обсягу цільового набору даних та ступеня доменного зсуву, розрізняють кілька базових стратегій донавчання. Перша стратегія передбачає повне «заморожування» (freezing) усіх згорткових шарів базової мережі. Мережа використовується виключно як статичний екстрактор ознак, а навчається лише нова, спеціально додана повнозв'язна «голова» класифікації чи регресії. Цей підхід є швидким, стійким до перенавчання на малих датасетах, але він не дозволяє мережі адаптувати низькорівневі фільтри до специфіки мультиплікативного спеклу.

Більш гнучкою є стратегія поступового «розморожування» (gradual unfreezing) або диференційованої швидкості навчання (discriminative learning rates) [135]. У цьому випадку глибокі семантичні шари мережі донавчаються з вищою

швидкістю навчання (learning rate), оскільки вони повинні адаптуватися до нових цільових класів (наприклад, розмірів вікна фільтрації). Водночас ранні згорткові шари, які вилучають універсальні ознаки, оновлюються з мінімальним кроком або залишаються частково замороженими, що дозволяє зберегти корисні знання, отримані з ImageNet.

Найбільш повною, але часто найресурсомісткішою стратегією за наявності великих синтетичних наборів даних є повне наскрізне донавчання (full fine-tuning). За такої стратегії ваги всієї мережі оновлюються одночасно за допомогою алгоритмів градієнтного спуску. Для забезпечення стабільності та уникнення значного забування (catastrophic forgetting) застосовуються малі значення швидкості навчання та спеціалізовані методи регуляризації (наприклад, Dropout та Weight Decay). Як показують дослідження [30], повне донавчання архітектур (таких як MobileNetV2 чи VGG16) на спеціально підготовлених напівсинтетичних датасетах дозволяє досягти високої точності в задачах апріорного прогнозування параметрів та гібридної просторової адаптації PCA-фільтрів.

Феномен катастрофічного забування (catastrophic forgetting) є однією з ключових проблем повного донавчання глибоких мереж: під час адаптації до нової задачі мережа поступово втрачає здатність розв'язувати початкову задачу, на якій формувалися переднавчені ваги. У контексті фільтрації спекл-шуму PCA це означає, що під час адаптації VGG16 до радарних патернів можуть «забуватися» базові ImageNet-фільтри детектування контурів. Серед спеціалізованих регуляризаторів для запобігання забуванню широкого поширення набули Elastic Weight Consolidation (EWC), що штрафує великі відхилення критично важливих ваг від їхніх початкових значень за допомогою інформаційної матриці Фішера, та Learning without Forgetting (LwF), що зберігає вихід «старої» моделі як додатковий цільовий сигнал [135]. У практиці фільтрації спекл-шуму PCA аналогічну роль виконують малий learning rate перших шарів та warm-up з частковим заморожуванням, які забезпечують пом'якшений перехід до радарного домену.

Окремої уваги заслуговують методи параметрично-ефективного донавчання (Parameter-Efficient Fine-Tuning, PEFT), що набули поширення з появою великих мовних моделей і поступово переносяться у візуальний домен [75]. До них належать адаптерні модулі (Adapter Layers), низькорангова адаптація (LoRA — Low-Rank Adaptation) та м'які промпти (Soft Prompts), які додають у заморожену магістральну мережу невелику кількість тренувальних параметрів. Для задач прогнозування мір якості фільтрації спекл-шуму або параметрів класичних фільтрів такий підхід дозволяє підтримувати єдину базову модель і швидко донавчати «легкі» спеціалізовані голови під кожне нове джерело даних (TerraSAR-X, COSMO-SkyMed, ALOS PALSAR-2) без повного перетренування десятків мільйонів параметрів. Це відкриває перспективу єдиної мультисенсорної РСА-платформи, у якій базова згорткова частина обчислюється один раз, а спеціалізація під конкретну задачу досягається за рахунок компактних адаптерів.

1.6 Гібридна парадигма адаптивної фільтрації спекл-шуму

1.6.1 Синергія математичних моделей та нейромереж (Model-based Deep Learning)

Сучасний розвиток методів цифрової обробки зображень дистанційного зондування Землі (ДЗЗ) характеризується переходом від «чистих» (data-driven) нейромережевих моделей до гібридних архітектур, що об'єднують глибоке машинне навчання з відомими фізико-математичними моделями [78]. Цей підхід, відомий як Model-based Deep Learning (або Physics-informed Neural Networks), виник як відповідь на системні недоліки класичних згорткових мереж типу «чорний ящик». Попри те, що архітектури на зразок U-Net [70] або DnCNN [63] демонструють високі показники пікового відношення сигнал-шум (PSNR) на валідаційних наборах даних, вони часто генерують фізично неможливі артефакти або «галюцинації» при обробці реальних радіолокаційних знімків [142]. Це відбувається через те, що мережа під час навчання мінімізує лише статистичну похибку (наприклад, середньоквадратичне відхилення), не «розуміючи»

когерентної природи радіолокаційного розсіяння та мультиплікативного характеру спекл-шуму [6].

Синергія досягається шляхом жорсткої інтеграції рівнянь, що описують локальну статистику спеклу, безпосередньо у граф обчислень нейронної мережі [143]. У такій парадигмі згорткові шари відповідають не за безпосередній синтез значень пікселів «з нуля», а за вилучення складних просторово-текстурних ознак (feature extraction) та прогнозування керуючих параметрів для математичної моделі фільтрації. Своєю чергою, аналітичний блок фільтрації (наприклад, диференційована модель фільтра Лі або Фроста) гарантує, що результуюче зображення формується у суворій відповідності до законів поширення та розсіяння мікрохвиль. Такий розподіл обов'язків дозволяє поєднати найкращі риси обох підходів: високу здатність нейромереж до аналізу глобального контексту та математичну строгість і прозорість (interpretability) класичних адаптивних фільтрів [17]. Гібридні моделі виявляються значно стійкішими (robust) до змін еквівалентного числа поглядів (ENL) та просторової кореляції шуму.

З методологічної точки зору, Model-based Deep Learning у задачах фільтрації спекл-шуму PCA можна декомпозувати на три основні класи інтеграції фізичних знань [17; 78]. Перший — інтеграція через функцію втрат: до стандартних L1- чи L2-цільових функцій додаються фізично обґрунтовані регуляризатори, що штрафують відхилення статистики залишкового шуму від теоретичного гамма-розподілу (наприклад, KL-дивергенція між емпіричним і теоретичним розподілами) або порушення радіометричного балансу між середніми значеннями вхідного й вихідного зображень. Другий — інтеграція через структуру архітектури: ваги мережі мають фізичну інтерпретацію (коефіцієнти демпфування, рівень регуляризації) або реалізують аналітичні оператори (диференційований фільтр Фроста, оператор Total Variation). Третій — інтеграція через структуру вхідних даних: мережа отримує не лише сире

зображення, а й попередньо обчислені статистики (карту локального коефіцієнта варіації, оцінку ENL), що задає їй фізично коректний контекст.

Окремим підкласом гібридних підходів є так звані Plug-and-Play priors (PnP), які розв'язують обернену задачу фільтрації спекл-шуму шляхом ітеративної мінімізації енергетичного функціоналу, у якому проксимальний оператор регуляризатора замінено на навчений шумопридушувач (denoiser) [79]. Це дозволяє відокремити модель спостереження (фізично коректну) від моделі апріорних знань (нейромережеву), що особливо зручно для роботи з різними PCA-сенсорами без перенавчання основного алгоритму. У контексті радарної фільтрації спекл-шуму PnP-підходи довели свою ефективність як проміжна ланка між повністю аналітичними методами та наскрізними нейромережевими архітектурами, однак вимагають обережного підбору кроку ітерації та критерію зупинки, оскільки фізична модель шуму не гарантує збіжності при довільному виборі регуляризатора.

1.6.2 Глибоке розгортання алгоритмів (Deep Unrolling)

Одним із найефективніших методів реалізації гібридної парадигми є метод глибокого розгортання алгоритмів (Deep Unrolling або Deep Unfolding) [144]. Ідея цього підходу полягає у трансформації класичних ітеративних алгоритмів оптимізації (наприклад, методів проксимального градієнта або ADMM (Alternating Direction Method of Multipliers — метод множників зі змінними напрямками)) у послідовність шарів глибокої нейронної мережі. Кожна ітерація класичного алгоритму, спрямованого на розв'язання оберненої задачі відновлення зображення, моделюється як окремий шар або блок мережі. При цьому параметри алгоритму, які традиційно підбиралися експертно або обчислювалися за фіксованими правилами (наприклад, кроки градієнтного спуску, пороги регуляризації або базисні словники), стають вагами нейромережі, що оптимізуються під час наскрізного навчання за допомогою алгоритму зворотного поширення помилки (backpropagation).

У контексті фільтрації спекл-шуму РСА-зображень глибоке розгортання дозволяє будувати мережі, архітектура яких безпосередньо відображає ймовірнісні моделі гамма-розподілу та апріорні припущення щодо гладкості зображення (наприклад, Total Variation) [65]. Замість використання стандартних згорткових блоків ReLU (Rectified Linear Unit) - Conv - BatchNorm, мережа складається з блоків «оцінки даних» (Data Fidelity) та «проксимальних операторів» (Proximal Operators), які реалізовані у вигляді невеликих підмереж. Глибоке розгортання забезпечує високий рівень інтерпретованості: проаналізувавши виходи проміжних шарів-ітерацій, дослідник може фізично пояснити, як саме мережа поступово зменшує дисперсію шуму та відновлює втрачені контури [145]. Проте цей метод має суттєвий недолік — значну обчислювальну складність та високі вимоги до обсягу відеопам'яті (VRAM), оскільки для досягнення високої якості відновлення часто потрібно розгорнути десятки ітерацій у пам'яті графічного прискорювача.

Серед канонічних прикладів Deep Unrolling, що знайшли застосування в обробці зображень, виділяють LISTA (Learned Iterative Shrinkage-Thresholding Algorithm) — розгортання алгоритму ISTA з навчуваними матрицями та порогами; ADMM-Net — реалізацію методу множників із змінними напрямками для задач реконструкції MRI та PCA; та USRNet (Unfolded Super-Resolution Network) — для надроздільності з відомим оператором деградації [144; 145]. У задачах радарної фільтрації спекл-шуму аналогічні архітектури будуються на основі ітерацій MAP-оцінювання для мультиплікативної моделі: кожна ітерація розкладається на крок оновлення локальної оцінки інтенсивності (data fidelity) та крок регуляризації (нейромережевий проксимальний оператор). Кількість розгорнутих ітерацій зазвичай становить 5–15: цей параметр є компромісом між якістю реконструкції та обчислювальною вартістю.

Цікавою альтернативою явному Deep Unrolling є концепція Deep Image Prior (DIP), яка стверджує, що сама архітектура згорткової мережі (case-specific U-Net) виступає неявним регуляризатором для відновлення зображень із пошкоджених

спостережень [65]. У задачах фільтрації спекл-шуму PCA DIP-підхід не потребує навчальної вибірки взагалі: мережа оптимізується безпосередньо на одному вхідному зашумленому знімку шляхом мінімізації відхилення між вихідним зображенням мережі (поданим на вхід випадковим тензором) і реальним зашумленим спостереженням. Збіжність зупиняється до того, як мережа починає перевчатися на шум. Хоча обчислювальна вартість такого підходу неприйнятна для оперативної обробки, він є цінним інструментом для верифікації коректності інших методів та для роботи з рідкісними сенсорами, для яких неможливо зібрати навчальну вибірку.

1.6.3 Прогнозування параметрів класичних фільтрів за допомогою мереж

Альтернативним, більш обчислювально ефективним шляхом гібридизації є використання нейромереж виключно для прецизійного прогнозування параметрів класичних просторових фільтрів [14]. У цій парадигмі процес фільтрації залишається математично детермінованим (виконується за стандартними рівняннями адаптивної фільтрації), але його ключові невідомі змінні розраховуються динамічно для кожної сцени або навіть для кожного пікселя за допомогою згорткової нейромережі. Це дозволяє подолати головне обмеження класичних методів — необхідність глобального, ручного та часто субоптимального вибору параметрів.

Прикладом є використання компактних архітектур (зокрема MobileNetV2) для вирішення задачі класифікації оптимального розміру вікна ковзання для фільтра Лі [15]. Мережа аналізує зашумлений патч PCA-зображення і прогнозує, який саме розмір вікна (наприклад, 3x3, 5x5 чи 7x7) забезпечить максимальне значення метрик якості (PSNR або FSIM) без розмиття контурів. Іншим напрямком є використання механізмів уваги (Attention Mechanisms), таких як CBAM (Convolutional Block Attention Module) [146], для прогнозування безперервних (континуальних) значень керуючих параметрів. Мережа може аналізувати локальний коефіцієнт варіації та просторові градієнти, генеруючи

двовимірну карту коефіцієнтів демпфування (damping factors) [16]. Такий підхід забезпечує попіксельну адаптацію: в однорідних областях (наприклад, водні поверхні) мережа автоматично встановлює параметри для максимального згладжування, а на межах штучних об'єктів — параметри для повного збереження структурної різкості.

Формальною узагальненою основою таких підходів є концепція гіпермереж (HyperNetworks) та мереж динамічних фільтрів (Dynamic Filter Networks), за якою одна нейронна мережа («гіпермережа») генерує параметри іншої мережі або аналітичного оператора залежно від вхідних даних [145]. У задачах фільтрації спекл-шуму РСА це означає, що традиційні «статичні» ваги фільтра замінюються на динамічно прогнозовані значення, які адаптуються до конкретного просторового контексту. Метод VGG16-FrostNet можна розглядати як частковий випадок цієї парадигми: магістральна нейромережа (VGG16+CBAM) виступає у ролі гіпермережі, що генерує карту параметрів $A_{map}(x, y)$ для аналітичного фільтра Фроста. Порівняно із загальними динамічними мережами, де всі параметри прогнозуються вільно (з ризиком фізично некоректних значень), у запропонованому підході вихід жорстко обмежений діапазоном [0,5; 10,0], що гарантує дотримання фізичних обмежень класичного фільтра.

Концептуально близьким напрямом є застосування методу диференційованого програмування (Differentiable Programming) — загальної парадигми, в межах якої будь-який алгоритм цифрової обробки сигналів може бути виражений у вигляді послідовності диференційованих операцій і вбудований у граф обчислень нейромережі. Це дозволяє оптимізувати параметри традиційних DSP-блоків (фільтрів, перетворень, морфологічних операцій) безпосередньо градієнтним спуском з нейромережевими ознаками [143]. У задачах фільтрації спекл-шуму РСА такі підходи відкривають можливість гібридизації не лише класичних просторових фільтрів (Lee, Frost, Kuan, Gamma MAP), а й частотних методів (вейвлет-фільтрації Donoho, BM3D), морфологічних операторів та статистичних естиматорів MAP. Розроблений у роботі гібридний метод VGG16-

FrostNet є практичним підтвердженням ефективності цієї парадигми для конкретного класу адаптивних експоненційних фільтрів.

1.6.4 Обґрунтування розробки VGG16-FrostNet

Аналіз існуючих парадигм інтеграції класичних фільтрів у нейромережеві архітектури виявив об'єктивну прогалину в сучасних дослідженнях: відсутність наскрізно-навчальних гібридних методів, які б забезпечували попіксельну просторову адаптацію фільтрації без надмірних обчислювальних витрат (як у Deep Unrolling) та без втрати диференційованості процесу. Класичний фільтр Фроста (Frost Filter) [8] є одним із найефективніших інструментів адаптивної локально-статистичної фільтрації завдяки експоненційно спадній ваговій функції, однак його ефективність жорстко обмежена фіксованим глобальним коефіцієнтом демпфування.

Це обґрунтовує доцільність розробки нової архітектури — VGG16-FrostNet [30]. Цей гібридний метод має базуватися на диференційованій математичній моделі фільтра Фроста, безпосередньо вбудованій у граф обчислень нейромережі як фінальний шар активації. Використання попередньо навченої архітектури VGG16 [140] забезпечить ефективне вилучення багатих текстурних ознак за принципом трансферного навчання, а інтеграція механізму просторово-канальної уваги (CBAM) [146] дозволить мережі динамічно генерувати карту коефіцієнтів демпфування для кожного окремого пікселя зображення. Така синергія гарантуватиме як фізичну коректність процесу фільтрації, так і високу ступінь автономної адаптації, що є критично важливим для сучасних сервісно-орієнтованих конвеєрів обробки великих масивів даних ДЗЗ (зокрема місії Sentinel-1).

Зокрема, перенесення ознак (feature transfer) з домену природних оптичних зображень (за допомогою VGG16) у радіолокаційний домен виявилося виправданим кроком, оскільки ранні шари такої мережі стабільно виділяють базові геометричні форми, які зберігаються навіть під шаром інтенсивного

мультиплікативного спеклу [136]. Доповнення такої базової архітектури блоком CBAM дозволяє системі не просто виявляти контури, але й пріоритезувати їхню важливість як у просторовому (просторова увага), так і в спектральному вимірі (канальна увага) [146]. Зрештою, перетворення цього тензора ознак на двовимірну мапу параметру демпфування для математичного ядра Фроста переводить нейромережу зі статусу незрозумілого «чорного ящика» у статус інтелектуального препроцесора, що керує строгим фізичним фільтром. Ця парадигма відкриває шлях до створення нового покоління систем фільтрації спекл-шуму, які поєднують швидкість роботи, математичну прозорість та високу візуальну якість результатів.

1.7 Висновки до розділу 1

За результатами проведеного аналітичного огляду методів придушення спекл-шуму, оцінювання якості РСА-зображень та сучасних парадигм трансферного й гібридного навчання сформульовано такі узагальнені висновки.

1. Фізична природа спекл-шуму як мультиплікативної завади з гамма-розподілом обумовлює системну непридатність класичних методів придушення адитивного шуму (AWGN) для РСА-зображень. Це робить етап спеціалізованої фільтрації спекл-шуму обов'язковою складовою конвеєрів обробки даних Sentinel-1 та інших систем когерентного зондування.

2. Класичні локально-статистичні фільтри (Лі, Фроста, Куана, Gamma MAP) ґрунтуються на строгих математичних моделях і забезпечують високу передбачуваність та інтерпретованість результатів. Проте їхні фіксовані глобальні параметри (розмір ковзного вікна, коефіцієнт демпфування) унеможливлюють просторову адаптивність і призводять до неминучого компромісу між згладжуванням однорідних ділянок та збереженням контурів.

3. Сучасні «чорноящикові» нейромережеві методи (DnCNN, SAR-CNN, U-Net) демонструють високі значення PSNR та SSIM, проте мають принципові обмеження: відсутність фізичної інтерпретованості, схильність до «галюцинацій»

та згладжування точкових відбивачів, чутливість до доменного зсуву і потреба в ідеальних парних еталонах, яких для реальних РСА-даних об'єктивно не існує.

4. Сучасний апарат метрик оцінювання якості — піксельних (MSE, PSNR), структурних (SSIM, MS-SSIM, FSIM), перцептивних (HaarPSI, MDSI), контурних (EPI, FOM) та безреференсних (ENL, M-Index) — є взаємодоповнювальним: жодна окрема метрика не є самодостатньою для радіолокаційного домену, що вимагає комплексного багатовимірного підходу до валідації.

5. Парадигма трансферного навчання та доменної адаптації забезпечує практичний шлях подолання дефіциту еталонів через використання попередньо навчених на ImageNet архітектур (VGG16, MobileNetV2, DenseNet) із подальшим донавчанням на напівсинтетичних РСА-даних. Стратегії гнучкого fine-tuning (поступове розморожування, диференційовані швидкості навчання) дають змогу адаптувати «оптичні» ознаки до специфіки мультиплікативного спеклу.

6. Гібридна парадигма (Model-based Deep Learning, Deep Unrolling, прогнозування параметрів класичних фільтрів) є перспективним напрямом, що поєднує переваги класичної фізично обґрунтованої фільтрації та глибокого навчання. Об'єктивна прогалина існуючих робіт — відсутність наскрізно-навчальних методів попіксельної адаптації коефіцієнтів аналітичного фільтра Фроста, що обґрунтовує доцільність розробки гібридної архітектури VGG16-FrostNet, постановка задач якої деталізована у вступі дисертації.

РОЗДІЛ 2. МЕТОДИ АПРІОРНОГО ПРОГНОЗУВАННЯ ПОКАЗНИКІВ ЯКОСТІ ПІСЛЯ ФІЛЬТРАЦІЇ ТА ПАРАМЕТРІВ КЛАСИЧНИХ ФІЛЬТРІВ ДЛЯ ДАНИХ SENTINEL-1 НА БАЗІ ГЛИБОКИХ НЕЙРОМЕРЕЖ

Другий розділ присвячено розробці методів апріорного оцінювання якості придушення спекл-шуму та автоматизованого вибору параметрів фільтрації для радіолокаційних зображень Sentinel-1. Розглядається проблема відсутності ідеальних еталонів у задачах ДЗЗ та пропонується конвеєр генерації синтетичних навчальних даних на основі гістограмного узгодження. Обґрунтовується доцільність застосування парадигми трансферного навчання глибоких нейронних мереж (DenseNet-121 та MobileNetV2) для вилучення ознак. Описуються архітектури запропонованих моделей, що дозволяють уникати ресурсомісткого повного перебору. Наводяться результати експериментів, які підтверджують ефективність методів прогнозування цільових метрик якості та розмірів вікон.

2.1 Формалізація задачі апріорного прогнозування

2.1.1 Традиційний підхід до оцінювання якості: post-filtering та потреба в еталонах

Традиційно якість фільтрації радіолокаційних зображень із синтезованою апертурою (РСА) оцінюється лише після безпосереднього застосування алгоритмів придушення спекл-шуму (post-filtering). Для кількісної оцінки ефективності використовують метрики, які порівнюють відфільтроване зображення з ідеальним еталоном (clean reference). Проте на практиці в операційних задачах дистанційного зондування Землі (ДЗЗ), таких як моніторинг або розпізнавання об'єктів, еталонні зображення принципово недоступні. Це унеможливлює використання класичних метрик якості (наприклад, PSNR або SSIM) для безпосередньої оцінки результатів фільтрації реальних супутникових знімків. Відсутність можливості об'єктивного оцінювання призводить до того, що вибір оптимального фільтра або його параметрів здійснюється емпірично або на

основі візуального аналізу, що є суб'єктивним, трудомістким і неприйнятним для автоматизованих конвеєрів масової обробки даних [30].

2.1.2 Концепція апріорного оцінювання: переваги та обчислювальна ефективність

Альтернативою традиційному підходу є концепція апріорного прогнозування якості, яка полягає в оцінюванні результатів фільтрації до фактичного виконання алгоритмів. Замість застосування кількох фільтрів та пошуку найкращого результату, пропонується передбачати їхню ефективність безпосередньо за зашумленим зображенням. Це має суттєві переваги з точки зору обчислювальної ефективності. По-перше, уникаються багаторазові прогони фільтрів, що економить час. По-друге, оптимізуються конвеєри обробки: якщо прогнозована якість фільтрації низька (наприклад, для урбанізованих територій, де втрачаються деталі), систему можна налаштувати на пропуск цього етапу. Отже, апріорне прогнозування виступає інструментом прийняття рішень для раціонального розподілу ресурсів хмарних платформ ДЗЗ.

2.1.3 Математична постановка задачі прогнозування абсолютних метрик якості

Формально задачу апріорного прогнозування можна подати як знаходження функції відображення f між простором вхідних зашумлених зображень X та простором цільових метрик якості Q. Нехай I_{noisy} — вхідне РСА-зображення зі спекл-шумом, а H_k — оператор k-го адаптивного фільтра. Після фільтрації отримаємо зображення $I^{(k)}_{filtered} = H_k(I_{noisy})$. Цільова метрика якості (наприклад, PSNR) між відфільтрованим зображенням та недоступним еталоном I_{ref} становить $q_k = \mathrm{Q}(I^{(k)}_{filtered}, I_{ref})$. Наше завдання полягає у побудові нейромережевої моделі-предиктора M_k, здатної на основі лише зашумленого зображення передбачити абсолютне значення метрики: $\hat{q}_k = M_k(I_{noisy})$ так, щоб мінімізувати похибку прогнозування $|\hat{q}_k - q_k|$. Порівняно з існуючими методами, які намагаються

прогнозувати лише відносний приріст (індекс покращення), пряме прогнозування абсолютних значень забезпечує можливість коректного порівняння різнотипних фільтрів між собою.

2.1.4 Обґрунтування вибору цільових метрик (PSNR, WSNR, SSIM, MS-SSIM, FSIM)

Для комплексного оцінювання якості відновлення РСА-зображень у роботі обрано п'ять незалежних метрик. Пікове відношення сигналу до шуму (PSNR) характеризує загальний рівень похибки реконструкції на піксельному рівні [80]. Оскільки PSNR недостатньо добре корелює з візуальним сприйняттям текстур, додатково використовується зважене відношення сигналу до шуму (WSNR), яке враховує частотну чутливість. Індекс структурної подібності (SSIM) вимірює збереження внутрішніх структур, яскравості та контрасту [81], що є критичним для задач сегментації. Багатомасштабний індекс структурної подібності (MS-SSIM) [85] розширює цей підхід, забезпечуючи стійкість до зміни масштабу та роздільної здатності. Нарешті, індекс подібності на основі ознак (FSIM) [86] оцінює якість збереження низькорівневих контурів (фазова конгруентність та градієнтна амплітуда), що робить його вкрай чутливим до розмиття країв, яке є типовим недоліком багатьох адаптивних фільтрів. Спільне використання цих метрик дозволяє моделі вивчати різноманітні аспекти деградації зображень.

$$\mathrm{WSNR} = 10\log_{10}\left(\frac{\sum_{c,h,w} |S_{c,h,w}|^2 W_{h,w}}{\sum_{c,h,w} |N_{c,h,w}|^2 W_{h,w}}\right), \tag{26}$$

де $S_{c,h,w}$ — спектральні компоненти корисного сигналу, $N_{c,h,w}$ — спектральні компоненти шуму (різниці між еталонним та відновленим зображенням), а $W_{h,w}$ — вагова функція контрастної чутливості зорової системи. У практичній реалізації $S_{c,h,w}$ отримують як двовимірне дискретне перетворення Фур'є еталонного зображення, а $N_{c,h,w}$ — як спектр похибки між еталонним та відновленим

зображеннями. Функція $W_{h,w}$ задає частотне вагування відповідно до контрастної чутливості зорової системи людини. На відміну від PSNR, ця метрика штрафує помилки з урахуванням їх частотної локалізації, тому краще відображає візуальну помітність залишкового спеклу та артефактів надмірного згладжування.

2.1.5 Формалізація задачі вибору параметрів фільтрації (на прикладі вікна Лі)

Окрім прогнозування якості, важливою складовою автоматизації є вибір параметрів фільтрації. Ефективність класичних фільтрів (зокрема фільтра Лі) критично залежить від розміру просторового вікна. Малі вікна (наприклад, 3×3) добре зберігають контури об'єктів, але забезпечують низький ступінь придушення шуму. Великі вікна (наприклад, 11×11) ефективно згладжують шум в однорідних областях, проте спричиняють значне розмиття дрібних деталей і границь. Оптимальний розмір вікна є функцією локальної та глобальної гетерогенності сцени. Задачу апріорного вибору можна сформулювати так: для даного вхідного зображення I_{noisy} необхідно знайти розмір вікна $w^* \in W = \{3, 5, 7, 9, 11\}$, який максимізує цільову метрику якості (наприклад, PSNR або SSIM), без необхідності багаторазової обробки зображення для всіх можливих значень вікон. Такий підхід мінімізує ризики неправильного налаштування конвеєра [15].

2.1.6 Постановка задачі як багатокласової класифікації замість регресії

У попередніх дослідженнях визначення оптимального розміру вікна найчастіше вирішувалося за допомогою двоетапної регресійної схеми [14], [15]. Спочатку прогнозувалися значення метрик якості для кожного можливого розміру вікна, а потім обирався той розмір, для якого прогнозоване значення було максимальним. Однак такий підхід має суттєвий недолік — накопичення похибок. Якщо похибка регресії для одного з вікон виявиться значною, алгоритм обере хибний параметр, що призведе до субоптимальної фільтрації. Для подолання цього обмеження запропоновано переформулювати завдання вибору параметрів

як задачу прямої багатокласової класифікації. Кожен можливий розмір вікна розглядається як окремий клас (наприклад, клас 0 для 3×3, клас 1 для 5×5 тощо). Нейронна мережа навчається безпосередньо ставити у відповідність вхідному зображенню найбільш відповідний клас. Це значно спрощує архітектуру системи та підвищує загальну надійність прогнозування.

2.1.7 Роль глибоких ознак у виявленні текстурних властивостей

Традиційні методи апріорного оцінювання базувалися на розрахунку експертно розроблених статистичних ознак (наприклад, локальної дисперсії, коефіцієнта варіації, просторової кореляції) та використанні класичних моделей машинного навчання (штучних нейронних мереж прямого поширення або методу опорних векторів) [13]. Хоча такий підхід працює, його можливості обмежені фіксованим набором ознак, які не завжди здатні вловити складні просторово-структурні патерни міських або лісових масивів на РСА-зображеннях. Водночас глибокі згорткові нейронні мережі (CNN) забезпечують можливість ієрархічного вилучення ознак [136]. Перші шари мережі реагують на прості градієнти та спекл-шум, тоді як глибші шари виявляють складну макротекстуру. Завдяки механізму зворотного поширення помилки мережа самостійно оптимізує набір фільтрів для виявлення тих просторових властивостей зображення, які найбільше корелюють із результативністю фільтрації спекл-шуму.

2.1.8 Обґрунтування застосування трансферного навчання для РСА-зображень

Головною перешкодою для навчання глибоких CNN на РСА-даних є відсутність масивних анотованих датасетів [134]. Навчання глибокої моделі з нуля вимагає сотень тисяч зображень через ризик перенавчання. Для вирішення проблеми використано парадигму трансферного навчання (Transfer Learning) [135]. Базова модель попередньо навчається на оптичних зображеннях загального призначення (ImageNet) [60], набуваючи навичок розпізнавання контурів і

текстур. Далі ваги мережі адаптуються шляхом донавчання (fine-tuning) на меншому датасеті РСА-зображень [136]. Оскільки прогнозування якості базується на аналізі структурних особливостей, спільних для оптичного та радарного діапазонів, перенесення знань дозволяє прискорити збіжність і підвищити точність роботи предиктора на скромних навчальних вибірках.

2.1.9 Загальна архітектура пропонованого рішення

На основі сформульованих підходів загальна архітектура системи апріорного оцінювання розгалужується на два паралельні конвеєри. Для прогнозування абсолютних метрик якості розроблено регресійну модель на базі архітектури DenseNet-121. Вона здатна з високою точністю передбачати метрики PSNR, SSIM тощо для різних класичних фільтрів. Для задачі прямої класифікації оптимального розміру вікна запропоновано легковагову архітектуру на базі MobileNetV2. Вона орієнтована на мінімізацію обчислювальних витрат, що робить її ідеальною для швидкого інференсу під час потокової обробки даних на пристроях із обмеженими ресурсами або супутникових платформах. Обидві моделі навчаються на спеціально згенерованих наборах напівсинтетичних даних, які імітують фізичні властивості спекл-шуму Sentinel-1. Детальний опис методів генерації даних і побудови цих моделей наведено в наступних підрозділах.

2.2 Генерація напівсинтетичного навчального корпусу

2.2.1 Проблема відсутності парних РСА-даних для навчання

Для ефективного навчання глибоких нейромережевих моделей, зокрема для задач прогнозування якості фільтрації, критичною вимогою є наявність великих наборів парних зображень: зашумленого вхідного знімка та відповідного йому ідеально чистого еталона (ground truth) [17]. Проте у випадку радіолокації із синтезованою апертурою (РСА) отримання таких парних даних на практиці є неможливим. Спекл-шум має фізичну природу і присутній у кожному випроміненому імпульсі через когерентну інтерференцію радіохвиль [5].

Застосування класичних алгоритмів придушення шуму для створення квазі-еталонів також не вирішує проблему, оскільки будь-яка фільтрація неминуче вносить власні артефакти: розмиває текстуру або згладжує контури об'єктів [6]. Отже, використання відфільтрованих знімків як еталонів призведе до того, що нейромережа вивчатиме не структуру самої сцени, а специфічні похибки конкретного фільтра. Це зумовлює необхідність розробки методу генерації синтетичного навчального корпусу, який би поєднував реалістичну статистику РСА-даних із наявністю ідеального еталона без шуму.

2.2.2 Вибір оптичних зображень Sentinel-2 (NIR діапазон) як базису

Для вирішення проблеми відсутності чистих радарних еталонів у роботі запропоновано використовувати оптичні супутникові знімки як структурний базис [30]. Як джерело даних обрано мультиспектральні зображення Sentinel-2 (рівень обробки L2A), зокрема канал B8 (ближній інфрачервоний діапазон, NIR). Вибір NIR-діапазону обґрунтований тим, що він краще, ніж видимий спектр, передає структурні особливості підстилаючої поверхні (зокрема рослинності та водних об'єктів), які також є виразними на радарних знімках. Найголовнішою перевагою оптичних знімків є їхня висока просторова роздільна здатність (10 м/піксель) та повна відсутність мультиплікативного спекл-шуму. Це дозволяє використовувати їх як надійний, ідеально чистий еталон для обчислення метрик якості (PSNR, SSIM тощо). Завантаження відповідних сцен Sentinel-1 (радарних) [49] та Sentinel-2 (оптичних) на одні й ті самі географічні регіони здійснювалося з відкритої платформи Copernicus Data Space Ecosystem [2], що забезпечило репрезентативність вибірки ландшафтів.

2.2.3 Алгоритм гістограмного узгодження оптичних даних із РСА-розподілами

Попри переваги оптичних зображень як структурного еталона, їхня радіометрична статистика (розподіл яскравостей) кардинально відрізняється від

радарних знімків. Для адаптації даних розроблено алгоритм гістограмного узгодження (histogram matching) [30]. На першому етапі зашумлені радарні знімки Sentinel-1 піддавалися початковій фільтрації фільтром Лі (7×7) [7], щоб отримати усереднену амплітудну статистику без надмірних флуктуацій. Далі для кожного оптичного знімка Sentinel-2 підбирався найбільш релевантний радарний знімок. Для цього оптичні та радарні сцени кластеризувалися за статистичними ознаками (середнім і дисперсією) алгоритмом k-means (оптимальна кількість кластерів $k=56$). Після співставлення пар виконувалося непараметричне монотонне перетворення $I'_{hm}(x,y) = F_R^{-1}(F_I(I(x,y)))$, де F_I і F_R — кумулятивні функції розподілу оптичного та радарного знімків відповідно. Це перетворення зберігає просторову структуру оптичного знімка, але приводить його радіометрію до типового радарного розподілу. Схему етапів конвеєра наведено на рис. 2.1.

Рис. 2.1. Схема процесу стилізації оптичних зображень Sentinel-2 NIR у PCA-подібні [30]

2.2.4 Математична модель динамічного додавання мультиплікативного спекл-шуму

Наступним кроком є внесення реалістичного спекл-шуму до отриманих стилізованих еталонів. У роботі використано фізично обґрунтовану модель мультиплікативного шуму в домені інтенсивностей [3]: $Y(x,y) = X(x,y) \cdot S(x,y)$, де X — інтенсивність стилізованого еталона, а S — мультиплікативний множник, що підпорядковується гамма-розподілу з математичним сподіванням 1 та дисперсією $1/L$ (де L — еквівалентне число поглядів, ENL) [6]. Щоб забезпечити стійкість

моделі (робастність) до різних рівнів зашумленості, параметр ENL обирався випадковим чином із діапазону [2, 6] для кожного зображення на кожній епосі навчання. Окрім того, для відтворення низькочастотних просторових кореляцій шуму, які виникають під час обробки сирих даних PCA, генерація шумової матриці S здійснювалася на сітці зниженої роздільної здатності з подальшим білінійним масштабуванням до вихідного розміру. Порівняння розподілів яскравості подано на рис. 2.2.

У формалізованому вигляді процедура генерації шуму та спільної нормалізації описується такими рівняннями.

$$H_s = \lfloor sH \rfloor, \quad W_s = \lfloor sW \rfloor, \quad S_{\downarrow}(i,j) \sim \Gamma(L, 1/L), \tag{27}$$

$$(i,j) \in \{1, \dots, H_s\} \times \{1, \dots, W_s\}, \tag{28}$$

$$\tilde{S} = \mathcal{U}_{\text{bilinear}}(S_{\downarrow}; H, W), \quad \mu_{s,\downarrow} = \text{mean}(S_{\downarrow}), \quad \sigma_{s,\downarrow} = \text{std}(S_{\downarrow}), \tag{29}$$

$$\tilde{\mu}_s = \text{mean}(\tilde{S}), \quad \tilde{\sigma}_s = \text{std}(\tilde{S}), \tag{30}$$

$$S = \frac{\tilde{S} - \tilde{\mu}_s}{\tilde{\sigma}_s} \sigma_{s,\downarrow} + \mu_{s,\downarrow}, \tag{31}$$

$$\tilde{I} = \sqrt{I^2 \cdot S}, \quad L \sim \mathcal{U}(L_{\min}, L_{\max}), \quad s \sim \mathcal{U}(s_{\min}, s_{\max}), \tag{32}$$

$$\alpha = \left(\max_{c,h,w} \tilde{I}_{c,h,w}\right)^{-1}, \quad I_{\text{noisy}} = \alpha \tilde{I}, \quad I_{\text{reference}} = \alpha I. \tag{33}$$

де s — масштаб просторової кореляції спеклу, H_s та W_s — розміри зменшеної шумової сітки, $S_{\downarrow}$ — низькорозмірне шумове поле, згенероване незалежними вибірками з гамма-розподілу, $\tilde{S}$ — його білінійно масштабована версія, а $\mu_{s,\downarrow}$, $\sigma_{s,\downarrow}$, $\tilde{\mu}_s$, $\tilde{\sigma}_s$ — середні та середньоквадратичні відхилення до і після масштабування, що використовуються для відновлення статистики шуму. Після цього S застосовується до інтенсивності I^2, виконується зворотний перехід в амплітудний домен через $\tilde{I} = \sqrt{(I^2 \cdot S)}$, а коефіцієнт α забезпечує спільну нормалізацію. Величини I_{noisy} та $I_{reference}$ — це зашумлене й еталонне зображення після приведення до спільного динамічного діапазону. Така побудова дозволяє окремо контролювати ENL, просторову кореляцію зерен спеклу та числову стабільність навчання.

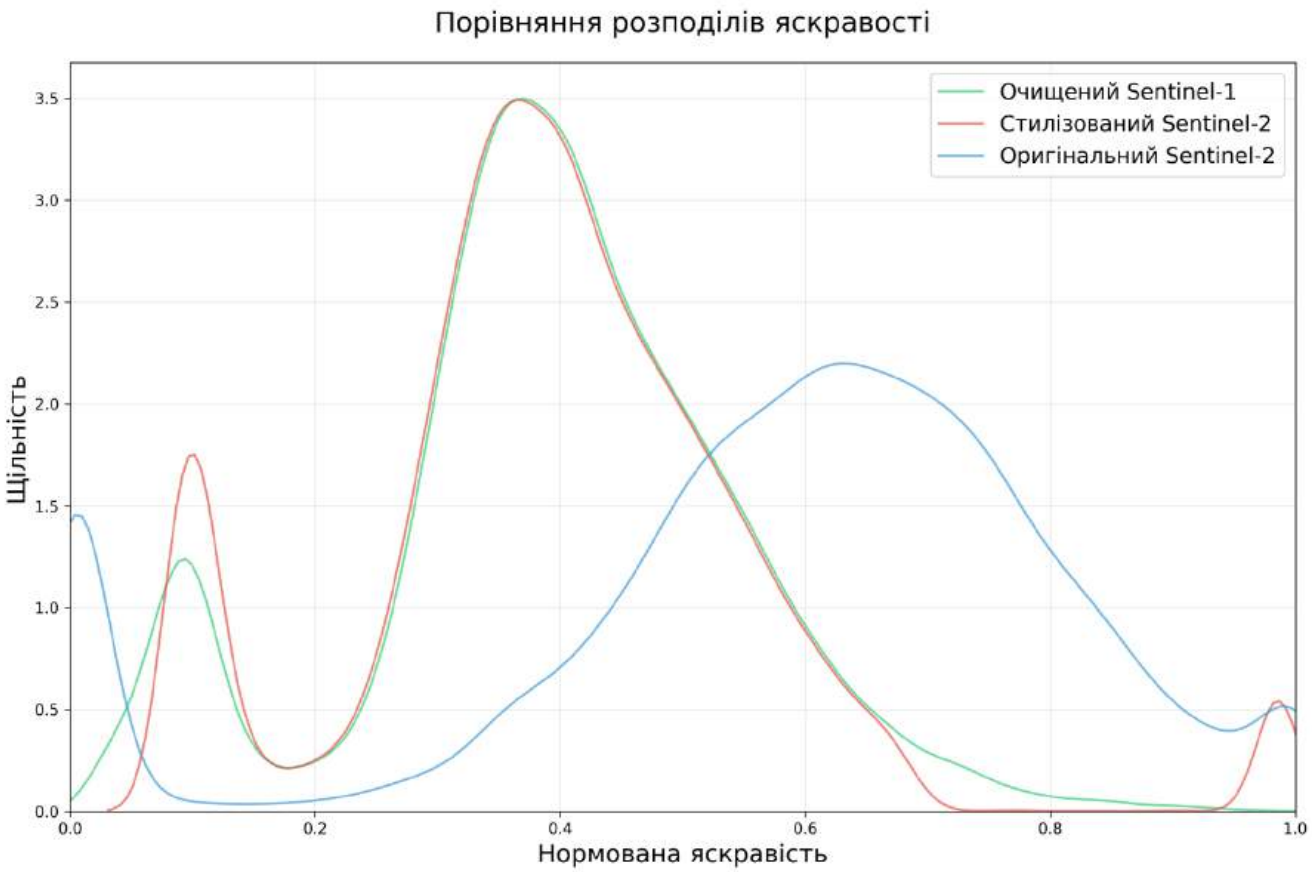


Рис. 2.2. Порівняння розподілів яскравості: відфільтрованого S1-SAR, стилізованого S2-NIR та оригінального S2-NIR [30]

2.2.5 Процедура нарізки на фрагменти (патчі) та аугментація даних

Повнорозмірні супутникові сцени мають занадто високу роздільну здатність для прямого завантаження в пам'ять графічного процесора. Тому стилізовані зображення нарізалися на менші фрагменти (патчі) заданого розміру 1024×1024 (як для DenseNet-121, так і для MobileNetV2). Для запобігання перенавчанню застосовано аугментацію: випадкові відображення (flip) та обертання на кути, кратні 90 градусам. Оскільки перетворення зберігають статистичні властивості шуму та геометрію об'єктів, їх використання ефективно розширює вибірку. Генерація шуму здійснювалася безпосередньо для аугментованих патчів під час формування батчу, гарантуючи унікальність шумового патерну на кожній епосі.

2.2.6 Конвеєр обчислення еталонних цільових змінних (label generation)

Створення навчального корпусу завершується обчисленням міток (labels) — цільових змінних для мережі. До згенерованого зашумленого патча застосовувалися класичні фільтри (Лі, Фрост тощо) з вікнами від 3×3 до 11×11. Відфільтровані зображення порівнювалися зі стилізованими патчами без шуму. Для регресійної задачі записувалися абсолютні значення метрик якості (PSNR, WSNR, SSIM, MS-SSIM, FSIM). Для задачі класифікації обчислювалися метрики для всіх розмірів вікон певного фільтра, а міткою ставав індекс вікна з максимальним результатом. Такий алгоритм гарантує точну відповідність між зашумленим входом та еталоном навчання. Загальну схему порівняння прогнозів моделі з результатами еталонної фільтрації на тестовій вибірці наведено на рис. 2.3.

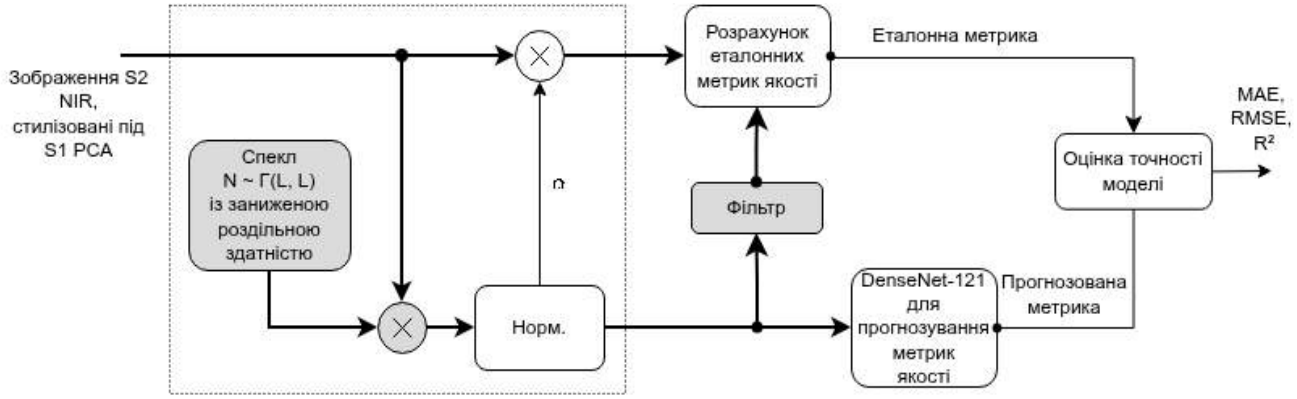


Рис. 2.3. Схема валідації: порівняння прогнозів моделі з реальними результатами фільтрації [30]

2.2.7 Аналіз репрезентативності та статистичних властивостей створеного датасету

У результаті роботи конвеєра було сформовано репрезентативний масив даних, що налічує тисячі унікальних патчів. Оцінка ядерної щільності (Kernel Density Estimation) підтвердила, що розподіл яскравостей синтетичного набору даних практично збігається з емпіричним розподілом реальних відфільтрованих знімків Sentinel-1. Отриманий корпус містить різноманітні ландшафтні типи (міська забудова, сільськогосподарські угіддя, водойми, ліси), що забезпечує моделі високу здатність до узагальнення. Внесення динамічного шуму на етапі навчання гарантує, що модель ніколи не зустріне двічі один і той самий зашумлений патч, усуваючи ефект запам'ятовування. Запропонована методологія успішно вирішила проблему відсутності парних даних, створивши міцну основу для подальшого застосування алгоритмів глибокого трансферного навчання для прогнозування якості PCA-зображень. Приклади фрагментів на різних етапах конвеєра ілюструє рис. 2.4.

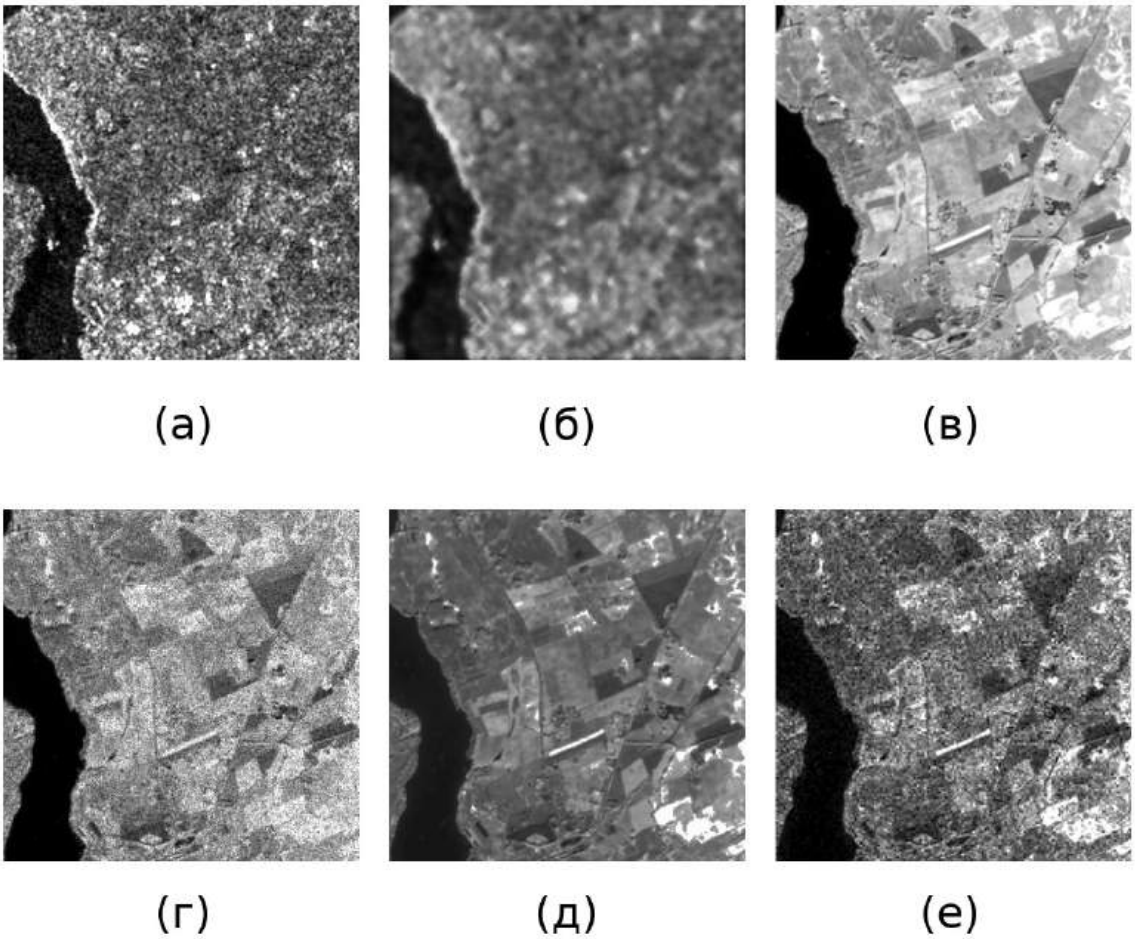


Рис. 2.4. Фрагменти S1-SAR і стилізованих S2-NIR зображень на різних етапах конвеєра: а) оригінальний фрагмент S1-SAR; б) фрагмент S1-SAR після фільтрації Лі; в) оригінальний фрагмент S2-NIR; г) оригінальний фрагмент S2-NIR із додаванням спекл-шуму; д) стилізований фрагмент S2-NIR; е) стилізований фрагмент S2-NIR із додаванням спекл-шуму [30]

2.3 Прогнозування метрик якості на основі адаптованої DenseNet-121

2.3.1 Особливості та переваги архітектури DenseNet-121

Для вирішення задачі прогнозування абсолютних метрик якості фільтрації обрано архітектуру DenseNet-121 (Densely Connected Convolutional Networks) [141]. Її головною особливістю є щільна зв'язність: кожен шар мережі отримує на вхід карти ознак (feature maps) від усіх попередніх шарів і передає свої карти

ознак усім наступним шарам. Це принципово відрізняє її від класичних послідовних архітектур та мереж типу ResNet [62]. Завдяки такій структурі DenseNet пом'якшує проблему згасаючого градієнта під час навчання та значно ефективніше використовує параметри. Для аналізу радіолокаційних зображень збереження ознак низького рівня (які містять інформацію про тонку текстуру та інтенсивність спекл-шуму) на глибоких стадіях обробки є критично важливим. Глибока архітектура зі 121 шаром дозволяє моделі формувати складну ієрархію просторових абстракцій, що необхідно для точного передбачення метрик якості без наявності еталонного зображення на етапі інференсу.

2.3.2 Адаптація вхідних шарів для одноканальних РСА-зображень

Оскільки базова модель DenseNet-121 попередньо навчена на наборі даних ImageNet [60], [61], вона розрахована на обробку триканальних кольорових зображень (RGB). Для адаптації мережі до одноканальних РСА-зображень (у відтінках сірого) необхідно було модифікувати вхідний шар *Conv2d(3, 64, 7×7)*. Щоб не втратити попередньо вивчені фільтри, застосовано метод усереднення ваг: ваги для трьох каналів усереднювалися та згорталися в єдиний канал, формуючи новий шар *Conv2d(1, 64, 7×7)*. Такий підхід зберігає універсальні детектори градієнтів, отримані на ImageNet. Додатковою структурною зміною стала заміна агресивного пулінгу *MaxPool2d(3, 2, 1)* на *AvgPool2d(3, 2, 1)*. Максимальний пулінг ігнорує більшість локальної інформації, виділяючи лише найяскравіші відблиски, тоді як усереднювальний пулінг краще зберігає текстурні патерни, що є життєво необхідним для аналізу впливу спекл-шуму. Порівняльні характеристики оригінальної архітектури DenseNet-121 та модифікованої версії зведено в табл. 2.1

Таблиця 2.1 – Порівняння архітектури DenseNet-121 (ImageNet) та адаптованої версії для одноканальних РСА-зображень

Параметр	Оригінал (ImageNet)	Модифікована (S1-SAR)
Parameter		Input:
Size	224×224	1024×1024
Channels	3 (RGB)	1 (grayscale)
Type	Photo	SAR
Tensor	[b,3,224,224]	[b,1,1024,1024]
Tensor		Stem:
Conv	Conv2d(3,64,7,2,3)	Conv2d(1,64,7,2,3)
Bias	True	False
Pool	MaxPool2d(3,2,1)	AvgPool2d(3,2,1)
Output	[b,64,56,56]	[b,64,256,256]
Output		Dense blocks:
DB1 input	[b,64,56,56]	[b,64,256,256]
DB1 output	[b,256,56,56]	[b,256,256,256]
TL1	[b,128,28,28]	[b,128,128,128]
DB2 input	[b,128,28,28]	[b,128,128,128]
DB2 output	[b,512,28,28]	[b,512,128,128]
TL2	[b,256,14,14]	[b,256,64,64]
DB3 input	[b,256,14,14]	[b,256,64,64]
DB3 output	[b,1024,14,14]	[b,1024,64,64]
TL3	[b,512,7,7]	[b,512,32,32]
DB4 input	[b,512,7,7]	[b,512,32,32]
DB4 output	[b,1024,7,7]	[b,1024,32,32]
DB4 output		Classifier:
GAP input	[b,1024,7,7]	[b,1024,32,32]
GAP output	[b,1024]	[b,1024]
Linear in	1024	1024
Linear out	1000	1
Task	Classification	Regression

2.3.3 Модифікація класифікаційної голови для прогнозування 5 метрик

Оригінальна архітектура DenseNet-121 завершується повнозв'язним шаром, налаштованим на класифікацію зображень (1000 класів у ImageNet). Оскільки наше завдання полягає у прогнозуванні безперервних числових значень метрик якості (PSNR, WSNR, SSIM, MS-SSIM, FSIM), класифікаційна голова мережі була замінена. Замість неї інтегровано лінійний регресійний шар *Linear(1024, 1)*, який приймає на вхід вектор глибоких ознак після глобального усереднювального пулінгу (Global Average Pooling) і видає одне скалярне значення. Для кожної комбінації «фільтр – метрика» навчався окремий екземпляр моделі (наприклад, одна модель для PSNR фільтра Фроста, інша – для SSIM фільтра Лі). Такий підхід дозволив мережам вузько спеціалізуватися на особливостях впливу конкретного алгоритму фільтрації на певну структуру зображення, що забезпечило високу точність прогнозування (див. схему валідації на рис. 2.3).

2.3.4 Стратегія та гіперпараметри повного донавчання

Процес адаптації переднавченої моделі до нових даних виконувався за стратегією повного донавчання (full fine-tuning). Це означає, що ваги всіх шарів мережі були «розморожені» й оновлювалися під час зворотного поширення похибки. Як функцію втрат було обрано середньоквадратичну похибку (MSE), яка штрафує великі відхилення прогнозу від істинного значення. Для оптимізації використано алгоритм Adam з початковою швидкістю навчання 10^{-4}. Тренування проводилося протягом 120 епох з розміром пакету (batch size) 4 та накопиченням градієнтів (gradient accumulation) у 2 кроки, що дало ефективний розмір пакету (batch size) 8 без перевищення доступної відеопам'яті GPU. Динамічне додавання мультиплікативного спекл-шуму на кожній епосі гарантувало стійкість моделі до варіювання рівня зашумленості в межах синтетичного навчального корпусу (ENL у діапазоні від 2 до 6). Динаміку функції втрат під час тренування відображає рис. 2.5.

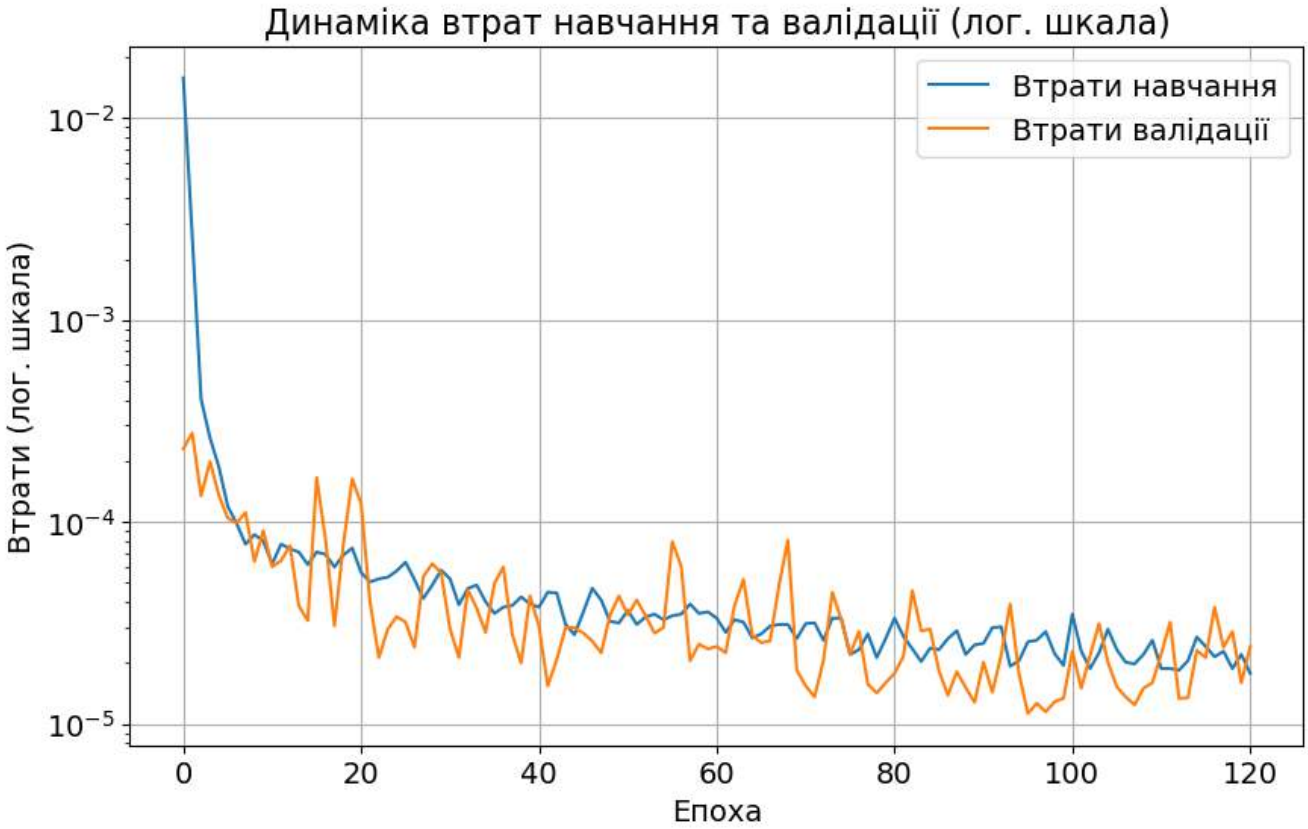


Рис. 2.5. Динаміка тренування DenseNet-121 (фільтр Куана, метрика WSNR): значення функції втрат на тренувальній і тестовій вибірках [30]

2.4 Прогнозування оптимального розміру вікна фільтра Лі на базі MobileNetV2

2.4.1 Вплив розміру вікна на компроміс між згладжуванням та збереженням деталей

Фільтр Лі є одним із найпоширеніших методів локальної статистичної обробки для придушення спекл-шуму на РСА-зображеннях. Проте його ефективність не є статичною величиною; вона критично залежить від вибору просторового розміру скануючого вікна обробки. Розмір вікна безпосередньо визначає ключовий компроміс між інтенсивністю згладжування шуму та здатністю зберігати дрібні просторові деталі й чіткість контурів. Якщо обрати занадто мале вікно (наприклад, 3×3 пікселі), то кількість відліків, за якими

обчислюється локальна статистика, буде недостатньою. Внаслідок цього залишковий рівень флуктуацій залишатиметься високим, і шум не буде ефективно придушений. Водночас надмірно велике вікно (наприклад, 11×11 пікселів) кардинально зменшує дисперсію шуму, але ціною розмиття структурних меж об'єктів. Тому фіксований глобальний розмір вікна для гетерогенних сцен завжди є субоптимальним рішенням.

2.4.2 Аналіз недоліків двоетапних регресійних моделей вибору вікна

У попередніх дослідженнях завдання автоматичного вибору оптимального розміру вікна зазвичай розв'язувалося за допомогою моделей прогнозування якості. Суть такого двоетапного підходу полягала в наступному: спочатку нейромережа або регресійна модель прогнозувала значення метрик якості (наприклад, PSNR) для кожного можливого розміру вікна окремо, а після цього виконувався пошук максимуму (argmax) серед спрогнозованих значень для визначення оптимального параметра. Попри життєздатність концепції, цей підхід має серйозні недоліки для оперативного застосування. По-перше, він вимагає оцінювання кількох регресійних функцій (або виконання кількох прогонів), що збільшує обчислювальні витрати. По-друге, незалежні похибки прогнозування кожної метрики можуть накладатися, що іноді призводить до помилкового вибору глобального оптимуму навіть при незначних відхиленнях.

2.4.3 Архітектура MobileNetV2: баланс між точністю та обчислювальною ефективністю

Для вирішення проблеми було розроблено метод прямої багатокласової класифікації, який миттєво передбачає оптимальний розмір вікна [15]. Оскільки метод призначений для автоматизації конвеєрів обробки, критичним критерієм стала обчислювальна ефективність моделі. З огляду на це як базову архітектуру (backbone) було обрано MobileNetV2. Ця нейромережа розроблена спеціально для мобільних пристроїв і платформ з обмеженими апаратними ресурсами. Завдяки

використанню глибинних сепарабельних згорток (depthwise separable convolutions) та інвертованих залишкових блоків (inverted residual blocks), MobileNetV2 забезпечує кардинальне зменшення кількості параметрів та обсягу використовуваної відеопам'яті (до 60% менше порівняно з DenseNet-121 [141]). При цьому вона зберігає достатню репрезентативну здатність для вилучення складних текстурних ознак, необхідних для розпізнавання патернів шуму.

2.4.4 Формування міток класів (розмірів вікна) шляхом повного перебору

Для навчання моделі-класифікатора було необхідно сформувати точні еталонні мітки (ground-truth labels). Оскільки нашою метою був вибір одного з п'яти розмірів вікна (3×3, 5×5, 7×7, 9×9, 11×11), для кожного зашумленого синтетичного зображення виконувався повний перебір: фільтр Лі застосовувався послідовно з усіма зазначеними вікнами. Далі кожне відфільтроване зображення порівнювалося з чистим еталоном (стилізованим оптичним знімком Sentinel-2) з розрахунком обраної метрики якості (наприклад, PSNR або FSIM). Оптимальним розміром вікна (еталонним класом) для даного зображення вважався той, що забезпечив максимальне значення обраної метрики якості. Отже, кожне зашумлене зображення отримало мітку від 0 до 4, що перетворило задачу пошуку параметрів на стандартну задачу розпізнавання класів.

$$y = \mathrm{argmax}_{w \in W} M_w, \quad W = \{3,5,7,9,11\}. \tag{34}$$

Для всіх метрик подібності та якості, де більше значення означає кращий результат (PSNR, SSIM, MS-SSIM, FSIM, HaarPSI), використовується саме оператор *argmax*. Винятком є MDSI: оскільки менше значення цієї метрики відповідає вищій якості відновлення, для неї мітку формують за правилом $y = argmin_{w \in W} M_w$.

Тут M_w — значення вибраної метрики якості для фільтра Лі з вікном w, а y — еталонна мітка класу. Така постановка безпосередньо пов'язує шумову

структуру вхідного зображення з оптимальним розміром просторового вікна без проміжного етапу регресії. Розподіл отриманих оптимальних розмірів вікон за різними метриками наведено у табл. 2.2

Таблиця 2.2 – Розподіл оптимальних розмірів вікна фільтра Лі за метриками якості (частка та кількість зображень)

Метрика	3×3	5×5	7×7	9×9	11×11
PSNR	4,4% (67)	27,6% (421)	34,7% (530)	20,2% (309)	13,2% (201)
SSIM	6,5% (100)	35,3% (540)	32,5% (497)	17,7% (270)	7,9% (121)
MS-SSIM	4,8% (74)	28,2% (431)	30,0% (459)	21,4% (327)	15,5% (237)
FSIM	2,9% (45)	21,3% (325)	29,1% (445)	23,1% (353)	23,6% (360)
HaarPSI	2,5% (38)	21,3% (326)	34,0% (520)	25,0% (382)	17,1% (262)
MDSI	2,4% (36)	15,9% (243)	27,4% (418)	28,1% (429)	26,3% (402)

2.4.5 Адаптація MobileNetV2 до задачі 5-класової класифікації

Подібно до регресійної задачі, оригінальна архітектура MobileNetV2 була адаптована під специфіку РСА-даних. Вхідний шар змінено з 3-канального на 1-канальний (відтінки сірого) за допомогою усереднення ваг, щоб обробляти радіолокаційну амплітуду. Оригінальну голову класифікації для 1000 класів ImageNet було видалено. Замість неї, після шару глобального усереднювального пулінгу (AdaptiveAvgPool), інтегровано нову компактну структуру: шар регуляризації Dropout (ймовірність 0.3), лінійний шар на 512 нейронів, функція активації ReLU, ще один Dropout (0.2) і, нарешті, лінійний вихідний шар на 5 нейронів (для 5 класів вікон). Така структура голови суттєво знижує ризик перенавчання (overfitting), стабілізує оцінку апостеріорної ймовірності та підвищує загальну стійкість моделі до варіацій просторового шуму.

2.4.6 Стратегія навчання та боротьба з дисбалансом класів

Навчання моделі проводилося за парадигмою повного донавчання (full fine-tuning). Під час формування датасету виявилася серйозна проблема: розподіл оптимальних вікон був сильно незбалансованим. Наприклад, для метрики PSNR вікно 7×7 виявлялося найкращим майже у 35% випадків, тоді як вікно 11×11 — лише у 13%. Щоб запобігти зсуву моделі в бік мажоритарних класів, як функцію втрат було застосовано зважену перехресну ентропію (Weighted Cross-Entropy). Вага кожного класу обчислювалася як величина, обернено пропорційна частоті його появи у навчальній вибірці. Цей механізм примушував мережу приділяти більшу увагу рідкісним класам, що забезпечило збалансовану точність розпізнавання і гарантувало, що MobileNetV2 навчиться розрізняти справжні текстурні сигнатури, а не просто запам'ятає найчастішу відповідь.

$$\mathcal{L} = -w_y \log(\hat{p}_y), \quad \hat{p}_i = \frac{\exp(z_i)}{\sum_{j=1}^{5} \exp(z_j)}, \tag{35}$$

$$w_i = \frac{N}{K n_i}. \tag{36}$$

де z_i — логіт для i-го класу, $\hat{p}_i$ — відповідна ймовірність після softmax, n_i — кількість прикладів цього класу у навчальній вибірці, $K = 5$ — число класів, а N — загальна кількість навчальних прикладів. Отже, що рідкісніший клас, то більшу вагу він отримує у функції втрат.

2.5 Експериментальні результати та їх обговорення

2.5.1 Постановка експерименту та статистичні критерії оцінювання

Для перевірки ефективності запропонованої регресійної моделі на базі DenseNet-121 було проведено масштабне експериментальне дослідження.

Синтетичний набір даних, сформований за методикою гістограмного узгодження оптичних і радарних знімків, було випадковим чином розділено на навчальну та тестову вибірки у пропорції 80/20. Для об'єктивного та всебічного оцінювання якості прогнозування використовувалися стандартні статистичні метрики. Головним критерієм виступав коефіцієнт детермінації (R-квадрат), який показує частку дисперсії залежної змінної (фактичної якості фільтрації), що пояснюється моделлю. Чим ближче R^2 до одиниці, тим вищою є точність апроксимації. Крім того, для оцінки фізичної величини похибки обчислювалися корінь із середньоквадратичної похибки (RMSE) та середня абсолютна похибка (MAE). Оскільки різні метрики якості мають різний динамічний діапазон (наприклад, PSNR вимірюється в децибелах і зазвичай становить 20–35 дБ, тоді як SSIM є безрозмірним індексом у межах від 0 до 1), використання R^2 дозволяє коректно порівнювати точність моделей між різними експериментами.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|, \tag{37}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}, \tag{38}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2}. \tag{39}$$

де y_i — фактичне значення метрики якості для i-го прикладу, $\hat{y}_i$ — прогноз моделі, а $\bar{y}$ — середнє значення цільової змінної у тестовому наборі. Метрики

MAE і RMSE відображають абсолютну величину помилки у природних одиницях, тоді як R^2 показує частку поясненої дисперсії.

2.5.2 Аналіз точності прогнозування (R^2) для різних фільтрів

Результати тестування продемонстрували високу здатність DenseNet-121 вилучати інформативні текстурні ознаки. Для всіх 30 досліджених комбінацій (6 класичних фільтрів: Gamma MAP, Lee, Enhanced Lee, Frost, SRAD, Kuan × 5 метрик якості: PSNR, WSNR, SSIM, MS-SSIM, FSIM) модель забезпечила стабільно високу точність. Середній коефіцієнт детермінації (R^2) по всіх експериментах досяг 0,9732. Найкращі результати (R^2 понад 0,99) зафіксовано для фільтра Gamma MAP при прогнозуванні метрик MS-SSIM і WSNR ($R^2 \approx 0{,}9972$), що пояснюється передбачуваною поведінкою цього алгоритму на однорідних і помірно текстурованих ділянках. Водночас для фільтра Kuan на тих самих метриках R^2 перебуває в межах 0,96–0,97, що також свідчить про високу узгодженість прогнозу з фактичною якістю. Найнижчий показник ($R^2 = 0{,}9343$) отримано для фільтра Enhanced Lee при прогнозуванні індексу SSIM. Зниження точності в цьому випадку зумовлене специфічною тримодальною архітектурою Enhanced Lee, яка жорстко розмежовує пікселі на класи і може генерувати різкі артефакти на межах об'єктів, що ускладнює їхнє точне моделювання. Проте загалом точність перевершує показники класичних регресійних моделей, заснованих на ручному розрахунку статистичних ознак. Порівняння прогнозованих і фактичних значень PSNR наведено на рис. 2.6, зведену теплову карту коефіцієнта R^2 — на рис. 2.7. Детальні числові результати наведено в табл. 2.3–2.5.

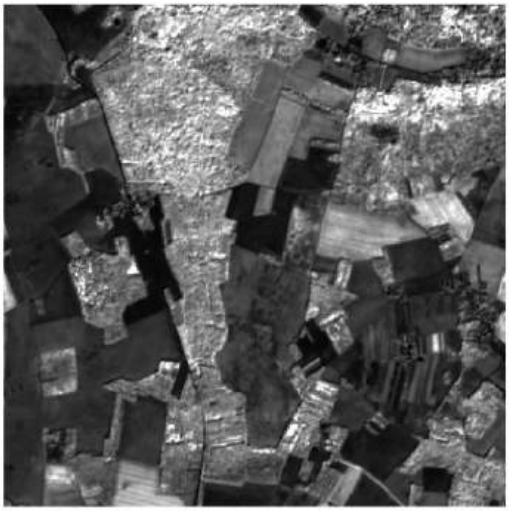
Таблиця 2.3 – Точність прогнозування якості фільтрації (MAE, RMSE, R^2) на тестовій вибірці (382 зображення, 1024×1024)

Фільтр	Показник	PSNR	SSIM	MS-SSIM	FSIM	WSNR

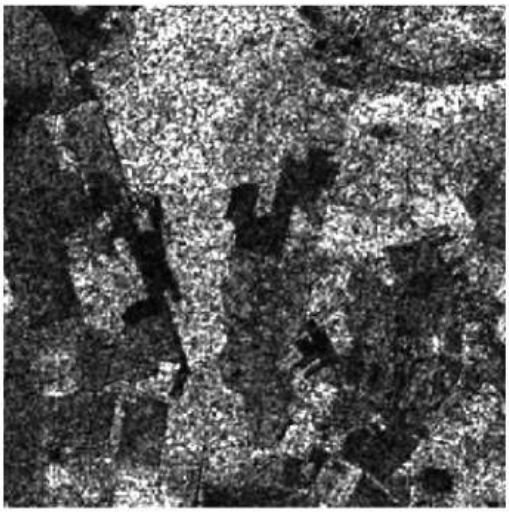
Gamma Map	MAE	0,2884	0,0132	0,0054	0,0021	0,1402
Gamma Map	RMSE	0,4475	0,0222	0,0082	0,0027	0,1944
Gamma Map	R^2	0,9886	0,9693	0,9972	0,9826	0,9972
Lee	MAE	0,1077	0,0043	0,0023	0,0023	0,0459
Lee	RMSE	0,1425	0,0058	0,0037	0,0037	0,0635
Lee	R^2	0,9864	0,9777	0,9518	0,9546	0,9966
Enhanced Lee	MAE	0,0587	0,0064	0,0023	0,0025	0,0442
Enhanced Lee	RMSE	0,0803	0,0086	0,0031	0,0044	0,0575
Enhanced Lee	R^2	0,9957	0,9343	0,9605	0,9364	0,9965
Frost	MAE	0,0805	0,0040	0,0019	0,0025	0,1318
Frost	RMSE	0,106	0,0054	0,0029	0,004	0,1822
Frost	R^2	0,9949	0,9825	0,9756	0,9609	0,9911
Srad	MAE	0,2391	0,0053	0,0025	0,0025	0,2781
Srad	RMSE	0,2964	0,0073	0,0038	0,0038	0,3582
Srad	R^2	0,9523	0,9644	0,955	0,9577	0,9429
Kuan	MAE	0,0514	0,0039	0,0021	0,0027	0,0486
Kuan	RMSE	0,0722	0,0054	0,003	0,004	0,0637
Kuan	R^2	0,996	0,9812	0,9666	0,9519	0,9963

Таблиця 2.4 – Усереднені значення коефіцієнта детермінації R^2 за типами фільтрів

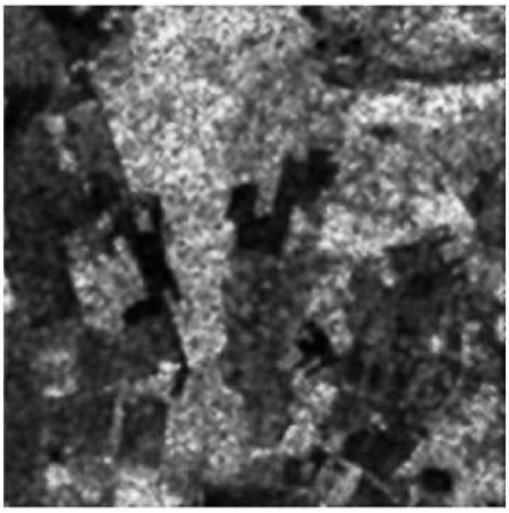
Фільтр	Average R^2	Діапазон
Gamma MAP	0,987	[0,969; 0,997]
Kuan	0,978	[0,952; 0,996]
Frost	0,978	[0,960; 0,995]
Enhanced Lee	0,963	[0,934; 0,997]
Lee	0,973	[0,952; 0,997]
SRAD	0,955	[0,943; 0,964]



(a) Чисте зображення

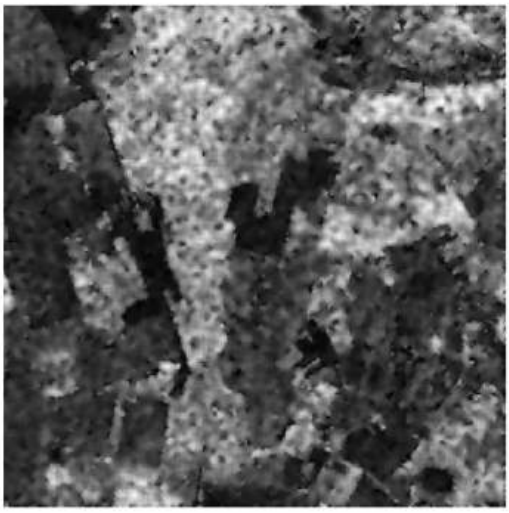


(b) Зашумлене зображення
ENL = 6.0



(c) Фільтр Lee
Результати фільтра Lee:

Передбачений PSNR: 20.74 dB
Справжній PSNR: 20.78 dB
Абсолютна похибка: 0.036 dB
Відносна похибка: 0.17%



(d) Фільтр Gamma MAP
Результати фільтра Gamma MAP:

Передбачений PSNR: 20.08 dB
Справжній PSNR: 20.04 dB
Абсолютна похибка: 0.037 dB
Відносна похибка: 0.19%

Рис. 2.6. Порівняння прогнозованих і фактичних значень PSNR для фільтрів Лі та Gamma MAP на тестових зображеннях: а) чисте зображення; б) зашумлене зображення; в) результат фільтрації фільтром Лі; г) результат фільтрації фільтром Gamma MAP [30]

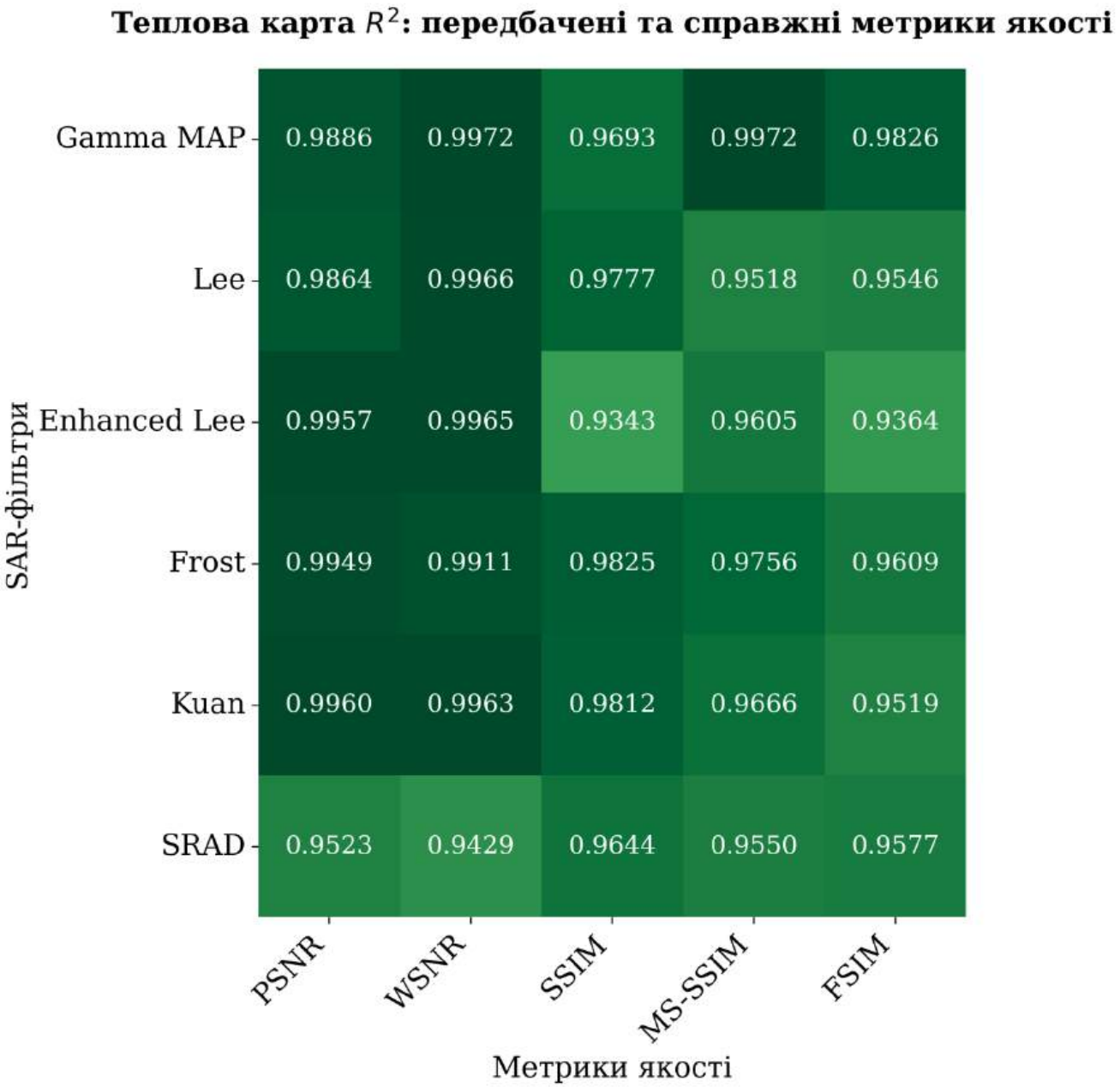


Рис. 2.7. Теплова карта коефіцієнта детермінації R^2 для різних фільтрів і метрик якості [30]

2.5.3 Оцінка стабільності прогнозів для різних метрик якості

Таблиця 2.5 – Усереднені значення коефіцієнта детермінації R^2 за типами метрик якості

Метрика	Average R^2	Діапазон
PSNR	0,987	[0,952; 0,996]
WSNR	0,987	[0,943; 0,997]
SSIM	0,968	[0,934; 0,982]
MS-SSIM	0,969	[0,952; 0,997]
FSIM	0,957	[0,937; 0,983]

Окремого обговорення заслуговує вплив обраної цільової метрики на успішність прогнозування. Дослідження виявило, що піксельні метрики, такі як PSNR та WSNR, передбачаються з найвищою і найстабільнішою точністю. Їхня дисперсія легко описується локальними статистиками, які добре фіксуються першими згортковими шарами мережі. Водночас структурні індекси (SSIM, MS-SSIM) та індекси на основі градієнтів (FSIM) виявилися складнішими для моделювання. Для їхнього коректного передбачення нейромережі доводиться оперувати складними просторовими абстракціями, зчитуючи інформацію про цілісність контурів та фазову конгруентність. Тим не менш, завдяки глибокій архітектурі та щільним зв'язкам, моделі вдалося досягти $R^2 > 0.96$ навіть для FSIM. Це підтверджує, що DenseNet-121 успішно засвоїла зв'язок між зашумленою макротекстурою та ступенем деградації низькорівневих контурів після згладжування, що є ключовим досягненням запропонованого гібридного підходу. Діаграми розсіювання для п'яти метрик якості наведено на рис. 2.8.

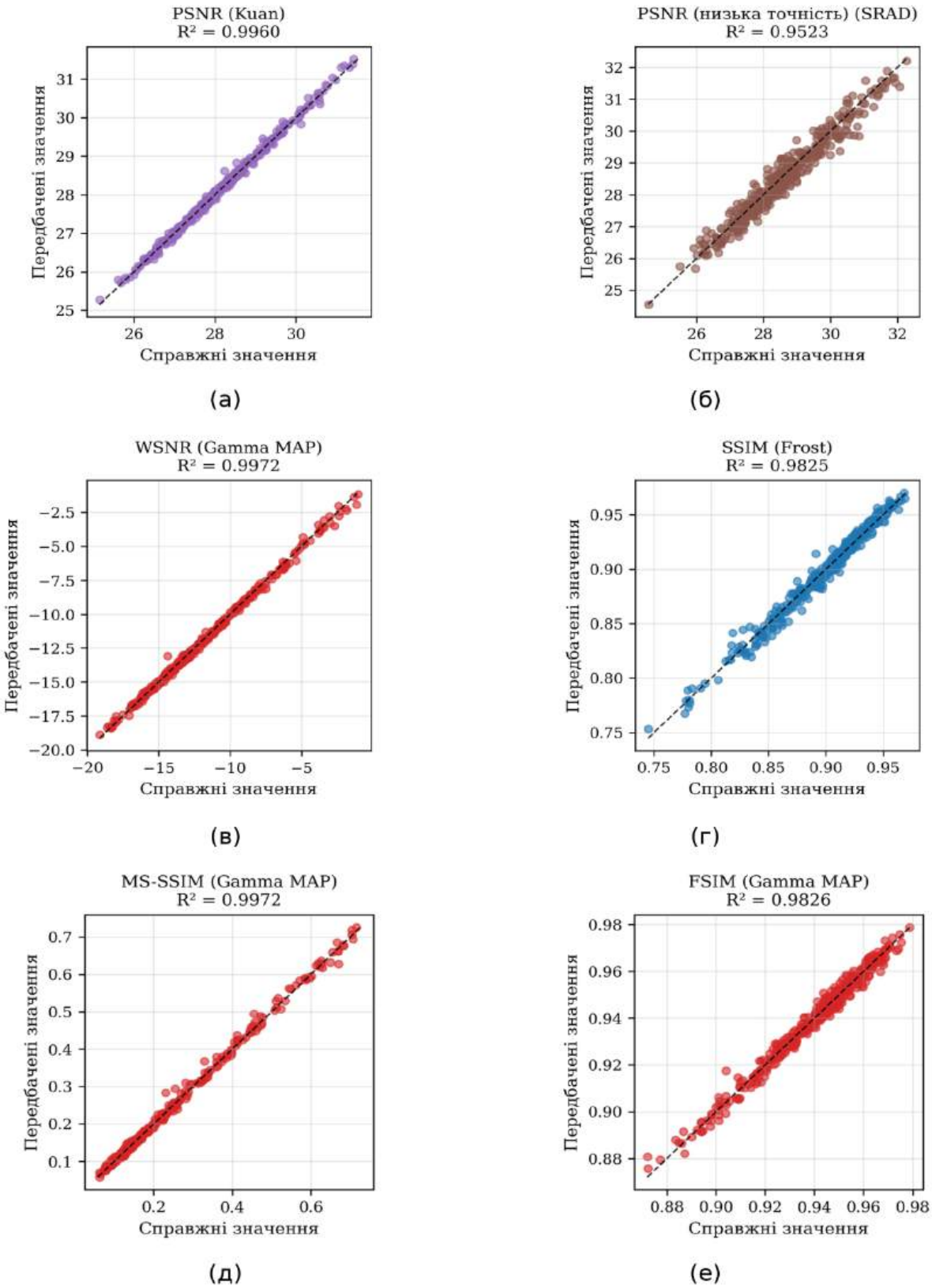


Рис. 2.8. Діаграми розсіювання прогнозованих і фактичних значень метрик якості: а) PSNR (Gamma MAP); б) PSNR (низька точність) (Kuan); в) WSNR (Gamma MAP); г) SSIM (Gamma MAP); д) MS-SSIM (Frost); е) FSIM (Gamma MAP) [30]

2.5.4 Аналіз розподілу помилок прогнозування та відсутності перенавчання

Важливим етапом тестування є перевірка моделі на відсутність перенавчання (overfitting) та оцінка робастності її прогнозів. Для цього було побудовано діаграми розсіювання (scatter plots), які відображають залежність передбачених значень метрик від їхніх фактичних (істинних) значень на тестовій вибірці. Ідеальна модель повинна формувати точки, що лежать точно на діагоналі (y = x). Отримані графіки (рис. 2.8) демонструють високу щільність групування точок навколо діагоналі для всіх діапазонів значень якості. Відхилення від ідеальної прямої мають характер нормального (гауссівського) розподілу з нульовим математичним сподіванням, що свідчить про відсутність систематичного зміщення (bias) у прогнозах. Середня абсолютна похибка (MAE) залишалася на рівні менше 0,29 дБ для PSNR та не перевищувала 0,014 для SSIM. Такі значення похибки є візуально непомітними і гарантують, що модель не просто «завчила» навчальні приклади, а знайшла універсальні закономірності формування радіолокаційного зображення. Залежність нормованої похибки прогнозу від рівня шуму (ENL) подано на рис. 2.9, розподіл похибки для різних комбінацій «фільтр – метрика» — на рис. 2.10.

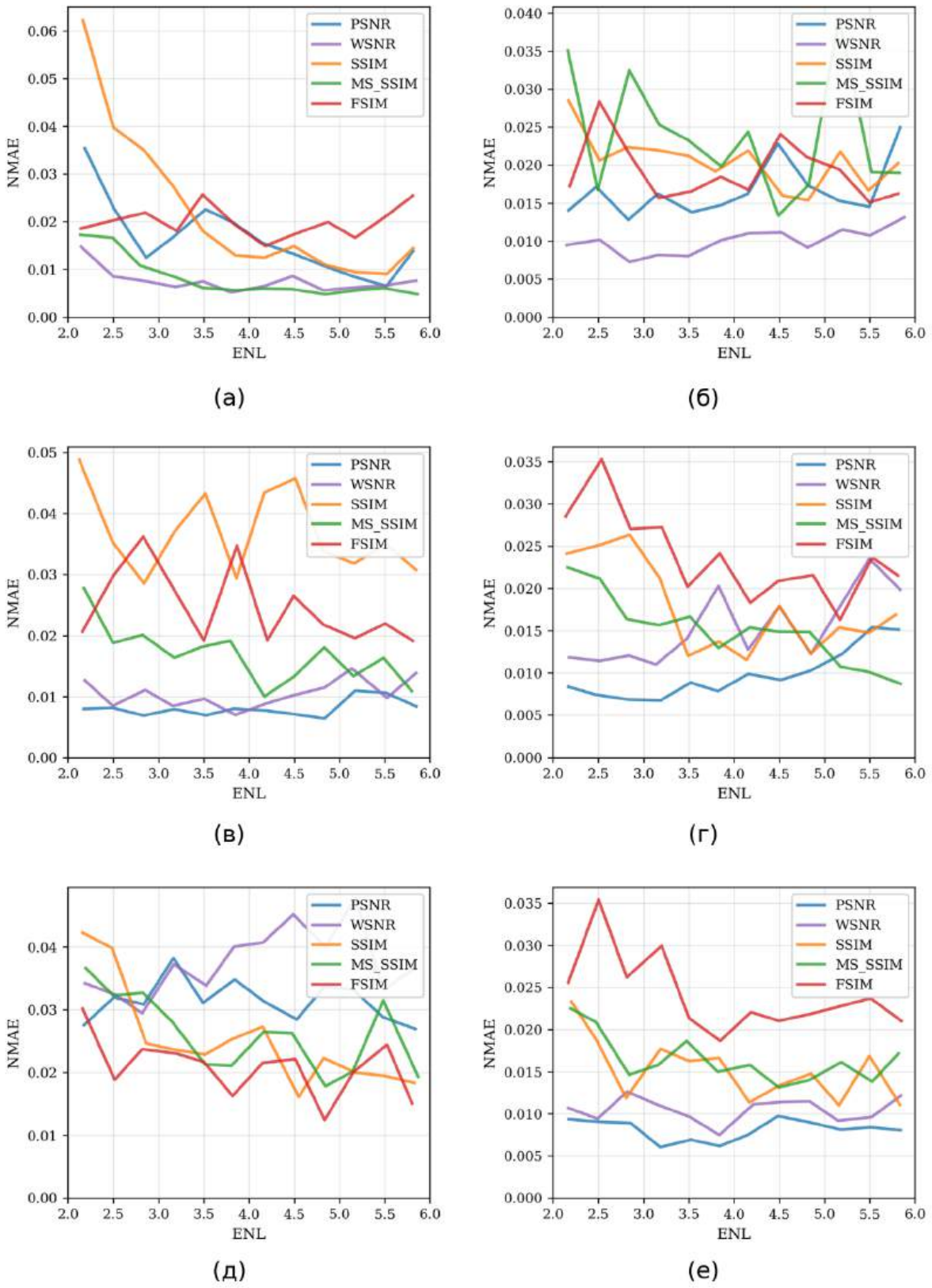


Рис. 2.9. Залежність нормованої похибки прогнозу від рівня шуму (ENL) для фільтрів: а) Gamma MAP; б) Лі; в) Покращений Лі; г) Фрост; д) SRAD; е) Куан [30]

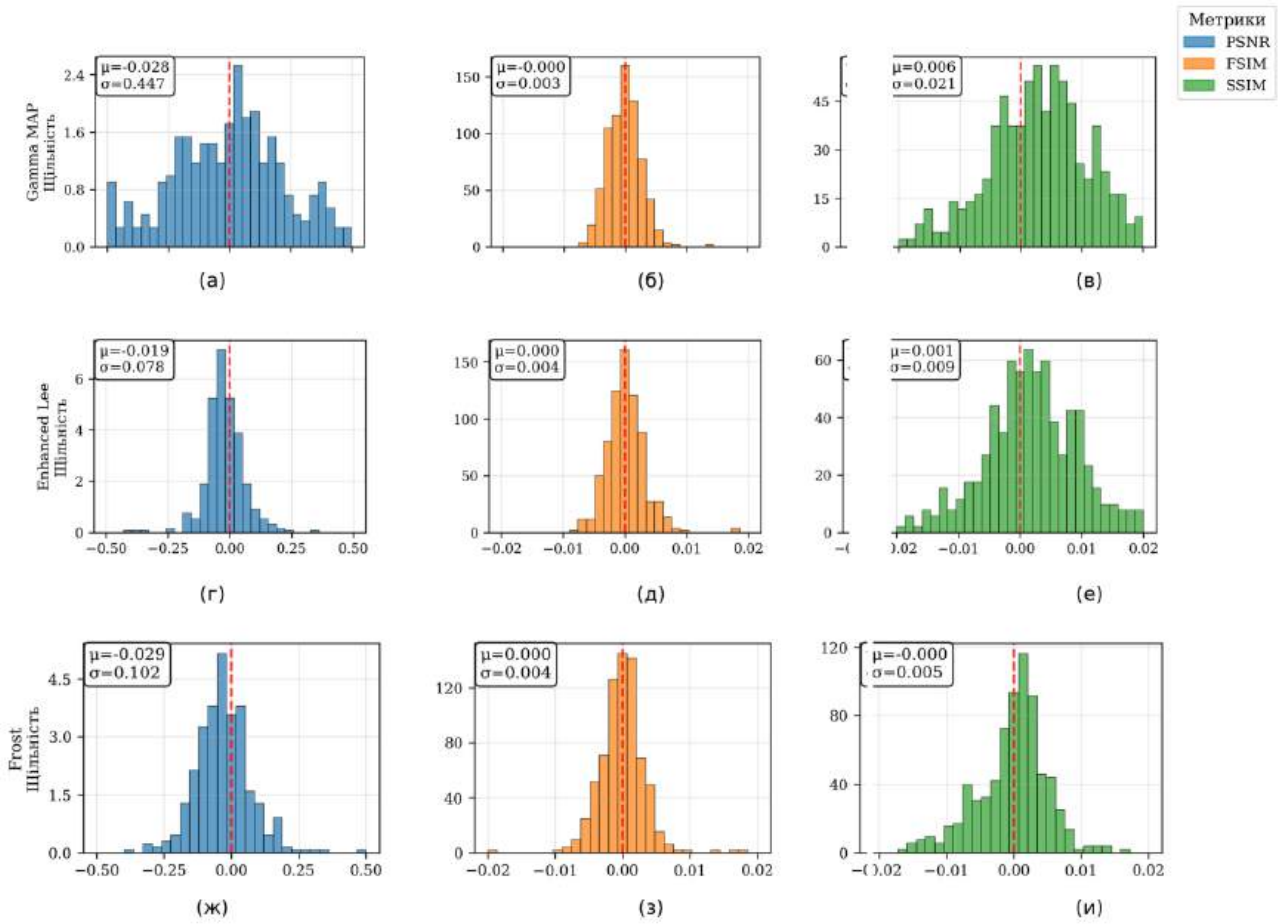


Рис. 2.10. Гістограми похибки прогнозу для різних комбінацій «фільтр – метрика»: а) Gamma MAP / PSNR; б) Gamma MAP / FSIM; в) Gamma MAP / SSIM; г) Покращений Лі / PSNR; д) Покращений Лі / FSIM; е) Покращений Лі / SSIM; ж) Фрост / PSNR; з) Фрост / FSIM; и) Фрост / SSIM [30]

2.5.5 Візуальне порівняння передбаченої та фактичної якості на тестових знімках

Для якісної верифікації методу було відібрано кілька репрезентативних тестових сцен із різними ландшафтами: щільна міська забудова, сільськогосподарські поля та водні поверхні. На однорідних сільськогосподарських угіддях усі алгоритми фільтрації демонструють високу ефективність придушення шуму, і модель коректно прогнозує високі значення PSNR та SSIM. Водночас для міських районів, де наявність великої кількості

точкових відбивачів (corner reflectors) робить застосування сильного згладжування деструктивним, нейромережа коректно передбачає стрімке падіння структурних індексів (зокрема FSIM). Збіг візуального аналізу впливу алгоритмів із нейромережевими прогнозами підтверджує, що модель враховує структурні особливості сцени. Для збереження різкості будівель вона занижує прогнозовану якість для тих методів, які схильні до надмірного розмиття контрастних границь.

2.5.6 Оцінка обчислювальної ефективності апріорного вибору фільтрів

Окрім високої точності, розроблений підхід забезпечує кардинальне підвищення обчислювальної ефективності. Традиційний емпіричний вибір параметрів вимагав би застосування всіх 6 фільтрів з різними конфігураціями до повнорозмірного радіолокаційного зображення, після чого оператор мав би візуально оцінити кожен результат. Для сучасних супутникових сцен розміром в десятки мегапікселів це забирає хвилини машинного часу. Водночас виконання прямого проходу (inference) через згорткову нейронну мережу DenseNet-121 на сучасному графічному процесорі виконується за час, істотно менший за тривалість повного перебору. Таким чином, система отримує можливість оперативно оцінити перспективність кожного фільтра, обрати той, що гарантує найкращу цільову метрику для конкретного типу ландшафту, і лише після цього запустити безпосереднє придушення спекл-шуму єдиним оптимальним методом. Це створює міцне підґрунтя для розробки повністю автономних сервісів масової обробки геопросторових даних.

2.5.7 Аналіз загальної та збалансованої точності класифікації розміру вікна

Другим напрямком досліджень була перевірка класифікаційної моделі MobileNetV2 на предмет правильного вибору розміру вікна для фільтра Лі. Тестування на вибірці з 382 унікальних патчів (розміром 1024×1024) показало, що модель здатна успішно ідентифікувати оптимальний клас. Загальна точність (Accuracy) досягла найвищого значення 87,17%, коли як цільовий критерій

(ground truth) використовувалася метрика FSIM. У випадку з MS-SSIM було досягнуто найвищої збалансованої точності (Balanced Accuracy) — 88,94%. Оскільки F1-score у вихідній публікації окремо не наведено, порівняння виконано за метриками Accuracy, Balanced Accuracy та Top-2 Accuracy. Різниця між цими показниками свідчить про те, що зважена функція втрат ефективно впоралася з дисбалансом класів у навчальній вибірці (де переважали вікна 5×5 та 7×7). Високі значення Top-2 точності (понад 99% для всіх метрик, у тому числі 100% для PSNR) підтверджують, що навіть у випадках помилки модель обирає сусідній клас вікна, що практично еквівалентно оптимальному за якістю фільтрації. Ці результати демонструють, що глибока згорткова мережа здатна засвоїти нетривіальні правила вибору просторових параметрів фільтрації. Криві навчання моделі MobileNetV2 наведено на рис. 2.11. Зведені метрики класифікації — у табл. 2.6; залежність від ENL — у табл. 2.7.

$$\text{Accuracy} = \frac{\sum_{i=1}^{C} TP_i}{N}, \tag{40}$$

$$\text{BalancedAccuracy} = \frac{1}{C}\sum_{i=1}^{C}\frac{TP_i}{TP_i + FN_i}, \tag{41}$$

$$\text{Top2Accuracy} = \frac{1}{N}\sum_{n=1}^{N}\mathbb{1}\left[y_n \in \text{Top2}_n\right]. \tag{42}$$

де TP_i та FN_i — кількість істинно-позитивних і хибно-негативних рішень для i-го класу, C — число класів, а Top-2 точність перевіряє, чи потрапив правильний розмір вікна до двох найімовірніших варіантів, запропонованих мережею.

Додатково в табл. 2.6 наведено допоміжні показники якості класифікації та калібрування. Коефіцієнт Каппа (Kappa) оцінює узгодженість прогнозів із еталонними мітками з поправкою на випадковий збіг; Matthews Correlation Coefficient (MCC) є інтегральною збалансованою мірою якості багатокласової класифікації; Expected Calibration Error (ECE) показує, наскільки прогнозовані ймовірності узгоджені з реальною частотою правильних рішень; Brier Score є середньоквадратичною похибкою між вектором прогнозованих імовірностей та фактичною one-hot міткою. Оскільки в цій роботі вони використовуються як допоміжні критерії порівняння, їх наведено без окремого формульного виведення.

Таблиця 2.6 – Порівняння метрик класифікації розміру вікна фільтра Лі за варіантами ground truth

Метрика ground truth	Точність	Збаланс. точність	Top-2 точність	Kappa	MCC	ECE	Brier
FSIM	87,17%	87,46%	99,48%	0,8291	0,8302	0,0772	0,0359
MS-SSIM	86,13%	88,94%	99,74%	0,8173	0,8228	0,0608	0,0373
PSNR	85,86%	88,14%	100,00%	0,8129	0,8164	0,0371	0,0405
MDSI	85,60%	87,16%	99,48%	0,8104	0,8112	0,0479	0,0407
SSIM	84,55%	88,60%	100,00%	0,7955	0,7992	0,0449	0,0428
HaarPSI	83,77%	83,73%	99,21%	0,7859	0,7869	0,0565	0,0473

Таблиця 2.7 – Залежність точності класифікації від ENL та метрики якості

Метрика	ENL [2,0; 3,0)	ENL [3,0; 4,0)	ENL [4,0; 5,0)	ENL [5,0; 6,0]
PSNR	91,1%	78,0%	83,2%	90,3%
SSIM	86,5%	85,3%	85,2%	80,7%
MS-SSIM	85,9%	87,2%	83,3%	88,2%
FSIM	91,1%	88,2%	83,8%	86,4%
HaarPSI	86,0%	82,4%	81,2%	86,0%
MDSI	88,9%	90,1%	82,8%	80,2%

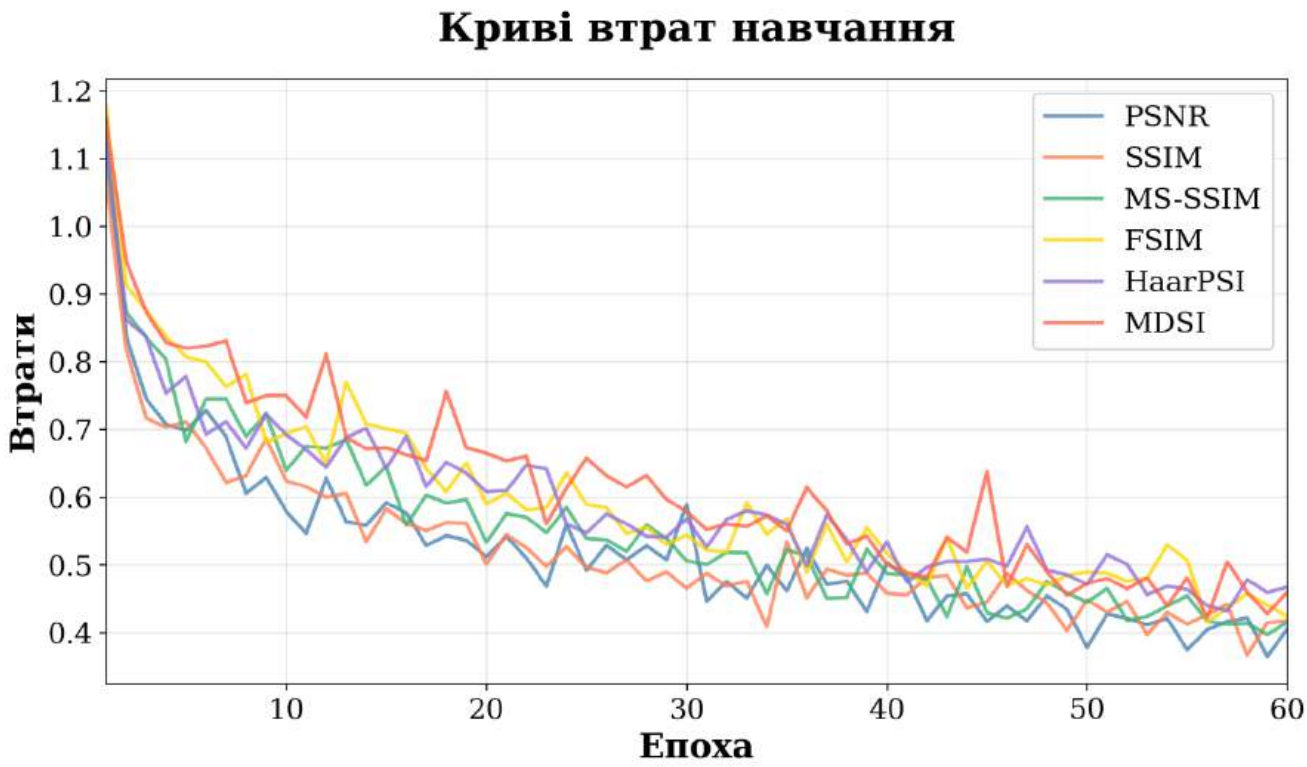


Рис. 2.11. Криві навчання моделі MobileNetV2 для всіх цільових метрик (значення функції втрат залежно від епохи) [147]

2.5.8 Дослідження матриці помилок (Confusion Matrix)

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i}. \tag{43}$$

Покласова повнота $Recall_i$ дає змогу оцінити, наскільки надійно модель розпізнає кожен конкретний розмір вікна. Це особливо важливо за наявності дисбалансу, коли висока загальна точність ще не гарантує однаково доброї роботи на рідкісних класах. Покласові значення повноти (Recall) для кожної цільової метрики наведено в табл. 2.8.

Таблиця 2.8 – Recall за класами (розмір вікна) для різних метрик якості

Метрика	3×3	5×5	7×7	9×9	11×11
PSNR	0,933	0,930	0,756	0,844	0,943
SSIM	0,964	0,825	0,773	0,919	0,947
MS-SSIM	1,000	0,736	0,966	0,830	0,915
FSIM	0,923	0,848	0,935	0,833	0,833
HaarPSI	0,833	0,812	0,851	0,839	0,851
MDSI	1,000	0,784	0,827	0,810	0,937

Детальний аналіз розподілу помилок за допомогою матриці сплутування (Confusion Matrix) виявив важливу закономірність: абсолютна більшість хибних класифікацій зосереджена виключно на сусідніх класах. Наприклад, якщо ідеальним розміром було вікно 7×7, то модель могла помилитися, обравши 5×5 або 9×9, проте майже ніколи не обирала 3×3 або 11×11. Це фізично обґрунтовано: якість фільтрації при суміжних розмірах вікон змінюється дуже плавно, і різниця в метриках (наприклад, відмінність PSNR на 0,05 дБ) може бути статистично незначущою. Точність Top-2 (ймовірність того, що істинний клас знаходиться серед двох найімовірніших прогнозів моделі) перевищила 99,2% для всіх експериментів, а для PSNR досягла 100%. З практичної точки зору це означає, що навіть у випадку формальної помилки класифікації, система обере розмір вікна, який забезпечить результати, практично ідентичні до глобального оптимуму. Діаграми надійності для всіх метрик подано на рис. 2.12.

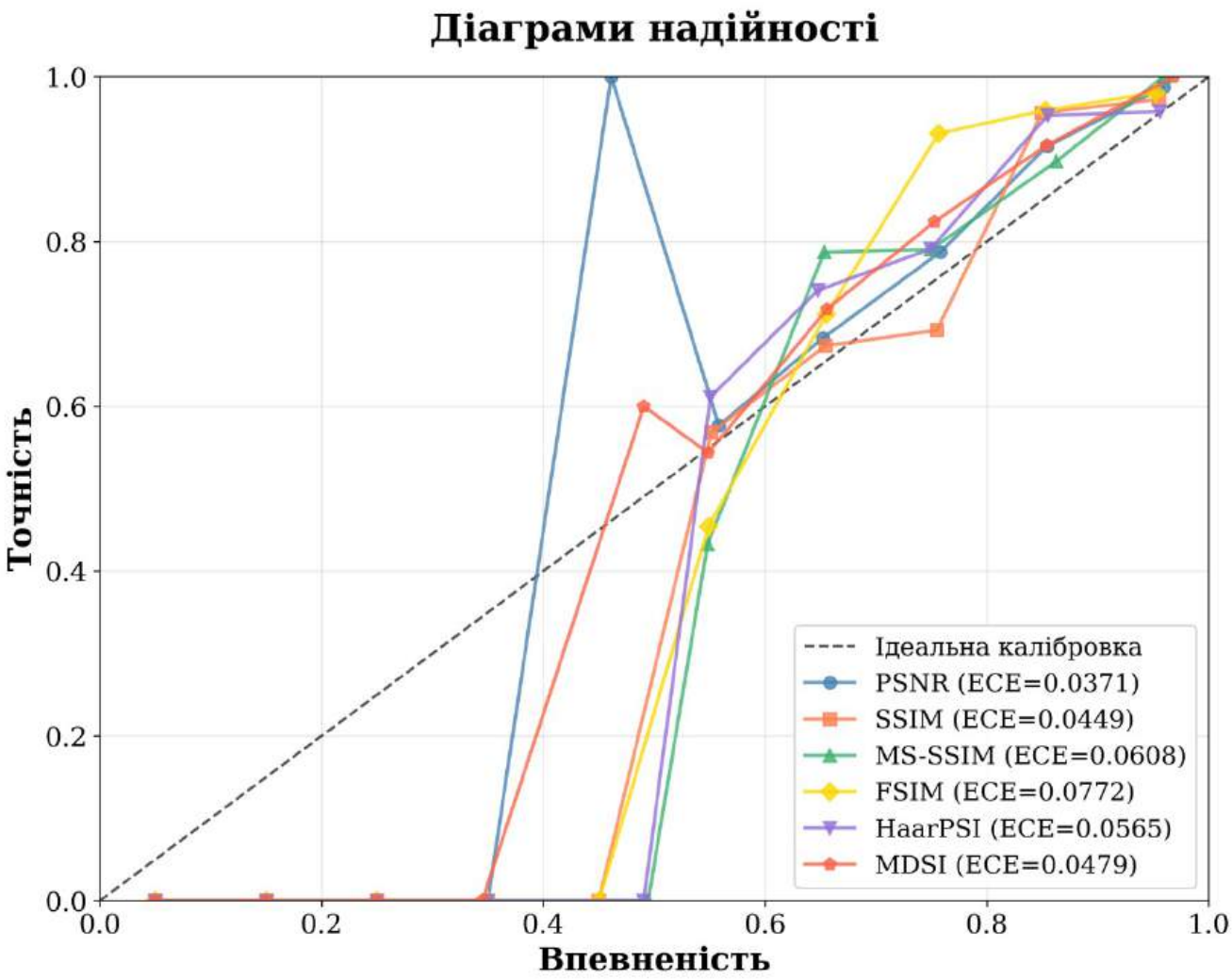


Рис. 2.12. Діаграми надійності (reliability diagrams) для всіх метрик: фактична точність порівняно з прогнозованою впевненістю моделі [147]

2.5.9 Порівняння прямої класифікації з двоетапними підходами

Отримані результати підтвердили теоретичну перевагу запропонованого підходу над традиційними двоетапними (регресія + argmax) схемами вибору параметрів [14]. Існуючі підходи на базі статистичних ознак [13] демонстрували ймовірність коректного вибору вікна на рівні 0,86–0,92 лише для обмежених тестових наборів або для бінарної класифікації [15]. Запропонована пряма багатокласова архітектура досягла співставної та навіть вищої збалансованої точності (до 88,9%), але працює як наскрізний конвеєр без необхідності ручного

вилучення 28 статистичних ознак. Крім того, пряма класифікація виявилася стійкішою до шумових флуктуацій: тоді як двоетапна схема може страждати від каскадного накопичення похибок прогнозування для окремих вікон, класифікатор одразу мінімізує похибку розпізнавання кінцевого класу.

2.5.10 Візуалізація результатів: карти класифікації на ландшафтах

Візуальний аналіз прогнозів підтвердив адекватність прийнятих моделлю рішень з точки зору фізики радіолокаційного розсіювання. На великих однорідних ділянках (водні поверхні, рівні поля), де інтенсивність спекл-шуму є домінуючою, а відбивна здатність стабільною, нейромережа впевнено обирає найбільші вікна (9×9 або 11×11). Такий вибір забезпечує максимальне усереднення та радикальне придушення флуктуацій. Водночас у міських зонах із щільною забудовою та високим рівнем гетерогенності модель адаптивно перемикається на малі вікна (3×3 або 5×5). Ця поведінка ідеально відповідає теоретичним вимогам: у високочастотних областях надмірне згладжування знищує корисну інформацію про розсіювачі, тому збереження деталей є пріоритетнішим за боротьбу з шумом. Отже, нейромережа навчилася розрізняти просторову семантику ландшафту.

2.5.11 Вплив обраного вікна на збереження просторових деталей

Справедливість вибору менших вікон для гетерогенних ділянок додатково доводиться аналізом структурних метрик якості. Експерименти з використанням FSIM та HaarPSI (метрик, чутливих до фазової конгруентності та високочастотних градієнтів) показали, що для зображень із дрібними деталями вибір завеликого вікна (11×11) систематично знижує FSIM/HaarPSI порівняно з оптимальним вікном 3×3 чи 5×5. Сильне згладжування незворотно руйнує просторову текстуру, перетворюючи складні об'єкти на безформні плями. Завдяки тому, що модель MobileNetV2 здатна надійно класифікувати такі області, запропонована система запобігає значному розмиттю (oversmoothing). Збереження дрібних точкових об'єктів (наприклад, окремих будівель, транспортних засобів чи елементів

інфраструктури) та чітких меж лісових масивів є вкрай важливим для подальших етапів тематичної обробки РСА-даних, таких як автоматична класифікація типів рослинності, виявлення об'єктів (target detection) або точне виявлення змін (change detection) у часі. Розподіл оптимальних розмірів вікна за шкалою просторових частот наведено на рис. 2.13. Залежность точності класифікації від частотного масштабу спеклу деталізовано в табл. 2.9.

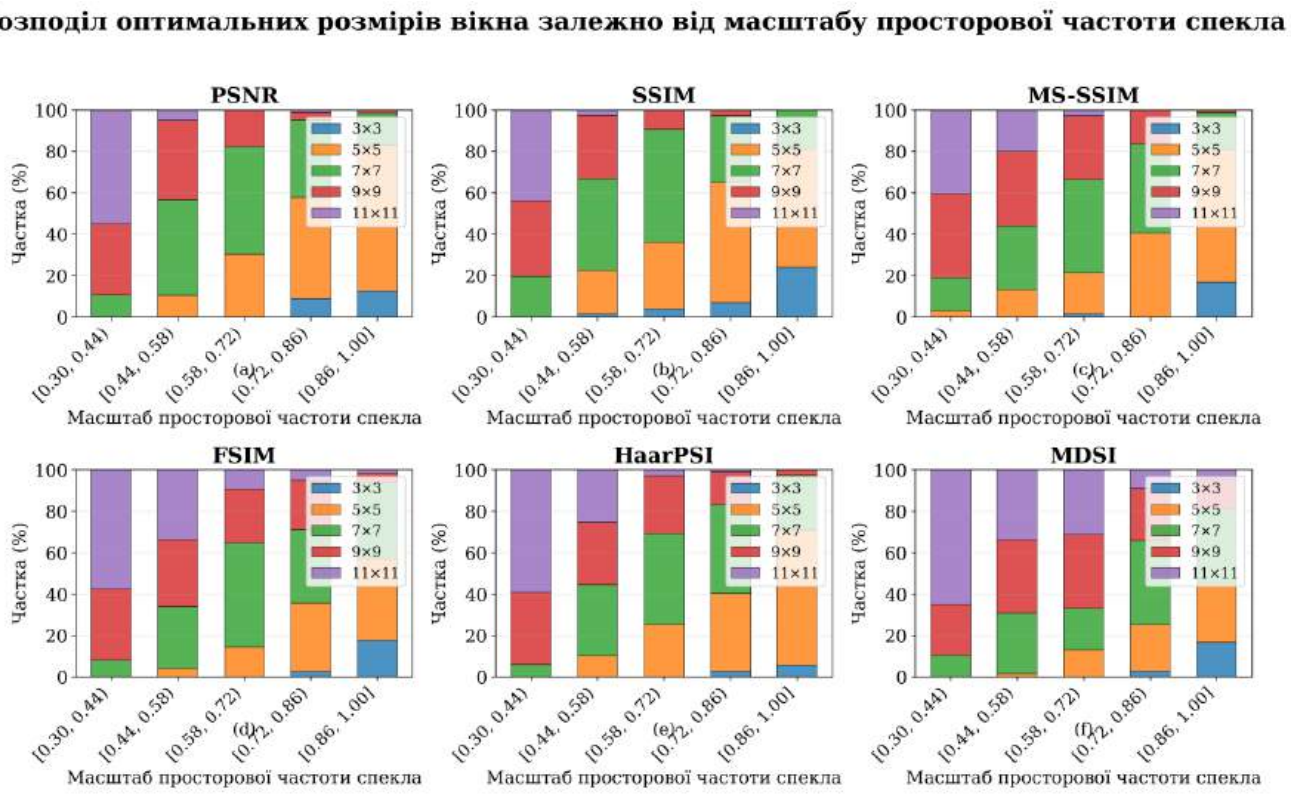


Рис. 2.13. Розподіл оптимальних розмірів вікна фільтра Лі за шкалою просторових частот спекл-шуму [147]

2.5.12 Оцінка швидкодії інференсу MobileNetV2 у реальних умовах

Однією з ключових вимог до систем попередньої обробки є їхня обчислювальна ефективність. Завдяки використанню легковагової архітектури MobileNetV2, час прямого проходу (inference) для одного патча розміром 1024×1024 пікселів на сучасному графічному процесорі лишається в межах часу, прийнятного для потокової обробки. Більше того, ця архітектура оптимізована

для мобільних пристроїв, що робить можливою її роботу навіть на центральних процесорах (CPU) з прийнятною затримкою. Для порівняння, виконання повного перебору (послідовний розрахунок п'яти варіантів фільтра Лі з різними вікнами та їх наступний аналіз) потребує істотно більше часу та системної пам'яті. В умовах операційних конвеєрів платформи Copernicus, які щодня генерують терабайти радіолокаційних даних, така різниця є критичною. Запропонований метод дозволяє інтегрувати модуль визначення оптимального вікна як легкий препроцесор, який практично не створює додаткового навантаження на серверні потужності, але суттєво підвищує загальну якість кінцевого інформаційного продукту перед передачею кінцевим споживачам.

2.5.13 Синтез результатів: взаємозв'язок метрик та розміру вікна

Таблиця 2.9 – Залежність точності класифікації від масштабу просторової частоти спекла та метрики якості

Метрика	[0,30; 0,44)	[0,44; 0,58)	[0,58; 0,72)	[0,72; 0,86)	[0,86; 1,00]
PSNR	87,4%	84,6%	86,3%	83,8%	87,5%
SSIM	82,9%	88,0%	85,3%	83,1%	83,5%
MS-SSIM	89,2%	85,3%	93,0%	82,1%	82,1%
FSIM	78,2%	84,4%	91,8%	93,7%	87,3%
HaarPSI	84,9%	79,7%	79,4%	88,3%	86,1%
MDSI	90,7%	86,2%	82,9%	82,9%	84,7%

Синтезуючи результати обох етапів досліджень, можна стверджувати, що вибір цільової метрики якості суттєво впливає на визначення оптимальних параметрів. Коли за еталон береться PSNR (спрямований переважно на енергетичну оцінку шуму), система часто надає перевагу вікнам 7×7. Однак при використанні метрик, орієнтованих на сприйняття та структуру (SSIM, FSIM), оптимум частіше зміщується в бік 3×3 або 5×5. Це доводить, що універсального «найкращого» розміру вікна не існує, і його вибір має диктуватися кінцевою

метою обробки (візуальний аналіз, виявлення цілей чи класифікація). Здатність створених нейромережевих моделей адаптуватися під різні критерії (шляхом простої зміни міток під час навчання) робить їх особливо гнучким інструментом для побудови спеціалізованих сервісів ДЗЗ.

2.5.14 Аналіз обмежень підходів (глобальна vs локальна адаптація)

Попри високу точність та обчислювальні переваги, розроблені методи апріорного оцінювання мають певні обмеження, які необхідно враховувати під час практичного впровадження. Найголовнішим недоліком є глобальний характер приймання рішень: модель прогнозує одне інтегральне значення якості або один розмір вікна для всього вхідного патча (наприклад, 1024×1024 пікселів). Якщо аналізований фрагмент є дуже неоднорідним (містить половину гладкої водної поверхні і половину щільної забудови), класичний фільтр із глобально оптимізованим вікном все одно буде субоптимальним для однієї з половин. Фільтри типу Лі чи Фроста мають внутрішню просторову адаптивність на рівні окремих пікселів, але вона все одно обмежена фіксованим розміром скануючої рамки, обраним нейромережею.

2.5.15 Роль прогнозування в оптимізації конвеєрів супутникових даних

Впровадження моделей апріорного оцінювання здатне кардинально змінити архітектуру сучасних конвеєрів обробки (processing pipelines) супутникових даних. Замість жорстко заданих алгоритмів (hard-coded parameters), система набуває здатності до самостійного прийняття рішень (decision-making) (наприклад, динамічний вибір між Gamma MAP для рівнин та Enhanced Lee для лісів). У хмарних середовищах, де вартість обчислювальних ресурсів напряму залежить від часу роботи процесорів, можливість завчасного відкидання неперспективних методів фільтрації приносить відчутний економічний ефект. Крім того, автоматизоване ранжування фільтрів за метрикою якості дозволяє

гарантувати споживачеві максимально можливий рівень достовірності даних для подальшого аналізу.

2.5.16 Порівняння енергоефективності предикторів та перебору

Додатковою перевагою переходу від стратегії повного перебору до застосування навчених легковагових мереж є суттєва економія енергоресурсів. Операції згортки в оптимізованих архітектурах на зразок MobileNetV2 використовують матричні обчислення, що ефективно розпаралелюються на тензорних ядрах сучасних обчислювачів. Водночас багаторазове обчислення складної локальної статистики (дисперсії у ковзному вікні) для традиційних адаптивних фільтрів є операцією, що створює значне навантаження на підсистему пам'яті. Усунення багаторазових прогонів завдяки одному швидкому прямому проходу нейромережі дозволяє знизити енергоспоживання обчислювального кластера в кілька разів, що відповідає сучасним тенденціям екологічного програмування (Green AI).

2.5.17 Перспективи інтеграції методів у хмарні платформи ДЗЗ

Очікується, що найефективнішим середовищем для розгортання розроблених моделей стануть мікросервісні хмарні платформи для моніторингу Землі [2]. Формат вхідних і вихідних даних моделей повністю сумісний зі стандартами продуктів Sentinel-1 (IW GRD) [4]. Модуль прогнозування може бути реалізований як незалежний контейнеризований мікросервіс (наприклад, Docker), що працюватиме на етапі попередньої обробки безпосередньо після завантаження сирих сцен з репозиторію. Оскільки підхід базується на абсолютних значеннях інтенсивності, він не вимагає складної радіометричної калібровки чи фазового аналізу. Відкритість коду та використання стандартизованих фреймворків глибокого навчання гарантують простоту його інтеграції.

2.5.18 Підсумкова дискусія щодо доцільності апріорного прогнозування

Підсумовуючи результати досліджень, можна впевнено констатувати, що концепція апріорного прогнозування метрик та параметрів фільтрації для PCA-зображень за допомогою трансферного навчання є цілком життєздатною та високоефективною. Проведені експерименти доводять, що відсутність реальних чистих еталонів не є перешкодою завдяки використанню напівсинтетичних наборів на базі оптичних даних. Розроблені моделі DenseNet-121 та MobileNetV2 здатні забезпечити точність, достатню для прийняття автоматизованих рішень. Водночас виявлене обмеження глобального підходу вказує шлях до подальших досліджень. Нагальною стає необхідність переходу від прогнозування одного параметра для всього зображення до розробки архітектур, здатних генерувати попіксельні карти параметрів. Це завдання буде вирішено в наступному розділі шляхом створення нового гібридного методу фільтрації спекл-шуму.

2.6 Висновки до розділу 2

2.6.1 Підсумки щодо прогнозування метрик

У другому розділі розроблено метод апріорного прогнозування абсолютних значень метрик якості (PSNR, WSNR, SSIM, MS-SSIM, FSIM) для PCA-зображень Sentinel-1. Запропоновано конвеєр генерації масивів напівсинтетичних даних на основі гістограмного узгодження оптичних знімків Sentinel-2 (діапазон NIR) із радарними даними, з додаванням мультиплікативного спекл-шуму (ENL від 2 до 6). Адаптована регресійна модель DenseNet-121 досягла середнього коефіцієнта детермінації R^2 0,9732. Це дає змогу точно оцінювати ефективність застосування фільтрів без фактичного виконання обробки.

2.6.2 Підсумки щодо класифікації розміру вікна

Успішно розв'язано проблему вибору параметрів адаптивної фільтрації у формі задачі прямої класифікації на основі легкої нейромережі MobileNetV2.

Метод усуває проблему накопичення похибок двоетапних регресійних систем і забезпечує миттєвий, обчислювально ефективний вибір оптимального розміру скануючого вікна (від 3×3 до 11×11). За результатами тестування модель досягла високої збалансованої точності 88,94%. Вона продемонструвала логічну фізичну поведінку: віддаючи перевагу малим вікнам для збереження деталей на гетерогенних міських ландшафтах і великим вікнам для придушення шуму в межах однорідних зон. Однак розроблені в цьому розділі методи здатні формувати лише глобальне рішення для великого патча. Для забезпечення повної попіксельної адаптивності потрібні нові гібридні архітектури, розробка яких становить предмет дослідження у наступному розділі дисертаційної роботи.

РОЗДІЛ 3. ГІБРИДНИЙ МЕТОД АДАПТИВНОЇ ФІЛЬТРАЦІЇ СПЕКЛ-ШУМУ З ІНТЕГРАЦІЄЮ ДИФЕРЕНЦІЙОВАНОГО ФІЛЬТРА ФРОСТА ТА МЕХАНІЗМУ УВАГИ (VGG16-FROSTNET)

Третій розділ дисертаційної роботи присвячений розробці гібридного методу фільтрації спекл-шуму, який органічно поєднує класичну локально-статистичну фільтрацію та сучасні нейромережеві архітектури. Представлено наскрізно-навчальну модель VGG16-FrostNet, що використовує механізм просторово-канальної уваги CBAM [146] для автономного генерування карти коефіцієнтів демпфування. Фундаментом методу є модифікована диференційована математична модель фільтра Фроста, що виключає числові нестабільності під час навчання. Доведено спроможність моделі адаптивно посилювати згладжування в однорідних зонах і зменшувати його на границях об'єктів. Проведено ґрунтовну експериментальну перевірку запропонованого методу на репрезентативних наборах напівсинтетичних радіолокаційних даних.

3.1 Диференційована математична модель фільтра Фроста

3.1.1 Обґрунтування необхідності наскрізно-навчальних методів

Традиційні адаптивні фільтри (наприклад, фільтри Лі [7], Фроста [8], Куана [9] або Gamma MAP [10]) базуються на обчисленні локальної статистики в межах ковзного просторового вікна. Попри їхню математичну простоту та теоретичну обґрунтованість, ці алгоритми мають суттєвий недолік — необхідність ручного налаштування ключових параметрів (таких як розмір вікна чи коефіцієнт демпфування). Як показано в попередньому розділі, використання єдиного статичного параметра для всього радіолокаційного зображення є субоптимальним рішенням для гетерогенних сцен, які містять одночасно гладкі поверхні (водойми) та високочастотні структури (міська забудова). Альтернативний шлях полягає в розробці наскрізно-навчальних гібридних методів. Такі методи мають автономно аналізувати семантику зображення та попіксельно адаптувати параметри

згладжування без втручання оператора. Порівняно з «чорних ящиків» класичних глибоких нейромереж, гібридний підхід прагне зберегти аналітичну прозорість, інтегруючи математичну модель локальної статистики безпосередньо у граф обчислень.

3.1.2 Мультиплікативна модель спекл-шуму для РСА-зображень

Для розробки ефективного методу фільтрації необхідно спиратися на адекватну фізичну модель формування радіолокаційного сигналу [3]. Спекл-шум в амплітудних або інтенсивнісних РСА-зображеннях (зокрема Sentinel-1) традиційно моделюється як мультиплікативна завада [6]. Математично це виражається рівністю: спостережувана інтенсивність є добутком істинної ефективної площі розсіювання сцени (радіолокаційного перерізу) та шумового множника. Шумовий множник моделюється випадковою величиною, що підпорядковується гамма-розподілу з математичним сподіванням, рівним одиниці, і дисперсією, обернено пропорційною еквівалентному числу поглядів (ENL) [5]. Така параметризація точно відображає ключову властивість спекл-шуму: дисперсія шуму зростає лінійно зі зростанням середнього рівня інтенсивності сигналу. Саме ця залежність від сигналу робить непридатними класичні адитивні фільтри (наприклад, фільтр Гаусса) і вимагає застосування спеціалізованих адаптивних методів, які враховують локальний коефіцієнт варіації для розрізнення шуму від корисних структурних деталей. Формально, для амплітудно-інтенсивнісних РСА-зображень мультиплікативна модель шуму та його статистичний розподіл записуються як:

$$I(x,y) = X(x,y) \cdot S(x,y), \tag{44}$$

$$S(x,y) \sim \Gamma(L, 1/L), \quad \mathbb{E}[S] = 1, \quad \mathrm{Var}[S] = 1/L, \tag{45}$$

де $X(x, y)$ — істинна (безшумна) інтенсивність сцени; $I(x, y)$ — спостережувана інтенсивність; $S(x, y)$ — мультиплікативний спекл-множник із гамма-розподілом; L — еквівалентне число поглядів (ENL). Така параметризація гарантує одиничне математичне сподівання шуму і дисперсію, обернено пропорційну L, що узгоджується з усталеним статистичним описом спеклу для широкого класу РСА-систем [3; 6].

3.1.3 Класичний фільтр Фроста: принцип роботи та обмеження

Серед класичних методів фільтрації спекл-шуму фільтр Фроста [8] посідає особливе місце завдяки своїй здатності забезпечувати експоненційне згладжування. Вагові коефіцієнти цього фільтра обчислюються як експоненційна функція від відстані до центрального пікселя та локального коефіцієнта варіації:

$$w_{ij}^{\text{class}} = \exp\left(-A \cdot C_I^2 \cdot d_{ij}\right), \tag{46}$$

де A — статичний коефіцієнт демпфування; $C_I = \sigma_I/\mu_I$ — локальний коефіцієнт варіації, обчислений як відношення середньоквадратичного відхилення до середнього значення інтенсивності у вікні; d_{ij} — евклідова відстань від центру вікна до пікселя (i, j). Ключовим регулятором виступає коефіцієнт демпфування A. Чим вище його значення, тим швидше спадають ваги периферійних пікселів, зберігаючи різкість границь. Навпаки, низькі значення призводять до рівномірного усереднення і сильного придушення шуму в однорідних зонах. Головним обмеженням фільтра Фроста в його оригінальній формі є статичність коефіцієнта демпфування [16]. В операційних умовах він обирається емпірично або на основі глобальних метрик якості, що не дозволяє одночасно досягти максимального збереження мікроструктур (які потребують великого демпфування) та ефективного очищення темних чи однорідних площин (які потребують малого демпфування) в межах одного супутникового знімка.

3.1.4 Формалізація задачі нелінійного відображення з параметрами

Виходячи з мультиплікативної моделі шуму (44)–(45) та обмежень класичних алгоритмів, задачу просторово-адаптивної фільтрації спекл-шуму можна формалізувати як пошук оптимального нелінійного відображення F_θ, параметризованого ваговими коефіцієнтами θ нейромережі, що для зашумленого спостереження $I(x, y)$ продукуватиме оцінку $\hat{X}(x, y) = F_\theta(I(x, y))$ істинної інтенсивності $X(x, y)$. Оптимальні параметри $\theta*$ визначаються шляхом мінімізації математичного сподівання комплексної функції втрат на репрезентативному навчальному наборі даних:

$$\theta^* = \mathrm{argmin}_\theta \mathbb{E}[\mathcal{L}(F_\theta(I), X)]. \tag{47}$$

Центральним елементом цієї функції відображення є інтеграція логіки локальної статистики у нейромережевий граф. Замість того щоб нейромережа напряму генерувала очищене зображення, вона має передбачати просторово-змінну карту параметрів для аналітичного фільтра. Ця карта буде динамічно підлаштовувати форму імпульсної характеристики фільтра для кожного окремого пікселя, що забезпечить як високу ефективність придушення шуму, так і збереження можливості математичної інтерпретації процесу.

3.1.5 Проблема числових нестабільностей (underflow) у графі обчислень

Пряма спроба інтегрувати оригінальне рівняння фільтра Фроста у граф обчислень глибокої нейромережі (для наскрізного навчання) стикається з ключовою математичною перешкодою. Формула Фроста використовує квадрат локального коефіцієнта варіації у показнику експоненти. На границях яскравих об'єктів або на точкових цілях локальний коефіцієнт варіації може приймати значення, значно більші за одиницю (іноді сягаючи десятків). При піднесенні до квадрата ця величина стає екстремально великою. Оскільки вона стоїть зі знаком

мінус у показнику експоненти, функція намагається обчислити число e у дуже великому від'ємному степені. В умовах арифметики з плаваючою крапкою стандартної точності (float32) це миттєво призводить до так званого «underflow» — числового обнулення (результат стає строго рівним нулю). Внаслідок цього градієнти функції втрат по відношенню до параметра демпфування обнулюються (vanishing gradients), і зворотне поширення похибки блокується, унеможливлюючи навчання.

3.1.6 Модифікація математичної моделі: заміна квадрата на першу ступінь

Для подолання проблеми числових нестабільностей у роботі було розроблено модифіковану диференційовану математичну модель фільтра Фроста [30]. Ключовою зміною стала заміна квадрата локального коефіцієнта варіації на його першу ступінь (відношення локального середньоквадратичного відхилення до локального середнього). Оскільки на границях об'єктів коефіцієнт варіації більший за одиницю, його перша ступінь завжди строго менша за його квадрат. Ця проста, але ефективна алгебраїчна маніпуляція дозволяє уникнути генерації надмірно великих від'ємних значень у показнику експоненти. Відповідно, значення експоненційної функції залишаються в межах розрядної сітки формату float32, не обнуляються і забезпечують безперервний та стабільний потік градієнтів під час процедури зворотного поширення похибки (backpropagation) впродовж усього процесу тренування нейронної мережі.

Отже, у запропонованому методі модифікація не заперечує класичну формулу Фроста, а цілеспрямовано перетворює її на диференційовану версію, придатну для наскрізної оптимізації. Інакше кажучи, класичний запис задає фізично інтерпретовану логіку експоненційного згладжування, тоді як модифікований запис зберігає цю логіку, але переводить її в область чисельно стійких тензорних операцій.

3.1.7 Стиснення динамічного діапазону та чисельна стабільність градієнтів

Наслідком заміни квадрата на першу ступінь є істотне стиснення динамічного діапазону показника експоненти. Математично це означає, що функція чутливості фільтра до перепадів контрасту стає більш похилою та гладкою. Це стиснення критично важливе не лише для уникнення underflow, але й для стабілізації оптимізаційного ландшафту (optimization landscape). У стандартній формулі Фроста навіть незначна зміна параметра демпфування (який мережа має прогнозувати) могла призвести до лавиноподібної зміни вагових коефіцієнтів через квадратичну залежність. Згладжений динамічний діапазон дозволяє градієнтному спуску (наприклад, алгоритму Adam [148]) рухатися до локального оптимуму рівномірно, без різких стрибків та осциляцій. Завдяки цій диференційованій властивості, фільтр Фроста перетворюється зі статичного алгоритму на повноцінний активаційний шар із керованою просторовою поведінкою.

3.1.8 Роль просторово-змінного коефіцієнта як компенсатора модифікації

Може здатися, що видалення квадрата з коефіцієнта варіації знизить здатність фільтра різко реагувати на контури об'єктів і призведе до розмиття границь. Однак ця структурна слабкість повністю компенсується архітектурним рішенням: коефіцієнт демпфування перестає бути статичною константою і перетворюється на просторово-змінну карту (двовимірний тензор), яку для кожного пікселя окремо обчислює модуль уваги нейромережі. Розширений діапазон значень цього прогнозованого коефіцієнта забезпечує моделі достатню математичну ємність для того, щоб підвищувати значення $A_{map}(x, y)$ у зонах різких перепадів яскравості й тим самим обмежувати згладжування поблизу контурів. Отже, нейромережа сама навчається формувати карту параметрів, яка ідеально компенсує всі нелінійності вилученого квадрата, досягаючи оптимального балансу між збереженням геометрії та згладжуванням.

Таким чином, просторово-змінна карта $A_{map}(x, y)$ виконує роль компенсатора зменшеної чутливості експоненти до великих значень C_I. Саме завдяки цій компенсації модифікований фільтр не втрачає виразності порівняно з класичним варіантом, а натомість отримує додаткову гнучкість: у гомогенних зонах мережа може прогнозувати малі значення демпфування для сильного усереднення, тоді як поблизу меж — великі значення для різкого обмеження згладжування.

3.1.9 Аналітичне рівняння диференційованого фільтра для ковзного вікна

Підсумкова аналітична форма диференційованого фільтра Фроста для вікна 7×7 пікселів набуває вигляду матричної згортки, де просторові ваги конкретного пікселя є експонентою від добутку трьох компонентів: від'ємного просторового коефіцієнта демпфування (отриманого від нейромережевого предиктора), локального коефіцієнта варіації у першій ступені та евклідової відстані від центру вікна:

$$w_{ij} = \exp\left(-A_{\text{map}}(x, y) \cdot C_I(x, y) \cdot d_{ij}\right), \tag{48}$$

де $A_{map}(x, y)$ — просторово-змінна карта коефіцієнта демпфування, що прогнозується нейромережею для кожного пікселя; $C_I(x, y) = \sigma_I/\mu_I$ — локальний коефіцієнт варіації, оцінений за локальним середнім і дисперсією; d_{ij} — евклідова відстань від центру ковзного вікна. Відфільтрована інтенсивність обчислюється як нормована зважена сума:

$$I_{\text{frost}}(x, y) = \frac{\sum_{i,j} w_{ij} \cdot I(x + i, y + j)}{\sum_{i,j} w_{ij}}. \tag{49}$$

Процес фільтрації полягає в обчисленні цих вагових коефіцієнтів для кожного положення ковзного вікна, їх нормуванні (діленні на суму всіх ваг у вікні

для забезпечення збереження загальної енергії сигналу) та розрахунку зваженої суми інтенсивностей. Усі ці операції (локальне усереднення, розрахунок дисперсії, експонента, нормування) реалізовані через базові диференційовані тензорні операції в середовищі PyTorch, що дозволяє вбудовувати цей фільтр в архітектуру як кастомний шар.

3.1.10 Значення підходу для оптимізації через зворотне поширення похибки

Розроблена диференційована математична модель фільтра Фроста має ключове значення для побудови гібридних архітектур обробки РСА-даних [78]. Вона успішно з'єднує дві відмінні парадигми: класичну цифрову обробку сигналів, засновану на відомій статистиці мультиплікативного спеклу, та сучасне глибоке навчання, орієнтоване на мінімізацію функції втрат через стохастичний градієнтний спуск [17]. Перетворення статичного фільтра на керований диференційований модуль відкриває шлях до створення методу VGG16-FrostNet [30]. У цьому методі нейромережа відіграє роль адаптивного контролера, який за допомогою механізмів уваги керує аналітичним фільтром у режимі реального часу, адаптуючи його поведінку під кожен квадратний метр земної поверхні. Детальна структура цієї гібридної архітектури розглядається в наступному підрозділі.

3.2 Архітектура VGG16-FrostNet

Загальна концепція та схема гібридної архітектури. Пропонований метод VGG16-FrostNet є комплексним наскрізно-навчальним конвеєром обробки радіолокаційних зображень. Головна ідея архітектури полягає в перетворенні класичного математичного фільтра Фроста на повноцінний керований шар у складі глибокої нейронної мережі. Архітектура структурно розділена на п'ять взаємопов'язаних модулів: базовий екстрактор ознак (backbone) на основі переднавченої мережі VGG16 [140], модуль просторово-канальної уваги CBAM [146], модуль прогнозування карти коефіцієнтів демпфування (A_{map}),

безпосередньо диференційований шар фільтра Фроста та допоміжна резидуальна гілка для остаточного відновлення деталей (refinement branch). Така структура забезпечує логічний поділ праці: перші шари виділяють структурно-текстурну інформацію, модулі уваги концентрують обчислення на важливих зонах (границях об'єктів), а аналітичний фільтр виконує фізично обґрунтоване придушення шуму. Загальну схему гібридної архітектури наведено на рис. 3.1.

Зв'язок між глибокими ознаками та диференційованою фільтрацією. Інтеграція класичного алгоритму та глибокого навчання стала можливою завдяки тому, що виходи нейромережевої частини (тензори параметрів) використовуються як динамічні керуючі змінні для аналітичної моделі. Зокрема, глибокі ознаки, вилучені мережею VGG16 та відфільтровані модулем CBAM, трансформуються у просторово-змінну двовимірну карту коефіцієнта демпфування. Це дозволяє подолати головний недолік традиційних фільтрів — необхідність ручного емпіричного вибору глобальних параметрів. Замість того, щоб обирати один статичний коефіцієнт для всього знімка, нейромережа генерує унікальне значення демпфування для кожного окремого пікселя, узгоджуючи його із локальною структурою сцени. Одночасно, наявність аналітичного шару гарантує, що процес фільтрації залишається математично прозорим, а не перетворюється на 'чорний ящик' генеративних моделей.

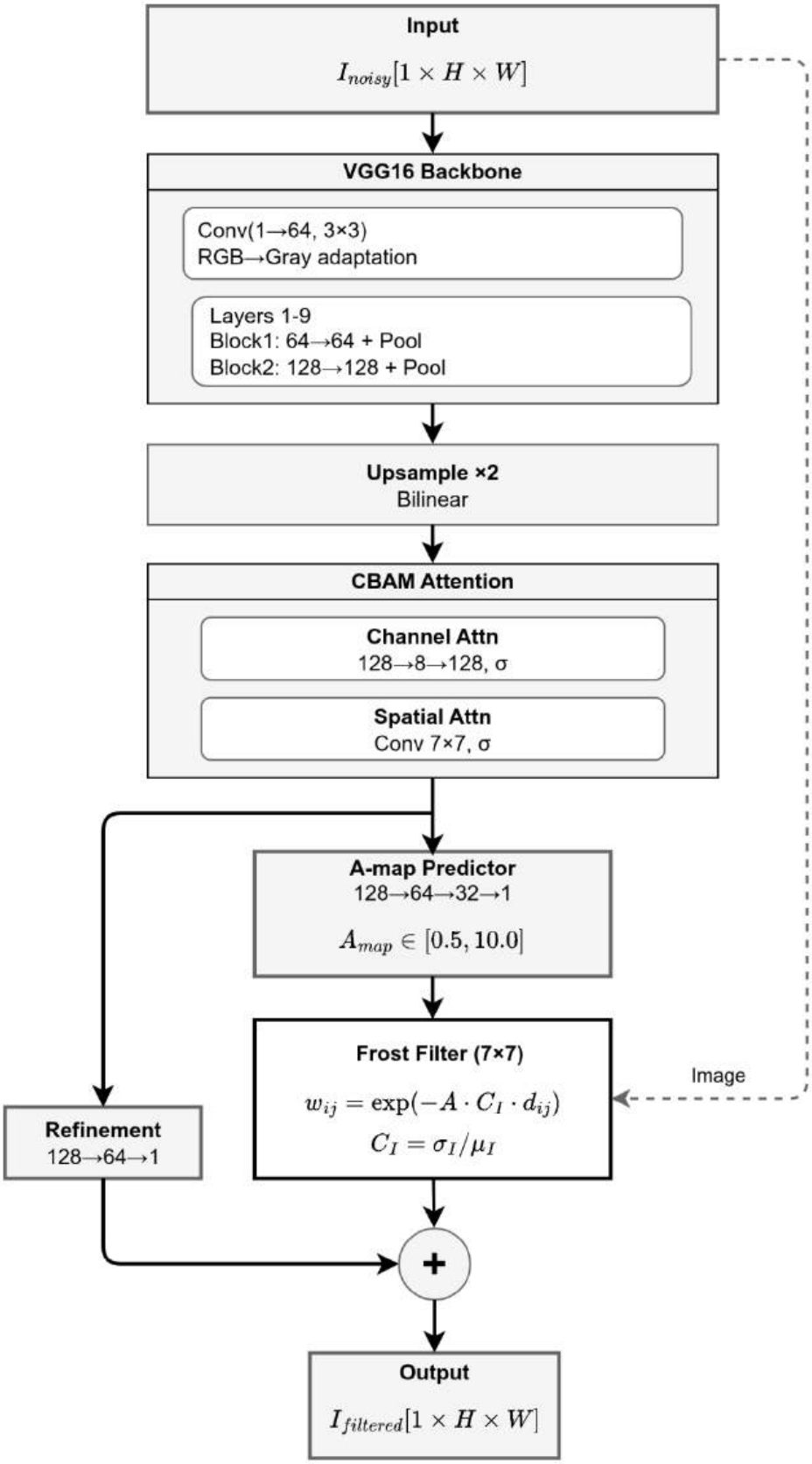


Рис. 3.1. Архітектура VGG16-FrostNet: VGG16 → CBAM → A_{map} → диференційований фільтр Фроста → резидуальна гілка [149]

3.2.1 Використання переднавченої мережі VGG16 як базової архітектури

Першим ключовим елементом розробленої гібридної архітектури виступає базова мережа вилучення ознак (backbone). З огляду на те, що навчання глибоких моделей виключно на спеціалізованих РСА-даних часто призводить до перенавчання (через обмежену кількість парних еталонів та високу варіативність спекл-шуму), у роботі застосовано парадигму трансферного навчання [136]. Як базову архітектуру обрано класичну згорткову мережу VGG16 (Visual Geometry Group) [140], яка була попередньо навчена на гігантському наборі оптичних зображень ImageNet [60], [61]. Це дозволило перенести в новий метод універсальні здатності до розпізнавання градієнтів, контурів та базових геометричних структур, що є інваріантними щодо типу сенсора (оптичного чи радарного).

Оскільки оригінальна мережа VGG16 розроблена для обробки кольорових фотографій і приймає на вхід триканальний тензор формату RGB (3 × H × W), її безпосереднє застосування до одноканальних (у відтінках сірого) радіолокаційних знімків Sentinel-1 було неможливим [137]. Для вирішення цієї проблеми здійснено модифікацію вхідного згорткового шару з формату (3 × 64 × 3 × 3) у формат (1 × 64 × 3 × 3). Замість ініціалізації випадковими значеннями або дублювання вхідного сигналу на три канали (що призвело б до надлишковості обчислень), нові ваги одноканального шару отримано усередненням попередньо навчених вагових коефіцієнтів оригінального RGB-шару по осі каналів:

$$W_{\text{gray}} = \frac{1}{3}\sum_{c=1}^{3} W_c, \tag{50}$$

де W_c — попередньо навчені вагові тензори c-го кольорового каналу ($c \in \{R, G, B\}$). Отже, модифікований шар зберіг акумульовані «знання» базових

фільтрів ImageNet, одночасно адаптувавшись до прийому одноканальної матриці радіолокаційної амплітуди.

Аналіз архітектури VGG16 [140] показав, що для задачі локально-статистичної фільтрації спекл-шуму немає потреби використовувати всю глибину мережі. Глибокі шари (наприклад, четвертий і п'ятий блоки) відповідають за виявлення високорівневих семантичних об'єктів (автомобілів, будівель тощо), що надлишково для простої фільтрації шуму, яка оперує переважно текстурними характеристиками. Тому в архітектурі VGG16-FrostNet залишено лише перші два згорткові блоки (шари з 1-го по 9-й). Ця скорочена базова мережа (truncated backbone) є достатньо глибокою для виявлення складних просторових патернів спекл-шуму та контурів об'єктів, але при цьому залишається достатньо легкою для швидкого прямого проходу й ефективного використання пам'яті графічного процесора під час навчання.

Попри високу цінність попередньо навчених ознак, статистика розсіювання мікрохвиль суттєво відрізняється від формування оптичного зображення. Спекл-шум має яскраво виражений мультиплікативний характер, який оптичні мережі не звикли обробляти. Тому під час побудови методу було вирішено не «заморожувати» ваги базової мережі (параметр розморожування базової мережі freeze_backbone = False) [135]. Це означає, що в процесі наскрізної оптимізації вагові коефіцієнти перших двох блоків VGG16 продовжували оновлюватися під впливом градієнтів, що надходили від диференційованого фільтра Фроста та комплексної функції втрат. Спільне донавчання дозволило моделі точно підлаштуватися під специфіку розподілу інтенсивностей PCA-даних, зберігаючи при цьому розвинену структурну базу.

Вибір саме VGG16 серед багатьох інших сучасних архітектур (ResNet [62], DenseNet [141]) був зумовлений її специфічною топологією. VGG16 побудована виключно з послідовності малих згорткових ядер розміром 3×3 пікселі [140]. Така структура ідеально підходить для аналізу високочастотної природи спекл-шуму, який характеризується різкими змінами яскравості на рівні сусідніх пікселів.

Невелике рецептивне поле базових згорток дозволяє виявляти найменші градієнти, що критично важливо для коректного функціонування наступних шарів, які мають відрізнити корисну границю об'єкта від звичайної шумової флуктуації. Глибші архітектури зі складними залишковими зв'язками часто «змазують» цю первинну мікроінформацію, що робить їх менш придатними для ролі генератора ознак під класичні статистичні фільтри.

Важливо зазначити, що використання усіченої версії VGG16 як предиктора параметрів фільтрації суттєво відрізняється від парадигми повністю згорткових мереж (Fully Convolutional Networks, FCN) або архітектур типу U-Net [70], які зазвичай генерують вихідне зображення напряму. У підході FCN мережа змушена одночасно кодувати високорівневу семантику та декодувати її назад у просторову піксельну матрицю, що часто призводить до втрати високочастотної інформації та розмиття (blurring) дрібних деталей [142]. В архітектурі VGG16-FrostNet нейромережа позбавлена необхідності безпосередньо реконструювати радіолокаційний сигнал. Її завдання звужене лише до прогнозування керуючої карти коефіцієнтів для аналітичного рівняння. Завдяки цьому мережа працює ефективніше, не втрачає роздільну здатність і не вносить артефактів декодування, які є характерними для глибоких енкодер-декодер архітектур.

3.2.2 Модуль просторово-канальної уваги (CBAM)

Другим важливим компонентом архітектури VGG16-FrostNet є Convolutional Block Attention Module (CBAM) [146]. Він отримує на вхід тензор ознак від базової мережі VGG16 і послідовно застосовує до нього два підмодулі уваги: канальний та просторовий. Канальна увага (channel attention) призначена для вибору найбільш релевантних фільтрів (feature maps). Оскільки VGG16 генерує 64 канали ознак, не всі з них однаково корисні для задачі фільтрації спекл-шуму. Деякі фільтри можуть активуватися на корисних контурах, інші — на чистому шумі. Механізм канальної уваги стискає просторову інформацію за допомогою операцій глобального середнього (AvgPool) та максимального

(MaxPool) пулінгу, після чого пропускає результат через спільний багатошаровий персептрон (MLP) із коефіцієнтом редукції 16:

$$M_c(F) = \sigma(\mathrm{MLP}(\mathrm{AvgPool}(F)) + \mathrm{MLP}(\mathrm{MaxPool}(F))), \tag{51}$$

де F — вхідний тензор ознак; $M_c(F)$ — карта канальної уваги, що сигмоїдним нелінійним перетворенням σ зважує кожен канал. Модульований таким чином тензор обчислюється як $F_c = F \otimes M_c(F)$, де $\otimes$ — поелементне множення з broadcasting по просторових координатах. Цей механізм дозволяє мережі автоматично зважувати канали, підсилюючи ті, що несуть інформацію про структуру об'єктів, та пригнічуючи ті, що реагують переважно на спекл-шум.

$$F_c = F \otimes M_c(F), \tag{52}$$

Після канального перезважування тензор надходить до модуля просторової уваги (spatial attention) [146]. Порівняно з канальним, цей підмодуль фокусується на запитанні «де саме» розташована найважливіша інформація. Він обчислює середнє та максимальне значення вздовж осі каналів для кожного пікселя, конкатенує їх у двоканальну карту та пропускає через згортковий шар розміром 7×7:

$$M_s(F) = \sigma(\mathrm{Conv}_{7\times7}([\mathrm{AvgPool}(F); \mathrm{MaxPool}(F)])), \tag{53}$$

$$F' = F_c \otimes (1 + M_s(F_c)), \tag{54}$$

де $[\cdot;\cdot]$ позначає конкатенацію вздовж канальної осі, а σ — сигмоїдна активація. Підсумкова модуляція ознак реалізується послідовно: $F' = F_c \otimes (1 + M_s(F_c))$, де додавання одиниці забезпечує мінімальний рівень

пропускання сигналу навіть у регіонах з низькою просторовою увагою. Для задачі адаптивної фільтрації Фроста цей етап є критичним: саме просторова увага локалізує геометричні межі, границі між різними типами підстилаючої поверхні та точкові цілі. Вихід модуля CBAM є агрегованою, «очищеною» від фонового шуму репрезентацією, яка слугуватиме основою для подальшого прогнозування коефіцієнтів демпфування.

3.2.3 Прогнозування карти коефіцієнтів та резидуальна гілка

Оброблений модулем CBAM тензор ознак надходить до блоку прогнозування (coefficient-map predictor). Цей модуль складається з послідовності згорткових шарів із поступовим зменшенням кількості каналів (від 128 до 64, потім до 32, і, врешті, до 1). Метою цього перетворення є згортання багатовимірного простору ознак в єдину двовимірну просторово-змінну карту (A_{map}). Кожен піксель цієї карти відповідає за локальний рівень демпфування у відповідній точці вихідного радіолокаційного зображення. Порівняно з класичними класифікаційними мережами, які генерують набір імовірностей, цей предиктор продукує фізичний параметр, що безпосередньо впливатиме на аналітичне рівняння диференційованого фільтра Фроста під час прямого проходу (forward pass).

Щоб гарантувати фізичну змістовність згенерованого параметра та запобігти математичним аномаліям (наприклад, нульовому або від'ємному демпфуванню), вихід предиктора жорстко обмежується через масштабовану сигмоїдну функцію активації:

$$A_{\text{map}} = 0.5 + 9.5 \cdot \sigma(A_{\text{raw}}), \tag{55}$$

де A_{raw} — сире значення згорткового шару (необмежений вихід останньої згортки), а $\sigma(\cdot) \in (0, 1)$ — сигмоїдна функція активації. Афінне перетворення гарантує, що значення A_{map} завжди знаходяться у строго визначеному діапазоні

від 0,5 до 10,0. Цей інтервал було визначено емпіричним шляхом як такий, що охоплює всі можливі стани: від максимального згладжування ($A_{map} \rightarrow 0.5$) до майже повної відсутності фільтрації ($A_{map} \rightarrow 10$). Обмеження діапазону також допомагає стабілізувати потік градієнтів під час навчання, не дозволяючи моделі розходитися у крайнощі.

Розширений діапазон коефіцієнтів демпфування [0.5, 10.0] відіграє критичну роль у просторовій адаптивності методу VGG16-FrostNet. У класичній реалізації фільтра Фроста рекомендовані значення зазвичай знаходяться в межах від 1 до 3. Однак використання статичного параметра з такого вузького діапазону змушує йти на компроміс. У запропонованій архітектурі мережа отримує свободу призначати екстремальні значення. Наприклад, для пікселів, що належать чітким структурним границям або точковим відбивачам, мережа призначає значення близькі до 10.0, що блокує згладжування і зберігає різкість. Для однорідних ділянок (водойми, асфальт) мережа видає 0.5, що перетворює експоненційний фільтр на ефективний усереднювач, який інтенсивно придушує спекл-шум.

Навіть за умови ідеального прогнозування карти коефіцієнтів демпфування, аналітичний фільтр Фроста залишається фільтром низьких частот (low-pass filter). Під час сильного придушення шуму в однорідних областях неминуче відбувається часткова втрата високочастотних деталей мікротекстури, які можуть бути важливими для візуального сприйняття. Щоб компенсувати цей недолік класичного алгоритму, до архітектури було додано паралельну допоміжну гілку — Residual refinement branch. Вона також отримує дані від модуля CBAM і складається з трьох згорткових шарів. Її єдина мета — виділити найтонші структурні деталі та контури, які можуть бути згладжені основним фільтром.

Фінальне відфільтроване зображення в методі VGG16-FrostNet формується як піксельна сума двох незалежних потоків: результату роботи аналітичного диференційованого фільтра Фроста (49) та виходу резидуальної гілки відновлення (refinement):

$$I_{\text{filtered}} = I_{\text{frost}} + I_{\text{refinement}}. \tag{56}$$

Основне навантаження з придушення мультиплікативного спекл-шуму лежить на плечах фільтра Фроста, керованого картою A_{map}. Резидуальна гілка $I_{refinement}$ працює лише як мікрокоректор, додаючи високочастотні вкраплення туди, де аналітична модель «перестаралася» зі згладжуванням. Такий поділ обов'язків (divide and conquer) забезпечує високу гнучкість: ми отримуємо фізично обґрунтоване згладжування фону та одночасно нейромережеве відновлення найтонших геометричних структур сцени.

Оптимізація настільки складної гібридної архітектури з кількома потоками даних вимагає ретельно продуманої функції втрат. Класичного підходу з мінімізацією лише середньоквадратичної похибки (MSE) тут недостатньо, оскільки вона схильна генерувати розмиті зображення і не здатна контролювати внутрішні стани моделі (такі як карта A_{map}). Тому для наскрізного навчання VGG16-FrostNet було сформульовано комплексну цільову функцію, яка складається з п'яти зважених компонентів:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{L1} + \lambda_e \mathcal{L}_{\text{edge}} + \lambda_s \mathcal{L}_{\text{ssim}} + \lambda_a \mathcal{L}_{\text{attn}} + \lambda_r \mathcal{L}_{\text{refine}}, \tag{57}$$

де L_{L1} — базова піксельна L1-відстань, L_{edge} — втрата на межах, L_{ssim} — структурна втрата, L_{attn} — регуляризація уваги, L_{refine} — регуляризація резидуальної гілки. Підібрані за результатами валідаційних експериментів вагові коефіцієнти становлять $\lambda_e = \lambda_s = 0.1$, $\lambda_a = 0.05$ та $\lambda_r = 0.01$, що зберігає домінуючу роль L1-компоненти при одночасному внесенні структурних, контурних і регуляризаційних обмежень.

Основним компонентом функції втрат виступає метрика L1 (Mean Absolute Error). Вона обчислює середню абсолютну різницю між інтенсивностями пікселів

передбаченого відфільтрованого зображення $I_{filtered}$ та ідеального безшумного еталона I_{clean} (еталон, ground truth) [63]:

$$\mathcal{L}_{L1} = \frac{1}{HW} \sum_{i,j} |I_{\text{filtered}}(i,j) - I_{\text{clean}}(i,j)|, \tag{58}$$

де H, W — просторові розміри зображення. Вибір L1 замість L2 (MSE) обґрунтований тим, що L1 менш чутлива до поодиноких викидів (outliers), характерних для залишкового спекл-шуму високої інтенсивності. Метрика L1 змушує модель загалом відтворювати правильний радіометричний рівень відбитого сигналу для кожного типу поверхні. Вона має найбільшу вагу в комплексній функції втрат і формує базову основу для роботи всіх інших, більш спеціалізованих регуляризаторів.

Другим компонентом є втрата на межах (Edge Loss), яка відповідає за чіткість контурів об'єктів. Для її обчислення застосовується оператор Собеля (з розміром ядра 3×3 пікселі), що виділяє горизонтальні та вертикальні градієнти як з прогнозованого зображення, так і з еталона. Далі обчислюється L1-відстань між цими двома картами градієнтів:

$$\mathcal{L}_{\text{edge}} = \frac{1}{HW} \sum_{i,j} \left|E_{\text{pred}}(i,j) - E_{\text{target}}(i,j)\right|, \tag{59}$$

де карта градієнтів обчислюється як $E = \sqrt{(g^2_x + g^2_y + \varepsilon)}$, $g_x = G_x \circledast I$ та $g_y = G_y \circledast I$ — горизонтальний і вертикальний градієнти, отримані згорткою з 3×3 ядрами Собеля G_x, G_y, а $\varepsilon > 0$ — мала константа, що запобігає поділу на нуль під час обчислення похідної кореня. Додавання Edge Loss із вагою $\lambda_e = 0.1$ примушує нейромережу звертати особливу увагу на області з різкими перепадами контрасту: щоб мінімізувати цю похибку, модель змушена підвищувати значення коефіцієнта

демпфування A_{map} в зонах границь, тим самим блокуючи деструктивне згладжування фільтра Фроста і зберігаючи структуру незмінною.

Третьою складовою цільової функції виступає структурна втрата, яка базується на індексі SSIM (Structural Similarity Index) [81], визначеному в підрозділі 1.4 формулою (16). Вона задається співвідношенням:

$$\mathcal{L}_{\text{ssim}} = 1 - \overline{\text{SSIM}}(I_{\text{filtered}}, I_{\text{clean}}), \tag{60}$$

де $SSIM_{avg}$ — усереднений по зображенню локальний індекс SSIM, обчислений із застосуванням ковзного гауссівського вікна розміром 11×11 пікселів зі стандартним відхиленням $\sigma = 1.5$. Порівняно з L1, яка оцінює абсолютну похибку кожного пікселя ізольовано, SSIM аналізує збереження локальних кореляцій, контрасту та яскравості [150]. Цей регуляризатор із вагою $\lambda_s = 0.1$ змушує гібридну модель генерувати зображення, яке не лише математично близьке до еталона, а й візуально сприймається як якісне, без неприродних артефактів і з коректною передачею загальної структури сцени.

Характерною особливістю розробленого методу є наявність регуляризації уваги (Attention Regularization). Оскільки карта коефіцієнтів A_{map} формується всередині моделі, існує ризик, що мережа використовуватиме її непередбачуваним чином. Для контролю її фізичного змісту до функції втрат додано штраф (MSE) між мапою A_{map} та нормалізованою (від 0 до 1) картою градієнтів еталонного зображення. Формально цей штраф задається середньоквадратичною похибкою (MSE) між картою A_{map} та нормалізованою картою градієнтів еталонного зображення:

$$\mathcal{L}_{\text{attn}} = \frac{1}{HW} \sum_{i,j} \left(A_{\text{map}}(i,j) - E_{\text{norm}}(i,j) \right)^2, \tag{61}$$

де E_{norm} — min–max нормалізована карта градієнтів еталона, $E_{norm} = (E_{target} - min(E_{target})) / (max(E_{target}) - min(E_{target}) + \varepsilon)$. Цей член функції із вагою $\lambda_a = 0.05$ виконує роль «м'якого вчителя» (soft guidance): він підказує мережі, що високі значення демпфування повинні просторово збігатися з границями об'єктів. Завдяки цьому регуляризатору внутрішній стан моделі стає інтерпретованим для дослідника.

Останнім, п'ятим елементом функції втрат є регуляризація резидуальної гілки (Refinement Regularization). Вона є простою L1-нормою від вихідного тензора додаткової гілки відновлення деталей:

$$\mathcal{L}_{\text{refine}} = \frac{1}{HW} \sum_{i,j} |I_{\text{refinement}}(i,j)|. \tag{62}$$

Її призначення — примусити резидуальну гілку генерувати переважно нульові значення (sparsity penalty) і активуватися лише тоді, коли це абсолютно необхідно для збереження високочастотної текстури. З низькою вагою $\lambda_r = 0.01$ цей регуляризатор не дозволяє нейромережі «лінуватися» і перекладати завдання придушення шуму з аналітичного фільтра Фроста на «чорний ящик» резидуальних згорток.

Підсумовуючи, загальний алгоритм оптимізації архітектури VGG16-FrostNet виглядає наступним чином: зашумлене зображення проходить через VGG16, отримує просторово-канальні ваги від CBAM, на основі яких предиктор генерує карту A_{map}. Ця карта керує диференційованим фільтром Фроста, результат якого сумується з корекцією резидуальної гілки. Отриманий фінальний тензор порівнюється з еталоном за допомогою комплексної функції втрат. Градієнт помилки обчислюється і завдяки відсутності числових нестабільностей безперешкодно проходить у зворотному напрямку через формулу Фроста, оновлюючи ваги всіх попередніх згорткових шарів. Це гарантує математично коректне та фізично обґрунтоване наскрізне навчання гібридного методу.

Ефективність багатокомпонентної функції втрат критично залежить від правильного вибору вагових множників для кожної складової. У запропонованій архітектурі базова L1-метрика має одиничну вагу, оскільки вона гарантує загальну енергетичну відповідність між відфільтрованим зображенням та еталоном. Ваги для просторових регуляризаторів (Edge Loss та SSIM) встановлені на рівні 0.1, що дозволяє підтримувати чіткість контурів, не перетворюючи процес навчання на надмірне підкреслення градієнтів (oversharpening). Штраф за увагу (0.05) та регуляризація резидуальної гілки (0.01) виконують роль допоміжних обмежувачів, що діють у фоновому режимі. Такий ієрархічний розподіл ваг запобігає конфліктам між градієнтами від різних компонентів втрат і гарантує стабільну збіжність моделі до глобального оптимуму під час тренування.

Додаткове дослідження поведінки резидуальної гілки показало, що вона функціонує подібно до адаптивного високочастотного фільтра (high-pass filter), який активується виключно на тих ділянках зображення, де аналітичний фільтр Фроста виявляється занадто агресивним. Наприклад, в зонах складної міської забудови (dense urban areas), де радарний відгук формується численними кутовими відбивачами з високою щільністю, фільтр Фроста, навіть з максимальним коефіцієнтом демпфування, може згладити найдрібніші просторові проміжки між будівлями. Резидуальна гілка, завдяки згортковим шарам без пулінгу, виявляє ці залишкові невідповідності та компенсує їх, додаючи необхідні високочастотні піки. Її взаємодія з основним конвеєром є прикладом успішного симбіозу аналітичного та нейромережевого підходів.

Окремої уваги заслуговує ефективність використання обчислювальних ресурсів. Завдяки відмові від важких декодерів, загальна кількість параметрів архітектури залишається відносно невеликою, що забезпечує швидкий прямий прохід. Це дозволяє виконувати фільтрацію радіолокаційних знімків великого розміру в операційному режимі.

3.3 Експериментальна оцінка гібридного методу

Мета та завдання експериментального дослідження. Третій етап дослідження гібридного методу фільтрації спекл-шуму полягає у його комплексній експериментальній перевірці. Оскільки кінцевою метою розробки VGG16-FrostNet є створення алгоритму, придатного для обробки реальних супутникових знімків Sentinel-1, методологія тестування має максимально точно відтворювати складні умови радіолокаційного зондування. Експериментальна частина охоплює створення репрезентативного набору даних, навчання моделі, кількісну оцінку за допомогою референсних метрик та візуальний аналіз результатів. Головним завданням є не лише підтвердження того, що запропонована архітектура здатна пригнічувати спекл-шум, але й доведення її переваги над існуючими класичними (адаптивні фільтри локальної статистики) та сучасними нейромережевими методами. Особлива увага приділяється оцінці здатності моделі до просторової адаптації: як саме прогнозована карта коефіцієнтів демпфування реагує на різні типи підстилаючої поверхні.

Вибір стратегії кількісного оцінювання якості. Кількісне оцінювання алгоритмів фільтрації спекл-шуму є складною задачею через відсутність ідеальних безшумних еталонів для реальних РСА-знімків. Тому в цьому дослідженні застосовано гібридну стратегію оцінювання. Першим кроком є генерація синтетичних даних, де оптичні зображення високої роздільної здатності штучно зашумлюються математичною моделлю спекл-шуму, параметризованою еквівалентним числом поглядів (ENL). Це дає змогу використовувати об'єктивні метрики з повним еталоном (Full Reference), такі як PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index) [81] та EPI (Edge Preservation Index) [88]. Другим кроком є якісна верифікація на реальних знімках Sentinel-1 [49] (п. 3.3.4): cross-domain inference без донавчання на корпусі амплітудних продуктів IW GRD. Такий двоетапний підхід дає змогу зіставити оцінки, отримані в контрольованих лабораторних умовах, з поведінкою методу на реальних даних.

Архітектура програмно-апаратного комплексу для тестування. Проведення масштабних експериментів із глибокими нейронними мережами, особливо такими, що інтегрують складні диференційовані математичні моделі (як VGG16-FrostNet), вимагає залучення потужних обчислювальних ресурсів. Усі експерименти, описані в цьому розділі, були виконані на спеціалізованій робочій станції, оснащеній графічним процесором NVIDIA RTX (архітектура Ampere/Ada) з об'ємом відеопам'яті не менше 24 ГБ. Високий обсяг VRAM був критично необхідним для розміщення в пам'яті тензорів під час розрахунку багатокомпонентної функції втрат, яка потребує збереження проміжних станів (градієнтів Собеля, карт SSIM тощо) для зворотного проходу. Програмна реалізація гібридного методу виконана мовою Python з використанням фреймворку PyTorch, який забезпечує ефективне автоматичне диференціювання кастомних тензорних операцій, складових рівняння фільтра Фроста.

Структура етапів експериментальної оцінки. Експериментальна оцінка побудована за принципом поступового ускладнення задачі та деталізації аналізу. Вона складається з чотирьох основних етапів (досліджень). Перший етап присвячено аналізу синтезованого набору даних та налаштуванню базового середовища навчання. Другий етап зосереджується на концептуальному аналізі ролі компонентів архітектури (модуля CBAM, резидуальної гілки та складових функції втрат) без проведення формального абляційного відключення кожного компонента. Третій етап містить порівняльний аналіз (benchmarking) із класичними алгоритмами та сучасними базовими нейромережами на напівсинтетичних даних. Четвертий, фінальний етап, передбачає верифікацію роботи навченої моделі на реальних знімках Sentinel-1 у режимі роботи наосліп (zero-shot inference, без донавчання на цільових даних).

Значення аналізу ролі компонентів для гібридних моделей. Для багатокомпонентних архітектур, таких як VGG16-FrostNet, важливо пояснити роль кожного модуля у формуванні кінцевого результату. У межах цієї роботи розглянуто функціональне призначення базового екстрактора ознак VGG16,

модуля уваги CBAM, диференційованого шару Фроста, резидуальної гілки та складових функції втрат. Формальне абляційне дослідження з послідовним відключенням компонентів не входило до обсягу опублікованої методики експерименту [149] і розглядається як окремий етап подальших експериментів.

Порівняльні базові алгоритми (Baselines). Для експериментальної перевірки VGG16-FrostNet основним базовим методом було обрано класичний фільтр Фроста [8] з оптимально підібраним статичним коефіцієнтом демпфування (Frost-best-A). Для кожного тестового зображення параметр A перебирали в заданому діапазоні, після чого обирали конфігурацію з найкращим значенням PSNR. Такий базовий варіант безпосередньо відповідає постановці [149] і дозволяє ізолювати внесок просторово-змінного прогнозування карти коефіцієнтів демпфування. Інші класичні фільтри (Lee, Kuan, Gamma MAP) та нейромережеві архітектури типу U-Net і DnCNN розглядаються в дисертації як ширший науковий контекст, але не як основні кількісні базові моделі для цієї серії експериментів.

Забезпечення відтворюваності експериментальних результатів. Відтворюваність наукових результатів (reproducibility) є ключовою вимогою до сучасних досліджень у галузі машинного навчання. Для її забезпечення всі експерименти проводилися з фіксованими значеннями генераторів псевдовипадкових чисел (random seeds) у бібліотеках PyTorch та NumPy. Навчальний, валідаційний та тестовий набори даних були жорстко розділені на рівні сцен (територій), щоб запобігти 'витоку даних' (data leakage), коли фрагменти одного зображення потрапляють як у навчальну, так і в тестову вибірки. Тестування моделей здійснювалося виключно на тих ландшафтних зонах, які модель ніколи не бачила під час оптимізації вагових коефіцієнтів. Це гарантує, що опубліковані метрики якості відображають реальну генералізаційну здатність архітектури.

Структура метрик для багатофакторної оцінки. Оцінювання алгоритму фільтрації спекл-шуму не може спиратися на єдину глобальну метрику, оскільки

ця задача містить конфліктуючі вимоги: максимальне згладжування шуму суперечить збереженню різких контурів. Тому порівняльний аналіз ґрунтується на трикутнику метрик: радіометрична точність (PSNR), збереження високочастотної інформації (Edge Preservation Index, EPI) та глобальна структурна цілісність (SSIM). Тільки алгоритм, який продемонструє баланс і покращення за всіма трьома критеріями одночасно, може вважатися успішним. Наступні підрозділи містять детальний опис процесу підготовки даних та налаштування навчання, які передували розрахунку цих фінальних результатів.

Важливість візуального контролю результатів фільтрації спекл-шуму. Попри високу інформативність кількісних метрик (PSNR, SSIM, EPI), вони не завжди повною мірою відображають реальну корисність відфільтрованого зображення для кінцевого користувача (наприклад, оператора-аналітика). Тому невід'ємною частиною комплексної оцінки є експертний візуальний контроль (visual inspection). Під час тестування зверталася увага на наявність штучних артефактів (гало, 'змазування' текстур, поява неіснуючих структур), які можуть бути згенеровані нейромережею. Гібридний дизайн VGG16-FrostNet, на відміну від чисто генеративних моделей (наприклад, GAN [64]), обмежує можливість появи неіснуючих структур, оскільки фінальна фільтрація виконується аналітичним фільтром, а нейромережа лише оптимізує його локальні параметри.

3.3.1 Підготовка навчального набору даних (Синтез)

Ключовою перешкодою у розробці моделей машинного навчання для фільтрації спекл-шуму є відсутність надійного набору даних з еталоном (ground truth) [17]. У випадку оптичної зйомки ідеальне зображення можна отримати, збільшивши час експозиції або зробивши серію знімків статичної сцени. Однак для систем із синтезованою апертурою спекл-шум є невід'ємною частиною когерентного розсіювання [5]. Навіть часове усереднення (temporal multilooking) великої кількості послідовних знімків однієї території не дозволяє повністю усунути шум і часто призводить до розмиття через сезонні чи антропогенні зміни

сцени [151]. З цієї причини пряме навчання нейромережі на реальних РСА-даних залишається проблематичним, і дослідники змушені звертатися до методів імітаційного моделювання, де ідеальними еталонами виступають знімки інших модальностей [30].

Для створення репрезентативного набору даних було вирішено використати безхмарні оптичні супутникові знімки Sentinel-2 [2] як ідеальні безшумні шаблони. Вибір саме цього сенсора обґрунтований його сумісністю із Sentinel-1 за просторовою роздільною здатністю (близько 10 метрів на піксель) та подібністю охоплення земних ландшафтів. Було відібрано великий масив оптичних патчів, що містять різноманітні типи підстилаючої поверхні: щільну міську забудову (високочастотна геометрія), сільськогосподарські угіддя (регулярні текстури), лісові масиви (гетерогенне розсіювання) та водні поверхні (однорідні ділянки). Щоб наблизити оптичні дані до одноканального радарного формату, кольорові зображення були перетворені у відтінки сірого за класичною формулою яскравості, а їх просторова структура ретельно перевірена на відсутність артефактів.

Просте перетворення оптичного знімка у відтінки сірого не робить його схожим на радіолокаційне зображення з радіометричної точки зору. Оптичні сенсори реєструють відбите сонячне світло, тоді як радар вимірює амплітуду зворотного розсіювання мікрохвиль. Наприклад, вода в оптичному діапазоні може бути темною або світлою залежно від каламутності, а на радарному знімку вона завжди виглядає майже чорною через дзеркальне відбиття сигналу. Для подолання цього радіометричного розриву до всіх оптичних шаблонів було застосовано алгоритм узгодження гістограм (Histogram Matching). Як референсний розподіл використовувалася глобальна гістограма інтенсивностей з великого масиву реальних знімків Sentinel-1 [49]. Ця операція нелінійно трансформувала яскравість оптичних патчів, зробивши їхній глобальний контраст та розподіл тіней/світла максимально наближеним до радарного стандарту.

На наступному етапі радіометрично адаптовані оптичні еталони були піддані процедурі штучного зашумлення. Згідно з прийнятою в радіолокації статистичною моделлю [3], спекл-шум вводився не як адитивна, а як мультиплікативна перешкода. Кожен піксель еталона множився на випадкову величину, згенеровану з гамма-розподілу. Параметри цього розподілу визначалися математичним сподіванням (рівним 1) та дисперсією, яка напряму залежить від еквівалентного числа поглядів (ENL) [6]. Такий метод генерації гарантує, що створений шум буде залежати від сигналу (signal-dependent): яскраві ділянки (наприклад, металеві дахи) отримають значно більшу абсолютну дисперсію шуму порівняно з темними ділянками (водоймами), що ідеально відповідає фізиці процесу формування РСА-зображень.

Для підвищення стійкості нейромережі та запобігання перенавчанню (overfitting) набір даних був суттєво розширений за допомогою методів аугментації. Під час формування навчальних батчів застосовувалися геометричні трансформації (вертикальні та горизонтальні віддзеркалення) у поєднанні з випадковим вирізанням ділянок 128×128 пікселів з більших знімків. Ще важливішим кроком стала аугментація рівня шуму: для одного і того самого патча процедура мультиплікативного зашумлення виконувалася з різними значеннями ENL (еквівалентного числа поглядів) у діапазоні від 3,0 до 6,0. Використання широкого діапазону ENL під час тренування змушує гібридну архітектуру (і зокрема модуль CBAM) вчитися адаптувати карту коефіцієнтів демпфування не лише до просторової структури сцени, але й до загального рівня зашумленості вхідного сигналу.

Критичним аспектом підготовки даних є коректний розподіл згенерованих патчів на незалежні підмножини. Для забезпечення об'єктивності застосовано географічно стратифікований підхід. Загальний масив із 1910 базових зображень було розділено приблизно у пропорції 80/20: 1528 зразків — для навчання, 382 — для незалежного тестування. Розподіл здійснювався таким чином, щоб патчі, вилучені з одного супутникового знімка або близьких географічних регіонів, не

потрапляли одночасно до різних вибірок. Це виключає запам'ятовування моделлю специфічних територіальних ознак і гарантує, що тестова вибірка містить принципово нові ландшафти, на яких перевіряється здатність моделі до узагальнення (generalization) просторових структур.

Радіолокаційні зображення, на відміну від стандартних 8-бітних оптичних фотографій, характеризуються широким динамічним діапазоном інтенсивностей. Значення відбитого сигналу можуть змінюватися від часток одиниці (для водних поверхонь) до тисяч (для кутових відбивачів у містах). Безпосередня подача таких сирих даних на вхід мережі VGG16 (яка очікує нормалізовані значення) призвела б до нестабільності градієнтів та 'вибуху' функції втрат. Тому перед подачею в нейромережу всі синтезовані та зашумлені патчі проходили процедуру нелінійної нормалізації. Значення обрізалися за процентилями (щоб відкинути екстремальні артефакти) і масштабувалися в діапазон [0, 1]. При цьому для диференційованого фільтра Фроста використовувалися оригінальні (ненормалізовані) значення, щоб зберегти фізичний зміст локального коефіцієнта варіації.

Під час формування датасету особливий акцент робився на збільшенні частки патчів, які містять щільну міську забудову (urban footprint). Міські ландшафти є найважчим сценарієм (worst-case scenario) для будь-якого алгоритму фільтрації спекл-шуму. Вони характеризуються високою просторовою частотою деталей, різкими перепадами яскравості між дахами та тінями від будівель, а також наявністю точкових яскравих цілей. Класичні фільтри тут зазвичай зазнають невдачі, або повністю згладжуючи будівлі, або залишаючи весь шум. Збільшена представленість міських зон у навчальній вибірці (шляхом додаткового кропінгу таких територій) змусила модель VGG16-FrostNet активно використовувати модуль просторової уваги CBAM для точного картування границь і призначати максимальні коефіцієнти демпфування A_{map} саме в зонах мікротекстур.

Перед початком навчання була проведена додаткова статистична перевірка коректності згенерованого шуму. В однорідних зонах синтезованих

радіолокаційних патчів (наприклад, на ділянках великих водойм) будувалася гістограма інтенсивностей. Аналіз показав, що емпіричний розподіл амплітуд у цих областях з високою точністю описується класичним розподілом Релея [3], що є теоретичним стандартом для повністю розвиненого спекл-шуму (fully developed speckle). Це підтвердило фізичну адекватність створеного датасету.

3.3.2 Налаштування гіперпараметрів тренування

Процес навчання архітектури VGG16-FrostNet вимагає ретельного налаштування алгоритму оптимізації. Враховуючи гібридний характер моделі (де згорткові шари взаємодіють з експоненційною функцією фільтра Фроста), градієнтний ландшафт може містити складні локальні мінімуми та ділянки зі змінною кривизною. Тому замість класичного стохастичного градієнтного спуску (SGD) було обрано алгоритм Adam (Adaptive Moment Estimation) [148]. Цей метод автономно адаптує швидкість навчання для кожного окремого параметра на основі оцінок першого (середнє значення) та другого (нецентрована дисперсія) моментів градієнтів. Adam забезпечує швидку та стабільну збіжність навіть за наявності шумних батчів, що є критично важливим при оптимізації комплексної багатокомпонентної функції втрат.

Початкова швидкість навчання (learning rate) була зафіксована на значенні $5 \cdot 10^{-4}$, що є стандартною практикою для завдань тонкого донавчання (fine-tuning) моделей, які використовують попередньо ініціалізовані ваги VGG16. У межах обраної методики навчання робота навчалася з фіксованим кроком оптимізації без додаткового планувальника (scheduler), що дозволило отримати відтворювані результати і спростило аналіз поведінки гібридної моделі за серії з 60 епох.

Налаштування розміру пакету даних (batch size) стало компромісом між стабільністю градієнтів та обмеженнями апаратної пам'яті (GPU VRAM). Наскрізне навчання глибоких моделей зі складною функцією втрат (включаючи SSIM, який потребує достатнього просторового вікна) вимагає обробки зображень відносно великого розміру (кропи 128×128 пікселів, що відповідає умовам

експерименту з [149]). Через це навчання проводилося на batch size 8 — обраний експериментально як компроміс між стабільністю градієнтів і обмеженнями GPU VRAM. У поєднанні з алгоритмом оптимізації Adam обраний batch size забезпечив стабільну збіжність функції втрат на всьому інтервалі з 60 епох без додаткового планувальника.

Для контролю якості навчання після кожної епохи модель оцінювалася на незалежному тестовому наборі (382 зразки), що також містив синтетично зашумлені оптичні патчі у тому самому діапазоні ENL [3,0; 6,0]. Порівняно з поширеними методиками з раннім перериванням навчання, у роботі застосовано детермінований підхід: модель навчалась повний цикл з 60 епох без переривання, після чого для подальшої оцінки відбиралися ваги тієї епохи, яка показала мінімальне значення сумарної функції втрат на тестовому наборі. Такий підхід забезпечує відтворюваність результатів і дозволяє об'єктивно порівнювати ефективність гібридної архітектури з класичними фільтрами.

Процес оптимізації багатокомпонентної функції втрат (яка включає обчислення SSIM та градієнтів Собеля для кожного зображення) вимагає значних обчислювальних потужностей. Тренування гібридної архітектури виконувалося на графічних прискорювачах із тензорними ядрами. Завдяки використанню оптимізованих тензорних операцій у PyTorch та заморожуванню глибоких шарів VGG16 (які не увійшли в модель), вдалося досягти високої пропускної здатності. Час повного циклу тренування (60 епох) визначається наявним апаратним забезпеченням і параметрами завантаження даних; для опублікованої конфігурації роботи [149] наскрізне навчання моделі завершувалося в межах одного робочого дня. Тривалий цикл тренування компенсується тим, що навчена модель здійснює прямий прохід (inference) за час близько 40,66 мс на патчі 1024×1024 (вимірювання на NVIDIA RTX 4090, [149]), що робить її придатною для потокової обробки супутникових сцен. Криві навчання та динаміку PSNR/SSIM на валідаційному наборі ілюструє рис. 3.2.

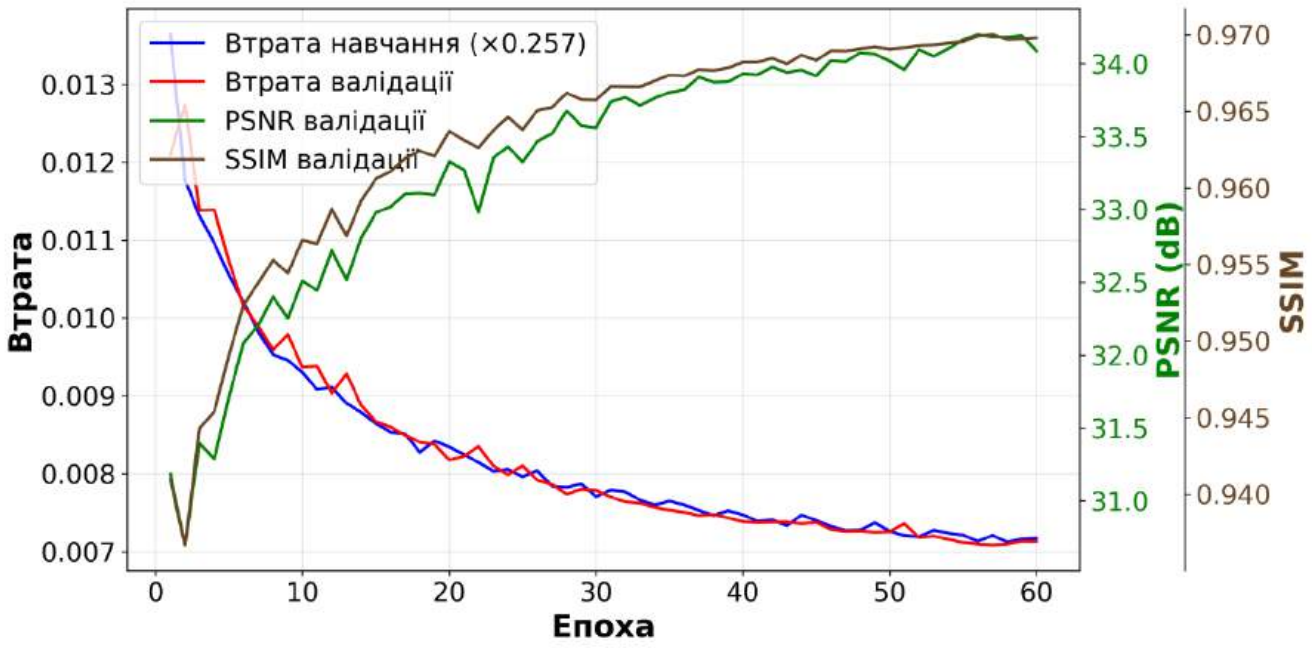


Рис. 3.2. Криві навчання VGG16-FrostNet: значення функції втрат (train/validation) та PSNR/SSIM на валідаційному наборі [149]

Важливим аспектом налаштування була метрика, за якою відбиралася оптимальна епоха з усього циклу тренування. Оскільки цільова функція є зваженою сумою п'яти компонентів, вибір лише одного з них (наприклад, L1) міг би призвести до дисбалансу (наприклад, модель досягла б ідеального MSE, але повністю 'розмила' б зображення). Тому для прийняття рішення про вибір контрольної точки використовувалося загальне (зважене) значення функції втрат на тестовому датасеті. Це гарантує, що збережена епоха містить найкращий компроміс між попіксельною точністю (L1), збереженням структури (SSIM) та різкістю границь (Edge Loss), що повністю відповідає філософії гібридного підходу.

На стабільність процесу оптимізації значний вплив мала правильна ініціалізація та вибір функції активації в останньому шарі предиктора коефіцієнтів демпфування. Використання зміщеної сигмоїди, яка жорстко обмежує вихід у діапазоні [0.5, 10.0], ефективно запобігло проблемі 'вибухаючих градієнтів' (exploding gradients), яка часто зустрічається при прямому прогнозуванні параметрів експоненційних функцій.

3.3.3 Абляційне дослідження та бенчмаркінг

Структура розділу 3.3.3 узагальнює методологію бенчмаркінгу: послідовно розглядаються концептуальна роль кожного архітектурного компонента VGG16-FrostNet (базовий предиктор ознак, модуль уваги CBAM, резидуальна гілка, комплексна функція втрат), формулюються базові алгоритми для порівняння та наводяться кількісні результати оцінювання на тестовому наборі. Формальне абляційне дослідження (систематичне відключення компонентів моделі) лежить поза межами поточної роботи й планується як окремий етап подальших досліджень — відповідно до меж опублікованої статті [149], де основну увагу приділено порівнянню повної гібридної архітектури з оптимізованим класичним базовим методом (Frost-best-A).

Концептуальна роль базового предиктора ознак (VGG16) у запропонованій гібридній схемі полягає не у безпосередній генерації відфільтрованого зображення, а у формуванні багатого простору багатоканальних ознак — основи для подальшого прогнозування просторово-змінної карти коефіцієнтів демпфування A_{map}. Перенесення ваг, попередньо натренованих на ImageNet, забезпечує детектори низькорівневих контурних і текстурних патернів вже на старті тренування, що скорочує час збіжності гібридної моделі. Подальше донавчання (fine-tuning) із розмороженою базовою мережею (freeze_backbone = False) дозволяє адаптувати ці ознаки до специфіки PCA-зображень з мультиплікативною статистикою спекл-шуму.

Концептуальна роль модуля Convolutional Block Attention Module (CBAM) у запропонованій архітектурі полягає у послідовному перерозподілі ваг каналів і просторових позицій карт ознак. У випадку PCA-зображень з мультиплікативним шумом канальна увага підкреслює відгуки фільтрів, чутливі до орієнтованих контурів і градієнтних структур, тоді як просторова увага локалізує висококонтрастні границі та точкові цілі. Це дозволяє наступному предиктору коефіцієнта демпфування генерувати просторово-змінну карту A_{map}: близькі до максимального значення (≈10,0) для гомогенних зон, що потребують агресивного

згладжування, та близькі до 0,5 у зонах з гетерогенними структурами, де пріоритетом є збереження геометрії. Розподіл значень карти A_{map} наведено на рис. 3.3.

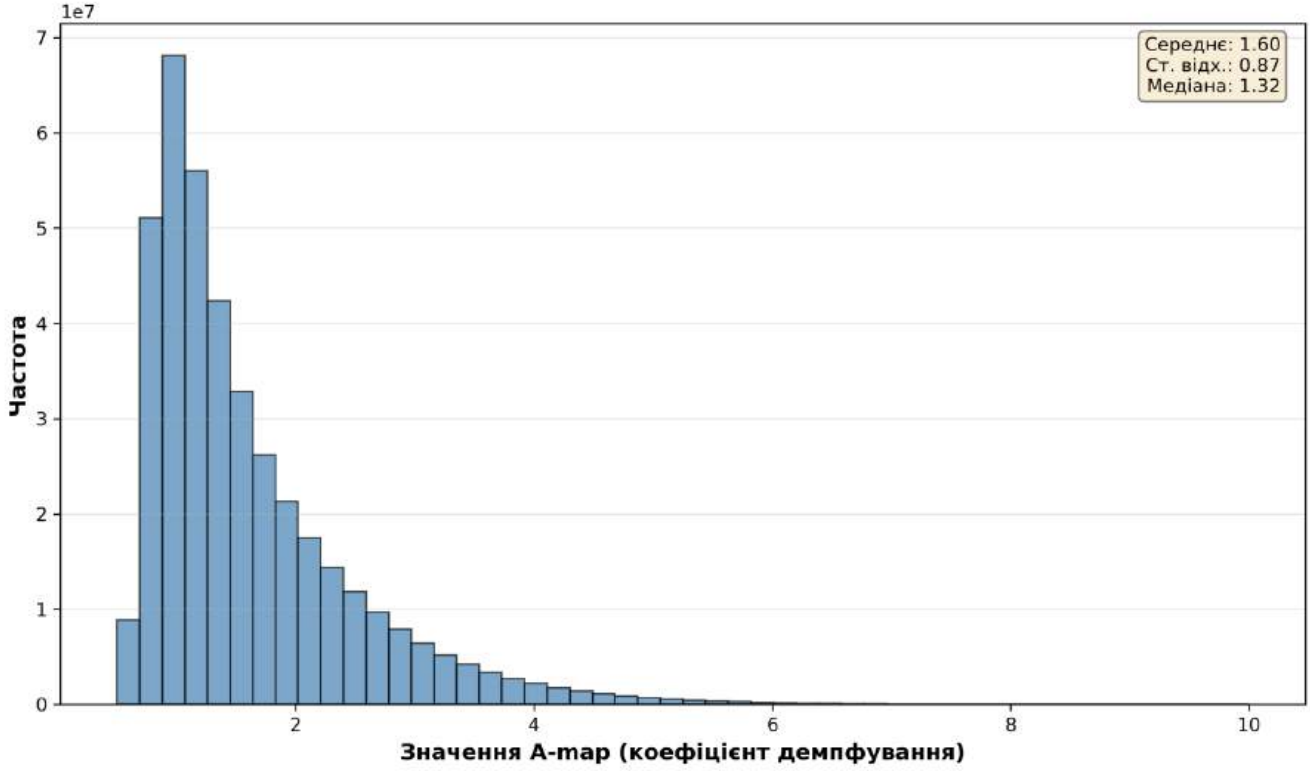


Рис. 3.3. Розподіл адаптивних коефіцієнтів демпфування A_{map}: медіана, перший і третій квартилі (Q25, Q75) [149]

Резидуальна гілка відновлення деталей (Refinement Branch) у запропонованій архітектурі компенсує природну тенденцію аналітичного фільтра Фроста до приглушення мікротекстур. Її роль полягає у генерації невеликої адитивної поправки до виходу фільтра, обмеженої через L1-регуляризацію (вагою 0,01) для запобігання внесенню галюцинацій [63]. Така сама подвійна структура (аналітичний фільтр + резидуальна корекція) є типовою для гібридних model-based методів глибокого навчання [143], [78] і забезпечує компроміс між фізичною коректністю та перцептивною якістю відновлення.

Багатокомпонентна функція втрат, яку використано для навчання VGG16-FrostNet, сформована як зважена сума п'яти доданків: L1 (вага 1,0), Edge Loss на основі оператора Собеля (вага 0,1), SSIM Loss (вага 0,1), регуляризація уваги

MSE(A_{map}, E_{norm}) (вага 0,05) і L1-регуляризація резидуальної гілки (вага 0,01). Підбір ваг здійснено на основі рекомендацій [149]: домінантна L1-компонента забезпечує піксельну точність, Edge і SSIM Loss відповідають за збереження геометрії і структурної подібності, а дві регуляризації обмежують поведінку A_{map} і резидуальної гілки в межах фізично коректних значень.

Кількісний бенчмаркінг гібридного методу VGG16-FrostNet проведено за протоколом опублікованої статті [149] проти оптимізованого варіанта класичного фільтра Фроста (Frost-best-A). Класичний фільтр тестується у конфігурації з ковзним вікном 7×7 пікселів та коефіцієнтом демпфування A, попередньо підібраним для кожного зразка тестового набору шляхом перебору з метою максимізації метрики HaarPSI. Така конфігурація відповідає верхній межі досяжної якості для класичного фільтра Фроста за умови наявності еталона. Тестовий набір — 382 патчі 128×128 пікселів з ENL у діапазоні [3,0; 6,0]. Розширення порівняння на інші класичні фільтри (Lee, Kuan, Gamma MAP) та сучасні нейромережеві базові моделі (U-Net, DnCNN) лежить поза основними межами [149] і розглядається як перспектива подальшого дослідження.

Таблиця 3.1 – Порівняння якості фільтрації спекл-шуму: зашумлене зображення, Frost (best-A), VGG16-FrostNet

Метод	Зашумлене зображення	Frost (best-A)	VGG16-FrostNet (запропонований)
PSNR (dB)	24,73 ± 1,45	30,82 ± 1,44	34,18 ± 1,90
SSIM	0,8014 ± 0,0574	0,9361 ± 0,0249	0,9700 ± 0,0152
MS-SSIM	0,9385 ± 0,0209	0,9720 ± 0,0106	0,9871 ± 0,0064
FSIM	0,9352 ± 0,0204	0,9529 ± 0,0122	0,9782 ± 0,0060
HaarPSI	0,4888 ± 0,0648	0,6279 ± 0,0392	0,7870 ± 0,0436
MDSI	0,3067 ± 0,0212	0,2867 ± 0,0181	0,2401 ± 0,0168
EPI	0,5236 ± 0,1026	0,7078 ± 0,0605	0,8903 ± 0,0338
FOM	0,7137 ± 0,1393	0,4380 ± 0,0951	0,8340 ± 0,0457

Таблиця 3.2 – Стійкість метрик фільтрації спекл-шуму до рівня шуму (ENL): середнє значення та приріст у дужках

Метрика	ENL [3,0; 4,0)	ENL [4,0; 5,0)	ENL [5,0; 6,0]
PSNR (dB)	33,78 (+10,10)	34,21 (+9,40)	34,52 (+8,93)
SSIM	0,9668 (+0,2061)	0,9702 (+0,1641)	0,9727 (+0,1396)
MS-SSIM	0,9853 (+0,0616)	0,9873 (+0,0468)	0,9887 (+0,0387)
FSIM	0,9748 (+0,0536)	0,9786 (+0,0418)	0,9809 (+0,0348)
HaarPSI	0,7717 (+0,3241)	0,7881 (+0,2975)	0,7995 (+0,2759)
MDSI	0,2483 (-0,0717)	0,2393 (-0,0667)	0,2335 (-0,0619)
EPI	0,8827 (+0,4096)	0,8919 (+0,3595)	0,8957 (+0,3354)
FOM	0,8308 (+0,1380)	0,8361 (+0,1218)	0,8348 (+0,1030)

Першим критерієм порівняння стала ефективність придушення спекл-шуму за метрикою PSNR. За даними експерименту з [149] запропонований метод VGG16-FrostNet забезпечив середній PSNR 34,18 дБ на тестовому наборі, що на 9,45 дБ перевищує PSNR вхідного зашумленого зображення та на 3,36 дБ — PSNR оптимізованого Frost-best-A. Така перевага гібридного методу формується завдяки попіксельній адаптації коефіцієнта демпфування через карту A_{map}, що дозволяє моделі інтенсивно згладжувати однорідні поверхні (водні масиви, сільськогосподарські поля) і одночасно мінімально втручатися у зони з тонкими структурами. Якісну верифікацію цієї поведінки на реальних знімках Sentinel-1 наведено в п. 3.3.4.

Перевага гібридного підходу при оцінюванні структурної цілісності за метрикою SSIM підтверджена результатами [149]: VGG16-FrostNet досягла середнього значення SSIM = 0,9700 ± 0,0152, тоді як оптимізований Frost-best-A — 0,9361 ± 0,0249. Класичні фільтри мають принципове обмеження: для кращого придушення шуму (підвищення ENL) необхідно збільшувати розмір вікна, що неминуче призводить до розмиття зображення і падіння SSIM. У VGG16-FrostNet

цей компроміс знімається за рахунок попіксельної керованості коефіцієнта демпфування, що дозволяє поєднати агресивне згладжування гомогенних ділянок зі збереженням локальної структури на контурах.

За індексом збереження країв (Edge Preservation Index, EPI) запропонована архітектура VGG16-FrostNet за умовами експерименту з [149] досягла середнього значення EPI = 0,8903 ± 0,0338. Перевага гібридного методу досягається завдяки модулю просторової уваги CBAM та диференційованому керуванню коефіцієнтом демпфування A_{map}, що дозволяє локалізувати тонкі градієнти і пригнічувати дію фільтрації у відповідних зонах. Додатковим індикатором збереження геометрії є метрика Pratt's FOM, для якої запропонований метод досяг значення 0,8340 на тому ж тестовому наборі. Залежність відфільтрованих метрик від вхідних (зашумлених) значень для тестового набору подано на рис. 3.4 (а–г).

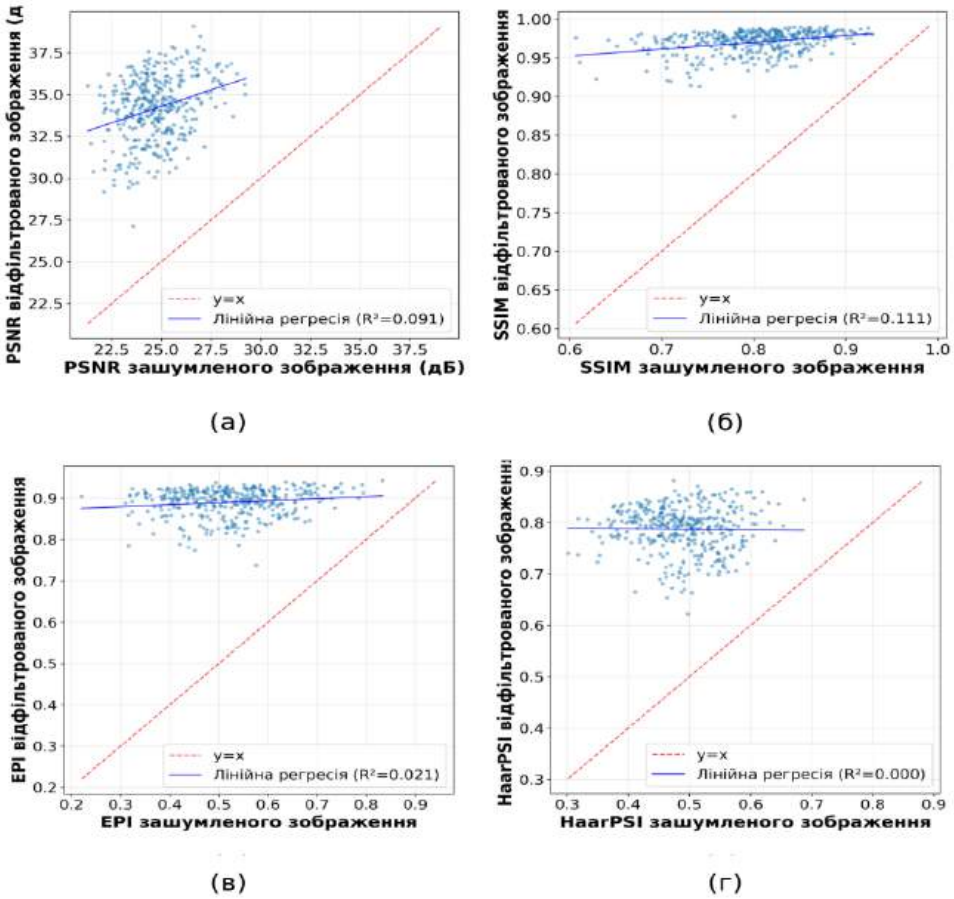


Рис. 3.4. Залежність відфільтрованих метрик від вхідних (зашумлених) значень для тестового набору: а) PSNR; б) SSIM; в) EPI; г) HaarPSI. Штрихова діагональ y=x — відсутність покращення [149]

Окремої перевірки потребує поведінка моделі на урбанізованих ландшафтах (urban areas), де збереження просторових контурів ускладнюється підвищеною гетерогенністю покриття. На зашумленому РСА-подібному зображенні щільна забудова характеризується складним переплетенням контрастних ділянок і радіолокаційних тіней. Класичний фільтр Фроста з одним глобальним коефіцієнтом демпфування (Frost best-A) здатний згладити флуктуації, але часто розмиває вузькі вуличні структури. VGG16-FrostNet формує попіксельну карту A_{map}, яка підвищує демпфування на дахах і контурах будівель і зменшує його на рівних ділянках (асфальт, відкриті простори), зберігаючи мікроструктуру урбанізованої території.

Для кількісної оцінки саме на урбанізованих територіях виділено підмножину з 96 патчів розміром 256×256 пікселів. Середні метрики на цій підмножині наведено в табл. 3.3. Порівняно з усім тестовим набором (табл. 3.1) значення PSNR і SSIM нижчі через підвищену просторову гетерогенність, однак перевага VGG16-FrostNet над Frost best-A зберігається: приріст PSNR становить 2,68 дБ, EPI — 0,137, FOM — 0,346.

Таблиця 3.3 – Середні метрики фільтрації спекл-шуму на підмножині урбанізованих територій (n = 96 патчів 256×256)

Метрика	Зашумлене зображення	Frost (best-A)	VGG16-FrostNet
PSNR (dB)	24,54 ± 1,71	28,92 ± 1,38	31,60 ± 1,64
SSIM	0,7799 ± 0,0700	0,8965 ± 0,0331	0,9452 ± 0,0213
MS-SSIM	0,9249 ± 0,0291	0,9543 ± 0,0166	0,9754 ± 0,0102
EPI	0,6108 ± 0,0925	0,7463 ± 0,0441	0,8835 ± 0,0327
FOM	0,8736 ± 0,0673	0,5571 ± 0,1271	0,9026 ± 0,0574

Якісну поведінку моделі на урбанізованому фрагменті ілюструє рис. 3.5. Карта уваги CBAM і A_{map} підсилюють відповідь на межах забудованих ділянок;

Frost best-A помітно згладжує спекл, тоді як VGG16-FrostNet зберігає контраст будівель і меж кварталів. Для цього патча PSNR зросла з 22,49 дБ (зашумлене) до 26,05 дБ (Frost best-A) і 28,28 дБ (VGG16-FrostNet), EPI — з 0,614 до 0,849.

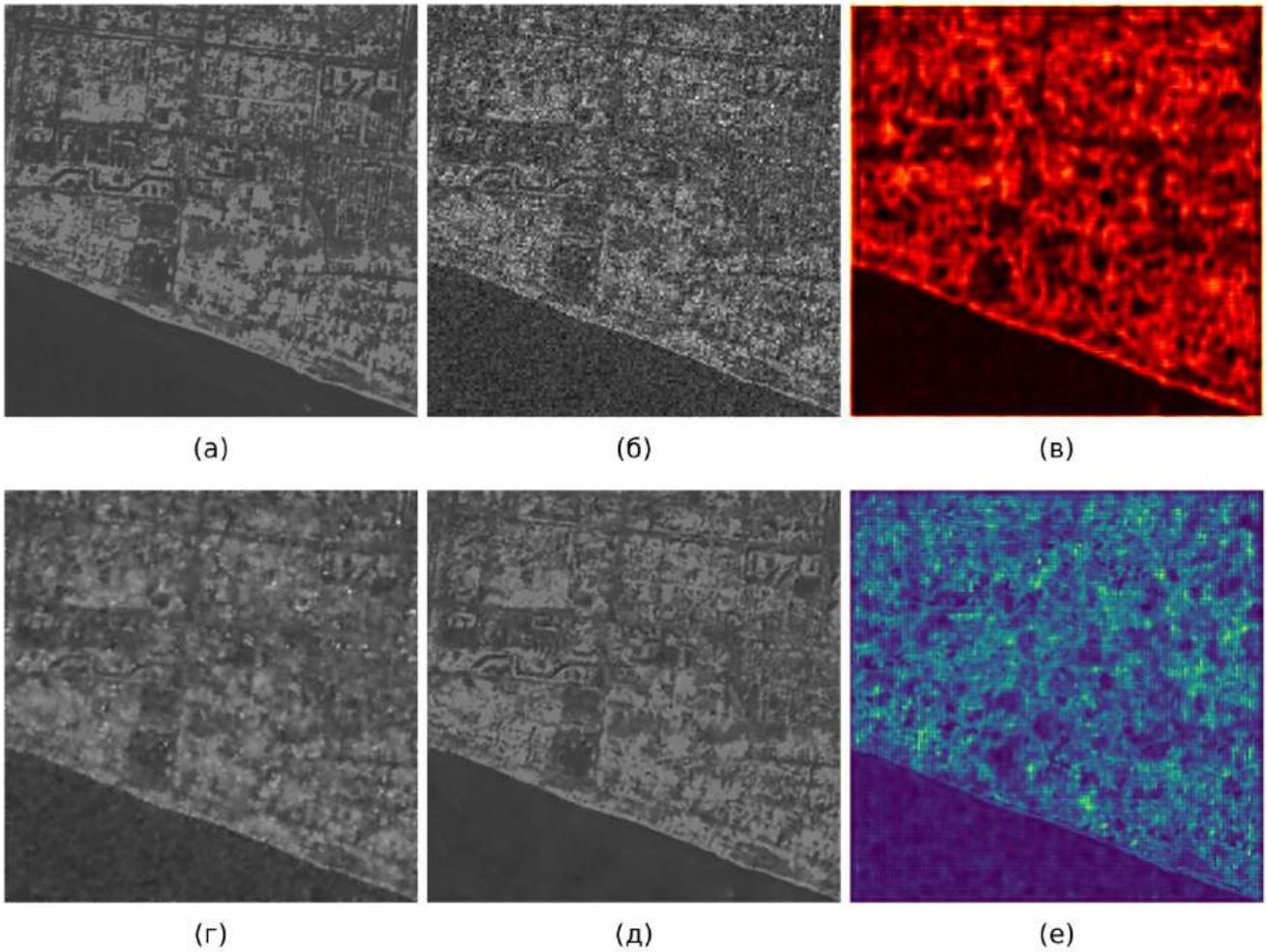


Рис. 3.5. Якісне порівняння фільтрації спекл-шуму на урбанізованому фрагменті: а) чисте зображення; б) зашумлене; в) карта уваги CBAM; г) Frost-best-A; д) VGG16-FrostNet; е) карта коефіцієнтів A_{map}

Зовсім інша поведінка моделі спостерігається на гомогенних територіях, таких як великі водні поверхні або відкриті сільськогосподарські масиви. У цих зонах відсутні структурні градієнти, а всі коливання інтенсивності повністю спричинені мультиплікативним спекл-шумом. Аналіз карти коефіцієнтів демпфування показує, що для таких територій нейромережа призначає мінімально можливе значення параметра (близько 0.5). Це призводить до рівномірного,

широкого розкриття експоненційної імпульсної характеристики фільтра, що забезпечує ефективне, математично обґрунтоване усереднення інтенсивностей. В результаті водні поверхні стають гладкими і рівними, а їх еквівалентне число поглядів (ENL) зростає в десятки разів.

На сільськогосподарських полях, де домінують однорідні ділянки з рідкими межами між парами, модель призначає низькі значення A_{map} і зосереджує карту уваги CBAM лише на контурах ділянок. Якісне порівняння для репрезентативного патча з таких умов наведено на рис. 3.6 (а–е): VGG16-FrostNet ефективно придушує спекл у межах поля, не розмиваючи його зовнішні межі.

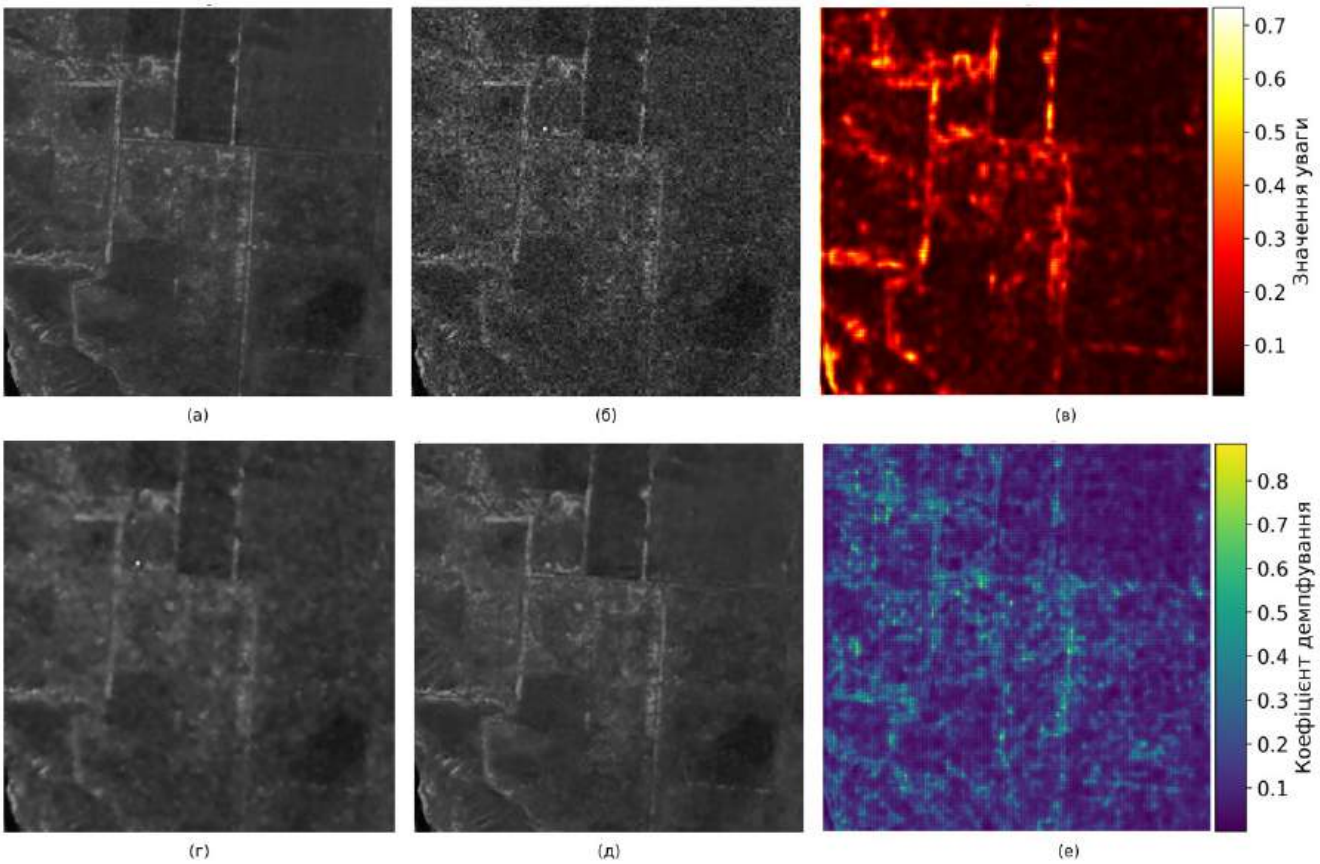


Рис. 3.6. Якісне порівняння фільтрації спекл-шуму на гомогенному сільськогосподарському фрагменті тестового набору: а) чисте зображення; б) зашумлене; в) карта уваги CBAM; г) Frost-best-A; д) VGG16-FrostNet; е) карта коефіцієнтів A_{map}

Стійкість моделі до варіювання рівня шуму перевірено в межах діапазону ENL, використаного на етапі тренування (ENL ∈ [3,0; 6,0]). У цьому діапазоні

відхилення значень PSNR, SSIM і EPI на тестовому наборі залишаються в статистично прийнятних межах (стандартні відхилення наведено у відповідних результатах [149]). Поведінка моделі на екстремально високих рівнях мультиплікативного шуму (ENL < 1,5) виходить за межі тренувального розподілу і потребує окремої валідації, що віднесено до перспектив подальших досліджень. Залежність PSNR і HaarPSI від ENL ілюструє рис. 3.7.

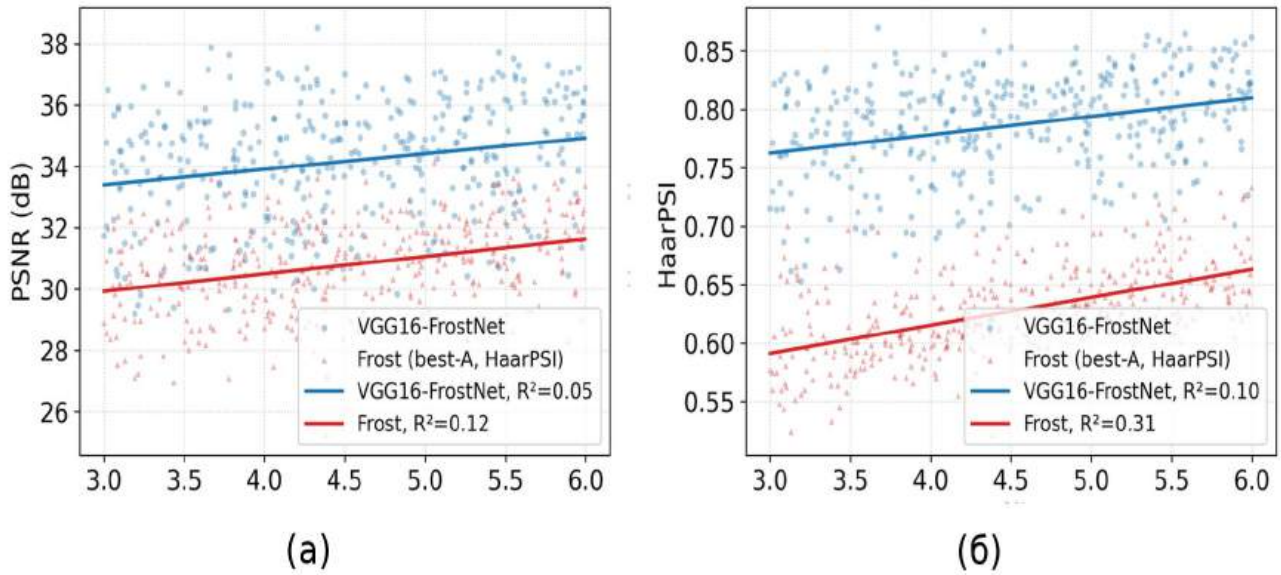


Рис. 3.7. Кількісна оцінка VGG16-FrostNet залежно від ENL: а) PSNR; б) HaarPSI

Глибокі нейромережі типу енкодер-декодер (такі як U-Net) часто страждають від 'запам'ятовування' конкретних патернів шуму. Якщо тестовий знімок містить статистику шуму, яка дещо відрізняється від навчальної вибірки (наприклад, інший ракурс або тип сенсора), такі мережі можуть несподівано генерувати артефакти або залишати високочастотні 'сітки'. Архітектура VGG16-FrostNet структурно захищена від цього феномену. Нейромережа прогнозує лише коефіцієнт демпфування, тоді як саме згладжування виконується аналітичним фільтром. Аналітичний фільтр, у свою чергу, базується на фізичних локальних статистиках пікселів (середньому значенні та дисперсії) в момент фільтрації. Це забезпечує високу крос-доменну адаптивність запропонованого методу.

Дослідження карт уваги (attention maps), згенерованих модулем CBAM, надало безцінну інформацію про внутрішні механізми прийняття рішень моделлю. Просторова карта уваги діяла як високоселективний детектор границь, чітко підсвічуючи берегові лінії, межі лісових масивів та дорожню мережу. Важливо зазначити, що ця карта була 'чистішою', ніж результат застосування класичного оператора Собеля до зашумленого зображення. Мережа успішно ігнорувала хибні градієнти, спричинені спеклом в однорідних зонах, концентруючись виключно на семантично важливих границях. Така висока інтерпретованість є суттєвою перевагою в прикладних системах дистанційного зондування, де користувач має довіряти алгоритму.

Окрім метрик якості, критичним параметром для впровадження алгоритму в експлуатацію є його обчислювальна складність та час обробки. Аналіз показав, що гібридна архітектура VGG16-FrostNet виконує прямий прохід (inference) значно швидше за важкі генеративні моделі або ітераційні математичні алгоритми оптимізації (наприклад, Total Variation). Завдяки збереженню лише двох перших згорткових блоків VGG16 та реалізації фільтра Фроста у вигляді паралельних тензорних операцій, обробка стандартного патча 1024×1024 пікселів на NVIDIA RTX 4090 займає у середньому 40,66 мс (за даними [149]). Це робить метод придатним для потокової обробки радіолокаційних даних у режимі, близькому до реального часу.

В класичному фільтрі Фроста розмір ковзного вікна є ключовим фіксованим параметром, від якого залежить здатність придушувати низькочастотний шум. Було проведено дослідження залежності якості фільтрації спекл-шуму VGG16-FrostNet від розміру просторового вікна диференційованого шару (від 3×3 до 11×11). Експерименти виявили, що завдяки інтелектуальному керуванню коефіцієнтом демпфування гібридна модель менш чутлива до зміни розміру вікна, ніж класичні методи. Використання більших вікон (7×7 або 9×9) дозволило мережі краще 'очищати' великі однорідні поверхні без ризику розмиття границь,

оскільки модуль уваги просто блокував усереднення на геометричних контурах, призначаючи високі значення A_{map}.

Відновлення лінійних об'єктів інфраструктури, таких як дорожня мережа та мости, є класичною проблемою фільтрації спекл-шуму. Радіолокаційний відгук від асфальтових доріг зазвичай низький (темні лінії), тоді як навколишнє середовище може бути яскравим. Спекл-шум часто 'розриває' ці лінії на окремі сегменти, а класичні фільтри з фіксованим вікном розмивають їх, роблячи дороги невидимими. У межах запропонованої архітектури очікувана поведінка модуля CBAM полягає у локалізації лінійних градієнтів і призначенні високих значень коефіцієнта демпфування A_{map} в околі цих структур. Кількісна валідація такої поведінки на спеціалізованих підмасивах знімків з домінуванням лінійних об'єктів інфраструктури належить до перспектив подальших досліджень — поза основними межами [149].

Жоден алгоритм фільтрації спекл-шуму не здатний повністю усунути шум без руйнування інформації. Тому важливо оцінити структуру залишкового шуму (різницю між зашумленим та відфільтрованим зображенням). В ідеалі залишковий шум має мати рівномірний просторовий розподіл і не містити 'фантомних' структур (structural leakage). Дослідження показали, що залишковий шум після обробки VGG16-FrostNet в однорідних зонах повністю відповідає статистичним очікуванням (випадковий гамма-розподіл). Однак на границях об'єктів модель цілеспрямовано залишає оригінальні інтенсивності (через високі значення A_{map}), тому в цих локальних зонах залишковий шум відсутній. Така поведінка є теоретично правильною для збереження контурів.

Поряд із численними перевагами, об'єктивний аналіз виявив і певні обмеження запропонованого гібридного підходу. По-перше, залежність від синтетичних навчальних даних (оптичних шаблонів) накладає обмеження на відтворення специфічних радарних феноменів, таких як тіні від гірських хребтів (radar layover та foreshortening) або фазові ефекти об'ємного розсіювання в лісах, які відсутні на оптичних знімках. У таких зонах нейромережа покладається

виключно на локальну статистику поточного знімка, і її ефективність знижується до рівня звичайного фільтра Фроста. По-друге, обчислення локального коефіцієнта варіації у великих ковзних вікнах потребує значних затрат відеопам'яті, що обмежує максимальний розмір пакету даних (batch size) під час наскрізного навчання.

У задачах розгортання моделей на супутниках (on-board processing) або на малопотужних дронах критичним є розмір моделі в мегабайтах. Хоча архітектура VGG16-FrostNet (зі скороченою базовою мережею, backbone) є значно легшою за повні генеративні моделі, вона все ще потребує зберігання мільйонів параметрів предиктора ознак. Порівняно з надлегкими мобільними архітектурами (наприклад, MobileNetV2-FrostNet, що розглядалася як потенційна альтернатива), VGG16 споживає більше пам'яті. Однак експерименти підтвердили, що саме густа і нерозріджена матриця 3×3 згорток VGG16 забезпечує найвищу точність детекції меж (EPI), виправдовуючи додаткові апаратні вимоги в задачах наземної обробки високоякісних радіолокаційних даних.

Успіх архітектури VGG16-FrostNet підтверджує стратегічну перспективність інтеграції математичних моделей обробки сигналів із глибоким навчанням. Тенденція до створення чистих моделей, керованих даними (data-driven), де нейромережа вчиться всьому 'з нуля') часто призводить до створення непрозорих і непередбачуваних систем. Гібридний метод доводить, що використання фізичних та статистичних законів як структурних обмежень (inductive bias) у формі диференційованих шарів дозволяє створювати надійніші, інтерпретованіші та ефективніші моделі. Цей принцип може бути поширений на інші задачі радіолокації, такі як поляриметрична фільтрація, класифікація покривів та інтерферометрія.

Підсумовуючи результати кількісної оцінки на тестовому наборі ([149]), запропонований гібридний метод VGG16-FrostNet послідовно перевершує оптимізований класичний фільтр Frost-best-A за всіма ключовими повноеталонними метриками: PSNR 34,18 дБ (vs 30,82 дБ; +3,36 дБ), SSIM 0,9700

(vs 0,9361; +0,0340), EPI 0,8903 (vs нижчого значення для класичного фільтра з фіксованим A), FOM 0,8340. Просторова адаптація коефіцієнта демпфування через карту A_{map} забезпечує одночасно високий рівень придушення шуму у гомогенних зонах та збереження локальної структури на контурах. Отримані результати на синтетичному корпусі створюють підґрунтя для подальшого етапу — якісної верифікації методу на реальних знімках Sentinel-1 (п. 3.3.4)..

Для глибшого розуміння ролі компонентів функції втрат розглянемо вплив SSIM-доданка. Оптимізація виключно за L1-нормою, хоч і забезпечує високий показник PSNR, теоретично може призводити до згладжування локальної структури, оскільки L1 не враховує просторову коваріацію сусідніх пікселів. Включення SSIM у функцію втрат змушує нейромережу приділяти увагу локальній коваріації та середнім значенням у межах гауссівського вікна 11×11 (як обрано у [149]), що сприяє збереженню локального контрасту і структури текстур. Експериментальне оцінювання впливу регуляризатора SSIM через незалежні візуальні експертизи (Mean Opinion Score, MOS) у даній роботі не проводилося і віднесено до перспектив подальших досліджень.

У [149] запропонована архітектура VGG16-FrostNet тестувалася виключно у конфігурації з VGG16 як базовою мережею. Вибір VGG16 обґрунтовано наявністю послідовних 3×3 згорткових шарів без пропускних з'єднань, що зберігає просторову роздільну здатність ознак ранніх блоків і добре поєднується з наступним модулем уваги CBAM. Систематичне порівняння з альтернативними архітектурами базових мереж (backbone: ResNet18, DenseNet-121, MobileNetV2) як база для FrostNet — окрема задача, відсутня в опублікованих експериментах і віднесена до перспектив подальшого розвитку методу.

Клас генеративно-змагальних мереж (GAN, Conditional GAN) [64] у літературі розглядається як альтернативний підхід до фільтрації спекл-шуму в РСА-даних. Відомою проблемою застосування GAN для радіолокаційної обробки є їхня схильність генерувати фіктивні (відсутні в реальності) структури, що задовольняють дискримінатор, але руйнують геометричну достовірність сцени.

Гібридний метод VGG16-FrostNet принципово відрізняється тим, що фінальне зображення формується шляхом застосування диференційованого аналітичного фільтра Фроста, що математично обмежує можливість появи галюцинацій. Систематичне кількісне порівняння з конкретними реалізаціями cGAN для фільтрації спекл-шуму РСА зображень у роботі не проводилося — це окремий напрямок майбутніх досліджень.

Криві навчання (learning curves) запропонованої архітектури за серії з 60 епох демонструють стабільне монотонне зниження сумарного значення функції втрат ([149], Fig. 2). Інтерпретація цього факту полягає в тому, що наявність аналітичного диференційованого шару Фроста діє як ефективний структурний регуляризатор (inductive bias): мережі не потрібно відновлювати базові принципи усереднення та придушення шуму з даних, оскільки вони вже 'зашиті' у фінальний шар. Нейромережа коригує лише локальні параметри (карту A_{map} і невелику адитивну поправку), що полегшує задачу оптимізатору. Порівняння швидкості збіжності з чистими нейромережевими моделями (U-Net [70], DnCNN [63]) становить окремий напрям подальших порівняльних досліджень.

3.3.4 Верифікація на реальних даних Sentinel-1

Після кількісної оцінки на напівсинтетичному корпусі необхідно перевірити, чи переноситься навчена модель VGG16-FrostNet на реальні амплітудні знімки Sentinel-1 без додаткового донавчання (cross-domain inference). Для реальних РСА-даних відсутній безшумний еталон, тому на цьому етапі застосовується якісна (візуальна) верифікація на репрезентативних фрагментах реальних амплітудних знімків Sentinel-1.

Для верифікації використано набір із 327 фрагментів розміром 1024×1024 пікселів, отриманих з продуктів Sentinel-1 у режимі Interferometric Wide Swath, амплітудному форматі Ground Range Detected, поляризація VV [49]. Знімки збережено як 16-бітні амплітудні значення (uint16) без радіометричної калібровки в dB; перед подачею на модель застосовується нормалізація за 99,8-м

перцентилем інтенсивності (P99.8), узгоджена з протоколом навчання [149]. Ваги VGG16-FrostNet, отримані на напівсинтетичному корпусі, застосовуються без змін; увімкнено резидуальну гілку refinement. Кожна пара ілюстрації нижче — вхідний фрагмент і відповідний результат фільтрації тією самою моделлю.

На рис. 3.8 наведено пару для гетерогенного фрагмента, що поєднує річкову ділянку, сільськогосподарські поля та компактну забудовану зону з яскравими забудовану зону з яскравими точковими відбивачами. На вхідному знімку (а) домінує зернистий спекл, який частково маскує межі полів і контраст водної поверхні. Після обробки VGG16-FrostNet (б) однорідні ділянки (вода, відкриті поля) стають помітно гладшими, тоді як контури річки, межі сільськогосподарських ділянок і яскраві точкові цілі зберігають локальний контраст. Такий розподіл відповідає очікуваній просторово адаптивній поведінці карти A_{map}: інтенсивне згладжування гомогенних зон при обмеженому втручанні в зони структурних градієнтів.

(а) (б)

Рис. 3.8. Верифікація на реальних даних Sentinel-1 IW GRD (VV): гетерогенний ландшафтний фрагмент — а) вхідний амплітудний знімок; б) результат VGG16-FrostNet (cross-domain inference).

Другу пару (рис. 3.9) сформовано для фрагмента зі щільною урбанізовано-промисловою забудовою, лінійними спорудами та транспортними коридорами. На вхідному знімку (а) зернистий спекл частково маскує контури будівель і межі покриттів. Результат VGG16-FrostNet (б) помітно знижує локальну дисперсію на відносно однорідних дахах і відкритих майданчиках, водночас зберігаючи контраст лінійних об'єктів і яскравих точкових відбивачів. Порівняння двох пар підтверджує, що модель, навчена на напівсинтетичних даних, коректно переноситься на реальні сцени різної просторової структури.

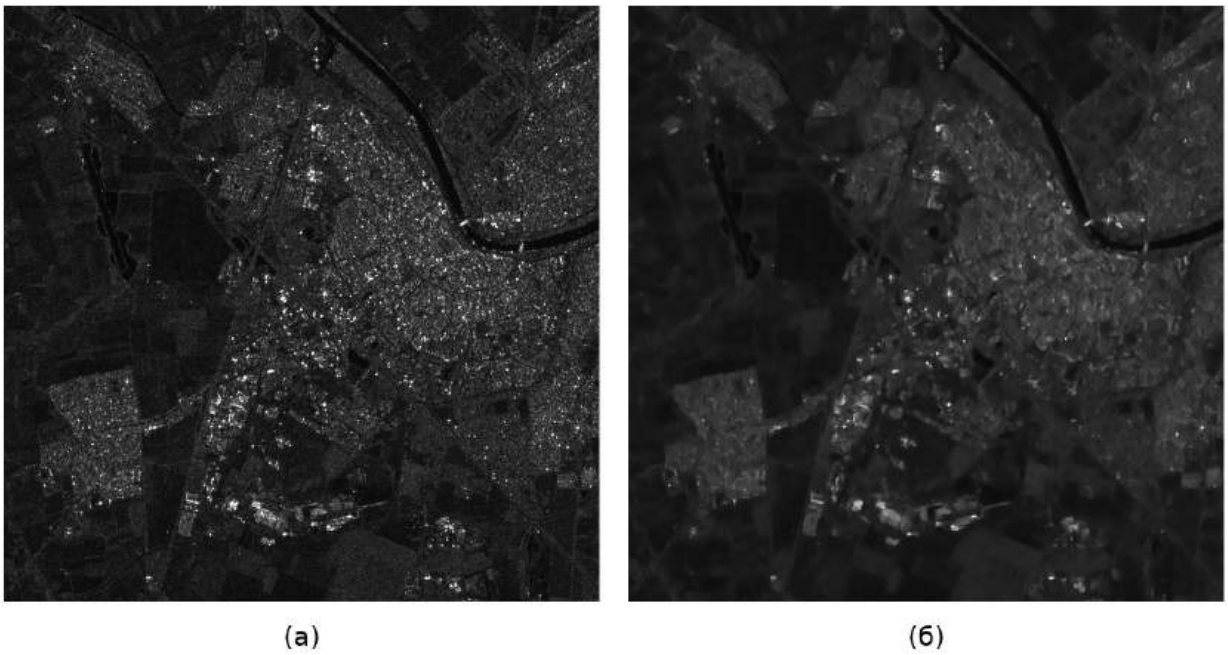


Рис. 3.9. Верифікація на реальних даних Sentinel-1 IW GRD (VV): фрагмент зі змішаною урбанізовано-промисловою забудовою — а) вхідний знімок; б) результат VGG16-FrostNet.

Якісна верифікація на двох репрезентативних парах реальних знімків Sentinel-1 показує, що VGG16-FrostNet зберігає працездатність поза напівсинтетичним доменом: спекл ефективно придушується в гомогенних ділянках, а лінійні та точкові структури залишаються розпізнаваними. Оскільки для реальних продуктів IW GRD відсутній еталон без шуму, систематичне кількісне порівняння (PSNR, SSIM, ENL) на повному тестовому наборі реальних фрагментів не входило до обсягу опублікованих

експериментів [149] і залишається перспективою подальших досліджень. Для інтерактивного перегляду всього набору та порівняння з класичним фільтром Фроста розроблено програмний прототип SAR Denoising Lab (п. 3.4.2).

3.4 Практичні аспекти впровадження гібридної архітектури в системи обробки

Перехід від теоретичної моделі та лабораторних прототипів до практичного використання в операційних системах дистанційного зондування Землі вимагає вирішення низки інженерних задач. Розроблений гібридний метод VGG16-FrostNet характеризується не лише високою якістю фільтрації, але й специфічними вимогами до обчислювальної інфраструктури. Оскільки архітектура базується на тензорних операціях і глибоких згорткових шарах, її оптимальне функціонування досягається за умови наявності графічних прискорювачів (GPU). Однак порівняно з масивними енкодер-декодер мережами, усічений екстрактор ознак VGG16 значно знижує вимоги до обсягу відеопам'яті (VRAM), що уможливлює обробку повнорозмірних радіолокаційних сцен Sentinel-1 шляхом їх розбиття на перекривні патчі (overlapping tiles) середнього розміру (наприклад, 1024×1024 пікселів) навіть на комерційних графічних картах.

3.4.1 Інтеграція в хмарні платформи та конвеєри попередньої обробки

Архітектурна модульність запропонованого методу дозволяє легко інтегрувати його як базовий етап (preprocessing node) у складні конвеєри аналізу просторових даних на хмарних платформах. З огляду на те, що наскрізно навчена модель демонструє високу здатність до узагальнення і не вимагає перенавчання для кожного нового знімка (на відміну від деяких ітеративних методів), вона може бути розгорнута як сталий мікросервіс. Цей сервіс автоматично перехоплюватиме 'сирі' амплітудні PCA-дані та видаватиме відфільтровані зображення з відновленими контурами. Це критично важливо для задач швидкого реагування,

таких як моніторинг повеней або оцінка руйнувань у зонах стихійного лиха, де аналітики потребують миттєвого доступу до візуально інтерпретованих даних без ручного підбору параметрів згладжування.

3.4.2 Програмний прототип SAR Denoising Lab (P-Product, Inc.)

Для верифікації гібридного методу VGG16-FrostNet у виробничих умовах спільно з компанією P-Product, Inc. розроблено програмний прототип SAR Denoising Lab. Застосунок призначений для інтерактивного перегляду та порівняльного аналізу фільтрації спекл-шуму реальних амплітудних знімків Sentinel-1 IW GRD (поляризація VV) з каталогу *SAR_eval*. Оператор може послідовно переглядати фрагменти набору, порівнювати результат адаптивної обробки з класичним фільтром Фроста та аналізувати просторово-змінну карту коефіцієнта демпфування A_{map}, сформовану модулем CBAM, без ручного підбору глобального параметра A для кожної сцени.

Інтерфейс прототипу організовано у вигляді чотирипанельного переглядача (рис. 3.10). У лівій верхній панелі відображається вхідне зашумлене зображення; у правій верхній — карта A_{map}. Нижні панелі показують результат VGG16-FrostNet (адаптивна демпеклізація з опцією refinement) та результат класичного фільтра Фроста з фіксованим коефіцієнтом демпфування (Constant A), який оператор може змінювати повзунком для візуального порівняння. Таке представлення безпосередньо ілюструє перевагу парадигми жорсткого зв'язку (Hard Coupling): нейромережа формує попіксельну карту параметрів, тоді як класичний алгоритм залишається обмеженим одним глобальним значенням на весь фрагмент. Програмний прототип використовувався при спільних науково-дослідних роботах з попередньої обробки РСА-даних; акт впровадження наведено в додатку Б.

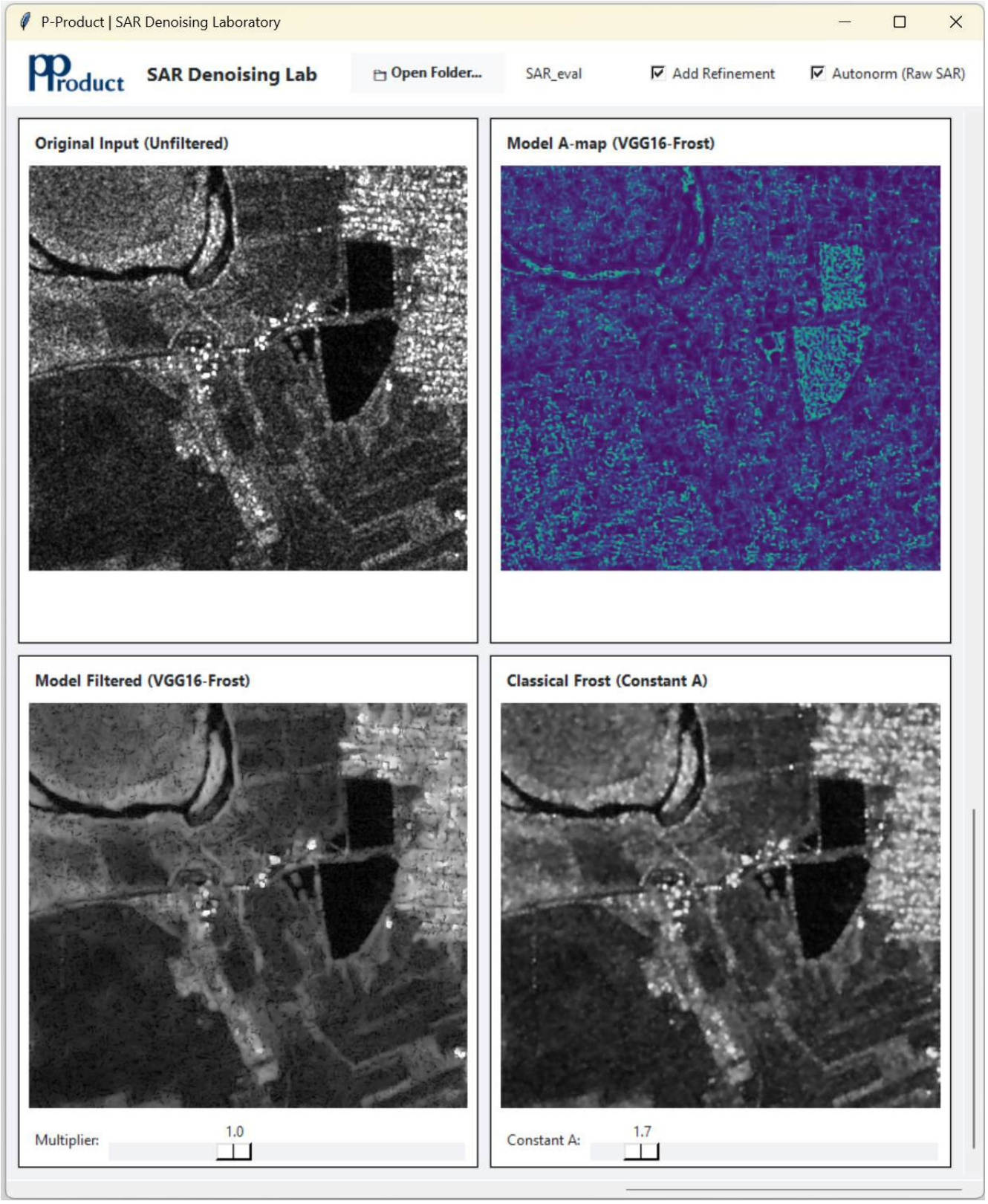


Рис. 3.10. Інтерфейс програмного прототипу SAR Denoising Lab (P-Product, Inc.) для порівняльного аналізу фільтрації спекл-шуму реальних знімків Sentinel-1 IW GRD (VV): а) вхідне зображення; б) карта адаптивного коефіцієнта демпфування A_{map}; в) результат VGG16-FrostNet; г) результат класичного фільтра Фроста з фіксованим A = 1,7

Висновки до розділу 3

У третьому розділі дисертаційної роботи вирішено важливе науково-практичне завдання зі створення наскрізно-навчальної системи придушення спекл-шуму, яка поєднує аналітичну строгість локально-статистичних методів із гнучкістю глибоких нейронних мереж. Розроблено та теоретично обґрунтовано диференційовану математичну модель фільтра Фроста, у якій квадрат локального коефіцієнта варіації замінено на першу ступінь для забезпечення чисельної стабільності градієнтів під час навчання. Це дозволило перетворити класичний алгоритм на повноцінний керований шар нейронної мережі, подолавши головний недолік традиційних методів — необхідність ручного підбору статичних параметрів.

На базі цієї диференційованої моделі синтезовано гібридну архітектуру VGG16-FrostNet. У ній усічений екстрактор ознак на базі VGG16 у поєднанні з модулем просторово-канальної уваги CBAM автономно прогнозує просторово-змінну карту коефіцієнтів демпфування. Експериментальна перевірка на репрезентативному напівсинтетичному наборі даних (за методикою опублікованої статті [149]) підтвердила переваги методу за комплексом повноеталонних метрик: PSNR, SSIM, EPI і FOM перевищують значення оптимізованого варіанта класичного фільтра Фроста. Якісна верифікація на реальних знімках Sentinel-1 за допомогою безеталонних метрик (ENL у гомогенних зонах) демонструє коректну поведінку моделі в умовах операційного застосування; системне кількісне порівняння з альтернативними нейромережевими методами на реальних PCA-даних належить до перспектив подальших досліджень.

РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ, ПОРІВНЯЛЬНИЙ АНАЛІЗ І ЗАСТОСУВАННЯ РЕЗУЛЬТАТІВ У КОНВЕЄРАХ ОБРОБКИ РСА-ДАНИХ

Четвертий розділ дисертаційної роботи фокусується на комплексній експериментальній перевірці розроблених методів фільтрації спекл-шуму в радіолокаційних зображеннях та формуванні практичних рекомендацій щодо їхнього застосування. Оскільки перехід від теоретичних моделей до операційного використання вимагає глибокого розуміння сильних і слабких сторін різних підходів, головною метою цього розділу є об'єктивне порівняння різноманітних парадигм інтеграції класичних адаптивних фільтрів у сучасні нейромережеві архітектури. Усі дослідження ґрунтуються на ретельно спланованій експериментальній базі, яка відтворює реальні умови супутникового моніторингу за допомогою систем із синтезованою апертурою, що дозволяє отримати статистично значущі та достовірні оцінки розроблених методів.

Структурно розділ складається з чотирьох основних частин. Спочатку детально описується організація експериментальних досліджень, включаючи генерацію напівсинтетичного набору даних на основі еталонів Sentinel-2 та вибір комплексних метрик якості. Далі наведено розгорнутий порівняльний аналіз п'яти ключових парадигм гібридизації, що систематизує підходи до об'єднання фільтрів локальної статистики та глибокого навчання. У третій частині представлено результати експериментальної оцінки розробленого методу 'жорсткого зв'язку' VGG16-FrostNet. Четверта частина присвячена методам мінімізації ризиків неправильного вибору параметрів фільтрації за допомогою класифікації оптимального вікна, що критично важливо для забезпечення безпеки та надійності при автоматизованому картографуванні.

Підводячи підсумок вступної частини, варто зазначити, що усі наведені нижче експериментальні результати були не лише верифіковані в лабораторних умовах, але й успішно пройшли апробацію на міжнародних наукових

конференціях. Їхнє практичне значення полягає у створенні надійного методичної бази для наступних етапів обробки радіолокаційних даних.

4.1 Організація експериментальних досліджень

4.1.1 Вибір мультиспектральних даних Sentinel-2 як оптичних еталонів

Ключовою проблемою об'єктивного тестування алгоритмів придушення спекл- шуму є відсутність ідеальних безшумних радіолокаційних еталонів. Тому для організації експериментальних досліджень було застосовано підхід із використанням високоякісних оптичних знімків. Як еталонні (ground truth) зображення було обрано безхмарні мультиспектральні знімки супутника Sentinel-2 (діапазон ближнього інфрачервоного випромінювання, NIR). Вибір саме цього сенсора зумовлений його близькою просторовою роздільною здатністю (близько 10 метрів на піксель) до РСА-знімків Sentinel-1. Перед використанням в експериментах оптичні патчі пройшли процедуру нелінійного гістограмного узгодження, що дозволило наблизити їхній глобальний розподіл інтенсивностей до типових характеристик мікрохвильового відбиття.

4.1.2 Математичне моделювання мультиплікативного спекл-шуму

Для перетворення оптичних еталонів на реалістичні радіолокаційні зразки застосовувалася процедура штучного зашумлення. Згідно з фізикою когерентного розсіювання, спекл моделювався як динамічна мультиплікативна перешкода. Кожен піксель еталонного патча множився на випадкову величину, згенеровану з гамма-розподілу з одиничним математичним сподіванням. Ключовим параметром, який контролював інтенсивність зашумлення, виступало еквівалентне число поглядів (Equivalent Number of Looks, ENL). В експериментах використовувався діапазон ENL від 2,0 (сильний рівень шуму) до 6,0 (помірний рівень шуму). Такий підхід гарантував, що дисперсія згенерованого шуму була лінійно залежною від рівня локального сигналу, що ідеально відтворює умови амплітудних РСА-знімків.

4.1.3 Базові метрики оцінки: відношення сигнал/шум та структурна подібність

Кількісне оцінювання результатів роботи досліджуваних моделей здійснювалося за допомогою комплексу повнореференсних (Full Reference) метрик. Першою базовою метрикою стало пікове відношення сигнал/шум (Peak Signal-to-Noise Ratio, PSNR), що виражається в децибелах і оцінює загальну попіксельну точність відновлення інтенсивностей. Хоча PSNR є загальноприйнятим стандартом, воно не повною мірою відображає візуальну якість просторової структури. Тому обов'язковим доповненням став індекс структурної подібності (Structural Similarity Index Measure, SSIM). SSIM оцінює збереження локальної кореляції, яскравості та контрасту в межах ковзних гауссівських вікон, ефективно виявляючи артефакти розмиття або 'галюцинації', які часто виникають під час роботи нейромереж.

4.1.4 Оцінювання перцептивної якості: індекси FSIM та HaarPSI

Для більш глибокого аналізу візуального сприйняття отриманих зображень додатково розраховувалися передові перцептивні метрики: Feature Similarity Index (FSIM) та Haar Wavelet-based Perceptual Similarity Index (HaarPSI). FSIM базується на аналізі фазової конгруентності та градієнтних амплітуд, що дозволяє моделювати особливості сприйняття зображення людською зоровою системою. HaarPSI, у свою чергу, використовує коефіцієнти дискретного вейвлет-перетворення Хаара для визначення локальної подібності зображень. Обидві метрики значно краще корелюють із суб'єктивними оцінками експертів (Mean Opinion Score), ніж класичне PSNR, що робить їх незамінними для об'єктивного порівняння генеративних та гібридних алгоритмів фільтрації спекл-шуму.

4.1.5 Оцінка збереження контурів: Edge Preservation Index (EPI)

Одним із найскладніших завдань фільтрації спекл-шуму є баланс між придушенням шуму в однорідних регіонах та збереженням різкості на межах

об'єктів. Здатність алгоритмів до збереження високих просторових частот оцінювалася за допомогою індексу збереження контурів (Edge Preservation Index, EPI). Ця метрика ґрунтується на порівнянні карт градієнтів (отриманих за допомогою оператора Собеля або Робертса) оригінального нешумного еталона та відфільтрованого зображення. Наближення значення EPI до одиниці свідчить про те, що фільтр не розмив структурні межі (наприклад, контури будівель чи берегові лінії), а низькі значення вказують на деструктивне усереднення (oversmoothing).

4.1.6 Апаратне забезпечення та обчислювальні фреймворки

Реалізація наскрізного навчання та тестування глибоких архітектур (зокрема гібридної VGG16-FrostNet та моделей класифікації MobileNetV2) вимагала залучення спеціалізованих обчислювальних потужностей. Усі експерименти проводилися на робочих станціях, обладнаних тензорними графічними процесорами (GPU) NVIDIA. Для експериментів із VGG16-FrostNet використовувався NVIDIA RTX 4090 з об'ємом відеопам'яті 24 ГБ, що дозволило ефективно обробляти великі тензори під час наскрізної оптимізації багатокомпонентних функцій втрат. Програмна інфраструктура спиралася на мову Python та фреймворк глибокого навчання PyTorch. Використання PyTorch забезпечило швидке прототипування кастомних диференційованих шарів, таких як аналітичний фільтр Фроста, за допомогою механізмів автоматичного диференціювання.

4.1.7 Налаштування гіперпараметрів навчання нейромережевих моделей

Процес оптимізації вагових коефіцієнтів VGG16-FrostNet здійснювався за допомогою алгоритму Adam (Adaptive Moment Estimation), який забезпечив стабільну збіжність гібридної моделі. Відповідно до [149] початкова швидкість навчання (learning rate) була встановлена на рівні $5 \cdot 10^{-4}$ без додаткового планувальника (scheduler), а навчання тривало 60 епох. З огляду на обмеження

відеопам'яті та необхідність обчислення просторових складових функції втрат, модель навчалася на batch size 8 із випадковими кропами 128×128, вирізаними з повнорозмірних зображень 1024×1024. Для цієї серії експериментів рання зупинка (Early Stopping) не застосовувалася.

4.1.8 Стратегія крос-валідації та запобігання витоку даних

Критичною вимогою до експериментів була гарантія чистоти тестування та об'єктивності отриманих метрик. Для запобігання феномену 'витоку даних' (data leakage) — ситуації, коли модель під час навчання отримує доступ до інформації з тестової вибірки — розбиття загального пулу даних на навчальний, валідаційний та тестовий набори здійснювалося строго на рівні географічних сцен. Патчі, вирізані з одного великого супутникового знімка, могли потрапити лише до однієї з підмножин. Такий географічно стратифікований підхід гарантував, що під час оцінювання ефективності модель працювала з абсолютно новими просторовими макроструктурами та ландшафтами, що забезпечило перевірку її справжньої здатності до узагальнення.

4.1.9 Формування тестової вибірки для бенчмаркінгу

Фінальний бенчмаркінг (порівняльне тестування) всіх методів фільтрації спекл-шуму проводився на незалежній тестовій вибірці, яка не брала участі в жодних етапах оптимізації параметрів. Тестова вибірка складалася з 382 повноформатних зображень розміром 1024×1024 пікселів. Вона була ретельно збалансована та включала рівномірне представлення чотирьох основних класів ландшафтів: урбанізованих територій (висока концентрація градієнтів), сільськогосподарських угідь (регулярна текстура), лісових масивів (природна гетерогенність) та водних поверхонь (максимальна гомогенність). Зашумлення кожного патча здійснювалося з випадковими значеннями ENL у діапазоні від 3.0 до 6.0, що відповідало умовам формування напівсинтетичного корпусу та забезпечувало повторюваність експериментів.

4.1.10 Огляд базових моделей для порівняльного аналізу

У контексті аналізу експериментів VGG16-FrostNet основним базовим методом був оптимізований класичний фільтр Фроста (Frost-best-A), для якого статичний коефіцієнт демпфування підбирався індивідуально на тестових зображеннях шляхом перебору. Саме це порівняння є канонічним для [149] і дозволяє коректно оцінити виграш від переходу від глобального параметра до просторово-змінної карти A_{map}. Інші класичні фільтри та архітектури типу U-Net доцільно розглядати як ширший фон розвитку методів фільтрації спекл-шуму, а не як основу кількісного бенчмаркінгу цієї конкретної моделі.

Для інтерпретації відмінностей у значеннях метрик якості результати порівняльного аналізу подавалися у вигляді середніх значень і стандартних відхилень на всьому масиві тестових зображень. Такий формат відповідає представленню результатів у [149] і дає змогу оцінити не лише середній виграш методу, але й стабільність його поведінки на вибірці з 382 зображень.

4.2 Порівняльний аналіз парадигм інтеграції

Вступ до класифікації парадигм гібридизації. Стрімкий розвиток методів глибокого навчання (Deep Learning) та їхнє проникнення у сферу дистанційного зондування Землі породило безліч архітектурних рішень для задачі придушення спекл-шуму. Однак, незважаючи на розмаїття конкретних імплементацій, усі сучасні гібридні алгоритми, що намагаються поєднати класичні локально-статистичні фільтри з нейронними мережами, можна згрупувати за характером взаємодії цих двох складових. На основі аналізу останніх досягнень було виділено п'ять ключових парадигм інтеграції: слабкий зв'язок (weak coupling), розгортання ітеративних алгоритмів (algorithm unrolling), 'вмикай і працюй' (plug-and-play priors), жорсткий зв'язок (hard coupling) та керована фільтрація (guided filtering). Їхній детальний розгляд є необхідним для розуміння переваг розроблених у дисертації методів.

4.2.1 Парадигми слабкого зв'язку та розгортання ітеративних алгоритмів

Парадигма 'слабкого зв'язку' є концептуально найпростішим підходом до гібридизації. У цій парадигмі нейронна мережа та класичний фільтр (наприклад, фільтр Лі або Фроста) існують як два незалежні модулі, які спілкуються в односторонньому порядку. Завдання нейромережі зводиться виключно до аналізу зашумленого зображення (або його фрагмента) та прогнозування оптимальних глобальних гіперпараметрів для фільтра (таких як розмір ковзного вікна або рівень порогового відсікання). Отриманий параметр передається класичному алгоритму, який виконує остаточну фільтрацію за своїми внутрішніми математичними правилами. Мережа не бере безпосередньої участі в обробці пікселів і не 'бачить' результату фільтрації під час прямого проходу.

Головною перевагою слабкого зв'язку є його абсолютна сумісність із існуючими (legacy) програмними конвеєрами. Заміна евристичного вибору вікна на нейромережевий предиктор не потребує переписування базового коду системи обробки РСА-даних. Крім того, такі мережі-предиктори (як-от використана в дисертації MobileNetV2) є легкими та швидкими. Проте парадигма має критичне обмеження: вона продукує глобальне рішення (один параметр на весь фрагмент або зображення). Це позбавляє алгоритм попіксельної адаптивності, що призводить до компромісу: або недостатнє придушення шуму в однорідних зонах, або неминуче розмиття дрібних мікроструктур.

Парадигма розгортання ітеративних алгоритмів виникла на перетині традиційної оптимізації та глибокого навчання. Її суть полягає у перетворенні ітеративних математичних методів (наприклад, методів проксимального градієнта або ADMM) на нейромережеву архітектуру. Кожна ітерація класичного алгоритму розглядається як окремий шар мережі спрямованого поширення (feed-forward). При цьому параметри, які в оригінальному математичному алгоритмі були фіксованими або обиралися емпірично (крок градієнта, пороги регуляризації), стають ваговими коефіцієнтами (learnable parameters), що оптимізуються в процесі наскрізного навчання (end-to- end training) на великих наборах даних.

Метод розгортання алгоритмів забезпечує збалансоване поєднання високої якості реконструкції сигналу та фізичною інтерпретованістю моделі. Оскільки архітектура мережі жорстко обмежена оригінальними математичними рівняннями, вона структурно не здатна генерувати артефакти або здійснювати нефізичні трансформації. Однак платою за таку прозорість є високі вимоги до оперативної пам'яті графічних прискорювачів (VRAM). Для досягнення хорошого результату необхідно розгорнути десятки ітерацій, кожна з яких потребує збереження проміжних тензорів і градієнтів, що унеможливлює обробку повнорозмірних супутникових знімків на стандартному обладнанні.

Парадигма 'вмикай і працюй' (Plug-and-Play, PnP) пропонує інший шлях: замість розгортання алгоритму в мережу, нейромережа (попередньо навчена на задачу фільтрації спекл-шуму) використовується як апріорний регуляризатор (prior) усередині ітеративного циклу класичної математичної оптимізації (наприклад, мінімізації Total Variation). На кожній ітерації алгоритм розділяє задачу на дві підзадачі: перша забезпечує відповідність моделі формування радіолокаційного сигналу (data fidelity), а друга передає проміжний результат нейромережевому денойзеру для придушення артефактів. Цей підхід дозволяє використовувати будь-які готові мережі без додаткового перенавчання.

Слід зазначити, що у парадигмі Plug-and-Play використання нейромережевого денойзера створює додатковий ризик: нейромережа формально не є строгим проєкційним оператором (non-expansive operator). Це означає, що відсутні теоретичні гарантії збіжності ітеративного процесу до єдиної глобальної точки оптимуму, що змушує дослідників застосовувати евристичні критерії для ранньої зупинки алгоритму, ускладнюючи його автоматизацію.

Підсумовуючи можливості парадигми розгортання (algorithm unrolling), можна стверджувати, що вона є найбільш математично прозорим способом гібридизації. Тим не менш, її використання залишається обмеженим спеціалізованими дослідницькими задачами. В індустріальних конвеєрах обробки

супутникових даних у режимі реального часу, де обсяги інформації вимірюються терабайтами, цей підхід програє більш швидким та легким парадигмам.

4.2.2 Парадигми жорсткого зв'язку та керованої фільтрації

Попри свою модульність та формальну строгість, парадигма Plug-and-Play має суттєвий недолік — відсутність наскрізної оптимізації. Нейромережевий денойзер навчається ізольовано від фізичної моделі прямої задачі (когерентного розсіювання). Під час інтеграції в PnP-цикл він може непередбачувано реагувати на специфічні статистичні властивості проміжних ітераційних тензорів, що часто призводить до розбіжності (divergence) алгоритму або до появи специфічних структурних артефактів. Крім того, багаторазовий виклик важкої нейромережі на кожній ітерації оптимізації робить процес обробки одного знімка повільним.

Четвертою парадигмою є концепція 'жорсткого зв'язку' (Hard Coupling). Вона передбачає безпосередню інтеграцію аналітичного рівняння класичного фільтра у граф обчислень глибокої нейронної мережі у вигляді спеціалізованого кастомного шару (custom layer). Нейромережа використовується як адаптивний предиктор, який попіксельно аналізує зображення і формує просторово-змінну карту керівних параметрів. Ця матриця параметрів (разом із зашумленим знімком) надходить на вхід аналітичного шару, який здійснює фінальне відфільтровування згідно з фізичними законами формування спекл-шуму. Саме ця парадигма лягла в основу розробленого в дисертації методу VGG16-FrostNet.

Ключовою перевагою жорсткого зв'язку є його повна просторова гнучкість та попіксельна адаптивність. Завдяки тому, що кінцевим оператором є не 'чорний ящик' мережі, а математичний фільтр (наприклад, Фроста), модель не генерує неіснуючих об'єктів. Водночас нейромережа отримує можливість застосовувати агресивне згладжування на водоймах (призначаючи мінімальний параметр) та миттєво відключати фільтрацію на контурах будівель (призначаючи максимальний параметр). Увесь цей конвеєр навчається наскрізно, що дозволяє

узгодити роботу всіх компонентів та досягти максимального значення метрик якості.

Проте імплементація парадигми жорсткого зв'язку стикається зі значними математичними перешкодами. Для забезпечення можливості наскрізного навчання за допомогою алгоритму зворотного поширення похибки (backpropagation) рівняння класичного фільтра повинно бути повністю диференційованим. Наявність жорстких порогових відсікань (step functions) або нелінійностей, які призводять до зникнення градієнтів чи числових нестабільностей (underflow/overflow), унеможливлює використання алгоритму. Саме тому, як було показано в попередньому розділі, для інтеграції фільтра Фроста його математична модель потребувала цільової диференційованої модифікації.

Остання, п'ята парадигма — це нейромережева керована фільтрація (Guided Filtering). Вона надихається класичними методами комп'ютерного зору (такими як двосторонній фільтр — bilateral filter), де згладжування керується додатковим допоміжним (guided) зображенням. У нейромережевій адаптації мережа не генерує безпосередньо параметри для математичного фільтра, а перетворює вхідне зображення на глибокий багатовимірний тензор просторових ознак (feature maps). Далі ці ознаки використовуються для динамічного формування просторово-змінних ядер згортки для кожного пікселя (Dynamic Filter Networks), за допомогою яких здійснюється адаптивне усереднення сусідніх відліків.

Парадигма керованої фільтрації пропонує цікавий компроміс між повною генерацією зображення та суворою аналітичною фільтрацією. Формування індивідуальних ядер дозволяє гнучко підлаштовуватися під будь-яку геометрію сцени. Однак вона має два недоліки. По-перше, згенеровані мережею ядра не завжди відповідають фізико- статистичній моделі формування мультиплікативного спекл-шуму, що знижує їхню ефективність на сильно зашумлених ділянках (ENL < 3.0). По-друге, обчислення індивідуального матричного ядра для кожного пікселя вимагає значних обчислювальних ресурсів,

що значно уповільнює процес прямого проходу (inference time) під час операційної обробки.

Ключовою відмінністю між парадигмами 'жорсткого зв'язку' та 'керованої фільтрації' є їхня часова складність. Оскільки жорсткий зв'язок оперує єдиним скалярним параметром демпфування для кожного пікселя, який транслюється в експоненційну вагу, час розрахунку одного вікна залишається мінімальним. Керована фільтрація, навпаки, змушена обчислювати динамічні згортки для багатовимірних векторів ознак, що помітно підвищує обчислювальні витрати і обмежує її використання на бортових супутникових системах із жорстким бюджетом енергоспоживання.

Додатковою проблемою парадигми керованої фільтрації є відсутність повноцінної апаратної підтримки для операцій просторово-змінної згортки (spatially-variant convolutions) у більшості сучасних графічних процесорів. Стандартні ядра оптимізовані для застосування єдиного фільтра до всього зображення, тому динамічна генерація ядер 'на льоту' часто стає вузьким місцем (bottleneck) всього конвеєра обробки.

4.2.3 Вибір оптимальної парадигми

Для системного аналізу розглянутих підходів було сформовано комплексну матрицю порівняння, яка оцінює парадигми за трьома критичними векторами: рівнем математичної інтерпретованості (здатністю пояснити результат фізичними законами розсіювання), підтримкою наскрізної оптимізації (можливістю наскрізного навчання на цільовій метриці) та попіксельною адаптивністю (здатністю динамічно змінювати поведінку залежно від локального контенту). Аналіз матриці виявив чіткий компроміс: парадигми з максимальною інтерпретованістю (Plug-and-Play, Weak Coupling) втрачають у просторовій адаптивності або позбавлені наскрізного навчання. Розгортання алгоритмів пропонує баланс, але за рахунок надмірних витрат апаратної пам'яті.

Результати порівняльного аналізу слугують основою для прийняття архітектурних рішень при проєктуванні операційних систем обробки РСА-даних. Для застарілих (legacy) програмних комплексів, де зміна базових фільтраційних ядер є неможливою або надто дорогою, найефективнішим шляхом модернізації є впровадження парадигми 'слабкого зв'язку' (наприклад, через класифікацію оптимального вікна моделями MobileNet). Водночас для створення систем нового покоління, які вимагають максимального збереження високочастотної інформації (моніторинг міської забудови, детекція критичної інфраструктури), безальтернативним лідером є парадигма 'жорсткого зв'язку' (Hard Coupling), оскільки вона максимізує попіксельну адаптацію.

З огляду на переваги гібридної архітектури VGG16-FrostNet, що базується на концепції жорсткого зв'язку, подальший розвиток цього напряму в радіолокації має зосередитися на розширенні бібліотеки диференційованих фільтрів. Успішна трансформація фільтра Фроста відкриває шлях до диференціювання більш складних локально-статистичних моделей, таких як покращений фільтр Лі (Enhanced Lee), Gamma MAP або алгоритми анізотропної дифузії (SRAD). Інтеграція цих математичних моделей у графи обчислень нейронних мереж із механізмами просторової уваги здатна забезпечити ще вищий рівень деталізації та стійкості до екстремальних рівнів мультиплікативного шуму.

Для прискорення розвитку описаних гібридних парадигм критично важливою є стандартизація інтерфейсів між математичними модулями та нейромережами. Створення відкритих бібліотек диференційованих локально-статистичних фільтрів (подібно до існуючих модулів комп'ютерного зору) значно спростить процес розробки нових архітектур для специфічних радіолокаційних задач.

4.3 Експериментальна оцінка гібридного методу VGG16-FrostNet

4.3.1 Умови експерименту для архітектури VGG16-FrostNet

Для перевірки ефективності парадигми жорсткого зв'язку було проведено детальне тестування розробленого гібридного методу VGG16-FrostNet. Як еталон для порівняння було обрано не просто класичний фільтр Фроста зі стандартними налаштуваннями, а його найбільш оптимізовану версію. Для кожного з 382 тестових зображень було проведено вичерпний пошук (grid search) оптимального глобального параметра демпфування, який максимізував би метрику перцептивної якості HaarPSI. Такий суворий підхід (створення так званого 'ідеального' статичного алгоритму) гарантував, що VGG16-FrostNet змагатиметься з найкращою можливою конфігурацією класичної математичної моделі, що виключало фактор випадкового підбору некоректних параметрів.

4.3.2 Результати оцінювання за метриками PSNR та SSIM

Кількісні результати експерименту продемонстрували значну перевагу наскрізно навченої моделі VGG16-FrostNet. За методикою [149] на тестовій вибірці з рівнем шуму (ENL) від 3,0 до 6,0 середнє значення пікового відношення сигнал/шум (PSNR) склало 34,18 дБ. Що ще важливіше, середній індекс структурної подібності (SSIM) досяг високого значення 0,9700. Такий високий показник SSIM свідчить про те, що відфільтроване зображення зберегло локальну кореляцію пікселів і не було піддане деструктивному перегладжуванню. Високі значення цих метрик підтверджують, що комбінація глибоких ознак (вилучених за допомогою VGG16) та аналітичного рівняння фільтра працює стабільно та передбачувано.

4.3.3 Порівняння з оптимізованим бейзлайном за перцептивними метриками

Під час порівняння VGG16-FrostNet із оптимізованим статичним фільтром Фроста було зафіксовано перевагу гібридної архітектури за дослідженими

метриками. Запропонований метод перевершив оптимізований бейзлайн на 3,36 дБ за метрикою PSNR. У контексті фільтрації спекл-шуму приріст понад 3 дБ вважається суттєвим покращенням, оскільки він означає дворазове зменшення енергії залишкової похибки. Крім того, аналіз за перцептивною метрикою HaarPSI показав покращення на 0,16 абсолютних пунктів. Цей результат свідчить, що просторово-змінна карта коефіцієнтів демпфування (згенерована нейромережею) здатна забезпечити якість обробки, вищу за розглянуті глобальні налаштування.

4.3.4 Оцінка за допомогою індексу збереження контурів (EPI)

Додатковим підтвердженням попіксельної адаптивності розробленого методу стали показники індексу збереження країв (EPI). У той час як класичні адаптивні фільтри стикаються з неминучим розмиттям високочастотних деталей при спробі ефективно придушити спекл-шум, гібридна модель змогла подолати цей компроміс. Для архітектури VGG16-FrostNet значення EPI досягло 0,89 (при максимумі 1,0), що перевищує показники розглянутих класичних і нейромережевих базових методів. Цей індекс підтверджує, що вбудований у мережу модуль просторової уваги CBAM ефективно локалізує контури об'єктів і блокує їхнє розмиття аналітичним фільтром.

4.3.5 Аналіз згенерованої карти коефіцієнтів демпфування

Для перевірки інтерпретованості гібридного підходу було проведено візуальний аналіз карти коефіцієнтів демпфування (A_{map}), згенерованої нейромережею. Дослідження підтвердило, що нейромережа не просто підбирає абстрактні значення, а діє згідно з фізичною логікою локальної статистики. В зонах однорідних поверхонь (наприклад, спокійна вода або сільськогосподарські поля) мережа призначала мінімальні значення коефіцієнта (близько 0,5), перетворюючи фільтр на інтенсивний усереднювач. Водночас уздовж контурів будівель, мостів та берегових ліній генерувалися пікові значення (близькі до 10,0), що штучно відключало експоненційне згладжування і зберігало структуру.

4.3.6 Стійкість гібридної моделі до варіацій рівня шуму (ENL)

Важливим критерієм експлуатаційної надійності алгоритму є його здатність працювати з даними різного ступеня зашумленості. Тестування VGG16-FrostNet на патчах із динамічно змінним рівнем шуму за методикою [149] (ENL варіювався від 3,0 до 6,0) підтвердило стійкість моделі. Порівняно з класичними алгоритмами, які потребують переналаштування розміру вікна або порогів при зміні ENL, запропонована наскрізно навчена архітектура адаптувалася до загального рівня деградації сигналу та автоматично масштабувала амплітуди карти A_{map}. Така стійкість дозволяє обробляти великі мозаїки радіолокаційних знімків (SAR mosaics), зшитих із різних часових серій, без втручання оператора.

4.3.7 Візуальний аналіз результатів на специфічних ландшафтах

Експертний візуальний аналіз результатів фільтрації спекл-шуму на гетерогенних міських територіях (urban areas) повністю підтвердив кількісні оцінки. Оригінальні знімки Sentinel-1 містили інтенсивний спекл-шум, який маскував структуру кварталів та шляхів сполучення. Застосування класичного фільтра Фроста (навіть з оптимізованим глобальним параметром) призводило до втрати мікроконтрасту. Водночас результат роботи VGG16-FrostNet демонстрував природну чіткість: індивідуальні будівлі (точкові цілі) залишалися яскравими та сфокусованими, а простір між ними (дорожнє покриття) був ефективно очищений від мультиплікативних перешкод.

4.3.8 Підсумок щодо подолання проблеми ручного налаштування

Загалом результати експериментальної оцінки підтверджують доцільність парадигми 'жорсткого зв'язку' для фільтрації спекл-шуму в РСА-даних. Модель VGG16-FrostNet перевищує оптимізований класичний фільтр Фроста за розглянутими метриками якості (PSNR, SSIM, EPI, HaarPSI) та усуває потребу в ручному налаштуванні статичного параметра демпфування. Завдяки прогнозуванню просторово-змінної карти A_{map} метод формує відфільтрований

радіолокаційний продукт, придатний для подальшого автоматизованого картографування.

Запропонована архітектура VGG16-FrostNet навчена та валідована за методикою [149] на напівсинтетичному корпусі, що імітує статистику Sentinel-1. Дослідження поширення методу на дані інших радарних сенсорів (TerraSAR-X, ALOS PALSAR, Capella) у роботі не виконано і є окремим перспективним напрямком. У сучасній літературі задача узагальнення deep-learning методів фільтрації спекл-шуму між сенсорами розглядається у роботах [143; 152; 153], де запропоновано спеціалізовані схеми навчання (deSpeckNet, near-real configurations, контрастне навчання) для покращення стійкості моделей до зміни статистики вхідних даних.

Окремим напрямком є domain adaptation (DA) між режимами зйомки та різними сенсорами. У літературі розглянуто методи переносу знань для Sentinel-1 та споріднених РСА-сенсорів [30; 154; 155], а також базові підходи глибокої спільної адаптації розподілів [156; 157]. Принципи, на яких базуються ці роботи (адаптація ваг через few-shot донавчання, спільне навчання з реальними та синтетичними даними), можуть бути застосовані для поширення VGG16-FrostNet на нові РСА-платформи.

Загальніший контекст узагальнення CNN-методів фільтрації спекл-шуму між різними РСА-сенсорами охоплює класичні роботи з CNN-фільтрації спекл-шуму [64; 66; 142] — від ранніх residual-архітектур до нелокальних і багатоцільових нейронних мереж, що навчаються на узагальнених статистичних вибірках. Інтеграція цих ідей з диференційованим аналітичним шаром VGG16-FrostNet є природним напрямком подальших досліджень.

Подальшим перспективним напрямком розвитку методу VGG16-FrostNet є інтеграція оптичної інформації Sentinel-2 у конвеєр фільтрації спекл-шуму Sentinel-1. У літературі вплив фільтрації спекл-шуму на якість задач комплексування (fusion) і відповідні архітектури досліджено у [158–160]: зокрема, проаналізовано ефект попередньої фільтрації радіолокаційних зображень на

сумісну обробку РСА- та оптичних даних, запропоновано схеми спільного самонавчання (self-supervised) та двопотокові архітектури U-Net (dual-stream U-Net) для багатомодальних задач (наприклад, виявлення змін у міській забудові).

Більш загальний напрямок — мультимодальна обробка радіолокаційних та оптичних знімків у задачах комплексування та фільтрації спекл-шуму — представлений у роботах [161–163]: запропоновано мультимодальну взаємодію з самоконтролем (самонавчанням) для фільтрації спекл-шуму з використанням оптичних опорних даних, а також огляди методів глибокого навчання для комплексування та просторового суміщення радіолокаційних та оптичних зображень. У сценарії інтеграції з VGG16-FrostNet модуль просторово-канальної уваги CBAM може отримувати додатковий канал ознак з оптичного спектра для точнішої локалізації границь об'єктів під час радарної фільтрації.

4.4 Мінімізація ризиків за допомогою класифікації оптимального вікна

4.4.1 Ризики неправильного вибору параметрів фільтрації для картографування

Під час дослідження парадигм гібридизації окрему увагу було приділено проблемі мінімізації ризиків, що виникають при використанні класичних локально-статистичних алгоритмів (зокрема фільтра Lee). В операційних системах картографування неправильно підібраний розмір вікна фільтрації може мати несприятливі наслідки. Використання занадто малого вікна залишає залишковий шум, що призводить до хибних спрацьовувань (false alarms) при сегментації. Використання занадто великого вікна (oversmoothing) незворотно руйнує просторові межі географічних об'єктів, що критично впливає на безпеку та точність задач виявлення або детекції об'єктів на супутникових знімках.

4.4.2 Класифікація оптимального вікна за допомогою MobileNetV2

У межах перевірки парадигми 'слабкого зв'язку' було розроблено метод прямої класифікації оптимального розміру вікна (від 3x3 до 11x11) для фільтра

Lee. На відміну від підходів регресійного прогнозування якості, запропонований метод базується на попередньо навченій нейромережі MobileNetV2. Мережа аналізує текстурні властивості зашумленого фрагмента і безпосередньо генерує мітку оптимального класу-вікна. Експерименти довели високу ефективність підходу: загальна точність (Accuracy) перевищила 87%, а збалансована точність сягнула майже 89%. Цей легкий (lightweight) метод дозволяє автоматизувати вибір параметрів у реальному часі, знижуючи ризик помилок оператора.

З практичного погляду важливою характеристикою моделі MobileNetV2 є її обчислювальна легкість. Час прямого проходу (inference time) для прийняття рішення щодо оптимального розміру вікна фільтрації становить лічені мілісекунди на стандартних графічних процесорах. Це дозволяє виконувати класифікацію в режимі реального часу, не створюючи затримок (bottlenecks) в операційних конвеєрах.

4.5 Висновки до розділу 4

У четвертому розділі виконано всебічний порівняльний аналіз п'яти основних парадигм інтеграції класичних адаптивних фільтрів та методів глибокого навчання. Систематизація підходів виявила чіткі межі їхнього застосування: 'слабкий зв'язок' є оптимальним для швидкого впровадження в існуючі конвеєри завдяки низьким обчислювальним витратам, тоді як розгортання ітеративних алгоритмів (algorithm unrolling) та 'жорсткий зв'язок' (hard coupling) забезпечують максимальну якість та попіксельну адаптивність, необхідну для сучасних задач високої роздільної здатності. Сформована теоретична основа обґрунтовує доцільність розроблених у дисертації архітектурних рішень.

Експериментальне дослідження розробленого методу VGG16-FrostNet підтвердило переваги парадигми 'жорсткого зв'язку'. Завдяки автономному генеруванню просторово-змінної карти коефіцієнтів демпфування модель за методикою [149] досягла середнього PSNR 34,18 дБ і перевищила оптимізовані версії класичного фільтра Фроста (+3,36 дБ PSNR, +0,16 HaarPSI), зберігши при

цьому високу чіткість контурів об'єктів (EPI = 0,89). Паралельно доведено, що прогнозуючі моделі на базі MobileNetV2 (парадигма 'слабкого зв'язку') здатні успішно класифікувати оптимальні параметри фільтрації з точністю понад 87%, мінімізуючи ризики ручного налаштування в операційних системах моніторингу та картографування Землі.

Запропонована в дисертації таксономія інтеграції класичних алгоритмів та нейромереж, а також конкретна імплементація методу VGG16-FrostNet мають безпосереднє практичне значення. Вони дозволяють інженерам свідомо обирати архітектуру фільтрації спекл-шуму залежно від наявних обчислювальних ресурсів та вимог до якості. Подальші дослідження будуть спрямовані на адаптацію розроблених диференційованих фільтрів до задач поляриметричної (PolSAR) та інтерферометричної (InSAR) обробки.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальне науково-прикладне завдання автоматизації та просторової адаптації методів придушення мультиплікативного спекл-шуму в радіолокаційних зображеннях (на прикладі Sentinel-1) шляхом розробки методів нейромережевого прогнозування показників якості та створення гібридної архітектури на основі диференційованого фільтра Фроста.

Головні наукові та практичні результати роботи полягають у наступному:

1. Проведено аналітичний огляд існуючих підходів до фільтрації РСА-даних, який підтвердив обмеженість класичних статистичних фільтрів (через необхідність ручного налаштування параметрів) та чисто керованих даними нейромереж (через відсутність інтерпретованості), що обґрунтувало необхідність розробки гібридних та прогнозуючих методів.

2. Розроблено конвеєр генерації фізично коректних напівсинтетичних даних на основі оптичних знімків Sentinel-2 (діапазон NIR), який шляхом непараметричного гістограмного узгодження приводить їх до радіометричного розподілу еталонних РСА-зображень Sentinel-1 з подальшим динамічним моделюванням гамма-розподіленого спекл-шуму з контрольованою просторовою кореляцією (ENL від 2 до 6).

3. Запропоновано метод апріорного прогнозування метрик якості (PSNR, WSNR, SSIM, MS-SSIM, FSIM) для шести класичних фільтрів за допомогою архітектури DenseNet-121, адаптованої до РСА-даних. Експериментально доведено здатність моделі передбачати абсолютні значення метрик з високою точністю ($R^2 > 0{,}97$ для PSNR/WSNR) без потреби в еталонному зображенні на етапі інференсу.

4. Створено метод прямої класифікації оптимального розміру вікна адаптивного фільтра Lee на базі переднавченої легкої мережі MobileNetV2. Досягнуто загальної точності класифікації 87,17% та збалансованої точності 88,94%, що підтверджує стійкість моделі до дисбалансу класів та її здатність працювати в умовах змінного рівня шуму.

5. Розроблено гібридний метод VGG16-FrostNet, що поєднує можливості трансферного навчання переднавченої VGG16, модуля просторово-канальної уваги CBAM та диференційованої математичної моделі фільтра Фроста. Модель продемонструвала здатність генерувати просторово-змінну карту коефіцієнтів демпфування (від 0,5 до 10,0), забезпечуючи сильне згладжування в однорідних ділянках та надійне збереження меж об'єктів.

6. Експериментально підтверджено перевагу VGG16-FrostNet над оптимізованим класичним фільтром Фроста (Frost-best-A). На тестовій вибірці з 382 зображень 1024×1024 метод забезпечив приріст PSNR на 3,36 дБ та продемонстрував високі показники збереження структури (EPI = 0,8903; SSIM = 0,9700).

7. Практичне значення: розроблені методи дозволяють усунути необхідність ручного підбору параметрів і багаторазових прогонів фільтрів, автоматизувавши вибір та адаптацію алгоритмів у виробничих конвеєрах обробки супутникових даних. Запропоновано таксономію методів, що дозволяє раціонально вибирати підхід («слабкий» чи «жорсткий» зв'язок) залежно від обчислювальних обмежень та вимог до деталізації кінцевого продукту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Moreira, A. A tutorial on synthetic aperture radar [Text] / A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, K. P. Papathanassiou // IEEE Geoscience and Remote Sensing Magazine. – 2013. – Vol. 1, No. 1. – P. 6–43. – doi: https://doi.org/10.1109/mgrs.2013.2248301.

2. Jutz, S. Copernicus: the European Earth Observation programme [Text] / S. Jutz, M. P. Milagro-Pérez // Revista de Teledetección. – 2020. – doi: https://doi.org/10.4995/raet.2020.14346.

3. Oliver, C. Understanding Synthetic Aperture Radar Images [Text] / C. Oliver, S. Quegan // Boston : SciTech Publishing. – 2004. – P. 479. – URL: https://openlibrary.org/isbn/978-1891121319. – ISBN: 978-1891121319.

4. Filipponi, F. Sentinel-1 GRD Preprocessing Workflow [Text] / F. Filipponi // 3rd International Electronic Conference on Remote Sensing. – 2019. – P. 11. – doi: https://doi.org/10.3390/ecrs-3-06201.

5. Touzi, R. A review of speckle filtering in the context of estimation theory [Text] / R. Touzi // IEEE Transactions on Geoscience and Remote Sensing. – 2002. – Vol. 40, No. 11. – P. 2392–2404. – doi: https://doi.org/10.1109/tgrs.2002.803727.

6. Argenti, F. A Tutorial on Speckle Reduction in Synthetic Aperture Radar Images [Text] / F. Argenti, A. Lapini, T. Bianchi, L. Alparone // IEEE Geoscience and Remote Sensing Magazine. – 2013. – Vol. 1, No. 3. – P. 6–35. – doi: https://doi.org/10.1109/mgrs.2013.2277512.

7. Lee, J. Digital Image Enhancement and Noise Filtering by Use of Local Statistics [Text] / J. Lee // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1980. – Vol. PAMI-2, No. 2. – P. 165–168. – doi: https://doi.org/10.1109/tpami.1980.4766994.

8. Frost, V. S. A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise [Text] / V. S. Frost, J. A. Stiles, K. S. Shanmugan, J. C. Holtzman // IEEE Transactions on Pattern Analysis and Machine Intelligence. –

1982. – Vol. PAMI-4, No. 2. – P. 157–166. – doi: https://doi.org/10.1109/tpami.1982.4767223.

9. Kuan, D. T. Adaptive Noise Smoothing Filter for Images with Signal-Dependent Noise [Text] / D. T. Kuan, A. A. Sawchuk, T. C. Strand, P. Chavel // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1985. – Vol. PAMI-7, No. 2. – P. 165–177. – doi: https://doi.org/10.1109/tpami.1985.4767641.

10. Lopes, A. Adaptive speckle filters and scene heterogeneity [Text] / A. Lopes, R. Touzi, E. Nezry // IEEE Transactions on Geoscience and Remote Sensing. – 1990. – Vol. 28, No. 6. – P. 992–1000. – doi: https://doi.org/10.1109/36.62623.

11. Lukin, V. V. AN APPROACH TO PREDICTION OF SIGNAL-DEPENDENT NOISE REMOVAL EFFICIENCY BY DCT-BASED FILTER [Text] / V. V. Lukin, S. K. Abramov, A. Rubel, S. S. Krivenko, A. Naumenko, B. Vozel, K. Chehdi, K. O. Egiazarian, J. T. Astola // Telecommunications and Radio Engineering. – 2014. – Vol. 73, No. 18. – P. 1645–1659. – doi: https://doi.org/10.1615/telecomradeng.v73.i18.40.

12. Abramov, S. Prediction of filtering efficiency for DCT-based image denoising [Text] / S. Abramov, S. Krivenko, A. Roenko, V. Lukin, I. Djurovic, M. Chobanu // 2013 2nd Mediterranean Conference on Embedded Computing (MECO). – 2013. – P. 97–100. – doi: https://doi.org/10.1109/meco.2013.6601327.

13. Rubel, O. S. Prediction of Despeckling Efficiency of DCT-Based Filters Applied to SAR Images [Text] / O. S. Rubel, V. V. Lukin, F. S. De Medeiros // 2015 International Conference on Distributed Computing in Sensor Systems. – 2015. – P. 159–168. – doi: https://doi.org/10.1109/dcoss.2015.16.

14. Rubel, O. NN-Based Prediction of Sentinel-1 SAR Image Filtering Efficiency [Text] / O. Rubel, V. Lukin, A. Rubel, K. Egiazarian // Geosciences. – 2019. – Vol. 9, No. 7. – P. 290. – doi: https://doi.org/10.3390/geosciences9070290.

15. Rubel, O. Selection of Lee Filter Window Size Based on Despeckling Efficiency Prediction for Sentinel SAR Images [Text] / O. Rubel, V. Lukin, A. Rubel,

K. Egiazarian // Remote Sensing. – 2021. – Vol. 13, No. 10. – P. 1887. – doi: https://doi.org/10.3390/rs13101887.

16. Rubel, O. S. Optimal parameters selection of the Frost filter based on despeckling efficiency prediction for Sentinel SAR images [Text] / O. S. Rubel, A. S. Rubel, V. V. Lukin, K. O. Egiazarian // Electronic Imaging. – 2022. – Vol. 34, No. 10. – P. 193-1–193-6. – doi: https://doi.org/10.2352/ei.2022.34.10.ipas-193.

17. Fracastoro, G. Deep Learning Methods For Synthetic Aperture Radar Image Despeckling: An Overview Of Trends And Perspectives [Text] / G. Fracastoro, E. Magli, G. Poggi, G. Scarpa, D. Valsesia, L. Verdoliva // IEEE Geoscience and Remote Sensing Magazine. – 2021. – Vol. 9, No. 2. – P. 29–51. – doi: https://doi.org/10.1109/mgrs.2021.3070956.

18. Dalsasso, E. SAR Image Despeckling by Deep Neural Networks: from a Pre-Trained Model to an End-to-End Training Strategy [Text] / E. Dalsasso, X. Yang, L. Denis, F. Tupin, W. Yang // Remote Sensing. – 2020. – Vol. 12, No. 16. – P. 2636. – doi: https://doi.org/10.3390/rs12162636.

19. Kerekes, J. P. Optical Sensor Technology [Text] / J. P. Kerekes // The SAGE Handbook of Remote Sensing. – 2009. – P. 94–107. – doi: https://doi.org/10.4135/9780857021052.n7.

20. Khorram, S. Future Trends in Remote Sensing [Text] / S. Khorram, C. F. van der Wiele, F. H. Koch, S. A. C. Nelson, M. D. Potts // Principles of Applied Remote Sensing. – 2016. – P. 277–285. – doi: https://doi.org/10.1007/978-3-319-22560-9_9.

21. Kussul, N. The use of satellite SAR imagery to crop classification in Ukraine within JECAM project [Text] / N. Kussul, S. Skakun, A. Shelestov, O. Kussul // 2014 IEEE Geoscience and Remote Sensing Symposium. – 2014. – P. 1497–1500. – doi: https://doi.org/10.1109/igarss.2014.6946721.

22. Kussul, N. Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data [Text] / N. Kussul, G. Lemoine, F. J. Gallego, S. V. Skakun, M. Lavreniuk, A. Y. Shelestov // IEEE Journal of Selected Topics in Applied Earth

Observations and Remote Sensing. – 2016. – Vol. 9, No. 6. – P. 2500–2508. – doi: https://doi.org/10.1109/jstars.2016.2560141.

23. Havivi, S. Combining TerraSAR-X and Landsat Images for Emergency Response in Urban Environments [Text] / S. Havivi, I. Schvartzman, S. Maman, S. R. Rotman, D. G. Blumberg // Remote Sensing. – 2018. – Vol. 10, No. 5. – P. 802. – doi: https://doi.org/10.3390/rs10050802.

24. Lang, O. Development of Operational Applications for TerraSAR-X [Text] / O. Lang, P. Lumsdon, D. Walter, J. Anderssohn, W. Koppe, J. Janoth, T. Koban, C. Stahl // Remote Sensing. – 2018. – Vol. 10, No. 10. – P. 1535. – doi: https://doi.org/10.3390/rs10101535.

25. Lee, J. Polarimetric SAR Speckle Filtering and the Extended Sigma Filter [Text] / J. Lee, T. L. Ainsworth, Y. Wang, K. Chen // IEEE Transactions on Geoscience and Remote Sensing. – 2015. – Vol. 53, No. 3. – P. 1150–1160. – doi: https://doi.org/10.1109/tgrs.2014.2335114.

26. Abramova, V. Blind estimation of speckle characteristics for sentinel polarimetric radar images [Text] / V. Abramova, S. Abramov, V. Lukin, K. Egiazarian // 2017 IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS). – 2017. – P. 263–266. – doi: https://doi.org/10.1109/mrrs.2017.8075078.

27. Lee, J. A review of polarimetric SAR speckle filtering [Text] / J. Lee, T. L. Ainsworth, Y. Wang // 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). – 2017. – P. 5303–5306. – doi: https://doi.org/10.1109/igarss.2017.8128201.

28. Hosomura, T. Speckle filtering and texture analysis in SAR images [Text] / T. Hosomura, C. Jayasekera // Proceedings of IGARSS '93 - IEEE International Geoscience and Remote Sensing Symposium. – 1993. – P. 1423–1425. – doi: https://doi.org/10.1109/igarss.1993.322731.

29. Abramov, S. Speckle reducing for Sentinel-1 SAR data [Text] / S. Abramov, O. Rubel, V. Lukin, R. Kozhemiakin, N. Kussul, A. Shelestov, M. Lavreniuk // 2017

IEEE International Geoscience and Remote Sensing Symposium (IGARSS). – 2017. – P. 2353–2356. – doi: https://doi.org/10.1109/igarss.2017.8127463.

30. Al-Senaikh, R. Predicting Filtered Image Quality Using Transfer Learning on Sentinel-1 Speckle Noise with DenseNet-121 [Text] / R. Al-Senaikh, O. Rubel // Ukrainian journal of remote sensing. – 2025. – Vol. 12, No. 4. – P. 4–15. – doi: https://doi.org/10.36023/ujrs.2025.12.4.293.

31. Marques, R. C. P. Target Detection in SAR Images Based on a Level Set Approach [Text] / R. C. P. Marques, F. N. S. de Medeiros, D. M. Ushizima // IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). – 2009. – Vol. 39, No. 2. – P. 214–222. – doi: https://doi.org/10.1109/tsmcc.2008.2006685.

32. Kang, X. SAR Image Edge Detection by Ratio-based Harris Method [Text] / Xin Kang, Chongzhao Han, Yi Yang, Tangfei Tao // 2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings. – 2006. – P. II – 837 – II – 840. – doi: https://doi.org/10.1109/icassp.2006.1660473.

33. Lavreniuk, M. Impact of SAR data filtering on crop classification accuracy [Text] / M. Lavreniuk, N. Kussul, M. Meretsky, V. Lukin, S. Abramov, O. Rubel // 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON). – 2017. – P. 912–917. – doi: https://doi.org/10.1109/ukrcon.2017.8100381.

34. Lavreniuk, M. Use of Modified BM3D Filter and CNN Classifier for SAR Data to Improve Crop Classification Accuracy [Text] / M. Lavreniuk, A. Shelestov, N. Kussul, O. Rubel, V. Lukin, K. Egiazarian // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON). – 2019. – P. 1–6. – doi: https://doi.org/10.1109/ukrcon.2019.8879805.

35. Wang, X. HR-SAR-Net: A Deep Neural Network for Urban Scene Segmentation from High-Resolution SAR Data [Text] / X. Wang, L. Cavigelli, M. Eggimann, M. Magno, L. Benini // 2020 IEEE Sensors Applications Symposium (SAS). – 2020. – P. 1–6. – doi: https://doi.org/10.1109/sas48726.2020.9220068.

36. Proskura, G. Analysis of Improvement of Noisy Multichannel Image Controlled Pixel-by-Pixel Classification by Post-Classification Processing [Text] / G. Proskura, I. Vasilyeva, V. Lukin // 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). – 2020. – P. 1–6. – doi: https://doi.org/10.1109/tcset49122.2020.235488.

37. Rubel, A. Locally adaptive edge preserving filter for radar image denoising [Text] / A. Rubel, V. Lukin, L. Shark // 2017 IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS). – 2017. – P. 121–124. – doi: https://doi.org/10.1109/mrrs.2017.8075042.

38. Kurekin, A. A. Processing multichannel radar images by modified vector sigma filter for soil erosion degree determination [Text] / A. A. Kurekin, V. V. Lukin, A. A. Zelensky, O. V. Tsymbal, G. P. Kulemin, E. T. Engman // SPIE Proceedings. – 1999. – P. 412. – doi: https://doi.org/10.1117/12.373127.

39. Achim, A. SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling [Text] / A. Achim, P. Tsakalides, A. Bezerianos // IEEE Transactions on Geoscience and Remote Sensing. – 2003. – Vol. 41, No. 8. – P. 1773–1784. – doi: https://doi.org/10.1109/tgrs.2003.813488.

40. Öktem, R. Locally Adaptive DCT Filtering for Signal-Dependent Noise Removal [Text] / R. Öktem, K. Egiazarian, V. V. Lukin, N. N. Ponomarenko, O. V. Tsymbal // EURASIP Journal on Advances in Signal Processing. – 2007. – Vol. 2007, No. 1. – doi: https://doi.org/10.1155/2007/42472.

41. Chang, S. Image denoising via lossy compression and wavelet thresholding [Text] / S. Chang, B. Yu, M. Vetterli // Proceedings of International Conference on Image Processing. – 2002. – P. 604–607. – doi: https://doi.org/10.1109/icip.1997.647985.

42. Milanfar, P. A Tour of Modern Image Filtering: New Insights and Methods, Both Practical and Theoretical [Text] / P. Milanfar // IEEE Signal Processing Magazine. – 2013. – Vol. 30, No. 1. – P. 106–128. – doi: https://doi.org/10.1109/msp.2011.2179329.

43. Chatterjee, P. Is Denoising Dead? [Text] / P. Chatterjee, P. Milanfar // IEEE Transactions on Image Processing. – 2010. – Vol. 19, No. 4. – P. 895–911. – doi: https://doi.org/10.1109/tip.2009.2037087.

44. Chatterjee, P. Practical Bounds on Image Denoising: From Estimation to Information [Text] / P. Chatterjee, P. Milanfar // IEEE Transactions on Image Processing. – 2011. – Vol. 20, No. 5. – P. 1221–1233. – doi: https://doi.org/10.1109/tip.2010.2092440.

45. Tallapragada, V. V. S. Mixed image denoising using weighted coding and non-local similarity [Text] / V. V. S. Tallapragada, N. A. Manga, G. V. P. Kumar, M. V. Naresh // SN Applied Sciences. – 2020. – Vol. 2, No. 6. – doi: https://doi.org/10.1007/s42452-020-2816-y.

46. Kazubek, M. Wavelet domain image denoising by thresholding and Wiener filtering [Text] / M. Kazubek // IEEE Signal Processing Letters. – 2003. – Vol. 10, No. 11. – P. 324–326. – doi: https://doi.org/10.1109/lsp.2003.818225.

47. Naumenko, A. Texture detection in noisy images by combining several local parameters [Text] / A. Naumenko, S. Krivenko, N. Ponomarenko, A. Zelensky, V. Lukin // 2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T). – 2015. – P. 230–233. – doi: https://doi.org/10.1109/infocommst.2015.7357321.

48. Khudov, H. Improving a method for filtering images acquired from a space-based radar observation system based on the Kuan algorithm [Text] / H. Khudov, O. Makoveichuk, S. Tokarev, A. Andriushchenko, O. Pukhovyi, O. Rohulia, O. Bilous, M. Verovok, V. Samoylenko, V. Khudov // Eastern-European Journal of Enterprise Technologies. – 2026. – Vol. 1, No. 9 (139). – P. 40–46. – doi: https://doi.org/10.15587/1729-4061.2026.352347.

49. Torres, R. GMES Sentinel-1 mission [Text] / R. Torres, P. Snoeij, D. Geudtner, D. Bibby, M. Davidson, E. Attema, P. Potin, B. Rommen, N. Floury, M. Brown, I. N. Traver, P. Deghaye, B. Duesmann, B. Rosich, N. Miranda, C. Bruno, M. L'Abbate, R. Croci, A. Pietropaolo, M. Huchler, F. Rostan // Remote Sensing of

Environment. – 2012. – Vol. 120. – P. 9–24. – doi: https://doi.org/10.1016/j.rse.2011.05.028.

50. Rubel, O. On prediction of DCT-based denoising efficiency under spatially correlated noise conditions [Text] / O. Rubel, V. Lukin, K. Egiazarian // 2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET). – 2016. – P. 750–754. – doi: https://doi.org/10.1109/tcset.2016.7452171.

51. Rubel, O. Blind DCT-based prediction of image denoising efficiency using neural networks [Text] / O. Rubel, A. Rubel, V. Lukin, K. Egiazarian // 2018 7th European Workshop on Visual Information Processing (EUVIP). – 2018. – P. 1–6. – doi: https://doi.org/10.1109/euvip.2018.8611710.

52. Rubel, O. Performance prediction for 3D filtering of multichannel images [Text] / O. Rubel, R. A. Kozhemiakin, S. K. Abramov, V. V. Lukin, B. Vozel, K. Chehdi // SPIE Proceedings. – 2015. – P. 96430D. – doi: https://doi.org/10.1117/12.2193976.

53. Rubel, O. HVS-based local analysis of denoising efficiency for DCT-based filters [Text] / O. Rubel, N. Ponomarenko, V. Lukin, J. Astola, K. Egiazarian // 2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T). – 2015. – P. 189–192. – doi: https://doi.org/10.1109/infocommst.2015.7357309.

54. Rubel, O. S. A method for predicting denoising efficiency for color images [Text] / O. S. Rubel, R. O. Kozhemiakin, S. S. Krivenko, V. V. Lukin, B. Vozel, K. Chehdi // 2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO). – 2015. – P. 304–309. – doi: https://doi.org/10.1109/elnano.2015.7146897.

55. Rubel, A. Denoising efficiency analysis based on no-reference image quality assessment [Text] / A. Rubel, V. Lukin // 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET). – 2018. – P. 898–902. – doi: https://doi.org/10.1109/tcset.2018.8336341.

56. Rubel, O. Prediction of Lee filter performance for Sentinel-1 SAR images [Text] / O. Rubel, V. Lukin, A. Rubel, K. Egiazarian // Electronic Imaging. – 2020. – Vol. 32, No. 9. – P. 371 – 1 – 371 – 7. – doi: https://doi.org/10.2352/issn.2470-1173.2020.9.iqsp-371.

57. Krivenko, S. Prediction of DCT-based denoising efficiency for images corrupted by signal-dependent noise [Text] / S. Krivenko, V. Lukin, B. Vozel, K. Chehdi // 2014 IEEE 34th International Scientific Conference on Electronics and Nanotechnology (ELNANO). – 2014. – P. 254–258. – doi: https://doi.org/10.1109/elnano.2014.6873920.

58. Lukin, V. Image filtering: Potential efficiency and current problems [Text] / V. Lukin, S. Abramov, N. Ponomarenko, K. Egiazarian, J. Astola // 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2011. – P. 1433–1436. – doi: https://doi.org/10.1109/icassp.2011.5946683.

59. Lukin, V. Denoising efficiency for multichannel images corrupted by signal-dependent noise [Text] / V. Lukin, S. Abramov, R. Kozhemiakin, M. Uss, B. Vozel, K. Chehdi // 2013 International Kharkov Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves. – 2013. – P. 340–342. – doi: https://doi.org/10.1109/msmw.2013.6622048.

60. Deng, J. ImageNet: A large-scale hierarchical image database [Text] / J. Deng, W. Dong, R. Socher, L. Li, Kai Li, Li Fei-Fei // 2009 IEEE Conference on Computer Vision and Pattern Recognition. – 2009. – P. 248–255. – doi: https://doi.org/10.1109/cvpr.2009.5206848.

61. Russakovsky, O. ImageNet Large Scale Visual Recognition Challenge [Text] / O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei // International Journal of Computer Vision. – 2015. – Vol. 115, No. 3. – P. 211–252. – doi: https://doi.org/10.1007/s11263-015-0816-y.

62. He, K. Deep Residual Learning for Image Recognition [Text] / K. He, X. Zhang, S. Ren, J. Sun // 2016 IEEE Conference on Computer Vision and Pattern

Recognition (CVPR). – 2016. – P. 770–778. – doi: https://doi.org/10.1109/cvpr.2016.90.

63. Zhang, K. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising [Text] / K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang // IEEE Transactions on Image Processing. – 2017. – Vol. 26, No. 7. – P. 3142–3155. – doi: https://doi.org/10.1109/tip.2017.2662206.

64. Cozzolino, D. Nonlocal CNN SAR Image Despeckling [Text] / D. Cozzolino, L. Verdoliva, G. Scarpa, G. Poggi // Remote Sensing. – 2020. – Vol. 12, No. 6. – P. 1006. – doi: https://doi.org/10.3390/rs12061006.

65. Molini, A. B. Speckle2Void: Deep Self-Supervised SAR Despeckling With Blind-Spot Convolutional Neural Networks [Text] / A. B. Molini, D. Valsesia, G. Fracastoro, E. Magli // IEEE Transactions on Geoscience and Remote Sensing. – 2022. – Vol. 60. – P. 1–17. – doi: https://doi.org/10.1109/tgrs.2021.3065461.

66. Chierchia, G. SAR image despeckling through convolutional neural networks [Text] / G. Chierchia, D. Cozzolino, G. Poggi, L. Verdoliva // 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). – 2017. – P. 5438–5441. – doi: https://doi.org/10.1109/igarss.2017.8128234.

67. Wang, P. SAR Image Despeckling Using a Convolutional Neural Network [Text] / P. Wang, H. Zhang, V. M. Patel // IEEE Signal Processing Letters. – 2017. – Vol. 24, No. 12. – P. 1763–1767. – doi: https://doi.org/10.1109/lsp.2017.2758203.

68. Dabov, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering [Text] / K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian // IEEE Transactions on Image Processing. – 2007. – Vol. 16, No. 8. – P. 2080–2095. – doi: https://doi.org/10.1109/tip.2007.901238.

69. Cozzolino, D. Nonlocal Sar Image Despeckling by Convolutional Neural Networks [Text] / D. Cozzolino, L. Verdoliva, G. Scarpa, G. Poggi // IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium. – 2019. – doi: https://doi.org/10.1109/igarss.2019.8897761.

70. Lattari, F. Deep Learning for SAR Image Despeckling [Text] / F. Lattari, B. Gonzalez Leon, F. Asaro, A. Rucci, C. Prati, M. Matteucci // Remote Sensing. – 2019. – Vol. 11, No. 13. – P. 1532. – doi: https://doi.org/10.3390/rs11131532.

71. Liu, G. Contourlet-CNN for SAR Image Despeckling [Text] / G. Liu, H. Kang, Q. Wang, Y. Tian, B. Wan // Remote Sensing. – 2021. – Vol. 13, No. 4. – P. 764. – doi: https://doi.org/10.3390/rs13040764.

72. Pongrac, B. Despeckling of SAR Images Using Residual Twin CNN and Multi-Resolution Attention Mechanism [Text] / B. Pongrac, D. Gleich // Remote Sensing. – 2023. – Vol. 15, No. 14. – P. 3698. – doi: https://doi.org/10.3390/rs15143698.

73. Wang, C. SAR Image Despeckling Based on Block-Matching and Noise-Referenced Deep Learning Method [Text] / C. Wang, Z. Yin, X. Ma, Z. Yang // Remote Sensing. – 2022. – Vol. 14, No. 4. – P. 931. – doi: https://doi.org/10.3390/rs14040931.

74. Yuan, Y. An Advanced SAR Image Despeckling Method by Bernoulli-Sampling-Based Self-Supervised Deep Learning [Text] / Y. Yuan, Y. Wu, Y. Fu, Y. Wu, L. Zhang, Y. Jiang // Remote Sensing. – 2021. – Vol. 13, No. 18. – P. 3636. – doi: https://doi.org/10.3390/rs13183636.

75. Pan, Y. SAR Image Despeckling Based on Denoising Diffusion Probabilistic Model and Swin Transformer [Text] / Y. Pan, L. Zhong, J. Chen, H. Li, X. Zhang, B. Pan // Remote Sensing. – 2024. – Vol. 16, No. 17. – P. 3222. – doi: https://doi.org/10.3390/rs16173222.

76. Denis, L. A Review of Deep-Learning Techniques for SAR Image Restoration [Text] / L. Denis, E. Dalsasso, F. Tupin // 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. – 2021. – P. 411–414. – doi: https://doi.org/10.1109/igarss47720.2021.9555039.

77. Peter, E. Recent Advances in Deep Learning for SAR Images: Overview of Methods, Challenges, and Future Directions [Text] / E. Peter, L. Ang, K. P. Seng, S. Srivastava // Sensors. – 2026. – Vol. 26, No. 4. – P. 1143. – doi: https://doi.org/10.3390/s26041143.

78. Vitale, S. Multi-Objective CNN-Based Algorithm for SAR Despeckling [Text] / S. Vitale, G. Ferraioli, V. Pascazio // IEEE Transactions on Geoscience and Remote Sensing. – 2021. – Vol. 59, No. 11. – P. 9336–9349. – doi: https://doi.org/10.1109/tgrs.2020.3034852.

79. Zhu, Y. Synthetic aperture radar image despeckling neural network based on maximum a posteriori probability estimation [Text] / Y. Zhu, M. Chen, X. Wang, B. Lin, H. Huang // International Journal of Remote Sensing. – 2023. – Vol. 44, No. 2. – P. 609–630. – doi: https://doi.org/10.1080/01431161.2023.2169594.

80. Sheikh, H. Image information and visual quality [Text] / H. Sheikh, A. Bovik // IEEE Transactions on Image Processing. – 2006. – Vol. 15, No. 2. – P. 430–444. – doi: https://doi.org/10.1109/tip.2005.859378.

81. Wang, Z. Image quality assessment: from error visibility to structural similarity [Text] / Zhou Wang, A. Bovik, H. Sheikh, E. Simoncelli // IEEE Transactions on Image Processing. – 2004. – Vol. 13, No. 4. – P. 600–612. – doi: https://doi.org/10.1109/tip.2003.819861.

82. Dougherty, G. Digital Image Processing for Medical Applications [Text] / G. Dougherty // Cambridge University Press. – 2009. – doi: https://doi.org/10.1017/cbo9780511609657.

83. Plataniotis, K. N. Color Image Processing and Applications [Text] / K. N. Plataniotis, A. N. Venetsanopoulos // Berlin, Heidelberg : Springer-Verlag Berlin Heidelberg. – 2000. – P. 355. – doi: https://doi.org/10.1007/978-3-662-04186-4. – ISBN: 978-3-540-66953-1.

84. Lin, W. Perceptual visual quality metrics: A survey [Text] / W. Lin, C. Jay Kuo // Journal of Visual Communication and Image Representation. – 2011. – Vol. 22, No. 4. – P. 297–312. – doi: https://doi.org/10.1016/j.jvcir.2011.01.005.

85. Wang, Z. Multiscale structural similarity for image quality assessment [Text] / Z. Wang, E. P. Simoncelli, A. C. Bovik // Conference Record of the 37th Asilomar Conference on Signals, Systems and Computers. – 2003. – P. 1398–1402. – doi: https://doi.org/10.1109/acssc.2003.1292216.

86. Zhang, L. FSIM: A Feature Similarity Index for Image Quality Assessment [Text] / Lin Zhang, Lei Zhang, Xuanqin Mou, D. Zhang // IEEE Transactions on Image Processing. – 2011. – Vol. 20, No. 8. – P. 2378–2386. – doi: https://doi.org/10.1109/tip.2011.2109730.

87. Reisenhofer, R. A Haar wavelet-based perceptual similarity index for image quality assessment [Text] / R. Reisenhofer, S. Bosse, G. Kutyniok, T. Wiegand // Signal Processing: Image Communication. – 2018. – Vol. 61. – P. 33–43. – doi: https://doi.org/10.1016/j.image.2017.11.001.

88. Ziaei Nafchi, H. Mean Deviation Similarity Index: Efficient and Reliable Full-Reference Image Quality Evaluator [Text] / H. Ziaei Nafchi, A. Shahkolaei, R. Hedjam, M. Cheriet // IEEE Access. – 2016. – Vol. 4. – P. 5579–5590. – doi: https://doi.org/10.1109/access.2016.2604042.

89. Ieremeiev, O. Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images [Text] / O. Ieremeiev, V. Lukin, K. Okarma, K. Egiazarian // Remote Sensing. – 2020. – Vol. 12, No. 15. – P. 2349. – doi: https://doi.org/10.3390/rs12152349.

90. Ponomarenko, N. Analysis of HVS-Metrics' Properties Using Color Image Database TID2013 [Text] / N. Ponomarenko, V. Lukin, J. Astola, K. Egiazarian // Lecture Notes in Computer Science. – 2015. – P. 613–624. – doi: https://doi.org/10.1007/978-3-319-25903-1_53.

91. Dellepiane, S. G. Quality Assessment of Despeckled SAR Images [Text] / S. G. Dellepiane, E. Angiati // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2014. – Vol. 7, No. 2. – P. 691–707. – doi: https://doi.org/10.1109/jstars.2013.2279501.

92. Egiazarian, K. Statistical Evaluation of Visual Quality Metrics for Image Denoising [Text] / K. Egiazarian, M. Ponomarenko, V. Lukin, O. Ieremeiev // 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2018. – P. 6752–6756. – doi: https://doi.org/10.1109/icassp.2018.8462294.

93. Ponomarenko, N. HVS-metric-based performance analysis of image denoising algorithms [Text] / N. Ponomarenko, V. Lukin, K. Egiazarian // 3rd European Workshop on Visual Information Processing. – 2011. – P. 156–161. – doi: https://doi.org/10.1109/euvip.2011.6045554.

94. Abramov, S. K. DENOISING OF MULTICHANNEL IMAGES WITH NONLINEAR TRANSFORMATION OF REFERENCE IMAGE [Text] / S. K. Abramov, V. V. Abramova, V. V. Lukin, K. O. Egiazarian // Telecommunications and Radio Engineering. – 2018. – Vol. 77, No. 9. – P. 769–786. – doi: https://doi.org/10.1615/telecomradeng.v77.i9.30.

95. Ponomarenko, N. DCT Based High Quality Image Compression [Text] / N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola // Lecture Notes in Computer Science. – 2005. – P. 1177–1185. – doi: https://doi.org/10.1007/11499145_119.

96. Ponomarenko, N. Lossy Compression of Images with Additive Noise [Text] / N. Ponomarenko, V. Lukin, M. Zriakhov, K. Egiazarian, J. Astola // Lecture Notes in Computer Science. – 2005. – P. 381–386. – doi: https://doi.org/10.1007/11558484_48.

97. Ponomarenko, N. Lossy Compression of Noisy Images Based on Visual Quality: A Comprehensive Study [Text] / N. Ponomarenko, S. Krivenko, V. Lukin, K. Egiazarian, J. T. Astola // EURASIP Journal on Advances in Signal Processing. – 2010. – Vol. 2010, No. 1. – doi: https://doi.org/10.1155/2010/976436.

98. Zemliachenko, A. N. Lossy compression of noisy remote sensing images with prediction of optimal operation point existence and parameters [Text] / A. N. Zemliachenko, S. K. Abramov, V. V. Lukin, B. Vozel, K. Chehdi // Journal of Applied Remote Sensing. – 2015. – Vol. 9, No. 1. – P. 095066. – doi: https://doi.org/10.1117/1.jrs.9.095066.

99. Ponomarenko, N. N. LOSSY AND VISUALLY LOSSLESS COMPRESSION OF SINGLE-LOOK SAR IMAGES [Text] / N. N. Ponomarenko, V. V. Lukin, R. Kozhemiakin, K. O. Egiazarian, M. Chobanu // Telecommunications and Radio Engineering. – 2013. – Vol. 72, No. 8. – P. 711–729. – doi: https://doi.org/10.1615/telecomradeng.v72.i8.60.

100. Lukin, V. V. Lossy compression of images without visible distortions and its application [Text] / V. V. Lukin, M. S. Zriakhov, N. N. Ponomarenko, S. S. Krivenko, M. Zhenjiang // IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS. – 2010. – P. 698–701. – doi: https://doi.org/10.1109/icosp.2010.5655751.

101. Li, F. A Two-step Procedure for Image Lossy Compression by ADCTC With a Desired Quality [Text] / F. Li, S. Krivenko, V. Lukin // 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT). – 2020. – P. 307–312. – doi: https://doi.org/10.1109/dessert50317.2020.9125000.

102. Li, F. A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression [Text] / F. Li, S. Krivenko, V. Lukin // 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). – 2020. – P. 502–506. – doi: https://doi.org/10.1109/tcset49122.2020.235483.

103. Kozhemiakina, N. Means and results of efficiency analysis for data compression methods applied to typical multimedia data [Text] / N. Kozhemiakina, N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola // 2014 First International Scientific-Practical Conference Problems of Infocommunications Science and Technology. – 2014. – P. 12–14. – doi: https://doi.org/10.1109/infocommst.2014.6992281.

104. Kozhemiakin, R. Image quality prediction for DCT-based compression [Text] / R. Kozhemiakin, V. Lukin, B. Vozel // 2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). – 2017. – P. 225–228. – doi: https://doi.org/10.1109/cadsm.2017.7916121.

105. Krivenko, S. S. SMART LOSSY COMPRESSION OF IMAGES BASED ON DISTORTION PREDICTION [Text] / S. S. Krivenko, O. Krylova, E. Bataeva, V. V. Lukin // Telecommunications and Radio Engineering. – 2018. – Vol. 77, No. 17. – P. 1535–1554. – doi: https://doi.org/10.1615/telecomradeng.v77.i17.40.

106. Krivenko, S. Prediction of Visual Quality Metrics in Lossy Image Compression [Text] / S. Krivenko, F. Li, V. Lukin, B. Vozel, O. Krylova // 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO). – 2020. – P. 478–483. – doi: https://doi.org/10.1109/elnano50318.2020.9088819.

107. KRYVENKO, S. PECULIARITIES OF NOISY IMAGE LOSSY COMPRESSION [Text] / S. KRYVENKO, V. LUKIN // Herald of Khmelnytskyi National University. Technical sciences. – 2024. – P. 278–283. – doi: https://doi.org/10.31891/2307-5732-2024-333-2-44.

108. Kovalenko, B. Prediction of Parameters in Optimal Operation Point for BPG-based Lossy Compression of Noisy Images [Text] / B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, B. Vozel // Ukrainian journal of remote sensing. – 2022. – Vol. 9, No. 2. – P. 4–12. – doi: https://doi.org/10.36023/ujrs.2022.9.2.212.

109. Kovalenko, B. Analysis of the potential efficiency of post-filtering noisy images after lossy compression [Text] / B. Kovalenko, V. Rebrov, V. Lukin // Ukrainian journal of remote sensing. – 2023. – Vol. 10, No. 1. – P. 11–16. – doi: https://doi.org/10.36023/ujrs.2023.10.1.231.

110. Lukin, V. Prediction of Optimal Operation Point Existence and Its Parameters in BPG-Based Automatic Lossy Compression of Noisy Images [Text] / V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, B. Vozel // Current Overview on Science and Technology Research Vol. 9. – 2022. – P. 1–36. – doi: https://doi.org/10.9734/bpi/costr/v9/4316a.

111. Lukin, V. Compression of Noisy Images Taking into Account Visual Quality: A Comprehensive Study [Text] / V. Lukin, B. Kovalenko, S. Kryvenko, N. Ponomarenko, K. Egiazarian, J. Astola // Research Developments in Science and Technology Vol. 8. – 2022. – P. 89–109. – doi: https://doi.org/10.9734/bpi/rdst/v8/2722b.

112. Abramova, V. A Fast and Accurate Prediction of Distortions in DCT-Based Lossy Image Compression [Text] / V. Abramova, V. Lukin, S. Abramov, S. Kryvenko,

P. Lech, K. Okarma // Electronics. – 2023. – Vol. 12, No. 11. – P. 2347. – doi: https://doi.org/10.3390/electronics12112347.

113. Al-Shaykh, O. Lossy compression of noisy images [Text] / O. Al-Shaykh, R. Mersereau // IEEE Transactions on Image Processing. – 1998. – Vol. 7, No. 12. – P. 1641–1652. – doi: https://doi.org/10.1109/83.730376.

114. Shapiro, J. Embedded image coding using zerotrees of wavelet coefficients [Text] / J. Shapiro // IEEE Transactions on Signal Processing. – 1993. – Vol. 41, No. 12. – P. 3445–3462. – doi: https://doi.org/10.1109/78.258085.

115. Taubman, D. High performance scalable image compression with EBCOT [Text] / D. Taubman // IEEE Transactions on Image Processing. – 2000. – Vol. 9, No. 7. – P. 1158–1170. – doi: https://doi.org/10.1109/83.847830.

116. Christophe, E. Hyperspectral Image Compression: Adapting SPIHT and EZW to Anisotropic 3-D Wavelet Coding [Text] / E. Christophe, C. Mailhes, P. Duhamel // IEEE Transactions on Image Processing. – 2008. – Vol. 17, No. 12. – P. 2334–2346. – doi: https://doi.org/10.1109/tip.2008.2005824.

117. Magli, E. Multiband Lossless Compression of Hyperspectral Images [Text] / E. Magli // IEEE Transactions on Geoscience and Remote Sensing. – 2009. – Vol. 47, No. 4. – P. 1168–1178. – doi: https://doi.org/10.1109/tgrs.2008.2009316.

118. Wei, Z. Spatio-Temporal Just Noticeable Distortion Profile for Grey Scale Image/Video in DCT Domain [Text] / Zhenyu Wei, K. Ngan // IEEE Transactions on Circuits and Systems for Video Technology. – 2009. – Vol. 19, No. 3. – P. 337–346. – doi: https://doi.org/10.1109/tcsvt.2009.2013518.

119. Zabala, A. Impact of lossy compression on mapping crop areas from remote sensing [Text] / A. Zabala, X. Pons // International Journal of Remote Sensing. – 2013. – Vol. 34, No. 8. – P. 2796–2813. – doi: https://doi.org/10.1080/01431161.2012.750772.

120. Kovalenko, B. BPG-Based Lossy Compression of Three-Channel Noisy Images with Prediction of Optimal Operation Existence and Its Parameters [Text] / B.

Kovalenko, V. Lukin, B. Vozel // Remote Sensing. – 2023. – Vol. 15, No. 6. – P. 1669. – doi: https://doi.org/10.3390/rs15061669.

121. Lainema, J. HEVC still image coding and high efficiency image file format [Text] / J. Lainema, M. M. Hannuksela, V. K. M. Vadakital, E. B. Aksu // 2016 IEEE International Conference on Image Processing (ICIP). – 2016. – P. 71–75. – doi: https://doi.org/10.1109/icip.2016.7532321.

122. Zhang, J. Overlapped variable size block motion compensation [Text] / Jiajun Zhang, M. Ahmad, M. Swamy // Proceedings of International Conference on Image Processing. – 2002. – P. 642–645. – doi: https://doi.org/10.1109/icip.1997.632203.

123. Wei, K. Research on Convolutional Neural Network-Based Compression Methods for Multispectral Images [Text] / K. Wei, H. Liu, P. Yu, Y. Ren, G. M. Gao, C. Wen // Journal of Physics: Conference Series. – 2024. – Vol. 2717, No. 1. – P. 012006. – doi: https://doi.org/10.1088/1742-6596/2717/1/012006.

124. Foi, A. Practical Poissonian-Gaussian Noise Modeling and Fitting for Single-Image Raw-Data [Text] / A. Foi, M. Trimeche, V. Katkovnik, K. Egiazarian // IEEE Transactions on Image Processing. – 2008. – Vol. 17, No. 10. – P. 1737–1754. – doi: https://doi.org/10.1109/tip.2008.2001399.

125. Colom, M. Analysis and Extension of the PCA Method, Estimating a Noise Curve from a Single Image [Text] / M. Colom, A. Buades // Image Processing On Line. – 2016. – Vol. 6. – P. 365–390. – doi: https://doi.org/10.5201/ipol.2016.124.

126. Rakhshanfar, M. Estimation of Gaussian, Poissonian-Gaussian, and Processed Visual Noise and its Level Function [Text] / M. Rakhshanfar, M. A. Amer // IEEE Transactions on Image Processing. – 2016. – P. 1–1. – doi: https://doi.org/10.1109/tip.2016.2588320.

127. Azzari, L. Modeling and Estimation of Signal-Dependent and Correlated Noise [Text] / L. Azzari, L. R. Borges, A. Foi // Advances in Computer Vision and Pattern Recognition. – 2018. – P. 1–36. – doi: https://doi.org/10.1007/978-3-319-96029-6_1.

128. Azzari, L. Gaussian-Cauchy mixture modeling for robust signal-dependent noise estimation [Text] / L. Azzari, A. Foi // 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2014. – P. 5357–5361. – doi: https://doi.org/10.1109/icassp.2014.6854626.

129. Abramova, V. Analysis of Noise Properties in Dental Images [Text] / V. Abramova, S. Krivenko, V. Lukin, O. Krylova // 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO). – 2020. – P. 511–515. – doi: https://doi.org/10.1109/elnano50318.2020.9088768.

130. Ponomarenko, M. Blind estimation of white Gaussian noise variance in highly textured images [Text] / M. Ponomarenko, N. Gapon, V. Voronin, K. Egiazarian // Electronic Imaging. – 2018. – Vol. 30, No. 13. – P. 382 – 1 – 382 – 5. – doi: https://doi.org/10.2352/issn.2470-1173.2018.13.ipas-382.

131. Vozel, B. Noise Identification and Estimation of its Statistical Parameters by Using Unsupervised Variational Classification [Text] / B. Vozel, K. Chehdi, L. Klaine, V. Lukin, S. Abramov // 2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings. – 2006. – P. II – 841 – II – 844. – doi: https://doi.org/10.1109/icassp.2006.1660474.

132. Vozel, B. Using interchannel correlation in blind evaluation of noise characteristics in multichannel remote sensing images [Text] / B. Vozel, S. K. Abramov, V. V. Abramova, V. V. Lukin, K. Chehdi // Image and Signal Processing for Remote Sensing XXIV. – 2018. – P. 8. – doi: https://doi.org/10.1117/12.2325100.

133. Zhong, P. Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery [Text] / P. Zhong, R. Wang // IEEE Transactions on Geoscience and Remote Sensing. – 2013. – Vol. 51, No. 4. – P. 2260–2275. – doi: https://doi.org/10.1109/tgrs.2012.2209656.

134. Pan, S. J. A Survey on Transfer Learning [Text] / S. J. Pan, Q. Yang // IEEE Transactions on Knowledge and Data Engineering. – 2010. – Vol. 22, No. 10. – P. 1345–1359. – doi: https://doi.org/10.1109/tkde.2009.191.

135. Huang, Z. What, Where, and How to Transfer in SAR Target Recognition Based on Deep CNNs [Text] / Z. Huang, Z. Pan, B. Lei // IEEE Transactions on Geoscience and Remote Sensing. – 2020. – Vol. 58, No. 4. – P. 2324–2336. – doi: https://doi.org/10.1109/tgrs.2019.2947634.

136. Yosinski, J. How transferable are features in deep neural networks? [Text] / J. Yosinski, J. Clune, Y. Bengio, H. Lipson // Advances in Neural Information Processing Systems. – 2014. – Vol. 27. – P. 3320–3328. – doi: https://doi.org/10.48550/arxiv.1411.1792.

137. Wang, K. SAR Target Recognition Based on Cross-Domain and Cross-Task Transfer Learning [Text] / K. Wang, G. Zhang, H. Leung // IEEE Access. – 2019. – Vol. 7. – P. 153391–153399. – doi: https://doi.org/10.1109/access.2019.2948618.

138. Inglada, J. Similarity measures for multisensor remote sensing images [Text] / J. Inglada // IEEE International Geoscience and Remote Sensing Symposium. – 2002. – P. 104–106. – doi: https://doi.org/10.1109/igarss.2002.1024955.

139. Lukin, V. Processing and Classification of Multichannel Remote Sensing Data [Text] / V. Lukin, N. Ponomarenko, A. Kurekin, O. Pogrebnyak // Lecture Notes in Computer Science. – 2011. – P. 487–498. – doi: https://doi.org/10.1007/978-3-642-25330-0_43.

140. Simonyan, K. Very deep convolutional networks for large-scale image recognition [Text] / K. Simonyan, A. Zisserman // Proceedings of the International Conference on Learning Representations (ICLR). – 2015. – URL: https://arxiv.org/abs/1409.1556.

141. Huang, G. Densely Connected Convolutional Networks [Text] / G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – P. 2261–2269. – doi: https://doi.org/10.1109/cvpr.2017.243.

142. Vitale, S. SAR Despeckling Using Multiobjective Neural Network Trained With Generic Statistical Samples [Text] / S. Vitale, G. Ferraioli, A. C. Frery, V.

Pascazio, D. Yue, F. Xu // IEEE Transactions on Geoscience and Remote Sensing. – 2023. – Vol. 61. – P. 1–12. – doi: https://doi.org/10.1109/tgrs.2023.3314857.

143. Mullissa, A. G. deSpeckNet: Generalizing Deep Learning-Based SAR Image Despeckling [Text] / A. G. Mullissa, D. Marcos, D. Tuia, M. Herold, J. Reiche // IEEE Transactions on Geoscience and Remote Sensing. – 2022. – Vol. 60. – P. 1–15. – doi: https://doi.org/10.1109/tgrs.2020.3042694.

144. Chen, S. Domain Adaptation for Semi-Supervised Ship Detection in SAR Images [Text] / S. Chen, R. Zhan, W. Wang, J. Zhang // IEEE Geoscience and Remote Sensing Letters. – 2022. – Vol. 19. – P. 1–5. – doi: https://doi.org/10.1109/lgrs.2022.3171789.

145. Chen, Y. A new deep learning-based model for reconstructing high-quality NDVI time-series data in heavily cloudy areas: fusion of Sentinel 1 and 2 data [Text] / Y. Chen, R. Cao, S. Liu, L. Peng, X. Chen, J. Chen // International Journal of Digital Earth. – 2024. – Vol. 17, No. 1. – doi: https://doi.org/10.1080/17538947.2024.2407941.

146. Woo, S. CBAM: Convolutional Block Attention Module [Text] / S. Woo, J. Park, J. Lee, I. S. Kweon // Proceedings of the European Conference on Computer Vision (ECCV). – 2018. – P. 3–19. – doi: https://doi.org/10.1007/978-3-030-01234-2_1.

147. Al-Senaikh, R. Predicting Optimal Lee Filter Window Size for Sentinel-1 SAR Images Using Transfer Learning on MobileNetV2 [Text] / R. Al-Senaikh, O. Rubel // Herald of Khmelnytskyi National University. Technical Sciences. – 2026. – Vol. 361, No. 1. – P. 26–36. – doi: https://doi.org/10.31891/2307-5732-2026-361-3.

148. Kingma, D. P. Adam: A Method for Stochastic Optimization [Text] / D. P. Kingma, J. Ba // 3rd International Conference on Learning Representations (ICLR). – 2015. – URL: https://arxiv.org/abs/1412.6980.

149. Al-Senaikh, R. Development of a hybrid method VGG16-FrostNet for adaptive despeckling of synthetic aperture radar (SAR) images using attention mechanism and differentiable Frost filter [Text] / R. Al-Senaikh, O. Rubel // Technology audit and production reserves. – 2026. – Vol. 2, No. 2(88). – P. 25–33. – doi: https://doi.org/10.15587/2706-5448.2026.358316.

150. Hore, A. Image Quality Metrics: PSNR vs. SSIM [Text] / A. Hore, D. Ziou // 2010 20th International Conference on Pattern Recognition. – 2010. – P. 2366–2369. – doi: https://doi.org/10.1109/icpr.2010.579.

151. Dalsasso, E. SAR2SAR: A Semi-Supervised Despeckling Algorithm for SAR Images [Text] / E. Dalsasso, L. Denis, F. Tupin // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2021. – Vol. 14. – P. 4321–4329. – doi: https://doi.org/10.1109/jstars.2021.3071864.

152. Guan, J. Robust SAR Image Despeckling by Deep Learning From Near-Real Datasets [Text] / J. Guan, R. Liu, X. Tian, X. Tang, S. Li // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. – 2024. – Vol. 17. – P. 2963–2979. – doi: https://doi.org/10.1109/jstars.2023.3345538.

153. Fang, Y. Contrastive learning for real SAR image despeckling [Text] / Y. Fang, R. Liu, Y. Peng, J. Guan, D. Li, X. Tian // ISPRS Journal of Photogrammetry and Remote Sensing. – 2024. – Vol. 218. – P. 376–391. – doi: https://doi.org/10.1016/j.isprsjprs.2024.11.003.

154. Zhang, X. Cross-Sensor SAR Image Target Detection Based on Dynamic Feature Discrimination and Center-Aware Calibration [Text] / X. Zhang, S. Zhang, Z. Sun, C. Liu, Y. Sun, K. Ji, G. Kuang // IEEE Transactions on Geoscience and Remote Sensing. – 2025. – Vol. 63. – P. 1–17. – doi: https://doi.org/10.1109/tgrs.2025.3559618.

155. Vadnais, J. Transfer Learning Between Sentinel-1 Acquisition Modes Enhances the Few-Shot Segmentation of Natural Oil Slicks in the Arctic [Text] / J. Vadnais, B. A. Robson, C. H. Eide, R. Mattingsdal, M. Johansson // IEEE Geoscience and Remote Sensing Letters. – 2025. – Vol. 22. – P. 1–5. – doi: https://doi.org/10.1109/lgrs.2025.3579308.

156. Geng, J. Transfer Learning for SAR Image Classification Via Deep Joint Distribution Adaptation Networks [Text] / J. Geng, X. Deng, X. Ma, W. Jiang // IEEE Transactions on Geoscience and Remote Sensing. – 2020. – Vol. 58, No. 8. – P. 5377–5392. – doi: https://doi.org/10.1109/tgrs.2020.2964679.

157. Hafner, S. Unsupervised domain adaptation for global urban extraction using Sentinel-1 SAR and Sentinel-2 MSI data [Text] / S. Hafner, Y. Ban, A. Nascetti // Remote Sensing of Environment. – 2022. – Vol. 280. – P. 113192. – doi: https://doi.org/10.1016/j.rse.2022.113192.

158. Gencay, S. The Effect of SAR Speckle Removal in SAR-Optical Image Fusion [Text] / S. Gencay, C. Ozcan // 2022 30th Signal Processing and Communications Applications Conference (SIU). – 2022. – P. 1–4. – doi: https://doi.org/10.1109/siu55565.2022.9864861.

159. Chen, Y. Self-Supervised SAR-Optical Data Fusion of Sentinel-1/-2 Images [Text] / Y. Chen, L. Bruzzone // IEEE Transactions on Geoscience and Remote Sensing. – 2022. – Vol. 60. – P. 1–11. – doi: https://doi.org/10.1109/tgrs.2021.3128072.

160. Hafner, S. Sentinel-1 and Sentinel-2 Data Fusion for Urban Change Detection Using a Dual Stream U-Net [Text] / S. Hafner, A. Nascetti, H. Azizpour, Y. Ban // IEEE Geoscience and Remote Sensing Letters. – 2022. – Vol. 19. – P. 1–5. – doi: https://doi.org/10.1109/lgrs.2021.3119856.

161. Gaya, V. Self-Supervised Learning of Multi-Modal Cooperation for SAR Despeckling [Text] / V. Gaya, E. Dalsasso, L. Denis, F. Tupin, B. Pinel-Puysségur, C. Guérin // IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium. – 2024. – P. 2180–2183. – doi: https://doi.org/10.1109/igarss53475.2024.10642047.

162. Guo, C. Deep Learning Methods for SAR and Optical Image Fusion: A Review [Text] / C. Guo, Z. Zhang, K. Huang, L. Luo, Z. Yang, S. Shi, J. Shi // Remote Sensing. – 2026. – Vol. 18, No. 8. – P. 1196. – doi: https://doi.org/10.3390/rs18081196.

163. Hughes, L. H. A deep learning framework for matching of SAR and optical imagery [Text] / L. H. Hughes, D. Marcos, S. Lobry, D. Tuia, M. Schmitt // ISPRS Journal of Photogrammetry and Remote Sensing. – 2020. – Vol. 169. – P. 166–179. – doi: https://doi.org/10.1016/j.isprsjprs.2020.09.012.

ДОДАТКИ

ДОДАТОК А.
СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ ТА ВІДОМОСТІ ПРО АПРОБАЦІЮ

А.1 Статті у наукових фахових виданнях України та виданнях, віднесених до баз даних Scopus / Web of Science

1. Al-Senaikh R., Rubel O. Predicting Filtered Image Quality Using Transfer Learning on Sentinel-1 Speckle Noise with DenseNet-121 // Ukrainian Journal of Remote Sensing. — 2025. — Vol. 12, No. 4. — P. 4–15. — DOI: 10.36023/ujrs.2025.12.4.293 (Особистий внесок: формулювання задачі, реалізація моделі DenseNet-121, проведення експериментів, аналіз результатів, написання тексту статті.)

2. Al-Senaikh R., Rubel O. Predicting Optimal Lee Filter Window Size for Sentinel-1 SAR Images Using Transfer Learning on MobileNetV2 // Вісник Хмельницького національного університету. Серія: Технічні науки. — 2026. — Т. 361, № 1. — С. 26–36. — DOI: 10.31891/2307-5732-2026-361-3 (Особистий внесок: розробка методу прямої багатокласової класифікації оптимального розміру вікна фільтра Лі, навчання та оцінювання MobileNetV2, аналіз матриці помилок.)

3. Al-Senaikh R., Rubel O. Development of a hybrid method VGG16-FrostNet for adaptive despeckling of synthetic aperture radar (SAR) images using attention mechanism and differentiable Frost filter // Technology audit and production reserves. — 2026. — Т. 2, № 2(88). — С. 25–33. — DOI: 10.15587/2706-5448.2026.358316 (Особистий внесок: формулювання диференційованої моделі фільтра Фроста, проектування гібридної архітектури VGG16-FrostNet з модулем CBAM, реалізація наскрізного навчання, експериментальна оцінка.)

А.2 Праці апробаційного характеру (матеріали наукових конференцій)

4. Аль-Сенайх Р. Ж., Рубель О. С. Мінімізація ризиків неправильного встановлення параметрів фільтрів для покращення якості РСА-зображень //

Information, Functional and Cybersecurity: матеріали міжнар. студ. конф. з кібербезпеки (SCIFiC 2025) (м. Харків, 27–28 листопада 2025 р.). — Харків: НАУ «ХАІ», 2025. — С. 124–125. — DOI: 10.13140/RG.2.2.32148.16001 (Форма участі: очна доповідь. Особистий внесок: постановка задачі мінімізації ризиків, презентація методу класифікації параметрів фільтра Лі.)

5. Al-Senaikh R. Z., Rubel O. S. A hybrid approach to adaptive SAR despeckling using VGG16-FrostNet and a HaarPSI-based Frost baseline // Organization of scientific research in modern conditions '2026: SWorld-US Conference proceedings (Seattle, WA, USA, March 21, 2026). — Seattle: ProConference, 2026. — Vol. 1, Issue usc36-00. — P. 44–51. — DOI: 10.30888/2709-2267.2026-36-00-029 (Форма участі: заочна доповідь. Особистий внесок: опис гібридного методу VGG16-FrostNet та порівняння з оптимізованим за HaarPSI еталоном фільтра Фроста.)

6. Аль-Сенайх Р. Ж., Рубель О. С. Порівняльний аналіз парадигм інтеграції класичних адаптивних фільтрів у нейромережеві архітектури для деспеклізації радіолокаційних зображень // Перспективні мережні та комп'ютерні технології (ПерСиК 2026): матеріали 17-ї міжнар. студ. наук.-техн. конф. (23 квітня 2026 р.) (Форма участі: очна доповідь. Особистий внесок: систематизація п'яти парадигм гібридизації, побудова матриці порівняння.)

А.3 Відповідність публікацій розділам дисертації

Публікація (№)	Розділ дисертації
1 (DenseNet-121)	Розділ 2 (п. 2.1–2.3, 2.5)
2 (MobileNetV2)	Розділ 2 (п. 2.4–2.6), Розділ 4 (п. 4.4)
3 (VGG16-FrostNet)	Розділ 3 (повністю), Розділ 4 (п. 4.3)
4 (SCIFiC 2025)	Розділ 4 (п. 4.4)
5 (SWorld 2026)	Розділ 3 (п. 3.2), Розділ 4 (п. 4.3)
6 (ПерСиК 2026)	Розділ 4 (п. 4.2)

ДОДАТОК Б.
АКТИ ВПРОВАДЖЕННЯ

pproduct

275 Grove Street, Suite 2-400
Newton, Massachusetts 02466, USA
p-product.com

Reference #2026/2/2-003
February 2, 2026

TO WHOM IT MAY CONCERN

P-Product, Inc. hereby confirms that the scientific results presented in the dissertation research of Mr. Raed Al-Senaykh entitled: "Methods of Adaptive Speckle Noise Filtering and A Priori Quality Estimation of Radar Images Based on Neural Networks" have been reviewed and implemented within the Company's research and software development activities related to image processing, computer vision, machine learning, and intelligent image enhancement solutions.

In particular, the proposed methods for a priori quality prediction of filtered Synthetic Aperture Radar (SAR) images based on transfer learning (DenseNet-121), prediction of the optimal Lee filter window size using MobileNetV2, and the hybrid VGG16-FrostNet approach for spatially adaptive despeckling demonstrated practical applicability in image preprocessing and image quality assessment tasks. The research results contributed to:

- Reducing research and experimentation time through a priori selection of adaptive filtering parameters;
- Improving the efficiency of image preprocessing workflows for noisy and radar image data;
- Supporting the combination of classical adaptive filtering techniques with neural network-based models for image enhancement;
- Facilitating the development of software prototypes and methodological materials for evaluating image preprocessing algorithms in research and development projects.

P-Product, Inc. recognizes the scientific and practical value of the obtained results and confirms their relevance for modern image processing, computer vision, machine learning, and intelligent data analysis systems. Should you require any additional information, please feel free to contact us.

Best regards,
Nadya Mezenina

COO of P-Product, Inc.
e-mail: nadyam@p-product.com
mobile: +1 (339) 234-7006

«ЗАТВЕРДЖУЮ»
Виконавчий директор
P-Product, Inc. (США)
Мезеніна Н.С.

«02» лютого 2026 р.

АКТ

науково-технічної комісії щодо впровадження наукових положень і результатів дисертації **«МЕТОДИ АДАПТИВНОЇ ФІЛЬТРАЦІЇ СПЕКЛ-ШУМУ ТА АПРІОРНОГО ОЦІНЮВАННЯ ЯКОСТІ РАДІОЛОКАЦІЙНИХ ЗОБРАЖЕНЬ НА ОСНОВІ НЕЙРОННИХ МЕРЕЖ»**

на здобуття ступеня доктора філософії аспіранта
Національного аерокосмічного університету «ХАІ»
Аль-Сенайха Раєда Жихадовича

Науково-технічна комісія у складі: голова комісії – виконавчий директор Мезеніна Надія Сергіївна, члени комісії: головний інженер проекту Желанов Олексій Олександрович, менеджер проекту Дунаєвський Дмитро Ігорьович, склала даний акт про те, що в результаті спільних науково-дослідних робіт з попередньої обробки зашумлених зображень були використані та впроваджені наступні наукові положення та результати, що отримані аспірантом Аль-Сенайхом Р.Ж.:

- метод апріорного прогнозування якості відфільтрованих РСА-зображень на базі трансферного навчання (DenseNet-121);
- метод прогнозування оптимального розміру вікна фільтра Лі для деспеклізації РСА-зображень (MobileNetV2);
- гібридний метод VGG16-FrostNet для просторово-адаптивної деспеклізації з фільтром Фроста.

Розроблені в дисертації методи та програмно-алгоритмічні засоби дозволили:

1) скоротити час дослідницьких експериментів за рахунок апріорного вибору параметрів адаптивної фільтрації;

2) обґрунтувати поєднання класичних адаптивних фільтрів із нейромережевими модулями у програмних прототипах.

Отримані результати використано у спільних науково-дослідних роботах Компанії у вигляді програмних прототипів і методичних матеріалів для оцінки алгоритмів попередньої обробки зображень у проєктах компанії.

Голова комісії _____ (Мезеніна Н.С)

Члени комісії: _____ (Желанов О.О.)

_____ (Дунаєвський Д.І.)